Semi-quantitative characterisation of mixed pollen samples using MinION sequencing and Reverse Metagenomics (RevMet)

Ned Peel[1,2], Lynn V. Dicks[2], Darren Heavens[1], Lawrence Percival-Alwyn[1], Chris Cooper[3], Matthew D. Clark[1,4], Richard G. Davies[2], Richard M. Leggett[1], Douglas W. Yu[2,5,6,*]

[1] Earlham Institute, Norwich Research Park, Norwich, UK

[2] University of East Anglia, Norwich Research Park, Norwich, UK

[3] University of Cambridge, Cambridge, UK

[4] Natural History Museum, London, UK

[5] State Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, China

[6] Center for Excellence in Animal Evolution and Genetics, Chinese Academy of Sciences, Kunming, China

*Corresponding author: Douglas.Yu@uea.ac.uk

Running title: Characterisation of mixed pollen with RevMet

Key-words: bees, diet analysis, genome skim, metabarcoding, metagenomics, MinION, pollen, quantitative

Abstract

1. The ability to identify and quantify the constituent plant species that make up a mixed-species sample of pollen has important applications in ecology, conservation, and agriculture. Recently, metabarcoding protocols have been developed for pollen that can identify constituent plant species, but there are strong reasons to doubt that metabarcoding can accurately quantify their relative abundances. A PCR-free, shotgun metagenomics approach has greater potential for accurately quantifying species relative abundances, but applying metagenomics to eukaryotes is challenging due to low numbers of reference genomes.

2. We have developed a pipeline, RevMet (Reverse Metagenomics), that allows reliable and semi-quantitative characterization of the species composition of mixed-species eukaryote samples, such as bee-collected pollen, without requiring reference genomes. Instead, reference species are represented only by 'genome skims': low-cost, low-coverage, short-read sequence datasets. The skims are mapped to individual long reads sequenced from mixed-species samples using the MinION, a portable nanopore sequencing device, and each long read is uniquely assigned to a plant species.

3. We genome-skimmed 49 wild UK plant species, validated our pipeline with mock DNA mixtures of known composition, and then applied RevMet to pollen loads collected from wild bees. We demonstrate that RevMet can identify plant species present in mixed-species samples at proportions ≥1%, with few false positives and false negatives, and reliably differentiate species represented by high versus low amounts of DNA in a sample.

4. The RevMet pipeline could readily be adapted to generate semi-quantitative datasets for a wide range of mixed eukaryote samples, which could include characterising diets, quantifying allergenic pollen from air samples, quantifying soil fauna, and identifying the

compositions of algal and diatom communities. Our per-sample costs were £90 per genome skim and £60 per pollen sample, and new versions of sequencers available now will further reduce these costs.

Introduction

Pollination is a key ecosystem service; almost 90% of all flowering plant species, including 75% of food crops (mainly fruits, nuts, and vegetables), rely on animal pollination (Ollerton, Winfree, & Tarrant, 2011; Klein *et al.,* 2007). The benefits of pollinators, and pollinator-dependent plants, also include the production of medicines, biofuels, fibres, and construction materials (Potts *et al.,* 2016). There is growing concern over the decline of wild and domesticated pollinators and the resulting decrease in pollination services and crop production (Potts *et al.,* 2010; Burkle, Marlin, & Knight, 2013). These declines are thought to be caused by multiple threats acting together, including habitat loss, climate change, and the spread of diseases (Vanbergen *et al.,* 2013).

To mitigate drivers of pollinator decline, the Intergovernmental Science - Policy Platform for Biodiversity and Ecosystem Services (IPBES) has suggested three complementary strategies: (1) ecological intensification, which involves boosting agricultural production by increasing the provision of supporting ecological processes such as biotic pest regulation, nutrient cycling, and pollination (Bommarco, Kleijn, & Potts, 2013; Tittonell, 2014); (2) strengthening existing diversified farming systems, including gardens and agroforestry, for the generation of ecosystem functions; and (3) investment in ecological infrastructure, to protect, restore, and connect natural and semi-natural habitats across agricultural landscapes, so that pollinator species can more easily disperse and find nesting and floral resources (IPBES 2016).

However, knowledge gaps limit the effectiveness of these strategies (Wood, Holland, &

Goulson, 2015; Dicks *et al.,* 2013). For instance, it is still not clear which plant species are the most valuable food resources and how plant species vary in value across pollinator species, over time, and in different environmental conditions. It is also not well understood whether the addition of floral resources might draw pollinators away from pollinator-dependent crop plants (Morandin & Kremen, 2013), or whether floral enhancement will alter levels of plant-target specialism, at the levels of insect species and of individual insects, resulting in changes in pollination efficiency (Lucas *et al.*, 2018; Morales & Traveset, 2008).

Therefore, a crucial technical challenge for understanding plant-pollinator interactions is a method to identify *and* quantify the species of pollen that are consumed by pollinators. Identifying and quantifying pollen has traditionally been carried out by using light microscopy to distinguish plant species by grain morphology, a labour-intensive technique that requires expert knowledge and lacks discriminatory power at lower taxonomic levels (Long & Krupke, 2016; Khansari *et al*., 2012). In contrast, high-throughput DNA sequencing now allows pollen identification without expert knowledge of pollen morphology and taxonomy.

The currently dominant sequence-based method is metabarcoding, which involves amplifying taxonomically informative marker genes from mixed samples via polymerase chain reaction (PCR) (Ji *et al.*, 2013). The resulting amplification products, known as amplicons, are sequenced, and the reads are assigned to taxonomies by matching against barcode databases, such as the Barcode of Life Data System (Ratnasingham & Hebert, 2007). Notably for plants, there is no single barcode gene that matches the resolving power and universality of 16S rRNA for prokaryotes and Cytochrome Oxidase (CO1) for animals (Hollingsworth, Li, Van Der Bank, & Twyford, 2016). Instead, plant-related barcoding studies rely on a combination of marker genes, which include plastid regions *rbcL* and *matK* and the internal transcribed spacer (ITS) regions of nuclear ribosomal DNA (Li *et al.,* 2015; Hollingsworth *et al.,* 2016). Metabarcoding of mixed-species pollen samples can reveal the presence and absence of constituent plant species (or genera), but there are strong reasons

110 to doubt that metabarcoding can accurately quantify their relative abundances, due to PCR
111 amplification biases and varying copy numbers of barcode loci (Keller *et al.,* 2015;
112 Richardson *et al.,* 2015; Sickel *et al.,* 2015; Bell *et al.,* 2017).
113
114 In contrast to the targeted sequencing approach of metabarcoding, 'shotgun metagenomics'
115 involves randomly sequencing short stretches of genomic DNA from mixed samples. In
116 standard metagenomics, these short reads ('queries') are mapped to either assembled
117 genomes or to collections of barcode genes ('references'), which creates a requirement for
118 large numbers of reference genomes (Sharpton, 2014) or barcodes, with the latter being
119 very inefficient. Species identification is obtained by first calculating a similarity metric
120 between each short read and each reference sequence (e.g. % identity) and then using an
121 algorithm to assign each short read to the most likely reference sequence (Quince, Walker,
122 Simpson, Loman, & Segata, 2017). The potential key advantages of shotgun metagenomics
123 are that it can avoid the PCR-induced biases seen with metabarcoding, especially if PCR-
124 free library preparation protocols are used (see Nayfach & Pollard, 2016; Jones *et al.,* 2015)
125 and that by sampling across the whole genome, variation in the copy numbers of a few loci
126 is rendered less important. However, the requirement for reference genomes means that
127 most shotgun metagenomics studies focus on prokaryotic organisms, since large numbers
128 of prokaryote reference genomes are available. In contrast, eukaryotes are not well
129 represented in sequence databases and as a result have mostly been neglected in
130 metagenomic studies (Escobar-Zepeda, De León, & Sanchez-Flores, 2015). The low
131 numbers of reference genomes for eukaryotic species is because they are more expensive
132 to sequence and assemble (Gilbert & Dupont, 2011).
133
134 Here we demonstrate a metagenomic pipeline for eukaryotes that avoids the need to
135 assemble reference genomes. Instead, each reference species is represented by a 'genome
136 skim,' which is a low-cost, low-coverage, shotgun dataset, i.e. simply a set of short reads.
137 These sets of short reads are used to identify individual *long reads* from pollen that have

been generated by sequencing mixed-species pollen loads with the Oxford Nanopore Technologies' (ONT) MinION, a nanopore sequencing device (for a review of MinION applications and performance, see Leggett & Clark, 2017). Here, we generate reference genome skims for 49 wild UK plant species, and we use them to identify and quantify plant species in two kinds of query samples: mock, mixed-plant-species DNA mixtures of known composition and mixed-species pollen samples collected from wild bees. Since the long reads in each query sample are individually identified, we show that the frequency of long reads assigned to a plant species is a reasonably accurate estimate of that species' biomass frequency in a mixed-species sample. We call this pipeline Reverse Metagenomics, or RevMet, because we map reference sequences to query sequences, which is the reverse of the normal metagenomic protocol.

Methods

Sampling of bees and plant tissue

Sample collection took place in the Pensthorpe Natural Park area (52°49'23"N, 0°53'14"E) of Norfolk, UK, during June and July 2016. Leaf samples were collected from all plant species with open flowers, including grasses and trees, within a 100 m radius of the collection site (n = 49 species). Leaf tissue was preserved on dry ice in the field followed by storage at -80 °C. Foraging wild bees (n = 48: 9 *Apis mellifera*, 27 *Bombus terrestris/lucorum complex*, 12 *Bombus lapidarius*) were collected with hand nets or into falcon tubes directly from flowers and euthanized in falcon tubes containing ethanol-soaked tissue paper. Pollen loads were scraped from bee corbiculae using a mounted needle and stored in absolute ethanol. The plant species on which each bee was foraging when collected was recorded.

165

166 **Leaf tissue DNA extraction, library preparation, and Illumina sequencing**

167

168 Leaf tissue from each of the 49 plant species was disrupted by bead-beating using a 4-mm
169 stainless steel bead with a Qiagen TissueLyser II running at 22.5 Hz for 4 min, rotating the
170 adapter sets after 2 min. DNA was extracted using the DNeasy Plant Kit (Qiagen, Hilden,
171 Germany) following manufacturer's instructions. DNA concentrations were measured on a
172 Qubit 2.0 fluorometer (ThermoFisher, Waltham, USA) using the dsDNA HS assay kit, and
173 fragment size distribution was checked with a Genomic DNA Analysis ScreenTape on the
174 TapeStation 2200 (Agilent, Santa Clara, USA).

175

176 The Earlham Institute (Norwich, UK) applied a modified version of Illumina's Nextera
177 protocol, known as Low Input Transposase Enabled (LITE) protocol (Beier *et al.,* 2017), to
178 generate a separate sequencing library for each leaf sample, targeting an average insert
179 size of 500 bp. The LITE libraries were then pooled based on estimated genome sizes
180 (Supplementary Table S1), obtained from the Royal Botanic Gardens Kew Plant DNA C-
181 values database (Bennett and Leitch, 2012), in order to achieve 0.5x coverage of each
182 species genome. The pooled libraries were sequenced on one lane of Illumina HiSeq 2500
183 in Rapid Run mode (250 bp PE).

184

185 **Construction and sequencing of mock pollen samples**

186

187 DNA from twelve of the 49 plant species were used to construct six mock communities.
188 Each mock was made using 200 ng DNA in total, with species added at different
189 proportions: 0.08% to 45.25% (Table 1). For each mock, technical-replicate pairs were
190 prepared using ONT's (Oxford, UK) Rapid Barcoding Sequencing Kit (SQK-RBK001),
191 following the RBK_9031_v2_revl_09Mar2017 version of the manufacturer's protocol. The 12
192 libraries (six mocks, duplicated) were sequenced on a single MinION R9.5 flow cell (FLO-

193 MIN107).

194

195 **Bee-collected pollen DNA extraction, library preparation, and MinION sequencing**

196

197 After removing storage ethanol from the 48 bee-collected pollen loads, the pollen was
198 disrupted with ca. five 1-mm stainless steel beads for 2 min at 22.5 Hz using a Qiagen
199 TissueLyser II, rotating the adapter sets after 1 min. The pollen samples were resuspended
200 in 600 µl CTAB extraction buffer (2% CTAB, 1.4 M NaCl, 20 mM EDTA, pH 8.0, 100 mM
201 Tris-HCl pH 8.0), 0.5 µl of β-Mercaptoethanol, 4 µl of proteinase K, and vortexed for 5 s.
202 Following a 1 hr incubation at 55 °C, the tubes were centrifuged for 6 min at 18,000 x g. The
203 ≈500 µl of supernatant was extracted to a clean 1.5 ml tube before an equal volume of
204 chilled (2-8 °C) Phenol:Chloroform:Isoamyl Alcohol (25:24:1, v/v) was added to the lysate.
205 The samples were vortexed for 10 s (5 x 2 s bursts), centrifuged for 5 min at 14,000 x g, and
206 the upper aqueous phase (≈ 420 µl) was extracted by pipette and transferred into a clean
207 1.5 ml tube.

208

209 An equal volume of Agencourt AMPure XP beads was added to each sample, vortexed for
210 20 s (10 x 2 s bursts), and then incubated for 10 min at room temperature. By placing the
211 samples onto a magnetic tube rack for 5 min, the beads were separated from the solution,
212 and the cleared supernatant was removed by aspiration. The beads were washed twice
213 using the following protocol: 1 ml of 80% ethanol was added, incubated at room temperature
214 for 30 s, and then removed, followed by air drying for ≈ 3 min. The magnetic beads were
215 resuspended in 55 µl of EB (Elution Buffer: 10 mM Tris-HCl) and incubated at 37 °C for 10
216 min. The tubes were placed back onto the magnetic rack to bind the beads, and the eluted
217 DNA (≈ 50 µl) was transferred into fresh tubes. A 1 µl aliquot of 1-in-10 diluted Qiagen
218 RNase A was added to each DNA sample before being incubated for 30 min at 37 °C. The
219 concentration of the eluted DNA was assessed using the dsDNA HS assay on a Qubit 2.0
220 fluorometer. To check the DNA for degradation, fragment size distributions were checked

221 with a TapeStation 2200 using the Genomic DNA Analysis ScreenTape.

222

223 Finally, the extracted DNA was prepared and sequenced using the same protocol as used
224 for the DNA mocks above, except that only one library was prepared for each sample.
225 Twelve samples can be multiplexed using the Rapid Barcoding Sequencing Kit; we thus
226 required four flow cells. Due to continuous software upgrades by ONT, the specific software
227 versions of *MinKNOW* varied across runs and is recorded in the final sequence files (fast5
228 format), which are available from the EBI's European Nucleotide Archive (see Data
229 accessibility).

230

231 **Illumina and MinION read pre-processing**

232

233 Duplicate reads were removed from the 49 plant-reference Illumina datasets using *NextClip*
234 *1.3.2* (Leggett, Clavijo, Clissold, Clark, & Caccamo, 2014), and then *cutadapt 1.10* (Martin,
235 2011) was used to trim Illumina adaptors and filter out reads shorter than 100 bp. The
236 resulting unmerged FASTQ files constitute our 49 *reference skims*.

237

238 The MinION datasets from the 12 mocks and the 48 pollen loads were basecalled and
239 demultiplexed with *albacore 2.1.10* (ONT). The resulting FASTQ files were converted to
240 FASTA format. We removed long reads deriving from plant organelles because they are
241 highly conserved across plant species and in pilot tests we observed that mapping to
242 organellar long reads resulted in a higher rate of incorrect assignments than mapping to
243 nuclear long reads (data not shown). NCBI Entrez
244 (https://www.ncbi.nlm.nih.gov/sites/batchentrez) was used to download 2,583 Land Plant
245 organelle genomes, including 1,852 chloroplast, 226 mitochondrial, and 505 plastid
246 genomes. Organelle reads were identified by aligning each of the MinION datasets to the
247 organellar genomes using *minimap2 2.7* (Li, 2018) and removed from the FASTA files. The
248 resulting 60 (= 12 + 48) organelle-filtered FASTA files constitute our mock and pollen *query*

249 datasets, and in the next step, we used the 49 plant reference skims to assign a taxonomy
250 to each long read in the mock and pollen query datasets (Fig. 1c).
251
252 **Taxonomic assignment of mock-sample and bee-collected pollen MinION reads**
253
254 We used *bwa mem 0.7.17* (Li, 2013) to map the Illumina reads from each of the 49
255 reference skims against every individual long read in each of the mock and bee-collected
256 pollen datasets. *SAMtools 1.7* (Li *et al.,* 2009) was used to remove unmapped reads and
257 secondary and supplementary alignments. After SAMtools indexing, the depth of mapping
258 coverage at each long-read position was calculated using the SAMtools *depth* function,
259 followed by the calculation of 'percent coverage' for each long read (using a custom python
260 script), defined as the fraction of nucleotide positions that were mapped to by one or more
261 reference-skim Illumina reads.
262
263 We assigned each long read to the plant species that mapped with the highest percentage
264 coverage, unless the highest percent coverage was <15%, in which case the long read's
265 identity was judged ambiguous and left unassigned. Additionally, for clarity of presentation,
266 we implemented a 1% minimum-abundance filter, removing plant species represented by
267 fewer than 1% of the total assigned long reads in each sample.
268
269 **Reference-skim subsampling**
270
271 To estimate a minimum recommended depth of coverage needed per reference skim, we
272 subsampled one of the genome skims, *Knautia arvensis*, which is a major constituent
273 species in mock mixes MM1 and MM2. We randomly subsampled this skim from its
274 maximum of 0.65x down to 0.05x, in steps of 0.05x using a custom script. For each
275 subsample, the whole pipeline was re-run along with the full reference skims of the other 48
276 plant species. The number of mock reads assigned to *Knautia arvensis*, and the number of

277 unassigned reads, at each level of coverage was recorded. This subsampling was repeated
278 three times (Supplementary Fig. S1).
279
280 **Network construction**
281
282 We constructed a pollinator-plant network diagram for the 48 wild-bee pollen samples, using
283 the *bipartite 2.11* package (Dormann, Frund, Bluthgen, & Gruber, 2009) for the *R* statistical
284 language (R Core Team, 2018). For presentational clarity and to avoid any pollen from
285 previous foraging trips, we only show plant species represented by more than 10% of the
286 assigned reads in each sample.
287
288 **Results**
289
290 **A reference set of plant genome skims**
291
292 Low genome-coverage, short-read, shotgun-sequencing datasets ('reference skims') were
293 successfully generated for all 49 plant species (Fig. 1a). After pre-processing, the mean
294 estimated coverage was 0.6x (0.1 to 1x, details in Supplementary Table S1).
295
296 **Mock DNA mixes**
297
298 The six mock communities, each with two technical replicates, were sequenced on a
299 MinION. These produced reads with mean length 1914 bp (longest 41,058 bp). After
300 demultiplexing, 88.8% of the reads could be assigned to one of the 12 mocks, with the
301 remaining reads left unclassified. Sequences originating from organellar genomes made up
302 between 5.1% (MM4.2) to 10.2% (MM3.2) of the reads in the mocks and were removed. The
303 remaining number of reads per mock ranged from 733 (MM2.1) to 2174 (MM4.1), mean

1347.

Taxonomic assignment of mock-sample MinION reads

The 49 reference skims were separately mapped to each long read in each of the 12 mocks, and each long read was assigned to the plant species that mapped with the highest percent coverage, or left unassigned if the highest coverage was <15%. In total, 65.5% of the mock reads were assigned to a plant species, with 94.7% of those reads being assigned to a species known to be present in that mock sample. Almost all (93.4%) of the 563 false-positive read assignments were made to one species, *Ranunculus acris*, and all these assignments occurred in the mock samples that contained the very closely related species *Ranunculus repens*. We return to this in the Discussion. The few other false-positive assignments all occurred at a rate of less than 1% of the assigned long reads in their mixes and for presentational clarity are not shown in Fig. 2. The full results are in Table S2.

All of the plant species that had been added to the mock compositions at proportions ≥1% were detected by our method in at least one of the two replicates, and in general, the technical replicates showed a high level of repeatability (Fig. 2). In two of the mocks, there was one species each (*Lotus corniculatus* in MM2 and *Digitalis purpurea* in MM3*)* that were detected in only one of the two replicates. In these two replicates, *Lotus corniculatus* and *Digitalis purpurea* were expected to be present at only 3.0% and 4.6%, respectively. Both of these species were consistently underrepresented across our mock data sets.

Reference-skim subsampling

As expected, the larger the reference-skim dataset size for *Knautia arvensis*, the more reads in the MM1 and MM2 mocks were assigned to this species and the fewer reads left unassigned. Importantly, the rate of increase was decelerating (Supplementary Fig. S1);

332 over half of the MinION reads that were assigned to *Knautia arvensis* with a 0.65x genome
333 skim could also be assigned with just a 0.1x skim, even though all the other reference skims
334 in the mapping run were kept at their original sizes.
335
336 **Taxonomic assignment of bee-collected pollen MinION reads**
337
338 The 48 bee-collected pollen loads harvested from the corbiculae of three species, *Apis*
339 *mellifera*, *Bombus terrestris/lucorum* complex, and *Bombus lapidarius*, yielded DNA
340 quantities ranging from 191 to 3750 ng, and all successfully produced libraries,
341 demonstrating that pollen carried by individual bees can provide sufficient DNA for MinION
342 sequencing.
343
344 As with the 12 mocks, each of the reference skims was aligned to each long read in each of
345 the 48 pollen samples, the long reads were either assigned to the plant species achieving
346 the highest percent coverage or left unassigned, and any plant species assigned fewer than
347 1% of the long reads in each bee-collected pollen sample was filtered out (Supplementary
348 Table S3). In total, 49.7% of the long reads were assigned to one of the reference plant
349 species. In 38 of the 48 bees (79.2%), pollen from the plant species on which each bee was
350 captured was found to be present in that bee's pollen load (Supplementary Table S3).
351
352 Each of the 48 pollen loads was found to contain one or two major plant species (defined as
353 read frequency ≥10%) (Fig. 3a). All nine of the *Apis mellifera* pollen loads contained a single
354 major species, whereas 16 of 27 *Bombus terrestris/locorum complex* and 6 of 12 *Bombus*
355 *lapidarius* pollen loads were comprised of two major species (Fig. 3a). These differences in
356 mean number of major species were statistically significant (*Apis mellifera* versus *Bombus*
357 *terrestris/locorum complex* (Welch's t-test, $t = -6.15$, $df = 26$, p-value < 0.0001) and versus
358 *Bombus lapidarius* ($t = -3.32$, $df = 11$, p-value < 0.01)) (Fig. 3b). Another way of visualising
359 the wild-bee results is as a plant-pollinator network graph (Fig. 3c). Overall, 6 of the 49

360 reference plant species were identified as major components in the 48 pollen loads, and the
361 majority of bee-collected pollen samples were dominated by one plant species.
362

363 **Discussion**

364

365 Using light microscopy to identify plant species from pollen requires expert knowledge and is
366 costly when applied to many samples (Khansari *et al.,* 2012). There is a need for a quick
367 and low-cost method that can be scaled to large numbers of pollen samples. Metabarcoding
368 is the current leading candidate, but there are concerns over its discriminatory power at
369 lower taxonomic levels, and there is good reason to believe that metabarcoding does not
370 return reliable quantitative data (Keller *et al.,* 2015; Richardson *et al.,* 2015; Sickel *et al.,*
371 2015; Bell *et al.,* 2017). A PCR-free shotgun-metagenomics approach has greater potential
372 for providing reliable quantitative analysis with high power for resolving species. However,
373 applying shotgun metagenomics to eukaryotes is challenging due to the lack of reference
374 genomes (Gilbert & Dupont, 2011). We have developed a metagenomics method that avoids
375 the need for reference genomes. Instead, each reference species is represented by just a
376 low-cost genome skim, and we use a set of such skims to identify individual long reads from
377 pollen samples, produced by the MinION sequencer.
378

379 We evaluated our RevMet pipeline with mock DNA mixtures of known composition and then
380 applied the pipeline to pollen collected from wild bees. Our main findings are:
381

382 1) RevMet can identify plant species present in mixed-species samples at proportions
383 ≥1%, with few false positives and false negatives, and can reliably differentiate
384 species represented by high versus low amounts of DNA in a sample (Fig. 2,
385 Supplementary Table S2).

2) Genome skims with sequence coverage as low as 0.05x can be used for detecting species presence and for estimating relative abundance in terms of DNA mass. Increasing skim coverage increases detection power, at a decelerating rate (Supplementary Fig. S1).
3) Individual pollen loads collected from wild *Apis* and *Bombus* bees yield enough DNA for MinION sequencing (Supplementary Table S3) and generate plausible plant-pollinator networks, as evidenced by the fact that (a) 56.3% of the plant species on which the bees were collected were also the dominant constituent of the corresponding pollen sample (and 79% of plant species on which the bees were collected were detected in the corresponding pollen sample) (Supplementary Table S3), and (b) pollen species richnesses and compositions were more similar within bee species than across bee species (Fig. 3).
4) Our per-plant-species cost of a reference skim was £90, and our per-pollen-sample cost was £61, including DNA extraction, library preparation, and sequencing. Sequencing costs will likely drop further, given the new Illumina NovaSeq and new MinION 'Flongle'.

Semi-quantitative species compositions. – We were able to assign roughly 65% of the mock-mix MinION reads and just under 50% of the pollen-load MinION reads to our reference plant species. Importantly, the frequencies of MinION reads that were assigned to each reference plant species were reliably *'semi-quantitative*,' that is, able to differentiate low- and high-frequency plant species (Fig. 2). Within low- and high-abundance categories, accuracy was lower. For example, in mock sample MM1, *Knautia arvensis*, *Galium verum*, and *Crepis capillaris* were the three high-abundance species (each representing 30.3% of total input DNA mass each), and *Papaver somniferum*, *Anagallis arvensis*, and *Sambucus nigra* were the three low-abundance species (each representing 3.0% of total input DNA mass each). The RevMet pipeline estimated the three high-abundance frequencies at means of 34.0%,

413 14.7%, and 44.0%, and the three low-abundance species at 1.4%, 3.0%, and 3.0%,
414 respectively (Supplementary Table S2).
415
416 There are at least three reasons for the remaining quantitative error. First, although we
417 targeted 0.5x per reference skim, coverage still varied across species (Table S1), resulting
418 in different powers of discrimination, as shown by the experiment with subsamples of
419 *Knautia arvensis* (Supplementary Fig. S1). Fortunately, we found that even very low-depth
420 skims of 0.05x are useful for species detection and are probably still useful for differentiating
421 rare from abundant species (albeit with more error) (Supplementary Fig. S1). Genome sizes
422 are also estimated with error, so it is also helpful that the subsampling experiment suggests
423 that detection power asymptotes with higher sequencing depth (Supplementary Fig. S1),
424 and as sequencing costs fall further, we expect that the most robust protocol will be to target
425 1x coverage.
426
427 Second, very closely related species can generate false positives. Our reference-skim
428 database included six congener pairs, and we included two of the pairs (*Papaver* and
429 *Ranunculus*) in the mock mixes. In the case of *Papaver*, there were no *P. rhoeas* false-
430 positives greater than the 1% minimum-abundance filter in the mocks that contained *P.*
431 *somniferum* (MM1 and MM6) (Supplementary Table S2). In contrast, *Ranunculus acris* was
432 regularly incorrectly assigned to reads in mock mixes that contained the closely related
433 congener *Ranunculus repens*. In fact, almost all the false-positive assignments (93.4%)
434 were to *R. acris*. In retrospect, this result is expected because these two species are not
435 easily differentiated by pollen morphology (Forup & Memmott, 2005), floral morphology, or
436 even DNA barcodes (*rbcL* (99.1% similarity), *matK* (96.9%), *ITS2* (95.5%)). In other words,
437 the RevMet results are correctly telling us that the two *Ranunculus* species are very closely
438 related.
439

440 Third, MinION reads have relatively high error rates of roughly 5 to 10% depending on the
441 flow cell and kit used (Leggett & Clark, 2017). Although this is dropping over time, this error
442 rate unavoidably obscures differences between species (although not enough to confound
443 the two *Papaver* species). We note that one of the advantages of the RevMet approach is
444 that we use both sequence similarity and percent coverage as predictors of species
445 presence (Fig. 1C) . Using sequence similarity alone, we observed several instances of low
446 numbers of mapped reads being given false-positive assignments (data not shown). The
447 percent-coverage filter requires many reference-skim reads to independently identify a
448 species before an assignment is made.
449
450 *Reference-skim cost*. – The RevMet pipeline is relatively low cost. In our study, we
451 generated skims for 49 plant species, with genome sizes ranging from roughly 290 Mb
452 (*Epilobium hirsutum*) to just under 15 Gb (*Sambucus nigra*), targeting 0.5x coverage. All
453 skims were produced on a single lane of Illumina HiSeq 2500 (250 PE) at a mean coverage
454 of 0.57x. The average cost per skim in this study was just under £90, which includes the
455 DNA extraction, LITE library preparation, sequencing, and data QC. The per skim cost will
456 be lower in studies with smaller eukaryotic genomes and should be lower with Illumina's new
457 sequencer, the NovaSeq 6000. Genome assembly projects are likely to produce many such
458 datasets for free download in the future.
459
460 *Long-read MinION cost*. – We used ONT's first iteration of the Rapid Barcoding Kit (RBK-
461 001), which relies on transposase to randomly fragment DNA and simultaneously add
462 barcoded adapters. Longer read lengths have an increased likelihood of accurate species
463 assignment because they carry more sequence information. The two main ways to obtain
464 longer reads with transposase-based preparations are to: (1) increase the ratio of DNA to
465 transposase e.g. by increasing the input material or by heat killing a proportion of the
466 transposase (which also lowers sequencing yields); and (2) use higher molecular weight
467 input DNA. Since the release of RBK-001, ONT's chemistry has evolved, and their Rapid-

468 based kits have seen greater sequencing yields. However, the recommended input for the
469 latest iteration of the Rapid Barcoding Kit (RBK-004) is now higher, 400 ng of DNA per
470 sample. That said, we anticipate that input biomasses similar to those used in this study, 200
471 ng, will still be adequate. Also, even 400 ng is achievable, as 36 of 48 of our wild-bee pollen
472 samples samples yielded >400 ng (Supplementary Table S3). ONT have also recently
473 released the Flongle (~$90), which is a disposable nanopore that allows prolonged reuse of
474 the MinION. Our results suggest that ONT's target yield of 1 Gb per Flongle will be more
475 than enough for multiplexing twelve bee-collected pollen loads, reducing per-sample costs
476 from the £61 in this study to just under £16.
477
478 *Application to pollen collected from wild bees*. – The RevMet pipeline detected consistent
479 differences in plant species preference (Fig. 3C) and in foraging behavior between
480 honeybees *Apis mellifera* and the two bumblebees *Bombus terrestris/lucorum* and *B.*
481 *lapidarius* (Fig. 3AB). Importantly, because our data are semi-quantitative, we are able to
482 conclude that even the bumblebees showed fidelity to one plant species (Fig. 3C), a result
483 that would be less reliably concluded from metabarcoding data.
484
485 The RevMet pipeline can readily be applied to a wide range of research questions. Most
486 obviously, ReMet could be used to compare pollination networks across large-scale spatial
487 and biogeographical gradients (Pornon *et al.,* 2016) and quantify pollen-carrying capacities
488 of insects (Escaravage & Wagner, 2004). RevMet could potentially also be used to quantify
489 the degree to which co-attraction of pollinators leads not to benefits of increased pollinator
490 numbers but to loss of pollination service via competition (Pornon *et al.,* 2016). Outside of
491 pollination ecology, there is potential for semi-quantitative assessments of many other
492 eukaryotic species mixtures, including herbivore diets (Bhattacharyya, Dawson, Hipperson,
493 & Ishtiaq, 2018; Kress, García-Robledo, Uriarte, & Erickson, 2015); plant-fungus interactions
494 (Schröter *et al.,* 2018); allergenic pollen species from air samples (Kraaijeveld *et al.,* 2015);
495 and algal and diatom communities (Keller *et al.,* 2015). Furthermore, due to the portability

496 and real-time nature of the MinION platform, the method could be optimised for analysis in
497 the field alongside sample collection.
498
499 **Acknowledgements**
500
501 This work was supported by the Norwich Research Park Science Links Seed Fund, the
502 BBSRC Norwich Research Park Biosciences Doctoral Training Partnership (grant number
503 BB/M011216/1) and BBSRC Core Strategic Programme Grant BB/CSP17270/1 to the
504 Earlham Institute. We are grateful for the support of the Earlham Institute Genomics
505 Pipelines group and the NBI Computing infrastructure for Science (CiS) group. We thank
506 Iain Barr for help in fieldwork and Pensthorpe Natural Park for access.
507
508 **Authors' contributions**
509
510 MDC, RML, DWY, LVD, and RGD conceived and designed the study. LVD, RGD, and CC
511 collected the samples. NP, DH, and LP performed the experiments. NP, RML, and DWY
512 analyzed the data. NP and DWY led the writing of the manuscript. All authors gave final
513 approval for publication.
514
515 **Data accessibility**
516
517 The Illumina and MinION datasets are available in the European Nucleotide Archive
518 (http://www.ebi.ac.uk/ena) under study accession PRJEB30946. RevMet scripts are
519 available from https://github.com/nedpeel/RevMet and a tutorial using an example dataset
520 can be found at https://revmet.readthedocs.io/en/latest/.
521
522
523

References

Beier, S., Himmelbach, A., Colmsee, C., Zhang, X. Q., Barrero, R. A., Zhang, Q., … Mascher, M. (2017). Construction of a map-based reference genome sequence for barley, Hordeum vulgare L. *Scientific Data*, **4**, 170044.

Bell, K. L., Fowler, J., Burgess, K. S., Dobbs, E. K., Gruenewald, D., Lawley, B., … Brosi, B. J. (2017). Applying pollen DNA metabarcoding to the study of plant-pollinator interactions. *Applications in plant sciences*, **5**, 1600124.

Bennett, M., & Leitch, I. (2012). Plant DNA C-values Database. Available at http://data.kew.org/cvalues/

Bhattacharyya, S., Dawson, D. A., Hipperson, H., & Ishtiaq F. (2018). A diet rich in C3 plants reveals the sensitivity of an alpine mammal to climate change. *Molecular Ecology*, **28**, 250-265.

Bommarco, R., Kleijn, D., & Potts, S. G. (2013). Ecological intensification: Harnessing ecosystem services for food security. *Trends in Ecology and Evolution*, **28**, 230-238.

Burkle, L. A., Marlin, J. C., & Knight, T. M. (2013). Plant-pollinator interactions over 120 years: Loss of species, co-occurrence, and function. *Science*, **339**, 1611-1615.

Dicks, L. V., Abrahams, A., Atkinson, J., Biesmeijer, J., Bourn, N., Brown, C., … Sutherland, W. J. (2013). Identifying key knowledge needs for evidence-based conservation of wild insect pollinators: A collaborative cross-sectoral exercise. *Insect Conservation and Diversity*, **6**, 435-446.

Dormann, C. F., Frund, J., Bluthgen, N., & Gruber, B. (2009). Indices, Graphs and Null Models: Analyzing Bipartite Ecological Networks. *The Open Ecology Journal*, **2,** 7-24.

Escaravage, N., & Wagner, J. (2004). Pollination effectiveness and pollen dispersal in a Rhododendron ferrugineum (Ericaceae) population. *Plant Biol*, **6**, 606-615.

Escobar-Zepeda, A., De León, A. V. P., & Sanchez-Flores, A. (2015). The road to metagenomics: From microbiology to DNA sequencing technologies and bioinformatics. *Frontiers in Genetics*, **6**, 348.

Gilbert, J. A., & Dupont, C. L. (2011). Microbial Metagenomics: Beyond the Genome. *Annual Review of Marine Science*, **3**, 347-371.

Hollingsworth, P. M., Li, D. Z., Van Der Bank, M., & Twyford, A. D. (2016). Telling plant species apart with DNA: From barcodes to genomes. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **371**, 20150338.

IPBES. (2016). The assessment report on Pollinators, pollination and food production.

Ji, Y., Ashton, L., Pedley, S. M., Edwards, D. P., Tang, Y., Nakamura, A., … Yu, D. W. (2013). Reliable, verifiable and efficient monitoring of biodiversity via metabarcoding. *Ecology Letters*, **97**, 1966-1979.

Jones, M. B., Highlander, S. K., Anderson, E. L., Li, W., Dayrit, M., Klitgord, N., … Venter, J. C. (2015). Library preparation methodology can influence genomic and functional predictions in human microbiome research. *Proceedings of the National Academy of Sciences*, **112**, 14024-14029.

Keller, A., Danner, N., Grimmer, G., Ankenbrand, M., von der Ohe, K., von der Ohe, W., … Steffan-Dewenter, I. (2015). Evaluating multiplexed next-generation sequencing as a method in palynology for mixed pollen samples. *Plant Biology*, **17**, 558-566.

Khansari, E., Zarre, S., Alizadeh, K., Attar, F., Aghabeigi, F., & Salmaki, Y. (2012). Pollen morphology of Campanula (Campanulaceae) and allied genera in Iran with special focus on its systematic implication. *Flora: Morphology, Distribution, Functional Ecology of Plants*, **207**, 203-211.

Klein, A. M., Vaissière, B. E., Cane, J. H., Steffan-Dewenter, I., Cunningham, S. A., Kremen, C., & Tscharntke, T. (2007). Importance of pollinators in changing landscapes for world crops. *Proceedings of the Royal Society B: Biological Sciences*, **274**, 303-133.

Kraaijeveld, K., De Weger, L. A., Ventayol García, M., Buermans, H., Frank, J., Hiemstra, P. S., & Den Dunnen, J. T. (2015). Efficient and sensitive identification and quantification of airborne pollen using next-generation DNA sequencing. *Mol Ecol Resour*, **15**, 8-16.

Kress, W. J., García-Robledo, C., Uriarte, M., & Erickson, D. L. (2015). DNA barcodes for ecology, evolution, and conservation. *Trends in Ecology & Evolution*, **30**, 25-35.

Leggett, R. M., Clavijo, B. J., Clissold, L., Clark, M. D., & Caccamo, M. (2014). Next clip: An analysis and read preparation tool for nextera long mate pair libraries. *Bioinformatics*, **30**, 566-568.

Leggett, R. M., & Clark, M. D. (2017). A world of opportunities with nanopore sequencing. *Journal of Experimental Botany*, **68**, 5419-5429.

Li H. (2013). Aligning sequence reads, clone sequences and assembly contigs with bwa-mem. arXiv:1303.3997.

Li, H. (2018). Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics*, **34**, 3094-3100.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., … Durbin, R. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, **25**, 2078-2079.

Li, X., Yang, Y., Henry, R. J., Rossetto, M., Wang, Y., & Chen, S. (2015). Plant DNA barcoding: from gene to genome. *Biological Reviews of the Cambridge Philosophical Society*, **90**, 157-166.

Long, E. Y., & Krupke, C. H. (2016). Non-cultivated plants present a season-long route of pesticide exposure for honey bees. *Nature Communications*, **7**, 11629.

Lucas, A., Bodger, O., Brosi, B. J., Ford, C. R., Forman, D. W., Greig, C., ... de Vere, N. (2018). Generalisation and specialisation in hoverfly (Syrphidae) grassland pollen transport networks revealed by DNA metabarcoding. *Journal of Animal Ecology*, **87**, 1008-1021.

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.Journal*, **17**, 10.

Morales, C. L., & Traveset, A. (2008). Interspecific pollen transfer: magnitude, prevalence and consequences for plant fitness. *Critical Reviews in Plant Sciences*, **27**, 221–238.

633 Morandin, L. A., & Kremen, C. (2013). Hedgerow restoration promotes pollinator populations
634 and exports native bees to adjacent fields. *Ecological Applications*, **23**, 829–839.
635
636 Nayfach, S., & Pollard, K. S. (2016). Toward Accurate and Quantitative Comparative
637 Metagenomics. *Cell*, **166**, 1103-1116.
638
639 Ollerton, J., Winfree, R., & Tarrant, S. (2011). How many flowering plants are pollinated by
640 animals? *Oikos*, **120**, 321-326.
641
642 Pornon, A., Escaravage, N., Burrus, M., Holota, H., Khimoun, A., Mariette, J., … Vidal, M.
643 (2016). Using metabarcoding to reveal and quantify plant-pollinator interactions. *Scientific*
644 *Reports*, **6**, 27282.
645
646 Potts, S. G., Biesmeijer, J. C., Kremen, C., Neumann, P., Schweiger, O., & Kunin, W. E.
647 (2010). Global pollinator declines: Trends, impacts and drivers. *Trends in Ecology and*
648 *Evolution*, **25**, 345-353.
649
650 Potts, S. G., Imperatriz-Fonseca, V., Ngo, H. T., Aizen, M. A., Biesmeijer, J. C., Breeze, T.
651 D., … Vanbergen, A. J. (2016). Safeguarding pollinators and their values to human well-
652 being. *Nature*, **540**, 220-229.
653
654 Quince, C., Walker, A. W., Simpson, J. T., Loman, N. J., & Segata, N. (2017). Shotgun
655 metagenomics, from sampling to analysis. *Nature Biotechnology*, **35**, 833-844.
656
657 R Core Team (2018). R: A language and environment for statistical computing. R
658 Foundation for Statistical Computing. Available at http://www.R-project.org/.
659
660 Ratnasingham, S., & Hebert, P. D. (2007). BOLD: The Barcode of Life Data System
661 (http://www. barcodinglife.org). *Molecular ecology notes*, **7**, 355-364.
662
663 Richardson, R. T., Lin, C. H., Sponsler, D. B., Quijia, J. O., Goodell, K., & Johnson, R. M.
664 (2015). Application of ITS2 Metabarcoding to Determine the Provenance of Pollen Collected
665 by Honey Bees in an Agroecosystem. *Applications in Plant Sciences*, **3**, 1400066.
666
667 Schlumbaum A., Tensen M., & Jaenicke-Després V. (2008). Ancient plant DNA in
668 archaeobotany. *Vegetation History and Archaeobotany*, **17**, 233-244.
669

Schröter, K., Wemheuer, B., Pena, R., Schoning, I., Ehbrecht, M., Schall, P., … Polle, A. (2018). Assembly processes of trophic guilds in the root mycobiome of temperate forests. *Molecular Ecology*, **28**, 348-364.

Sickel, W., Ankenbrand, M. J., Grimmer, G., Holzschuh, A., Härtel, S., Lanzen, J., … Keller, A. (2015). Increased efficiency in identifying mixed pollen samples by meta-barcoding with a dual-indexing approach. *BMC Ecology*, **15**, 20.

Sharpton, T. J. (2014). An introduction to the analysis of shotgun metagenomic data. *Frontiers in Plant Science*, **5**, 209.

Tittonell, P. (2014). Ecological intensification of agriculture-sustainable by nature. *Current Opinion in Environmental Sustainability*, **8**, 53-61.

Vanbergen, A. J., Baude, M., Biesmeijer, J. C., Britton, N. F., Brown, M. J. F., Bryden, J., … Wright, G. A. (2013). Threats to an ecosystem service: Pressures on pollinators. *Frontiers in Ecology and the Environment*, **11**, 251-259.

Wood, T. J., Holland, J. M., & Goulson, D. (2015). Pollinator-friendly management does not increase the diversity of farmland bees and wasps. *Biological Conservation*, **187**, 120-126.

703 **Figure legends and table titles**

704

705 Figure 1. RevMet pipeline overview. a) Low coverage, short-read, reference datasets were
706 generated for 49 wild plant species. b) Bee-collected pollen loads were sequenced on a
707 MinION, generating long read datasets. c) The 49 short-read reference datasets were
708 separately mapped to the long-read pollen datasets, and each pollen read was assigned to
709 the plant species that mapped with the highest percent coverage, or left unassigned if the
710 highest coverage was <15%. d) Binned pollen reads were counted, noise was reduced by
711 implementing a 1% minimum-abundance filter, and then the remaining bin counts were
712 converted to percentages.

713

714 Figure 2. Expected vs observed mock mix compositions. Six mock plant DNA mixes, each
715 with two technical replicates, were sequenced on a MinION and the RevMet method was
716 applied. The first stacked bar of each triplet represents the expected proportions based on
717 input DNA. The second and third bar of each triplet reflect the observed MinION read
718 assignments resulting from this pipeline.

719

720 Figure 3. Bee-collected pollen compositions and plant-pollinator interactions. a) The number
721 of individual pollen loads sequenced from three different species of bee. The proportion of
722 pollen loads that contained a single major plant species are represented by green bars,
723 while those with two major plant species are shown in blue. b) Mean number of plant
724 species per pollen load for each of three different species of bee; ** $p<0.01$, *** $p<0.001$. c)
725 Bipartite plant-pollinator network. The upper bars represent individual pollen loads from
726 three different bee species, *Apis mellifera* (red), *Bombus terrestris/lucorum complex* (blue),
727 and *Bombus lapidarius* (purple). The lower bars (grey) represent plant species. Link width
728 indicates the MinION read proportion of each major plant species within each pollen load.
729 Table 1. DNA mock community compositions.

730

731 Figure S1. Numbers of mock-mix reads assigned to *Knautia arvensis*, and declines in the
732 number of unassigned reads, at different reference-skim coverage levels. The subsampling
733 was repeated three times.
734
735 Table S1. Estimated genome sizes, read counts and coverage for the genome skim
736 references.
737
738 Table S2. RevMet taxonomic assignments of mock-sample MinION reads.
739
740 Table S3. RevMet taxonomic assignments of bee-collected pollen MinION reads and pollen
741 sample information.

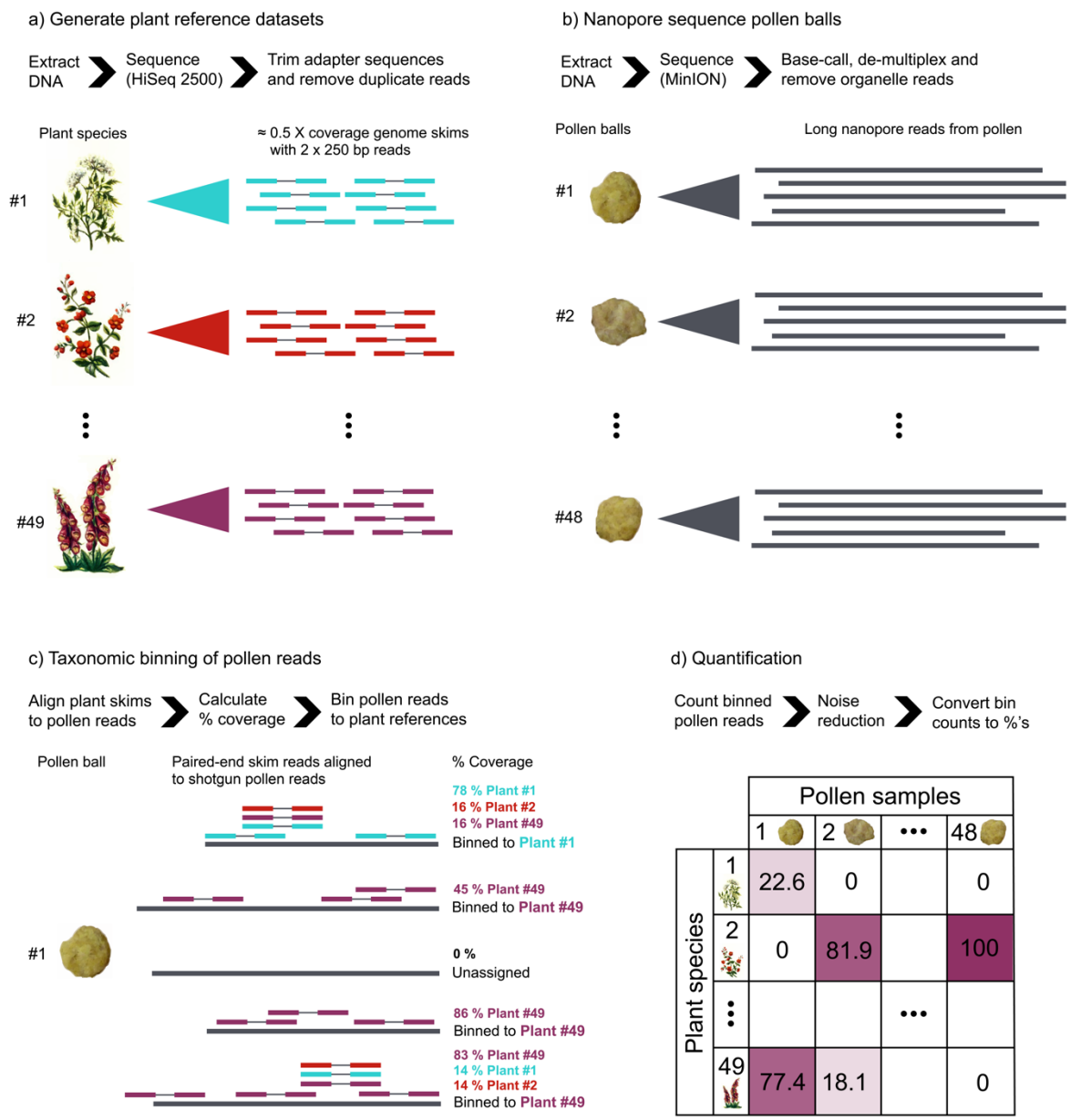

Figure 1. RevMet pipeline overview. a) Low coverage, short-read, reference datasets were generated for 49 wild plant species. b) Bee-collected pollen loads were sequenced on a MinION, generating long read datasets. c) The 49 short-read reference datasets were separately mapped to the long-read pollen datasets, and each pollen read was assigned to the plant species that mapped with the highest percent coverage, or left unassigned if the highest coverage was <15%. d) Binned pollen reads were counted, noise was reduced by implementing a 1% minimum-abundance filter, and then the remaining bin counts were converted to percentages.

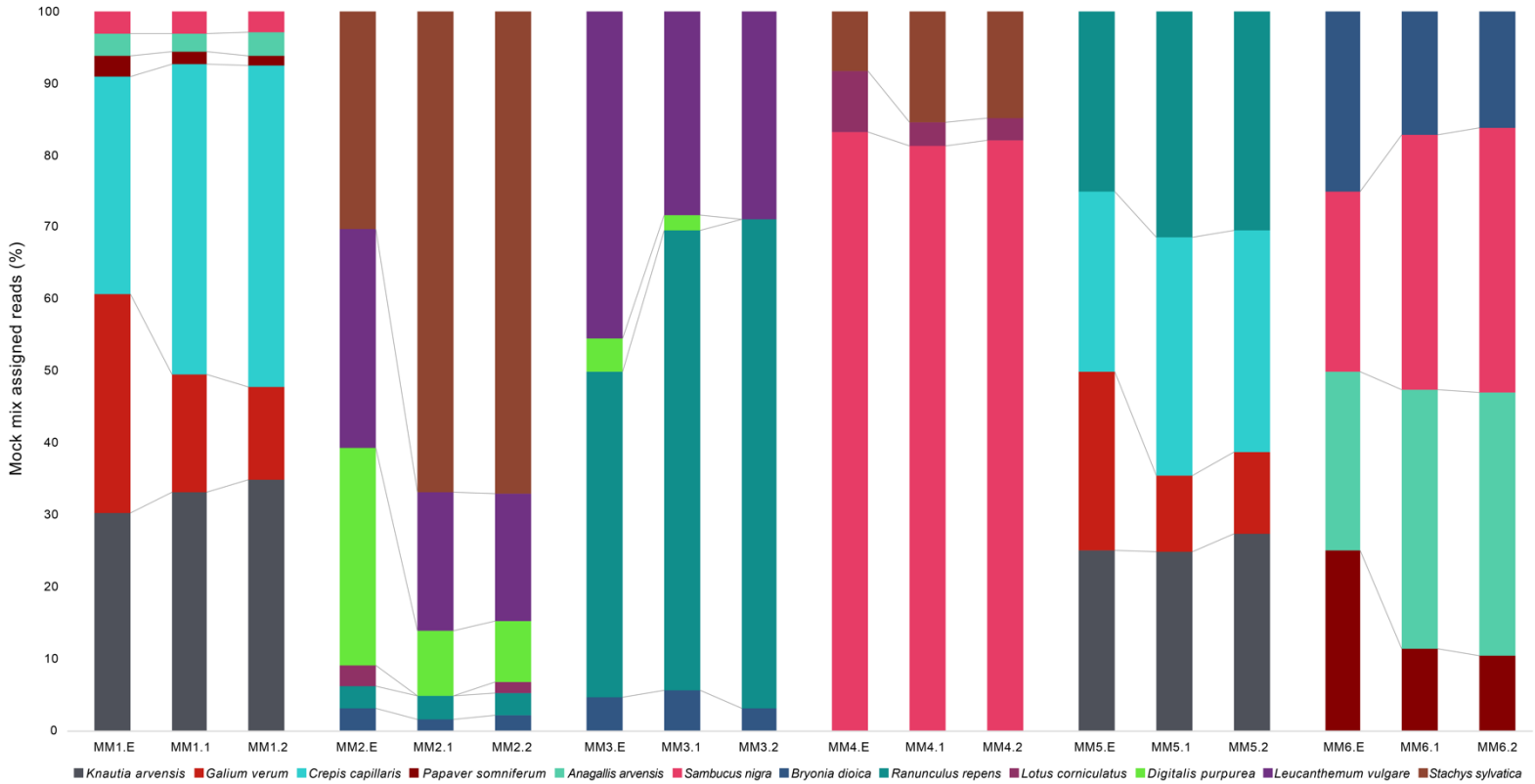

Figure 2. Expected vs observed mock mix compositions. Six mock plant DNA mixes, each with two technical replicates, were sequenced on a MinION and the RevMet method was applied. The first stacked bar of each triplet represents the expected proportions based on input DNA. The second and third bar of each triplet reflect the observed MinION read assignments resulting from this pipeline.

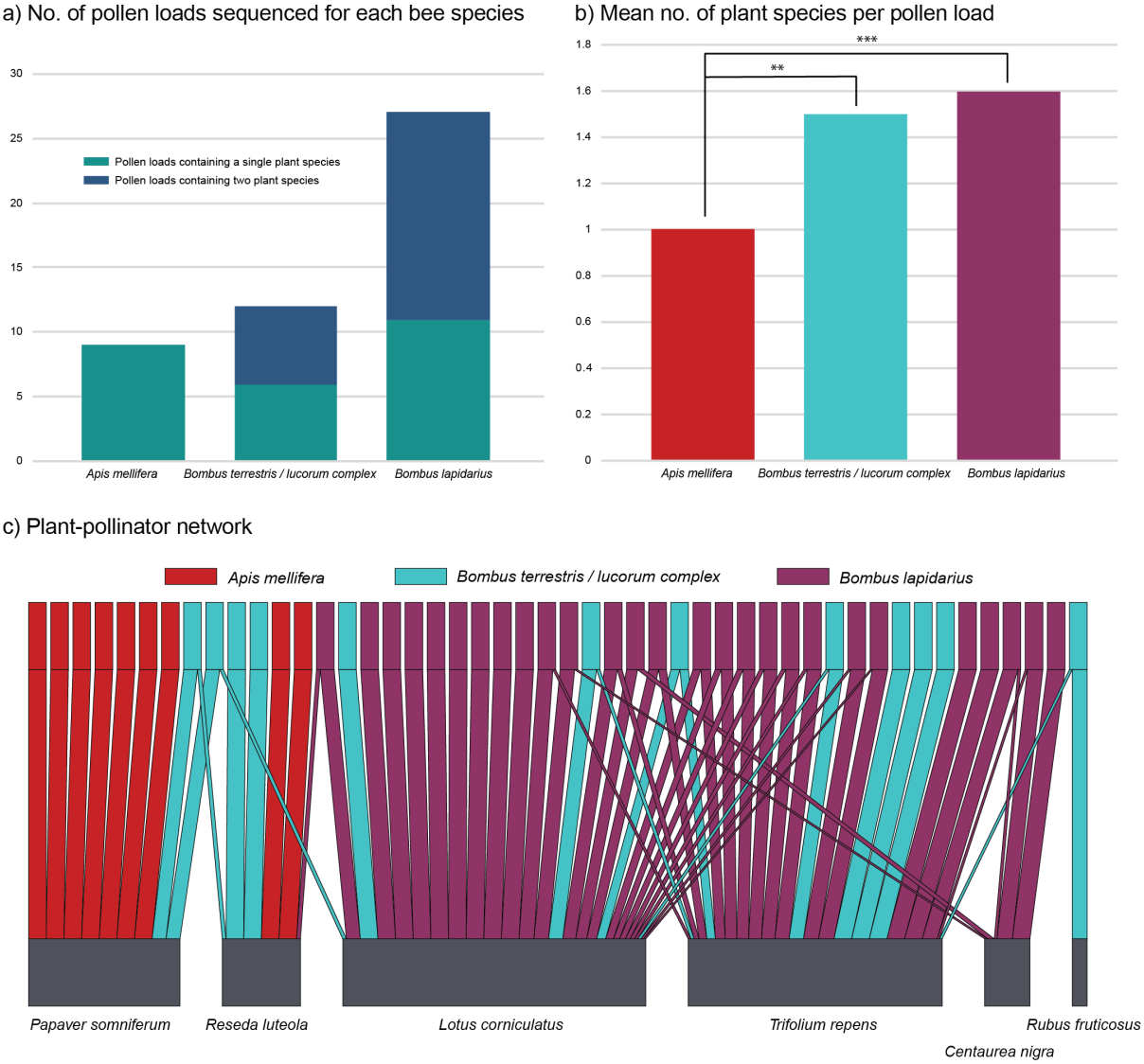

Figure 3. Bee-collected pollen compositions and plant-pollinator interactions. a) The number of individual pollen loads sequenced from three different species of bee. The proportion of pollen loads that contained a single major plant species are represented by green bars, while those with two major plant species are shown in blue. b) Mean number of plant species per pollen load for each of three different species of bee; ** $p<0.01$, *** $p<0.001$. c) Bipartite plant-pollinator network. The upper bars represent individual pollen loads from three different bee species, *Apis mellifera* (red), *Bombus terrestris/lucorum complex* (blue), and *Bombus lapidarius* (purple). The lower bars (grey) represent plant species. Link width indicates the MinION read proportion of each major plant species within each pollen load.

774 Table 1. DNA mock community compositions.

775

	Knautia arvensis	*Galium verum*	*Crepis capillaris*	*Papaver somniferum*	*Anagallis arvensis*	*Sambucus nigra*	*Bryonia dioica*	*Ranunculus repens*	*Lotus corniculatus*	*Digitalis purpurea*	*Leucanthemum vulgare*	*Stachys sylvatica*
MM1.Ratios	100	100	100	10	10	10	1	1	1	0	0	0
MM2.Ratios	0	0	0	1	1	1	10	10	10	100	100	100
MM3.Ratios	0	0	0	0	0	0	10	100	0	0	0	0
MM4.Ratios	0	0	0	1	0	1000	0	0	100	0	0	100
MM5.Ratios	100	100	100	0	1	1	1	100	0	0	0	1
MM6.Ratios	1	1	1	100	100	100	100	0	0	0	0	1
MM1.DNA (ng)	60.1	60.1	60.1	6.0	6.0	6.0	0.6	0.6	0.6	0.0	0.0	0.0
MM2.DNA (ng)	0.0	0.0	0.0	0.6	0.6	0.6	6.0	6.0	6.0	60.1	60.1	60.1
MM3.DNA (ng)	0.0	0.0	0.0	0.0	0.0	0.0	9.1	90.5	0.0	9.1	90.5	0.9
MM4.DNA (ng)	0.0	0.0	0.0	0.2	0.0	166.5	0.0	0.0	16.7	0.0	0.0	16.7
MM5.DNA (ng)	49.5	49.5	49.5	0.0	0.5	0.5	0.5	49.5	0.0	0.0	0.0	0.5
MM6.DNA (ng)	0.5	0.5	0.5	49.5	49.5	49.5	49.5	0.0	0.0	0.0	0.0	0.5
MM1.Percentages	30.0%	30.0%	30.0%	3.0%	3.0%	3.0%	0.3%	0.3%	0.3%	0.0%	0.0%	0.0%
MM2.Percentages	0.0%	0.0%	0.0%	0.3%	0.3%	0.3%	3.0%	3.0%	3.0%	30.0%	30.0%	30.0%
MM3.Percentages	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	4.5%	45.3%	0.0%	4.5%	45.3%	0.5%
MM4.Percentages	0.0%	0.0%	0.0%	0.1%	0.0%	83.3%	0.0%	0.0%	8.3%	0.0%	0.0%	8.3%
MM5.Percentages	24.8%	24.8%	24.8%	0.0%	0.3%	0.3%	0.3%	24.8%	0.0%	0.0%	0.0%	0.3%
MM6.Percentages	0.3%	0.3%	0.3%	24.8%	24.8%	24.8%	24.8%	0.0%	0.0%	0.0%	0.0%	0.3%

776