Disentangling bias for non-destructive insect metabarcoding

Francesco Martoni [Corresp., 1] , **Alexander M Piper** [1, 2] , **Brendan C Rodoni** [1, 2] , **Mark J Blacket** [1]

[1] Agriculture Victoria Research, AgriBio Centre for AgriBio Science, State Government Victoria, Bundoora, Victoria, Australia

[2] School of Applied Systems Biology, La Trobe University, Bundoora, Victoria, Australia

Corresponding Author: Francesco Martoni
Email address: francesco.martoni@ecodev.vic.gov.au

A fast and reliable method for obtaining a species-level identification is a fundamental requirement for a wide range of activities, from plant protection and invasive species management to biodiversity assessments and ecological studies. For insects, novel molecular techniques such as DNA metabarcoding have emerged as a rapid alternative to traditional morphological identification, reducing the dependence on limited taxonomic experts. Until recently, molecular techniques have required a destructive DNA extraction, precluding the possibility of preserving voucher specimens for future studies, or species descriptions. Here we paired insect metabarcoding with two recent non-destructive DNA extraction protocols, to obtain a rapid and high-throughput taxonomic identification of diverse insect taxa while retaining a physical voucher specimen. The aim of this work was to explore how non-destructive extraction protocols impact the semi-quantitative nature of metabarcoding, which alongside species presence/absence also provides a quantitative, but biased, representation of their relative abundances. By using a series of mock communities representing each stage of a typical metabarcoding workflow we were able to determine how different morphological (i.e., insect size and hardness) and molecular traits (i.e., DNA extraction and PCR), interact with different protocol steps to introduce quantitative bias into non-destructive metabarcoding results. We discuss the relevance of taxonomic bias to metabarcoding identification of insects and potential approaches to account for it.

1 **Disentangling bias for non-destructive insect**
2 **metabarcoding.**

4 Francesco Martoni[1], Alexander M. Piper[1,2], Brendan C. Rodoni[1,2], Mark J. Blacket[1]
5
6 [1] Agriculture Victoria Research, AgriBio Centre for AgriBio Science, Bundoora 3083, Victoria,
7 Australia
8 [2] School of Applied Systems Biology, La Trobe University, Bundoora 3083, Victoria, Australia
9
10 Corresponding Author:
11 Francesco Martoni
12 5 Ring Road, Bundoora, Victoria, 3083, Australia.
13 Email address: Francesco.Martoni@agriculture.vic.gov.au.

Abstract

16 A fast and reliable method for obtaining a species-level identification is a fundamental
17 requirement for a wide range of activities, from plant protection and invasive species
18 management to biodiversity assessments and ecological studies. For insects, novel molecular
19 techniques such as DNA metabarcoding have emerged as a rapid alternative to traditional
20 morphological identification, reducing the dependence on limited taxonomic experts. Until
21 recently, molecular techniques have required a destructive DNA extraction, precluding the
22 possibility of preserving voucher specimens for future studies, or species descriptions.
23 Here we paired insect metabarcoding with two recent non-destructive DNA extraction protocols,
24 to obtain a rapid and high-throughput taxonomic identification of diverse insect taxa while
25 retaining a physical voucher specimen. The aim of this work was to explore how non-destructive
26 extraction protocols impact the semi-quantitative nature of metabarcoding, which alongside
27 species presence/absence also provides a quantitative, but biased, representation of their relative
28 abundances. By using a series of mock communities representing each stage of a typical
29 metabarcoding workflow we were able to determine how different morphological (i.e., insect
30 size and hardness) and molecular traits (i.e., DNA extraction and PCR), interact with different
31 protocol steps to introduce quantitative bias into non-destructive metabarcoding results. We
32 discuss the relevance of taxonomic bias to metabarcoding identification of insects and potential
33 approaches to account for it.

34

Introduction

36   Species identification is a fundamental pre-requisite for basic and applied ecology. In the
37 field of entomology, species-level identification is required for biodiversity assessments and
38 checklists (Girón & Short 2021), understanding ecology and behavior (Lefort *et al.* 2020),
39 forensic investigation (Pohjoismäki *et al.* 2010), taxonomy (Schutze *et al.* 2017), and
40 management of agricultural pests. Invasive insect species are becoming a major threat to
41 agroecosystems (Paini *et al.* 2016), with biological invasions becoming one of the main menaces
42 to agricultural production (Meyerson & Mooney 2007; Hulme 2009; Chown *et al.* 2014).
43 Therefore, extensive trapping and monitoring activities to detect new insect invasions are carried
44 out both on agricultural properties (Low-Choy 2015) and protected environments, such as
45 National Parks (e.g., Davidovitch *et al.* 2009). This surveillance is leading to an increasing
46 demand of species-level identification for large volumes of insects trapped for plant protection,
47 biosecurity and agriculture (Piper *et al.* 2019).
48   However, availability of taxonomic expertise for insect identification is extremely
49 limited, with often only a few experts worldwide per taxonomic group. Therefore, a range of
50 molecular techniques have been developed to allow more standardized identification of insect
51 species by non-specialists (Piper *et al.* 2019). Most notably, the DNA barcoding technique
52 (Hebert *et al.* 2003) allows comparison of a short standardized genetic region from an
53 unidentified specimen to a vast number of known species deposited in reference databases.
54 Generally, invertebrate barcoding studies have targeted the subunit 1 of the mitochondrial
55 cytochrome oxidase gene (COI) for specimen identification (e.g., Andújar *et al.* 2018; Elbrecht
56 *et al.* 2017; Yu *et al.* 2012). DNA barcoding is regularly applied to identification of undescribed
57 species native to localised areas (Martoni *et al.* 2020), as well as to identification of invasive
58 insect pests (Armstrong & Ball 2005), and is now widely accepted within plant pest diagnostics
59 protocols (EPPO 2021; Ashfaq *et al.* 2016). However, difficulties remain scaling this approach to
60 the sheer numbers of specimens that can be caught in a surveillance trap, or in a field sampling
61 season. Therefore, with the advent of high throughput sequencing technologies, the focus is now
62 shifting from the single specimen sequencing of DNA barcoding to identifying entire
63 communities of species in parallel using DNA metabarcoding (Piper *et al.* 2019 and references
64 therein).
65   The semi-quantitative nature of metabarcoding data has led to concerns around the
66 appropriateness of this technology for surveillance, biomonitoring and assessment of pest
67 pressures (Darling *et al.* 2020; Martins *et al.* 2019). Metabarcoding assays only provide relative
68 abundance data, with sequence reads returned for a species only meaningful relative to the rest of
69 the taxa within the sample (Gloor *et al.* 2017). In addition to quantification issues relating to
70 sample composition, taxonomic biases can be introduced during laboratory processing of
71 samples, including during DNA extraction procedures and genetic marker isolation and
72 amplification. These biases arise due to species-specific differences in morphological and
73 molecular traits which interact with steps of the laboratory protocol to preferentially detect
74 certain taxa at the expense of others. DNA extraction from complex insect communities for

75 metabarcoding analysis has often involved destructive homogenisation of tissues, which results
76 in larger-sized organisms contributing a larger quantity of DNA molecules to the DNA
77 extraction pool than smaller bodied insects (Elbrecht *et al.* 2017). Nevertheless, when individual
78 specimen size is accounted for through size sorting prior to DNA extraction, the influence of
79 primer-template mismatch generally outweighs DNA extraction bias for macroinvertebrates
80 (Braukmann *et al.* 2019; Elbrecht & Leese 2015), at least for destructive homogenisation-based
81 DNA extraction.
82   Recently, non-destructive DNA extraction has emerged as an alternative to
83 homogenisation-based methods in order to retain voucher specimens for morphological
84 confirmation of metabarcoding detections (Carew *et al.* 2018; Martins *et al.* 2019; Batovska *et*
85 *al.* 2021). This is of particular importance in the context of agricultural biosecurity and other
86 regulatory applications of metabarcoding, allowing DNA sequences to be linked to an insect
87 sample, which can be preserved in entomological collections for future records (Martoni *et al.*
88 2021a). While it is well established for homogenised samples that large organisms will have a
89 higher abundance of DNA than small organisms, this may differ for non-destructive
90 metabarcoding, where contribution of DNA may instead depend on surface/volume ratio of
91 insect to extraction buffer (Marquina *et al.* 2019). Furthermore, differences in sclerotization of
92 exoskeletons could affect permeability of DNA and impact detection efficiency (Carew *et al.*
93 2018; Marquina *et al.* 2019). Moreover, with this move from destructive to non-destructive DNA
94 extraction it is unclear if earlier results and assumptions about the bias generating process still
95 hold.
96   Here we compared two different non-destructive DNA extraction methods previously
97 tested for their capability to extract DNA from preserved trapped insects (Martoni *et al.* 2021b).
98 We applied these to mock communities composed of a mix of possible insect pests and harmless
99 by-catch species, following a typical metabarcoding protocol from the recent literature, and using
100 two combinations of degenerate PCR primers. We aimed to measure the taxonomic bias
101 introduced by the DNA extraction, PCR, and library preparation stages, and evaluate the
102 downstream effects on the two main diagnostic-related aspects: sensitivity and quantitation.
103
104

105 ## Materials & Methods

106
107 **Samples and morphological traits**
108   For this study we used adult specimens belonging to 16 insect species (Table 1). Of these,
109 15 species were obtained from insect colonies reared at the AgriBio laboratory of Agriculture
110 Victoria, while another species (*Acizzia* sp.) was field-collected (Table 1). Insect specimens from
111 the colonies were preserved in absolute ethanol and deposited into the Victorian Agricultural
112 Insect Collection (VAIC). Measurements to obtain the volume size were taken from ten
113 individuals for each species and the result was averaged in Table 1. Measurements were taken
114 using the Leica Application Suite software v4.5.0, from five to 20 stacked images obtained using

115   a Leica stereo microscope M205C with a DFC450 camera. For each insect, images were taken
116   from dorsal, lateral, and frontal view in order to obtain a measurement of the volume (length *
117   width * depth) of head, thorax and abdomen. Hardness of the exoskeleton was estimated by
118   attempting to pierce it using forceps, and the 16 species were assigned to three discrete
119   categories (soft, intermediate, hard). Specimens from all taxa were then grouped into eight insect
120   pools, with six pools containing 100 individual insects and the same 15 species, while another
121   two pools also contained a single specimen of the *Bradysia* nr. o*cellaris,* for a total of 16 species,
122   and 101 individuals (Table SM1).
123
124   **Molecular analysis**
125   *Preparation of Insect mock communities.*
126   In order to partition total protocol bias into its constituent steps, mock communities of
127   bulk adult insect, genomic DNA and PCR amplicons were assembled to simulate the input of
128   each major metabarcoding laboratory step.
129   DNA was non-destructively extracted from the eight pools of insects using both the
130   QuickExtract kit (Biosearch Technologies, California, USA), for pools 1-4, and the DNEasy
131   Blood and Tissue kit (Qiagen, Germany; "DNEasy" hereafter), for pools 5-8, following the
132   methods used for single insects in Martoni *et al.* (2021b). These pools, while having a different
133   number of individuals for some of the species, were prepared with an almost identical Order-
134   level composition, taking into account insect biomass and exoskeleton hardness. Non-
135   destructive DNA extraction using QuickExtract was performed as follows: ethanol was removed
136   from the pooled insects and air-dried in tubes for 10 minutes. Five hundred microlitres of
137   QuickExtract were added to the pooled insects, ensuring all insects were submerged, vortexed
138   for 30 seconds, incubated at 65 °C for 6 minutes, vortexed for 15 seconds and incubated at 98 °C
139   for 2 minutes. The supernatant containing the extracted DNA was then transferred to a new
140   Eppendorf tube and stored at -20 °C until PCR amplification.
141   Non-destructive DNA extraction using DNEasy was performed following the first steps
142   of the protocol presented in Martoni *et al.* (2019) and Bahder *et al.* (2015). Briefly, ethanol was
143   removed from the insect pools (as above), insects were then submerged in an ATL
144   buffer/Proteinase K mix with a ratio of 9:1 and then incubated for approximately 17 hours
145   (overnight) at 56 °C. The supernatant was then removed from the insects (as above), and
146   processed further following the manufacturer instructions (i.e., filter column purification and
147   elution). Finally, pooled insect specimens were resuspended in absolute ethanol and preserved.
148   PCR amplification was conducted using either fwhF2
149   (GGDACWGGWTGAACWGTWTAYCCHCC)- fwhR2n
150   (GTRATWGCHCCDGCTARWACWGG) or fwhF2 – HexCOIR4
151   (TATDGTRATDGCHCCNGC), which amplify almost entirely overlapping regions of COI
152   (Vamos *et al.* 2017; Marquina *et al.* 2018). The primer pair fwhF2-fwhR2n targets a 205 bp
153   amplicon (excluding primers) from 346 bp to 551 bp within the conventional COI barcode
154   region, while fwhF2-HexCOIR4 target a 214 bp amplicon (excluding primers) from 346 bp to

155   560 bp. PCR amplification was performed using the Bioline MyFi DNA Polymerase kit
156   (Meridian Bioscience, Ohio, United States of America) using 2.5 µL of DNA template and 1 µL
157   each for the two primers (10 mM) in a 25 µL final volume. The PCR was run with the same
158   cycling conditions for both primer pairs, with an initial 5-minute denaturation at 95 °C, followed
159   by 30 cycles, of denaturation at 95 °C for 45 seconds, annealing at 50 °C for 30 seconds and
160   extension at 72 °C for 30 seconds, followed by a final extension at 72 °C for 7 minutes. PCR
161   amplification was verified on a 1% w/v agarose gel.
162   PCR amplicons were used as template for a second round of PCR to attach Illumina
163   sequencing adapters with unique dual indexes to each sample, using Phusion High-Fidelity DNA
164   Polymerase (New England Biolabs, Massachusetts, USA). PCR conditions were an initial
165   denaturation of 30 seconds at 98 °C followed by 8 cycles of denaturation at 98 °C for 10
166   seconds, annealing at 65 °C for 30 seconds and elongation at 72 °C for 30 seconds. The adapter
167   tailed and indexed amplicons were purified using AMPure XP beads (Beckman Coulter,
168   California, USA) following the manufacturer instructions. Library fragment size (amplicon +
169   adapters) and absence of primer dimers was verified on an Agilent TapeStation (Agilent
170   Technologies, California, USA) and all libraries were equimolarly pooled based on their
171   concentrations as determined by Qubit dsDNA HS Fluorometric Quantification (ThermoFisher
172   Scientific, Massachusetts, USA). As DNA concentrations in negative controls were too low to be
173   measured, they were pooled at the same volume of the lowest concentration mock community
174   library. The final pooled library was then diluted to 7 pM, spiked with 15% PhiX, and sequenced
175   on the Illumina MiSeq platform using the V2 reagent kit (2 x 250 bp reads) (Illumina, California,
176   USA).
177
178   *Preparation of DNA and PCR mock communities.*
179   To assemble mock communities representative of the post-DNA extraction stage (Figure
180   1; "DNA Pools"), DNA for each of the 16 species was destructively extracted separately from 5-
181   20 homogenised individuals (depending on their size) with the DNEasy kit. The DNA from each
182   insect species was then quantified using a Qubit 2.0 Fluorometer (Thermo Fisher Scientific,
183   Massachusetts, USA) and pooled together imitating the composition (% relative abundance) of
184   the original insect pools (Figure 1, "DNA pools"). In order to do this, an arbitrary 4 ng of DNA
185   was used as the unit corresponding to 1 specimen and PCR products were diluted to a
186   concentration of 4 ng/µL so to have a consistent volume for each pool (400 µL) at a consistent
187   concentration (4 ng/µL) (i.e., 4 ng of DNA for each insect composing the original pools).
188   Libraries were prepared from these pools following the remainder of the metabarcoding protocol
189   used for the whole insect pools.
190   To assemble mock communities representative of the post-amplification stage (Figure 1,
191   "PCR Pools"), DNA extracted from each insect species as above was then amplified separately
192   with each primer pair using the same PCR conditions as the whole insect mock communities.
193   The PCR amplicons were then quantified using Qubit 2.0 Fluorometer (Figure 2) and diluted to a
194   concentration of 2 ng/µL so to have a consistent volume for each pool (200 µL) at a consistent

195 concentration (2 ng/µL), as outlined above (i.e., 2 ng of DNA for each insect composing the
196 original pools, to reconstitute the original sample relative abundance), and libraries prepared as
197 above. A second MiSeq run ("Run 2") was performed three months after the first one ("Run 1",
198 which was conducted immediately) containing sequencing libraries generated from the DNA and
199 PCR pools, as well as repeating the original insect pools using the same MiSeq machine and
200 protocol. Run 2 of the insect pools was used to test how well the extracted DNA could be
201 amplified and sequenced following an extended period of storage at -20°C, a typical temperature
202 often used for long term storage of DNA samples, enabling future analysis on the same samples.
203
204   *Bioinformatics analysis*
205   Raw sequence reads were demultiplexed using bcl2fastq v2.2.0 allowing for no
206 mismatches to the expected index combinations (NCBI SRA acc no: PRJNA767112), then
207 trimmed of PCR primer sequences using BBDuK v38.9 (Bushnell *et al.* 2017). Sequence quality
208 profiles were used to further filter reads with >1 expected error (Edgar & Flyvbjerg 2015), or any
209 ambiguous 'N' Bases, then remaining sequences were denoised using DADA2 v1.16 (Callahan
210 *et al.* 2016) with the error model determined separately for each sequencing run. Following
211 denoising, amplicon sequence variants (ASVs) inferred separately from each sequencing run
212 were combined into a single table and any chimeric sequences removed de-novo using the
213 "removeBimeraDenovo" function in DADA2. To further filter any non-specific amplification
214 products and pseudogenes the ASV's were aligned to a profile hidden Markov model (Eddy
215 1998) of the COI barcode region from Piper *et al.* (2021) using the aphid v1.3.3 R package
216 (Wilkinson 2019), retaining sequences that met a minimum log-odds alignment score of 100 with
217 a minimum match length of 100 bp. Retained ASVs were then checked for frame shifts and stop
218 codons that commonly indicate pseudogenes (Roe & Sperling 2007). Taxonomy was determined
219 by aligning ASVs to reference sequences of each taxon used in the mock communities using
220 BLASTn v2.11.0 with a minimum percentage identity of 97% and minimum alignment coverage
221 of 95%. The 52 % of ASVs (0.1% of abundance) that couldn't be accurately mapped were
222 discarded using filtering functions contained in the phyloseq v1.36.0 (McMurdie & Holmes
223 2013) and tidyverse v1.3.1 (Wickham *et al.* 2019) R packages.
224
225   *Statistical analysis*
226   Differences in the number of ASVs detected between DNA extraction protocols, primer
227 sets, and mock communities representing each workflow step were tested for significance using
228 Analysis of Variance (ANOVA), followed by post-hoc pairwise comparisons with Tukey's
229 Honest Significant Difference (HSD) test (Tukey 1977). In order to compare the number of
230 ASVs between each protocol without the confounding effect of differing read depths all samples
231 were rarefied to 100,000 reads before ANOVAs were conducted. Differences in overall
232 quantitative performance between each protocol was measured using the Root Mean Square
233 Error (RMSE) between observed and expected relative abundances, an accuracy metric where
234 smaller values indicate a less biased protocol. Taxonomic bias for each primer and workflow

235   step (Insect pools, DNA pools, amplicon pools) was estimated using a linear model of the
236   compositional error (ratio between expected and observed abundances), with the results
237   geometrically centered to be relative to the 'average' taxon (McLaren *et al.* 2018) and standard
238   errors for each taxon coefficient generated from 1000 bootstrap resamples. The fit of the bias
239   model for the dataset was evaluated by its ability to predict the observed relative abundances
240   from sequencing using the known relative abundances in the mock communities. The separate
241   bias estimates from each workflow step were then used to partition the total protocol bias (that
242   represented by the Insect pools) into 3 different components for each primer set 1) DNA
243   extraction, 2) PCR amplification, and 3) sequencing and bioinformatics as per McLaren *et al.*
244   (2018). The change in abundance of each mock community taxon throughout the metabarcoding
245   protocols relative to its starting abundance (Figure SM1) was then calculated by sequentially
246   multiplying a starting abundance of 1 by the partitioned bias components for that taxon. Finally,
247   the effects of morphological traits (biomass, sclerotization) and molecular traits (GC% of the
248   whole amplicon, primer-template mismatch) on both the partitioned bias estimates and total
249   protocol bias (represented by the Insect pools) was tested for significance using a second linear
250   model fit to the bootstrapped bias estimates, and the relative influence of each trait on detection
251   efficiency determined from its coefficient in the regression model. All statistical analyses were
252   conducted within R 4.0.2 (R Core Team, 2020) using tidymodels v0.1.3 (Kuhn & Wickham
253   2020) packages.
254
255

256   **Results**

257

258   **Comparison between non-destructive DNA extraction methods**

259

260   A total of 7,241,574 and 8,676,552 reads were generated from the first and second MiSeq
261   sequencing runs respectively, consisting of 182 unique ASVs. Each sample received a mean
262   272,600 (± 3,762) reads, ranging from 615,687 reads for the highest sample to 619 reads for the
263   lowest negative control sample, and contained a mean of 26.7 ASVs (± 0.119, range: 11-34).
264   When reads that could not be classified to the known mock community members were removed
265   (e.g., chimeras), the mean reads per sample dropped slightly to 263,581 (± 4,076; range: 619-
266   614,641) and the mean number of ASVs dropped to a mean 14.9 per sample (± 0.03, range: 7-
267   16). Significant differences in the number of inferred ASVs were found between the different
268   protocols (ANOVA; $F_{(3, 60)}$ = 5.11, p = .003), primarily driven by the DNEasy treatments
269   showing significantly more unique ASVs than the QuickExtract (Tukey's HSD; p = .002). On
270   further exploration this difference was found to be due to a substantial dropout of taxa seen in the
271   replicated QuickExtract samples on Run 2, which was run three months later (Figure SM1).
272   When these analyses were repeated without the replicated QuickExtract samples, significant
273   differences were again found between the treatments (ANOVA; ($F_{(3, 52)}$ = 9.78, p < .001), driven
274   by both the QuickExtract and DNEasy treatments having higher numbers of ASVs than the DNA

275   pools (Tukey's HSD; p = .006) and the PCR pools (p = .002) with no significant difference
276   between the QuickExtract and DNEasy samples (p > 0.05). When considering only ASVs that
277   could be classified to species level, significant differences remained between treatments
278   (ANOVA; F(3, 52) = 6.18, p =.001), which pairwise comparisons revealed to be driven by the
279   QuickExtract having significantly higher ASVs than the DNA (Tukey's HSD; p = .002) or PCR
280   pools (p = .001), while no significant differences were found between the DNEasy and any other
281   treatment (p > 0.05). In contrast to the differences in protocols and pools, there were no
282   significant differences between the two primers sets in terms of both total ASVs (ANOVA; $F_{(1,}$
283    $_{54)}$ = 0.21, p > 0.05) or ASVs that could be classified to species level (ANOVA; $F_{(1, 54)}$ = 4.04, p
284   > 0.05).
285
286   **Identification of the sources of bias**
287   Both QuickExtract and DNEasy showed the highest deviation between expected and observed
288   (from sequencing) relative abundances at 16% and 17% RMSE respectively for both primer pairs
289   (Figure 3A) when compared to DNA pools (9% for the fwhF2-fwhR2n primer pair, and 7% for
290   the fwhF2-HexCOI4 primer pair) and the PCR pools (7% and 2%). There was a significant
291   association found between the compositional error (ratio between expected and observed
292   abundances) and species identity across all treatments and primer sets (Supplementary Table 1).
293   When the taxon coefficients estimated by the bias model (Supplementary Table 2) were used to
294   predict the observed relative abundances from the known number of individuals in the mock
295   communities, a substantial improvement in quantitative performance was seen compared to the
296   uncorrected sequencing data (Figure 3A). The model showed a good fit to the DNEasy pools,
297   reducing the RMSE to 2% and 3% for the fwhF2-fwhR2n and fwhF2-HexCOI4 primers
298   respectively, and a near perfect fit to the DNA and PCR pools with 1% RMSE for both primers.
299   While the bias model also reduced the RMSE for the QuickExtract samples, there was still
300   substantial variance seen with an RMSE of 9% for fwhF2-fwhR2n, and 7% for the fwhF2-
301   HexCOI4 (Figure SM2).
302
303   _Taxon-specific biases_
304   The taxonomic biases estimated by the model (Supplementary Table 2) revealed _A._
305   _solanicola_, _A. alternata_, _Acizzia_ sp. and _R. padi_ having the highest efficiency in the QuickExtract
306   and DNEasy pools, and to a lesser extend the DNA pool. On the other hand, the two _Carpophilus_
307   taxa showed the lowest relative efficiency (Figure 3B). The bias estimates for the QuickExtract
308   insect pools showed substantially higher variance compared to the DNEasy insect pools across
309   all taxa, with _Carpophilus truncatus_ showing the largest difference between the two extraction
310   methods (Figure 3B). When the total protocol bias was partitioned into the contribution of each
311   step (Supplementary Table 3), marked differences in bias were seen between protocol steps on
312   the same taxon (Figure 4A). For instance, _Carpophilus truncatus_ saw substantially lower DNA
313   extraction efficiency than the average taxon, but higher than average at the PCR stage. In
314   contrast, _Diuraphis noxia_ showed a higher DNA extraction efficiency at the DNA extraction

315 stage, but a lower efficiency in the later PCR step. The results for the PCR amplification are
316 similarly reflected in the DNA concentrations obtained when single species were amplified to
317 obtain mock communities representing post-PCR processes (Figure 2). Here we see that when
318 individual PCRs are conducted from the same starting DNA concentration, *Diuraphis noxia* has
319 much lower and *Carpophilus truncatus* has much higher DNA concentrations than the majority
320 of the other taxa after both 30 and 40 cycles of amplification. The competing effects of different
321 workflow stages on overall detection efficiency is particularly apparent when looking at the
322 change in abundances of molecules for each taxon throughout the workflow (Figure 4B). Some
323 taxa such as the psyllids *A. solanicola* and *A. alternata* consistently increase in abundance
324 throughout the workflow, while others such as the fruit fly *B. tryoni* saw an initial increase in
325 abundance, followed by a fall to almost equilibrium with the starting relative abundance (Figure
326 4B).
327
328   *Hardness- and biomass-associated bias*
329   Differences in whole protocol bias was significantly associated with insect traits for both
330 the QuickExtract ($F_{(5,54)}$=5.68, R^2_{adj}= 0.26, p<0.001) and DNEasy ($F_{(5,58)}$=20.50, adjusted
331 R^2=0.60, p<0.001) protocols (Figure 5A). The efficiency with which a species was detected in a
332 sample was positively influenced by whether that species had a soft (QuickExtract; $\beta = 35.53$,
333 95% CI [11.86, 22.10], DNEasy; $\beta = 15.88$, 95% CI [11.86, 22.10]), or intermediate hardness
334 exoskeleton (QuickExtract; $\beta = 19.13$, 95% CI [6.09, 62.83], DNEasy; $\beta = 8.55$, 95% CI[6.65,
335 11.46]), or high amplicon GC% (QuickExtract; $\beta = 31.93$, 95% CI [7.53, 82.098], DNEasy; $\beta =$
336 13.71, 95% CI [9.90, 19.57])(Figure 5A). Insects with a hard exoskeleton showed no increase in
337 detection efficiency for the DNEasy protocol ($\beta = 0.12$, 95% CI [-0.68, 0.86]), but increased
338 efficiency for the QuickExtract protocol although with a confidence interval overlapping zero (β
339 $= 2.29$, 95% CI [-0.43, 8.99]).Insect biomass unexpectedly had a strong negative effect on
340 detection efficiency for both protocols (QuickExtract; $\beta = -12.02$, 95% CI [-41.50, -3.07],
341 DNEasy; $\beta = -4.73$, 95% CI [-6.87, -3.40]), exceeded only by primer mismatch (QuickExtract; β
342 $= -23.20$, 95% CI [-82.01, -5.84], DNEasy; $\beta = -8.50$, 95% CI [-12.66, -5.63)(Figure 5A).
343 Generally, biases associated with each trait were more unpredictable (i.e., with larger variances
344 observed) when using QuickExtract ($R^2_{adj} = 0.26$) compared with DNEasy ($R^2_{adj} = 0.60$)
345 extraction method (Figure 5A). When considering just the bias contributed by the DNA
346 extraction stage (Figure 5B), insect traits were also associated with detection efficiency albeit to
347 a lesser degree than the whole protocol (QuickExtract; $F_{(5,54)} = 8.79$, $R^2_{adj} = 0.38$, p<0.001,
348 DNEasy; $F_{(5,58)} = 32.60$, $R^2_{adj} = 0.70$, p<0.001). Most notably, the relationship to insect biomass
349 was reversed when considering just the DNA extraction stage, with larger biomass showing a
350 slight positive effect on detection efficiency for both protocols (QuickExtract; $\beta = 0.94$, 95% CI
351 [0.20, 1.81], DNEasy; $\beta = 2.76$, 95% CI [2.15, 3.51]) (Figure 5B). A soft (QuickExtract; $\beta =$
352 1.78, 95% CI [1.11, 2.97], DNEasy; $\beta = 1.85$, 95% CI [1.56, 2.20]) or intermediate hardness
353 exoskeleton (QuickExtract; $\beta = 2.56$, 95% CI [1.85, 3.51], DNEasy; $\beta = 2.79$, 95% CI [2.50,
354 3.10]) again increased detection efficiency during DNA extraction, while a hard exoskeleton,

355 primer mismatch, or amplicon GC% all had confidence intervals overlapping zero for both
356 protocols (Figure 5B). When considering only the bias contributed by the PCR stage, the overall
357 effect of insect traits on detection efficiency was more comparable between protocols
358 (QuickExtract $F_{(5,54)}$=7.13, R^2_{adj} = 0.34, p<0.001, DNEasy; $F_{(5,58)}$7.65, R^2_{adj} = 0.35, p<0.001)
359 (Figure 5C). As expected, primer mismatch showed the strongest negative effect on detection
360 efficiency (QuickExtract; β = -3.18, 95% CI [-3.08, -2.83], DNEasy; β = -3.08, 95% CI [-3.44, -
361 2.83]), while amplicon GC% increased detection efficiency (QuickExtract; β = 3.15, 95% CI
362 [2.65, 3.83], DNEasy; β = 3.14, 95% CI [2.86, 3.50]). All values of exoskeleton hardness showed
363 a slight positive effect on detection efficiency at the PCR stage across both protocols (β 2.37 -
364 3.09), while biomass decreased efficiency (QuickExtract; β = -0.91, 95% CI [-1.16, -0.67],
365 DNEasy; β = -0.92, 95% CI [-1.06, -0.79]).

366

367

368 ## Discussion

369   Here, using non-destructive metabarcoding approaches, we successfully recorded all
370 insect species present in pools composed of 100-101 individuals, including those species that
371 were represented by just a single individual insect. At the same time, the use of these non-
372 destructive DNA extraction methods allowed morphological voucher specimens of the insects to
373 be preserved, as previously demonstrated (Martoni *et al.* 2019, Batovska *et al.* 2021). This is of
374 paramount importance for many regulatory applications of metabarcoding, where retaining
375 voucher specimens of potential pests or indicator taxa can be required for legal reasons, or to
376 simply provide a morphological specimen preserved in an entomological collection for future
377 taxonomic investigation (i.e., Martins *et al.* 2019, Martoni *et al.* 2019, Martoni *et al.* 2021a).
378   The results presented here show that non-destructive metabarcoding analysis can be
379 successfully applied to bulk samples of agriculturally relevant insects to obtain a species
380 identification, in agreement with recent studies (Carew *et al.* 2018; Nielsen *et al.* 2019; Batovska
381 *et al.* 2021). This suggests that non-destructive metabarcoding has potential applications not only
382 for biodiversity assessments but also for diagnostics and biosecurity purposes, to determine
383 presence/absence of pests in bulk traps. Our study further highlights the molecular and
384 morphological traits that affect quantification of individual insect species within non-
385 destructively extracted bulk insect pools.

386

387 **Metabarcoding bias and non-destructive DNA extractions**
388   The output generated by metabarcoding analyses is compositional data (Gloor *et al.*
389 2017). This means the relationship between the starting total abundance of a species and the
390 output counts of sequence reads is completely lost, and the sequence reads returned for a taxon
391 are only meaningful relative to the rest of the taxa within its sample (Gloor *et al.* 2017).
392 Furthermore, the relationship between counts of sequence reads and the individuals they arise
393 from is affected by a number of biases that systematically distort the measured sequence counts
394 of each species from their true abundances (McLaren *et al.* 2019). This bias can lead to

395 taxonomic dropouts by diluting certain taxa below the detection limit and brings additional
396 challenges for quantifying the number of individual specimens in a sample from metabarcoding
397 sequences alone.
398   Amongst the known sources of metabarcoding bias, a range of physical characteristics of
399 the insect community under study plays a key role, with perhaps the most obvious of these being
400 the large variation in specimen body sizes within insect community assemblages (Chown &
401 Gaston 2010). For example, even assuming that all available DNA is always extracted from a
402 sample, the amount of DNA (and the reads obtained from it) from a single insect will depend on
403 its biomass, which in turn depends on its species or life stage. Consequently, when aiming to use
404 metabarcoding as a semi-quantitative technique, it is important to remember that the reads
405 obtained from a single insect will vary depending on a number of its characteristics (Thomas *et*
406 *al.* 2016; McLaren *et al.* 2019), as well as on the composition of the insect pool analysed (Gloor
407 *et al.* 2017).
408   With non-destructive DNA extractions, an additional factor to consider is that non-
409 destructive protocols mostly act upon the external surface of the insect exoskeletons, as opposed
410 to destructive methods that can potentially access all of the DNA contained in the insect bodies.
411 Therefore, when comparing the number of reads obtained from the insect pools with that from
412 the DNA pools two additional factors had to be taken into consideration in addition to sample
413 composition: exoskeleton robustness and species biomass. When comparing these factors, a soft
414 exoskeleton had the strongest positive effect on relative efficiency, suggesting that soft-bodied
415 insects facilitate non-destructive DNA extraction, resulting in a higher representation than one
416 would expect from their relative biomass within the sample. Whether this proves beneficial or
417 problematic for species detection will depend on the specific communities a surveillance
418 programme is targeting. On the one hand, the smaller influence of species biomass could
419 alleviate the requirement for any morphological pre-sorting to ensure large specimens don't
420 drown out smaller ones (Elbrecht *et al.* 2017). However, the negative effect of hard exoskeletons
421 on detection efficiency could produce false negatives for low abundance taxa with a high level of
422 sclerotization. This is the most likely explanation for the differences in extraction efficiency
423 between taxa such as *Carpophilus truncatus,* having a highly sclerotized exoskeleton, in
424 comparison to the more soft-bodied insects such as *Diuraphis noxia* (Figure 5A).
425   These results were consistent between both the QuickExtract and DNEasy kit protocols,
426 suggesting that the bias towards softer bodied insects is an inherent aspect of the non-destructive
427 DNA extraction process itself, rather than a specific protocol or kit. Nevertheless, there are a
428 few key differences between the two extraction methods we evaluated that will influence their
429 practical application within future studies. Firstly, the QuickExtract method was substantially
430 faster, requiring only one hour of operator's time as opposed to the overnight incubation of the
431 DNEasy method. However, QuickExtract produced unpredictable variation in DNA extraction
432 performance and appeared less suitable for long-term DNA preservation. When the DNA
433 extracts from the insect mock communities were re-sequenced three months after extraction –
434 during which the templates were kept in -20 °C freezer – the QuickExtract template appeared to

435 have degraded dramatically, resulting in a number of species dropping out. On the other hand,
436 the presence/absence of species from the DNEasy kit was virtually identical to when they were
437 sequenced within days of being extracted. Therefore, DNA extraction products obtained with the
438 QuickExtract kit should be used immediately for analysis, or stored at lower temperatures, i.e., in
439 a -80 °C freezer (as suggested in the manufacturer's recommendations) if possible.
440   In addition, the results showed relative abundance measurements obtained through non-
441 destructive metabarcoding assays are most strongly influenced by the DNA extraction processes.
442 DNA extraction was the not just the largest contributor to protocol bias, but also had the highest
443 variance of all the tested workflow stages, making it less predictable. Nevertheless, it is unclear
444 if this aspect is specific to the non-destructive DNA extraction process or is applicable to DNA
445 extractions generally. Few metabarcoding studies have attempted to partition the total protocol
446 bias into separate components for each major protocol step. It would be valuable for future
447 research to compare the quantitative results of destructive and non-destructive extractions, using
448 a similar bias-partitioning approach to that implemented within our study. The non-destructive
449 DNA extraction bias issues highlighted here may turn out to be unavoidable when voucher
450 specimens are required, as destructive DNA extraction is not an option.
451
452 **Primer mismatch and PCR bias**
453   Beside DNA extraction, we wanted to assess if bias was introduced also at the PCR stage.
454 Mismatches between PCR primers and template molecules have been considered the primary
455 contributor to metabarcoding bias, particularly at the 3' end of the primer where nucleic acid
456 extension takes place (Piñol *et al.* 2015). Primer-template mismatch can be particularly
457 problematic for protein coding genes such as COI, due to natural degeneracy in the genetic code
458 leaving no strictly conserved gene regions for design of universal PCR primers (Deagle *et al.*
459 2014). This has necessitated the inclusion of multiple degenerate nucleotide bases in
460 metabarcoding primers in order to mitigate the effects of mismatch on detection efficiency and
461 abundance estimates (Elbrecht *et al.* 2019).
462   While much of the previous literature has highlighted that bias introduced during PCR
463 amplification is responsible for the semi-quantitative nature of metabarcoding results, our study
464 observed a smaller contribution of PCR compared to the DNA extraction process. While this
465 could potentially be attributed to non-destructive DNA extraction being more bias prone than its
466 homogenisation-based alternatives, this will require further experiments to demonstrate.
467 Alternatively, the smaller contribution of the PCR process to total protocol bias could have been
468 due to the two primer pairs employed in this study containing a number of degenerate bases,
469 having been designed to be generic for insects (Vamos *et al.* 2017; Marquina *et al.* 2018;
470 Elbrecht *et al.* 2019). Previous studies that have found high levels of PCR bias have mostly used
471 less-degenerate primers or primers which have since been identified to contain critical design
472 flaws (Elbrecht & Leese 2017; Piper *et al.* 2021). Nevertheless, some species included in our
473 study still showed mismatch to these generic primers (Table 1), and this was a primary driver of
474 reduced efficiency at the PCR stage (Figure 4C).

475   Additionally, primer bias was confirmed when quantifying the PCR product obtained
476   from each insect species in separate PCRs, where each primer pair amplified the DNA of some
477   species up to ten times more than others. Since this DNA had been normalized prior to PCR,
478   variation in subsequent concentration could be linked bias introduced during amplification. Such
479   variation in the concentration appeared to also depend on the number of PCR cycles performed,
480   with a higher number of cycles showing more similar concentrations across different insect
481   species. Nevertheless, this bias did not impact the presence/absence of targets, with both primer
482   pairs tested here recording all of the species present in all the pools, with a sensitivity of up to 1
483   in 101 for many of the species tested. Further work could be required to test if primer bias may
484   have greater effects on presence/absence in larger communities, or at a lower sequencing depth,
485   as well as if additional primer pairs might produce a lower primer bias. Nonetheless, it is
486   important to remember that, in a diagnostic context, priority is given to a precise
487   presence/absence assessment, especially when testing for the presence of unexpected pests or
488   rare species. Therefore, for general insect biodiversity monitoring and surveillance, the semi-
489   quantitative but taxonomically broader screening of agricultural traps or biodiversity assessment
490   samples that can be performed using MiSeq metabarcoding, could be considered the preferred
491   approach over a more quantitative but perhaps more time-consuming method, such as qPCR,
492   targeting each individual species separately.
493
494   **Possible strategies to account for metabarcoding bias.**
495   When considering the number of reads generated as output of the MiSeq metabarcoding,
496   the relative abundance of reads obtained for each insect species was correlated with the number
497   of individuals present in the pool but biased toward certain species. Following the multiplicative
498   model of McLaren *et al.* (2019), we used mock communities representing each step of the library
499   preparation pipeline to partition the total protocol bias into that explained by each component.
500   Indeed, the relative abundances obtained from different insect species appeared to be subject to
501   bias during the different steps of the laboratory workflow, including DNA extraction and PCR,
502   fundamentally due to both morphological and molecular traits differing across each insect group.
503   Amongst these, some of the morphological traits determined here to significantly affect the
504   metabarcoding results are surface, size and consistency of the insect tissues, as previously
505   reported for Coleoptera (Martins *et al.* 2019). On the other hand, molecular traits include PCR
506   primers mismatch, which was observed here for both the primer pairs used, despite these being
507   considered generic primers (Vamos *et al.* 2017; Marquina *et al.* 2018). In the literature, attempts
508   have been made to improve a quantitative output for metabarcoding, either by developing new
509   primers, or protocols (e.g., Elbrecht & Leese 2015; Deagle *et al.* 2019; Lamb *et al.* 2019). Our
510   results indicate that understanding the individual contributions of each laboratory stage, rather
511   than just the total overall protocol bias, should be a critical consideration for future efforts. This
512   was particularly apparent when measuring the effect of biomass on detection efficiency, as when
513   the total protocol bias was considered, larger biomass appeared to have a strong negative effect
514   on the number of reads for each given species. On the other hand, when considering just the bias

515   introduced by the DNA extraction stage, larger insect biomass increased the number of reads
516   produced, as expected from previous studies. Therefore, when optimizing protocols researchers
517   should be wary of the often-contrasting effects of the different laboratory steps, in order to not be
518   confounded into mistakenly attempting to optimize one aspect, when in reality the majority of
519   the bias is being introduced through a different laboratory step.
520   While efforts to further optimise protocols will no doubt prove important for increasing
521   the quantitative performance of metabarcoding, a less explored but complementary approach is
522   to use statistical models to actively correct for bias during analysis (Thomas *et al.* 2016;
523   Krehenwinkel *et al.* 2017; McLaren *et al.* 2019). Promisingly, we found bias to be highly
524   predictable with our model, demonstrating the potential for developing correction factors to
525   improve the quantitative results of metabarcoding assay for specific important target species,
526   e.g., agricultural pests. However, even with our consistently treated mock communities, DNA
527   extraction bias had the highest variance, and thus would be the most difficult factor to model and
528   correct for. In a surveillance situation, this could be exacerbated by differences in environmental
529   conditions between real trap samples and the mock communities used to derive the correction
530   factors introducing further bias (Krehenwinkel *et al.* 2018). Furthermore, developing correction
531   factors from mock communities requires knowing the organisms that will likely be encountered
532   *a priori*, as well as being able to acquire specimens of them. Unfortunately, in both contexts of
533   biodiversity monitoring and surveillance, knowing *a priori* what organisms will be recorded in
534   order to prepare targeted spike-ins or tailored mocks community could be challenging and, often,
535   defy the very purpose of the study.
536

Conclusions

538   Based on the results obtained here, different insect species, even belonging to the same
539   order or genus, are subject to different degrees of bias during a non-destructive metabarcoding
540   workflow. These biases are driven by species-specific morphological and molecular traits that
541   interact with different protocol steps to either increase or decrease detection efficiency. If bias
542   can be measured for that taxon *a priori* using mock communities, correction factors could
543   therefore be used to calibrate the results to better reflect the actual abundances, but requires
544   knowledge of the sample composition prior to analysis.
545   Ultimately, if the surveillance program or the biodiversity assessment are conducted
546   long-term using a non-destructive DNA extraction and thus retaining voucher specimens,
547   obtaining specimens to develop correction factors covering a wider diversity may eventually be
548   achievable. Inclusion of mock communities in the metabarcoding analysis would not only be
549   possible, but would sensibly improve in accuracy based on an ongoing monitoring of the
550   geographical area of interest. This highlights the importance of long-term monitoring and
551   biodiversity data collection projects not only for agricultural and industrial areas but also for
552   adjacent ecosystems that can contribute to the species diversity recorded during insect trapping
553   programs.
554

555
556 **Acknowledgements**

557   The authors thank PeerJ editor, Dr Sheila Colla, and two anonymous reviewers for their
558 useful comments and suggestions. The authors would like to thank the iMapPESTS project,
559 supported by Horticulture Innovation Australia (ST16010) through funding from the Australian
560 Government Department of Agriculture as part of its Rural R&D for Profit program and Grains
561 Research and Development Corporation, for funding these experiments and the publication of
562 this article. We thank Isabel Valenzuela, Jessi Henneken, and Will Boston (Agriculture Victoria)
563 for providing many of the laboratory colony insect specimens used in this study.

564
565
566 **Data availability**

567   Raw sequence data is available at NCBI SRA acc no: PRJNA767112. All code and
568 additional files required to reproduce the bioinformatic and statistical analyses presented in this
569 manuscript are available at the following GitHub repository:
570 https://github.com/alexpiper/disentangling_metabarcoding_bias_MS
571

References

- Andújar C, Arribas P, Yu DW, Vogler AP & Emerson BC (2018) Why the COI barcode should be the community DNA metabarcode for the metazoa. *Molecular Ecology*, 27: 3968–3975. https://doi.org/10.1111/mec.14844.

- Armstrong KF. & Ball SL. (2005) DNA barcodes for biosecurity: invasive species identification. *Philosophical transactions of the Royal Society B,* 360(1462): 1813–1823. https://doi.org/10.1098/rstb.2005.1713.

- Ashfaq M, Hebert PDN & Naaum A. (2016) DNA barcodes for bio-surveillance: Regulated and economically important arthropod plant pests. *Genome*, 59(11): 933–945. https://doi.org/10.1139/gen-2016-0024.

- Bahder BW, Bollinger L, Sudarshana MR & Zalom FG. (2015) Preparation of mealybugs (Hemiptera: Pseudococcidae) for genetic characterization and morphological examination. *Journal of Insect Science,* 15: 104. https://doi.org/10.1093/jisesa/iev086.

- Batovska J, Piper AM, Valenzuela I, Cunningham JP & Blacket MJ. (2021) Developing a Non-destructive Metabarcoding Protocol for Detection of Pest Insects in Bulk Trap Catches. *Scientific Reports*, 11, 7946. doi:10.1038/s41598-021-85855-6

- Braukmann TW, Ivanova NV, Prosser SW, Elbrecht V, Steinke D, Ratnasingham S, de Waard JR, Sones JE, Zakharov EV & Hebert PDN (2019) Metabarcoding a diverse arthropod mock community. *Molecular ecology resources*, 19(3): 711–727. https://doi.org/10.1111/1755-0998.13008.

- Bushnell B, Rood J, & Singer E. (2017) BBMerge – Accurate paired shotgun read merging via overlap. *PLOS ONE,* 12: 1–15. https://doi.org/10.1371/journal.pone.0185056.

- Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP (2016) DADA2: High-resolution sample inference from Illumina amplicon data. *Nature Methods,* 13: 581–583. https://doi.org/10.1038/nmeth.3869.

- Carew ME, Coleman RA & Hoffmann AA (2018) Can non-destructive DNA extraction of bulk invertebrate samples be used for metabarcoding? *PeerJ,* 6: e4980. https://doi.org/10.7717/peerj.4980.

- Chown SL, Hodgins KA, Griffin PC, Oakeshott, J.G., Byrne, M. & Hoffmann, A.A. (2014) Biological invasions, climate change and genomics. *Evolutionary Applications*, 8(1): 23–46. https://doi.org/10.1111/eva.12234.

- Chown SL & Gaston KJ. (2010) Body size variation in insects: a macroecological perspective. *Biological Reviews*, 85(1): 139–169. https://doi.org/10.1111/j.1469-185X.2009.00097.x.

- Darling JA, Pochon X, Abbott CL, Inglis GJ & Zaiko A. (2020) The risks of using molecular biodiversity data for incidental detection of species of concern. *Diversity and Distribution,* 26(9): 1116–1121. https://doi.org/10.1111/ddi.13108.

610 • Davidovitch L, Stoklosa R, Majer J, Nietrzeba A, Whittle P, Mengersen K & Ben-Haim
611 Y. (2009) Info-gap theory and robust design of surveillance for invasive species: The
612 case study of Barrow Island. *Journal of Environmental Management*, 90(8): 2785–2793.
613 https://doi.org/10.1016/j.jenvman.2009.03.011.

614 • Deagle BE., Jarman SN, Coissac E, Pompanon F & Taberlet P. (2014) DNA
615 metabarcoding and the cytochrome c oxidase subunit I marker: Not a perfect match.
616 *Biology Letters*, 10(9), 20140562. https://doi.org/10.1098/rsbl.2014.0562.

617 • Eddy SR. (1998) Profile hidden Markov models. *Bioinformatics* (Oxford, England),
618 14(9): 755–763.

619 • Edgar RC & Flyvbjerg H. (2015) Error filtering, pair assembly and error correction for
620 next-generation sequencing reads. *Bioinformatics*, 31(21): 3476–3482.
621 https://doi.org/10.1093/bioinformatics/btv4.

622 • Elbrecht, V., & Leese, F. (2017). Validation and Development of COI Metabarcoding
623 Primers for Freshwater Macroinvertebrate Bioassessment. Frontiers in Environmental
624 Science, 5, 11. https://doi.org/10.3389/fenvs.2017.00011

625 • Elbrecht V & Leese F. (2015) Can DNA-based ecosystem assessments quantify species
626 abundance? Testing primer bias and biomass—sequence relationships with an innovative
627 metabarcoding protocol. *PLOS ONE*, 10(7).
628 https://doi.org/10.1371/journal.pone.0130324.

629 • Elbrecht V, Peinert B & Leese F. (2017). Sorting things out: Assessing effects of unequal
630 specimen biomass on DNA metabarcoding. *Ecology and evolution,* 7(17): 6918–6926.
631 https://doi.org/10.1002/ece3.3192.

632 • Elbrecht V, Braukmann TWA, Ivanova NV, Prosser SWJ, Hajibabaei M, Wright M,
633 Zakharov EV, Hebert PDN & Steinke D. (2019) Validation of COI metabarcoding
634 primers for terrestrial arthropods. *PeerJ,* 7, e7745. https://doi.org/10.7717/peerj.7745.

635 • EPPO. (2021). PM 7/129 (2) DNA barcoding as an identification tool for a number of
636 regulated pests. EPPO Bulletin, 51(1), 100–143. https://doi.org/10.1111/epp.12724

637 • Girón JC & Short AEZ (2021) The Acidocerinae (Coleoptera, Hydrophilidae): taxonomy,
638 classification, and catalog of species. *ZooKeys*, 1045: 1–236.
639 https://doi.org/10.3897/zookeys.1045.63810.

640 • Gloor GB, Macklaim JM, Pawlowsky-Glahn V & Egozcue JJ (2017) Microbiome
641 datasets are compositional: and this is not optional. *Frontiers in microbiology*, 8: 2224.
642 https://doi.org/10.3389/fmicb.2017.02224.

643 • Hebert PDN, Cywinska A, Ball SL & DeWaard JR (2003) Biological identifications
644 through DNA barcodes. *Procedings of the Roayal Society B -Biological Sciences,* 270:
645 313–321. https://doi.org/10.1098/rspb.2002.2218.

646 • Hulme PE (2009) Trade, transport and trouble: Managing invasive species pathways in
647 an era of globalization. *Journal of Applied Ecology*, 46: 10–18.
648 https://doi.org/10.1111/j.1365-2664.2008.01600.x.

649 • Krehenwinkel H, Wolf M, Lim JY, Rominger AJ, Simison WB & Gillespie RG. (2017)
650 Estimating and mitigating amplification bias in qualitative and quantitative arthropod
651 metabarcoding. *Scientific Reports*, 7, 17668. https://doi.org/10.1038/s41598-017-17333-
652 x.

653 • Krehenwinkel H, Fong M, Kennedy S, Huang EG, Noriyuki S, Cayetano L & Gillespie
654 R. (2018) The effect of DNA degradation bias in passive sampling devices on
655 metabarcoding studies of arthropod communities and their associated microbiota. *PLOS
656 ONE*, 13(1), e0189188. https://doi.org/10.1371/journal.pone.0189188.

657 • Kuhn M & Wickham H. (2020) Tidymodels: a collection of packages for modeling and
658 machine learning using tidyverse principles. Available at: https://www.tidymodels.org.

659 • Lamb PD, Hunter E, Pinnegar JK, Creer S, Davies RG & Taylor MI. (2019) How
660 quantitative is metabarcoding: A meta-analytical approach. *Molecular ecology,* 28(2):
661 420–430. https://doi.org/10.1111/mec.14920.

662 • Lefort M-C, Beggs JR, Glare TR, Saunders TE, Doyle EJ & Boyer S (2020) A molecular
663 approach to study Hymenoptera diets using wasp nests. *NeoBiota*, 63: 57–79.
664 https://doi.org/10.3897/neobiota.63.58640.

665 • Low-Choy S. (2015*) Getting the story straight: Laying the foundations for statistical
666 evaluation of the performance of surveillance*. In: Jarrad F, Low-Choy S, Mengersen K,
667 eds. Biosecurity Surveillance: Quantitative Approaches. 6th ed.CABI; 43–73.

668 • Marquina D, Andersson AF & Ronquist, F. (2018) New mitochondrial primers for
669 metabarcoding of insects, designed and evaluated using in silico methods. *Molecular
670 Ecology Resources,* 19(1): 90–104. https://doi.org/10.1111/1755-0998.12942.

671 • Marquina D, Esparza-Salas R, Roslin T & Ronquist F. (2019) Establishing arthropod
672 community composition using metabarcoding: Surprising inconsistencies between soil
673 samples and preservative ethanol and homogenate from Malaise trap catches. *Molecular
674 Ecology Resources,* 19(6): 1516–1530. https://doi.org/10.1111/1755-0998.13071.

675 • Martins FMS, Galhardo M, Filipe AF, Teixeira A, Pinheiro P, Paupério J, Alves PC &
676 Beja P. (2019) Have the cake and eat it: Optimizing nondestructive DNA metabarcoding
677 of macroinvertebrate samples for freshwater biomonitoring. *Molecular ecology
678 resources,* 19(4): 863–876. https://doi.org/10.1111/1755-0998.13012.

679 • Martoni F, Valenzuela I & Blacket M. (2019) Non-destructive DNA extractions from fly
680 larvae (Diptera: Muscidae) enable molecular identification of species and enhance
681 morphological features. *Austral Entomology,* 56(4): 848–856.
682 https://doi.org/10.1111/aen.12419.

683 • Martoni F, Taylor GS, Blacket M. (2020) Illuminating insights into the biodiversity of the
684 Australian psyllids (Hemiptera: Psylloidea) collected using light trapping. *Insects*, 11(6):
685 354. https://doi.org/10.3390/insects11060354.

686 • Martoni F, Valenzuela I & Blacket M. (2021a) On the complementarity of DNA
687 barcoding and morphology to distinguish benign endemic insects from possible pests: the

688   case of *Dirioxa pornia* and the tribe Acanthonevrini (Diptera: Tephritidae: Phytalmiinae)
689   in Australia. *Insect Science*, 28(1): 261-270 https://doi.org/10.1111/1744-7917.12769.

690   • Martoni F, Nogarotto E, Piper AM, Mann R, Valenzuela I, Eow L, Rako L, Rodoni BC,
691   Blacket MJ. (2021b) Propylene Glycol and Non-Destructive DNA Extractions Enable
692   Preservation and Isolation of Insect and Hosted Bacterial DNA. Agriculture. 2021;
693   11(1):77. https://doi.org/10.3390/agriculture11010077

694   • McLaren MR, Willis AD & Callahan BJ (2019) Consistent and correctable bias in
695   metagenomic sequencing experiments. *eLife*, 8:e46923.
696   https://doi.org/10.7554/eLife.46923.

697   • McMurdie PJ & Holmes S. (2013) phyloseq: An R Package for Reproducible Interactive
698   Analysis and Graphics of Microbiome Census Data. *PLOS ONE,* 8, e61217.
699   https://doi.org/10.1371/journal.pone.0061217.

700   • Meyerson LA & Mooney HA. (2007) Invasive alien species in an era of globalization.
701   *Frontiers in Ecology and the Environment*, 5: 199–208. https://doi.org/10.1890/1540-
702   9295(2007)5[199:IASIAE]2.0.CO;2.

703   • Nielsen M, Gilbert MTP, Pape T, Bohmann K (2019) A simplified DNA extraction
704   protocol for unsorted bulk arthropod samples that maintains exoskeletal integrity.
705   *Environmental DNA*, 1(2), 144–154. https://doi.org/10.1002/edn3.16.

706   • Paini DR, Sheppard AW, Cook DC, De Barro, PJ, Worner S & Thomas MB (2016)
707   Global threat to agriculture from invasive species. *Proceedings of the National Academy
708   of Science USA,* 113: 7575–9. https://doi.org/10.1073/pnas.1602205113.

709   • Piñol J, Mir G, Gomez-Polo P & Agustí N. (2015) Universal and blocking primer
710   mismatches limit the use of high-throughput DNA sequencing for the quantitative
711   metabarcoding of arthropods. *Molecular Ecology Resources*, 15(4): 819–830.
712   https://doi.org/10.1111/1755-0998.12355.

713   • Piper A, Batovska J, Cogan NOI, Weiss J, Cunningham JP, Rodoni BC & Blacket MJ.
714   (2019) Prospects and challenges of implementing DNA metabarcoding for high-
715   throughput insect surveillance. *GigaScience*, 8: 1–22.
716   https://doi.org/10.1093/gigascience/giz092.

717   • Piper AM, Cogan NOI, Cunningham JP & Blacket MJ. (2021) Computational Evaluation
718   of DNA Metabarcoding for Universal Diagnostics of Invasive Insect Pests. BioRxiv.
719   https://doi.org/10.1101/2021.03.16.435710.

720   • Pohjoismäki JLO, Karhunen PJ, Goebeler S, Saukko P & Sääksjärvi IE. (2010) Indoors
721   forensic entomology: colonization of human remains in closed environments by specific
722   species of sarcosaprophagous flies. *Forensic Science International*, 199 (1–3): 38–42.
723   https://doi.org/10.1016/j.forsciint.2010.02.033.

724   • R Core Team (2020) R: A language and environment for statistical computing. R
725   Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

726   • Roe AD & Sperling FA. (2007) Patterns of evolution of mitochondrial cytochrome c
727   oxidase I and II DNA and implications for DNA barcoding. *Molecular Phylogenetics and*
728   *evolution,* 44(1): 325–345. https://doi.org/10.1016/j.ympev.2006.12.005.

729   • Schutze MK, Bourtzis K, Cameron SL, Clarke AR, De Meyer M, Hee AKW, Hendrichs
730   J, Krosch MN & Mwatawala M. (2017). Integrative taxonomy versus taxonomic
731   authority without peer review: the case of the Oriental fruit fly, *Bactrocera dorsalis*
732   (Tephritidae). *Systematic Entomology*, 42: 609–620. https://doi.org/10.1111/syen.12250.

733   • Thomas AC, Deagle BE, Eveson JP, Harsch CH & Trites AW. (2016) Quantitative DNA
734   metabarcoding: Improved estimates of species proportional biomass using correction
735   factors derived from control material. *Molecular Ecology Resources*, 16(3): 714–726.
736   https://doi.org/10.1111/1755-0998.12490.

737   • Tukey JW (1977) Exploratory data analysis. Addison-Wesley, Reading.

738   • Vamos EE, Elbrecht V & Leese F. (2017) Short COI markers for freshwater
739   macroinvertebrate metabarcoding. *Metabarcoding and Metagenomics*, 1: e14625.
740   https://doi.org/10.3897/mbmg.1.14625.

741   • Wickham H, Averick M, Bryan J, Chang W, D'Agostino McGowan L, François R,
742   Grolemund G, Hayes A, Henry L, Hester J, Kuhn M, Pedersen TL, Miller E, Bache SM,
743   Müller K, Ooms J, Robinson D, Seide DP, Spinu V, Takahashi K, Vaughan D, Wilke C,
744   Woo K & Yutani, H.(2019). Welcome to the Tidyverse. *Journal of Open Source*
745   *Software*, 4(43): 1686. https://doi.org/10.21105/joss.01686.

746   • Wilkinson, S. (2019). aphid: an R package for analysis with profile hidden Markov
747   models. Bioinformatics, 35(19), 3829–3830.
748   https://doi.org/10.1093/bioinformatics/btz159

749   • Yu DW, Ji Y, Emerson BC, Wang X, Ye C, Yang C & Ding Z. (2012) Biodiversity soup:
750   metabarcoding of arthropods for rapid biodiversity assessment and biomonitoring.
751   *Methods in Ecology and Evolution*, 3: 613-623. https://doi.org/10.1111/j.2041-
752   210X.2012.00198.x.

753
754
755

756 **Figures**

757

758 **Figure 1:** Workflow of the experiment for the three types of pools. Insect pools are the

759 result of non-destructive DNA extractions from the pooled insect specimens. DNA pools

760 are the result of pooled DNA that was destructively extracted from each insect species.

761 PCR pools are the result of DNA that was extracted, quantified, and amplified separately

762 from each insect species then pooled before indexing qPCR. All pools were indexed,

763 sequenced and analysed following the same protocol. Each pool was amplified using two

764 primer sets.

765

766 **Figure 2:** DNA concentration measured from the PCR amplification obtained for each of

767 the two primer pairs after 30 PCR cycles **(A)** or 40 PCR cycles **(B)** on 4 ng of DNA

768 extracted from each single insect species.

769

770 **Figure 3**: **A)** Comparison of observed relative abundances to expected relative

771 abundances of each taxon **B)** Estimated bias for each taxon displayed relative to the

772 geometric mean bias, with 95% and 50% confidence intervals displayed. RMSE; Root

773 Mean Square Error.

774

775 **Figure 4**: **A)** Metabarcoding bias partitioned to each protocol and primer step, with 95%

776 and 50% confidence intervals displayed. Change in composition of each taxon

777 throughout the workflow relative to the geometric mean taxon with standard errors

778 displayed, for **B)** QuickExtract, and **C)** DNEasy obtained by sequentially multiplying a

779 starting abundance of 1 by the partitioned bias estimates from panel A. Insect species

780 represented in graphs B and C are arranged and colour-coded based on graph A.

781

782 **Figure 5**: **A)** Coefficients of model predicting the estimated taxon-specific bias from the

783 traits in panel A for the unpartitioned bias and the 3 partitioned bias steps. Coefficients

784 are displayed on a pseudo-log scale to avoid compressing around zero.

785

786 **Supplementary figures**

787

788 **Figure SM1:** Comparison of quantitative **A)** and qualitative **B)** results for DNA extraction

789 methods on whole insect pools in Run 1, sequenced immediately, compared with Run

790 2, after three-months storage at -20 $^{\circ}$C.

791

792 **Figure SM2:** Fit of the bias model to each set of pools. Proportions are displayed on a

793 pseudo-log scale to avoid compressing variation around zero. RMSE; Root Mean

794 Square Error.

795

796 **Tables**

797

798 **Table 1: Composition of the eight pools used for this study.** The number of

799 individual insects is reported for each pool, as well as the total number of individuals (in

800 bold). DNA and PCR pools were assembled with the same proportions reported here.

801

802 **Table 2: Molecular and morphological characteristics of the insects used in the**

803 **pools.** Molecular traits reported include primer mismatch and GC% of the amplified

804 sequence for each of the two primer pairs used. Morphological traits include

805 measurements of insect volume, obtained by averaging measures across 10 specimens

806 per species, as well as a scale of exoskeleton hardness. Hardness values are 1=soft,

807 2=intermediate, 3=hard.

808

809 **Supplementary Tables**

810

811 **Supplementary Table 1: Overall model fits across the primers and mock**
812 **community types.**

813

814 **Supplementary Table 2: Estimated coefficients (bias) for each taxon, with**
815 **bootstrap standard errors across the primers and mock community types.**

816

817 **Supplementary Table 3: Partitioned bias, and bootstrap standard errors across**
818 **the primers and mock community types.**

Table 1: Composition of the eight pools used for this study.

The number of individual insects is reported for each pool, as well as the total number of individuals (in bold). DNA and PCR pools were assembled with the same proportions reported here.

1 **Table 1: Composition of the eight pools used for this study.** The number of

2 individual insects is reported for each pool, as well as the total number of individuals (in

3 bold). DNA and PCR pools were assembled with the same proportions reported here.

Species	Order	Pool							
		1	2	3	4	5	6	7	8
Carpophilus davidsoni	Coleoptera	25	50	5	10	25	50	5	10
C. truncatus	Coleoptera	1	1	1	1	1	1	1	1
Bactrocera tryoni	Diptera	1	1	1	1	1	1	1	1
Bradysia nr. ocellaris	Diptera	0	0	0	0	1	0	1	0
Drosophila hydei	Diptera	1	1	1	1	1	1	1	1
D. melanogaster	Diptera	5	10	25	50	5	10	25	50
D. simulans	Diptera	1	1	1	1	1	1	1	1
Scaptodrosophila lativittata	Diptera	1	1	1	1	1	1	1	1
Acizzia alternata	Hemiptera	1	1	1	1	1	1	1	1
A. solanicola	Hemiptera	10	25	50	5	10	25	50	5
Acizzia sp.	Hemiptera	1	1	1	1	1	1	1	1
Diuraphis noxia	Hemiptera	1	1	1	1	1	1	1	1
Metopolophium dirhodum	Hemiptera	1	1	1	1	1	1	1	1
Rhopalosiphum padi	Hemiptera	1	1	1	1	1	1	1	1
Aphidius colemani	Hymenoptera	36	5	6	12	36	4	9	20
Lysiphlebus testaceipes	Hymenoptera	14	0	4	13	14	1	1	5
	Total individuals	100	100	100	100	101	100	101	100

Table 2: Molecular and morphological characteristics of the insects used in the pools.

Molecular traits reported include primer mismatch and GC% of the amplified sequence for each of the two primer pairs used. Morphological traits include measurements of insect volume, obtained by averaging measures across 10 specimens per species, as well as a scale of exoskeleton hardness. Hardness values are 1=soft, 2=intermediate, 3=hard.

1 **Table 2: Molecular and morphological characteristics of the insects used in the pools.** Molecular traits reported
2 include primer mismatch and GC% of the amplified sequence for each of the two primer pairs used. Morphological traits
3 include measurements of insect volume, obtained by averaging measures across 10 specimens per species, as well as a
4 scale of exoskeleton hardness. Hardness values are 1=soft, 2=intermediate, 3=hard.

		Molecular traits						Morphological traits	
Species	Order	fwhF2 mismatch	fwhR2n mismatch	Amplicon GC%	fwhF2 mismatch	HexCOIR4 mismatch	Amplicon GC%	Volume (mm³)	Hardness
Carpophilus davidsoni	Coleoptera	0.12	0.09	0.37	0.12	0.06	0.38	3.8	3
Carpophilus truncatus	Coleoptera	0.04	0.04	0.35	0.04	0	0.35	3.91	3
Bactrocera tryoni	Diptera	0	0.04	0.37	0	0	0.37	26.51	2
Bradysia nr. *ocellaris*	Diptera	0.08	0	0.34	0.08	0	0.34	0.25	2
Drosophila hydei	Diptera	0.04	0	0.33	0.04	0.06	0.33	2.19	2
Drosophila melanogaster	Diptera	0	0	0.3	0	0	0.3	1.14	2
Drosophila simulans	Diptera	0	0	0.32	0	0	0.32	0.79	2
Scaptodrosophila lativittata	Diptera	0	0	0.28	0	0	0.28	3.09	2
Acizzia alternata	Hemiptera	0	0.04	0.3	0	0	0.31	0.22	1
Acizzia solanicola	Hemiptera	0	0.04	0.32	0	0	0.33	0.47	1
Acizzia sp.	Hemiptera	0.08	0.09	0.32	0.08	0	0.33	1.18	1
Diuraphis noxia	Hemiptera	0.04	0	0.2	0.04	0	0.2	0.5	1
Metopolophium dirhodum	Hemiptera	0.04	0	0.22	0.04	0	0.22	0.65	1
Rhopalosiphum padi	Hemiptera	0.04	0	0.19	0.04	0	0.2	0.18	1
Aphidius colemani	Hymenoptera	0	0	0.25	0	0	0.25	0.32	2
Lysiphlebus testaceipes	Hymenoptera	0	0	0.24	0	0	0.25	0.17	2

Figure 1

Figure 1: Workflow of the experiment for the three types of pools.

Insect pools are the result of non-destructive DNA extractions from the pooled insect specimens. DNA pools are the result of pooled DNA that was destructively extracted from each insect species. PCR pools are the result of DNA that was extracted, quantified, and amplified separately from each insect species then pooled before indexing qPCR. All pools were indexed, sequenced and analysed following the same protocol. Each pool was amplified using two primer sets.
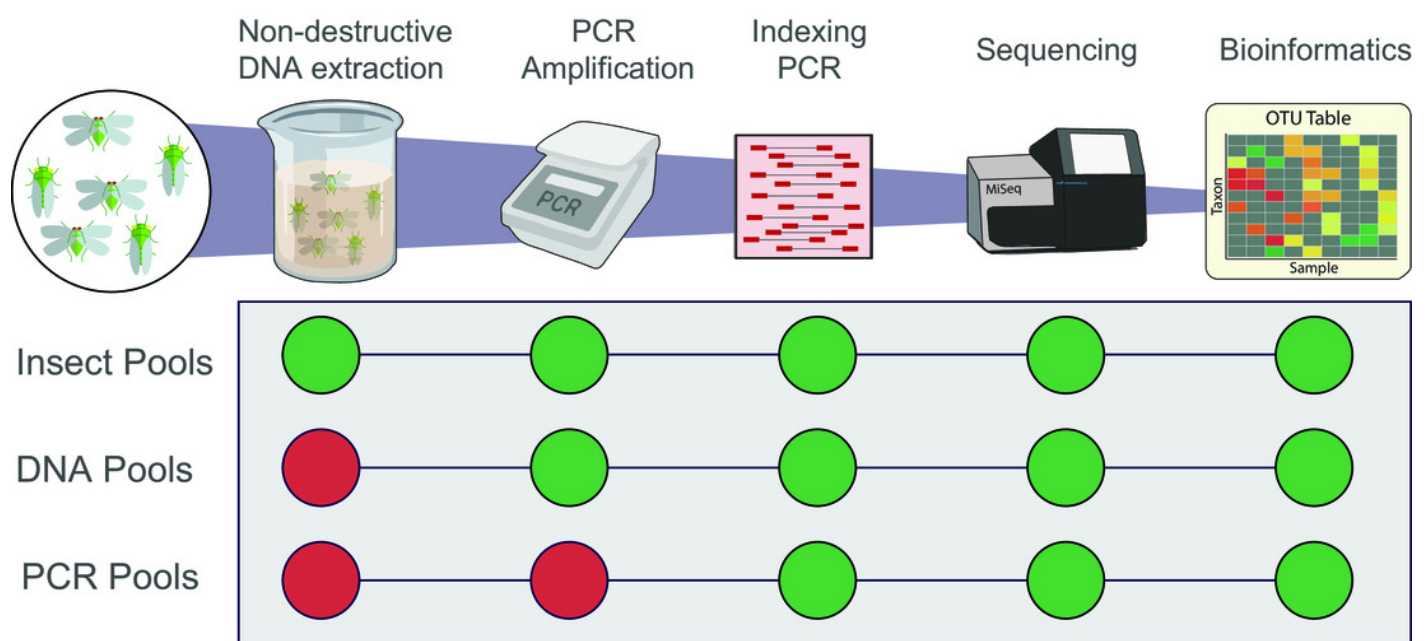
Figure 2

Figure 2: DNA concentration of PCR amplification for each species.

DNA concentration was measured from the PCR amplification obtained for each of the two primer pairs after 30 PCR cycles **(A)** or 40 PCR cycles **(B)** on DNA extracted from each single insect species.
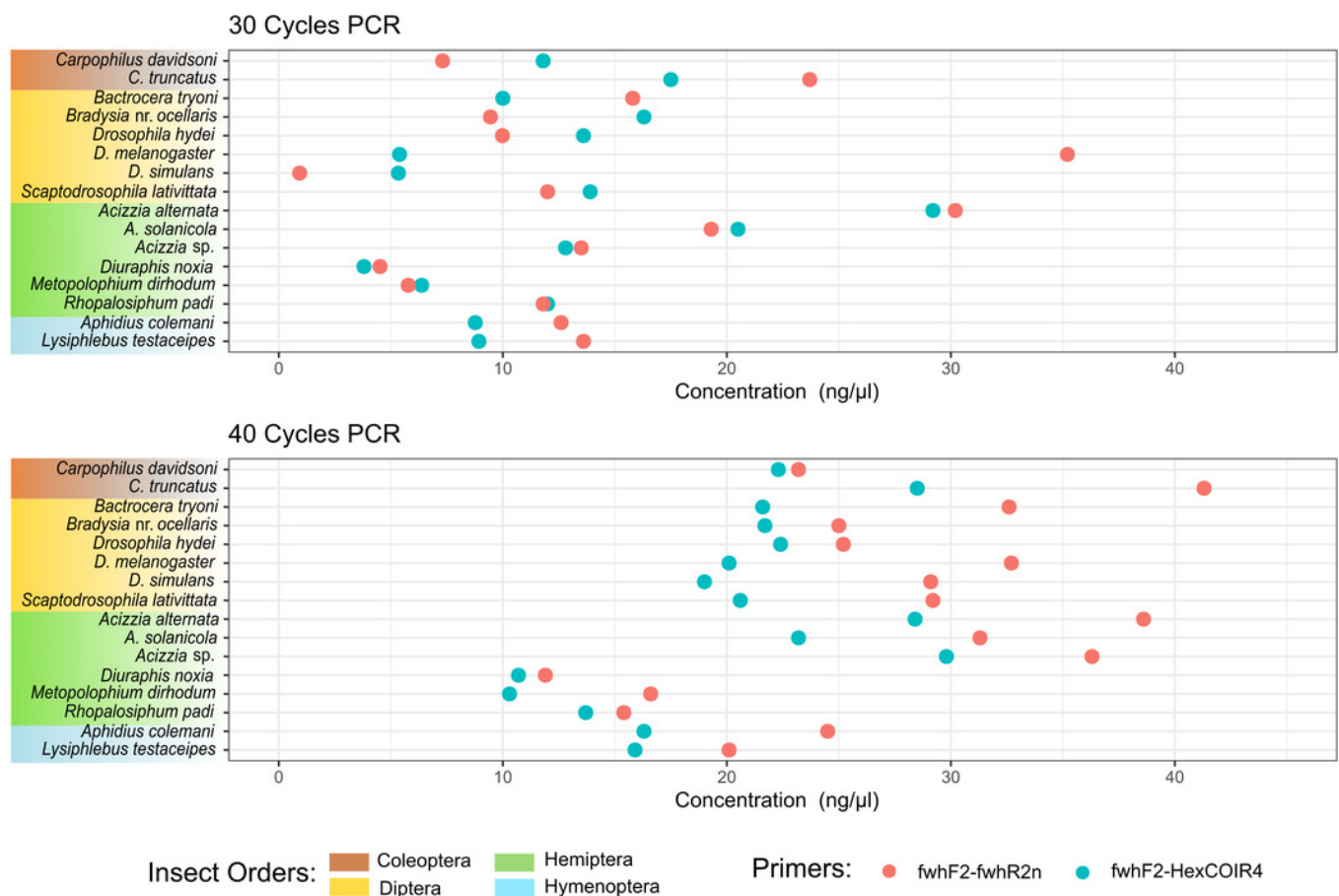
Figure 3

Figure 3: Comparison of estimated and observed proportion of reads.

A) Comparison of observed relative abundances to expected relative abundances of each taxon **B)** Estimated bias for each taxon displayed relative to the geometric mean bias, with 95% and 50% confidence intervals displayed. RMSE; Root Mean Square Error.
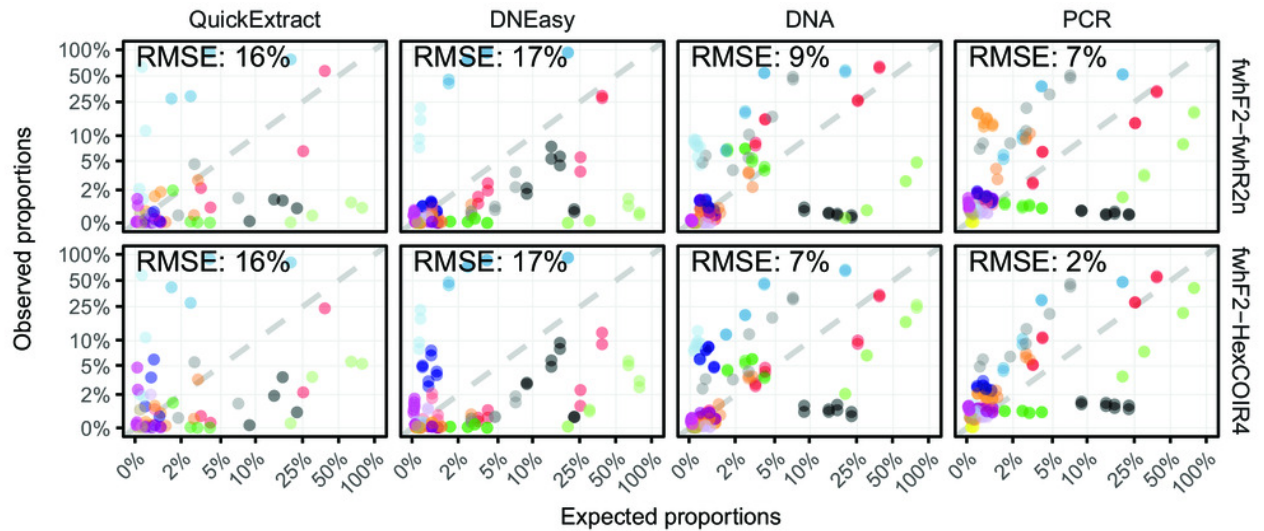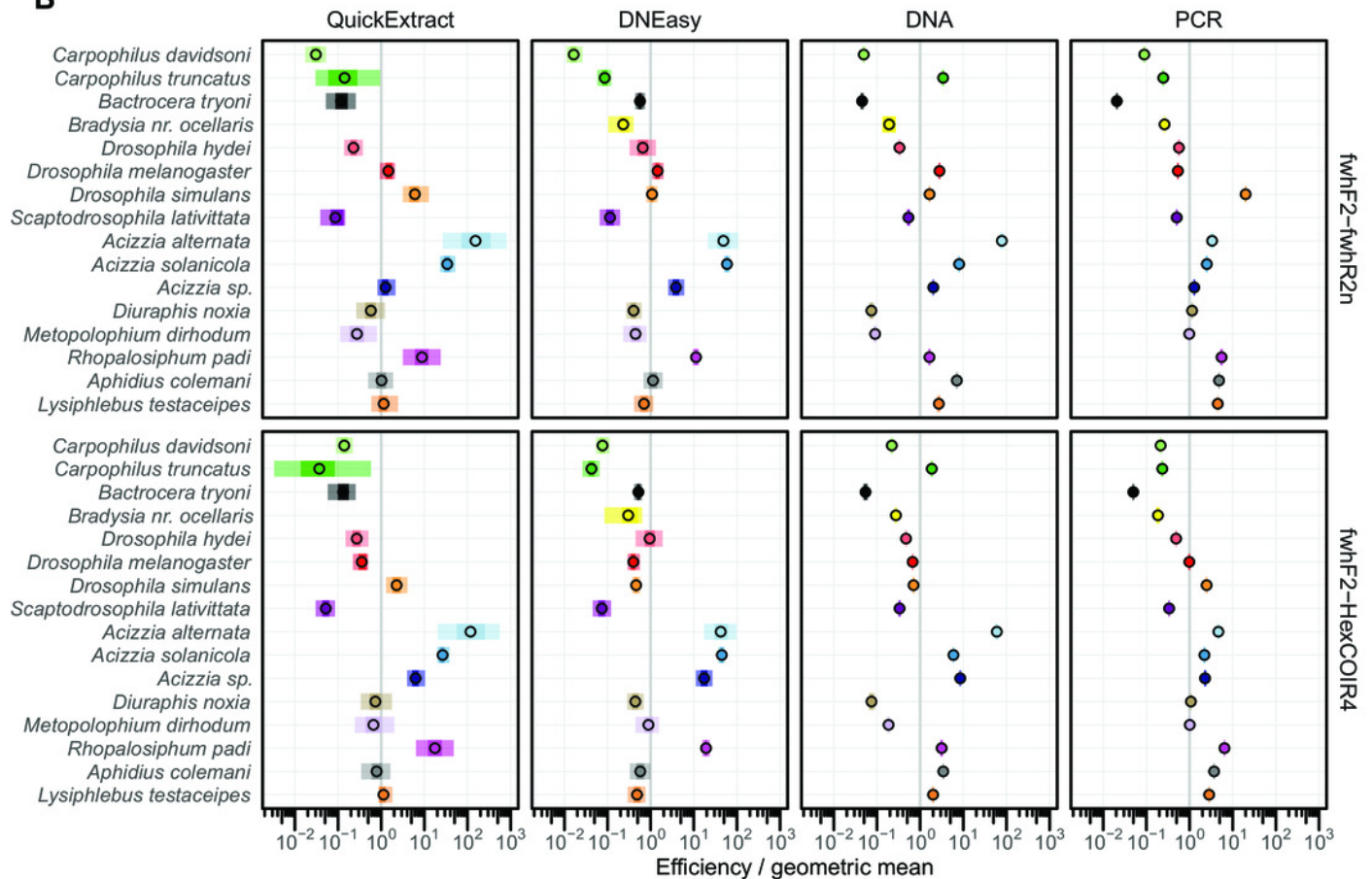
Figure 4

Figure 4: Metabarcoding bias partitioned

A) Metabarcoding bias partitioned to each protocol and primer step. Change in composition of each taxon throughout the workflow relative to the geometric mean taxon, for **B)** QuickExtract, and **C)** DNEasy obtained by sequentially multiplying a starting abundance of 1 by the partitioned bias estimates from panel A.
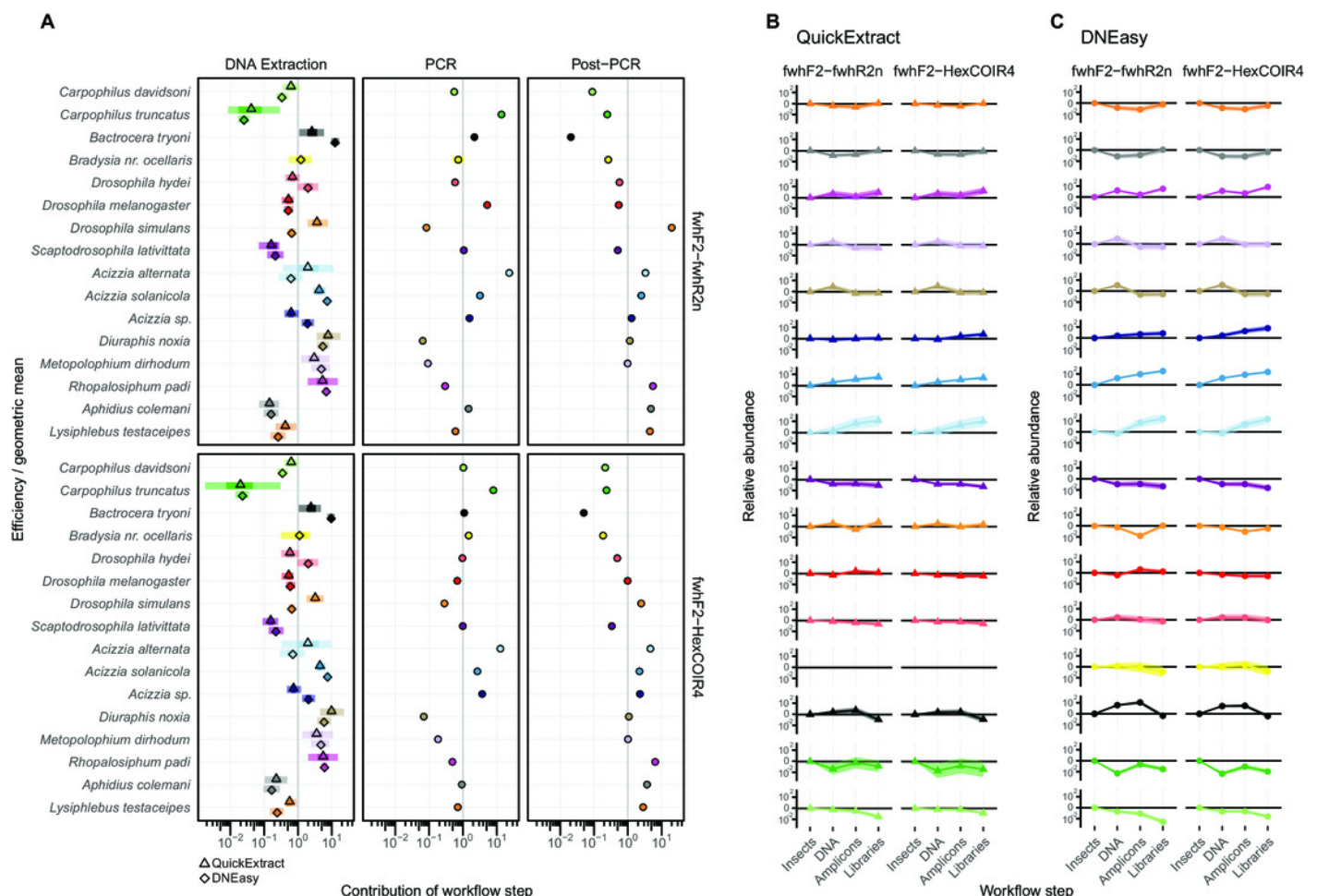
Figure 5

Figure 5: Model predicting the estimated taxon-specific bias.

Coefficients of model predicting the estimated taxon-specific bias from the traits in panel A for the unpartitioned bias and the 3 partitioned bias steps. Coefficients are displayed on a pseudo-log scale to avoid compressing around zero.