SETAC PRESS

Environmental Toxicology and Chemistry, Vol. 35, No. 8, pp. 1970–1977, 2016
© 2015 SETAC
Printed in the USA

Effects of Ivermectin

IS QUALITATIVE AND QUANTITATIVE METABARCODING OF DUNG FAUNA BIODIVERSITY FEASIBLE?

WOLF U. BLANCKENHORN,*† PATRICK T. ROHNER,† MARCO V. BERNASCONI,†‡ JOHANNES HAUGSTETTER,§ and ANDRES BUSER§
†Department of Evolutionary Biology and Environmental Studies, Zürich, Switzerland
‡Natur-Museum Luzern, Luzern, Switzerland
§Ecogenics GmbH, Schlieren, Switzerland

(Submitted 19 March 2015; Returned for Revision 27 May 2015; Accepted 6 October 2015)

Abstract: In biodiversity assessments, especially of small-bodied organisms for which taxonomic expertise is lacking, identification by genetic barcoding may be a cost-effective and efficient alternative to traditional identification of species by morphology, ecology, and behavior. The authors tested the feasibility and accuracy of such an approach using dung insects of practical relevance in ecotoxicological assessments of veterinary pharmaceutical residues in the environment. They produced 8 known mixtures that varied in absolute and relative composition of small-bodied and large-bodied species to see whether mitochondrial cytochrome *c* oxidase subunit 1 barcoding picks up all species qualitatively and quantitatively. As demonstrated before in other contexts, such metabarcoding of large numbers of dung insect specimens is principally possible using next-generation sequencing. The authors recovered most species in a sample (low type I error), at minimum permitting analysis of species richness. They obtained even quantitative responses reflecting the body size of the species, although the number of specimens was not well detected. The latter is problematic when calculating diversity indices. Nevertheless, the method yielded too many closely related false positives (type II error), thus generally overestimating species diversity and richness. These errors can be reduced by refining methods and data filtering, although this requires bioinformatics expertise often unavailable where such research is carried out. Identification by barcoding foremost hinges on a good reference database, which does not yet exist for dung organisms but would be worth developing for practical applications. *Environ Toxicol Chem* 2016;35:1970–1977. © 2015 SETAC

Keywords: Biodiversity assessment Barcode Species detection Survival Taxonomic resolution

INTRODUCTION

Biodiversity assessments are central to evolutionary biology, biogeography, and (community) ecology [1–4] but are also widely used in applied fields such as conservation biology and toxicology [5,6]. The central argument is that greater biodiversity is perceived by humans as indicating more pristine and natural, often more stable, and essentially better habitats [7,8]. Any disturbance in terms of natural habitat destruction (e.g., through climate change or flooding) or human-mediated pollution typically reduces biodiversity because some species might become locally or even globally extinct. Various biodiversity estimators exist and are extensively used, the most prominent being species richness (i.e., simply counting the total number of species present in a sample) and the Shannon-Weaver and Gini-Simpson diversity indices with their associated Hill numbers, which in addition take into account the relative abundances of all species present (reviewed by Gotelli and Chao [9]).

Although common and applied widely, the identification of all or even only a part of the organisms present at a particular site or in a sample is time-consuming and therefore costly. Such work requires expert knowledge because of the huge number of species existing on earth, know-how that unfortunately is increasingly being lost in the community of biologists. Identification efforts can be substantial and crucial in commercial practical applications. An opportunity for reducing the cost of biodiversity assessments has recently appeared in the form of genetic barcoding [10–12]. Ideally, every species can be identified by a unique genetic barcode, for which production costs have dropped dramatically in recent years and will continue to do so [13]. In practice, some species share barcodes, although at times cryptic subspecies may have different barcodes, so barcoding may not be successful in an estimated but debated 5% to 20% of species for various reasons [14]. Regardless, these developments have given rise to the idea of identifying many species at once from bulk samples, especially in ecological diversity contexts, so-called metabarcoding [15–18]. Provided an appropriate reference data bank such as GenBank or BOLD exists or can be generated ad hoc, identification by genetic barcoding may eventually become feasible and potentially less expensive than traditional identification of species based on morphology, ecology, and/or behavior, at least for many small-bodied taxa for which few classic identification guides are available [19]. In the present study we tested the feasibility and accuracy of such identification by metabarcoding with an example of concrete practical relevance to the field of toxicology, focusing on the commonly used mitochondrial cytochrome *c* oxidase subunit 1 gene.

Dung fauna as a practical and cost-sensitive test case

Ecotoxicology is a prominent field in which biodiversity assessments are applied. A particular example of an ecotoxicological subfield with current relevance to regulators concerns nontarget effects of veterinary pharmaceutical product residues on dung-dwelling organisms [20,21]. Vertebrate dung represents a compact ecosystem or community by itself. A multitude

This article includes online-only Supplemental Data.
* Address correspondence to wolf.blanckenhorn@ieu.uzh.ch
Published online 9 October 2015 in Wiley Online Library (wileyonlinelibrary.com).
DOI: 10.1002/etc.3275

of often specialized dung organisms break down the dung; these include primarily insects (beetles, flies, and springtails), earthworms, nematodes, fungi, and bacteria, most being decomposers, but the community also includes predators and parasitoids [22,23]. This important ecosystem service may be threatened by livestock medications, which can have nontarget lethal and nonlethal effects on various dung organisms. As a consequence, there are worldwide regulations mandating toxicological tests of new veterinary medical products on dung organisms [24,25].

Toxicological 50% lethal dosage (sometimes called 50% effective concentration) tests are typically conducted on single species in the laboratory. As representatives of the dung community, 2 flies (*Scathophaga stercoraria* and *Musca autumnalis*) and 2 beetles (*Onthophagus taurus* and *Aphodius constans*), 1 tropical and 1 temperate, are suggested test organisms, in addition to earthworms and possibly springtails [26–28]. It is doubtful that any single species is sufficiently representative in terms of encompassing the degree of sensitivities expected in nature [6,29]. Instead, higher-tier tests assessing the whole dung community in the field, which would be more comprehensive and realistic, are currently being evaluated. Such replicated tests generate (tens of) thousands of specimens of hundreds of dung organism species that need to be identified to assess their biodiversity.

In recent years, an international consortium of practitioners and regulators has been involved in investigating the validity of higher-tier ecotoxicological field tests. Such assessments of the entire dung biodiversity were in principle found to be repeatable and hence feasible in practice and are reported in this special section [22,23,30–32]. The present study involved processing thousands of adult insect specimens that emerged from experimental dung pats, which were identified to various taxonomic levels (species, genera, family). To date, unfortunately, only a small proportion of dung organisms are found in genetic data banks; and although highly desirable, a complete data bank for dung organisms will not be realized any time soon. Therefore, identification of dung organisms by barcoding may not yet be practical [15,33]. Nevertheless, we asked whether in principle such an approach would be scientifically (and economically) feasible in this context, by subjecting mixtures of known and already barcoded dung insects to bulk metabarcoding via modern next-generation sequencing techniques. Because dung organisms feature various body sizes and concomitant DNA amounts, we particularly wondered whether we could obtain not only a qualitative list of the various species present in a given sample but also a quantitative estimate reflecting their numbers and sizes, as would be required to compute more complex measures such as the Shannon-Weaver and Gini-Simpson diversity indices [9]. In so doing, we essentially followed an approach already taken in other contexts with other organisms, including nematodes [17,18], river benthos [34], freshwater diatoms [35], terrestrial insects [36], and chironomids [37]. Our main purpose was not to further develop such methods—we leave this to the experts [19,38–43]. Instead, we assessed this approach in the very specific practical context of our broader study reported in this special section. We are aware that barcode diversity can in principle be assessed without necessary reference to identified species by instead scoring so-called operational taxonomic units (OTUs) [33], but this was not our focus because we are ultimately interested in the functional consequences of specific dung organisms [31].

MATERIALS AND METHODS

To address the feasibility of identification by metabarcoding, we produced 8 known mixtures of dung beetles and flies to determine whether the technique picks up all species qualitatively and quantitatively (including *Drosophila melanogaster* as a well-represented, positive control). We used only freshly killed, ethanol-preserved adult specimens, as opposed to larvae or pupae, to mimic the identification of insects emerging from medication-treated experimental dung pats colonized in nature in so-called higher-tier tests [22,23]. The mixtures were varied in absolute and relative composition of approximately 8 species (Table 1). We assessed small-bodied and large-bodied species separately, as well as mixes of body sizes. We used species common in Switzerland, many of which were taken from our own laboratory cultures (in general iso-female lines). All recognized species (Table 1) were known to have been barcoded before and therefore should be listed in existing data banks. We deliberately included a few specimens not identified to the species level, as occurred frequently in our field tests [22,23], to determine whether and how these would be picked up (cf. Hajibabaei et al. [34]).

For practical reasons, we focused on the cytochrome *c* oxidase subunit 1 gene, arguably the most commonly available barcode in data banks at this time. The insect cytochrome *c* oxidase subunit 1 gene is approximately 1530 nucleotides long [44–46], has been used extensively to resolve phylogenetic relationships at many taxonomic levels, and is informative across a broad range of insect taxa (Supplemental Data, Figure S1). As a consequence, its use as a standard for insect phylogenetics has been strongly advocated [47]. Because of the availability of several so-called universal primers [48,49], various regions of the cytochrome *c* oxidase subunit 1 gene have been widely used as genetic markers for insect phylogenies. With the advent of the DNA barcoding era, approximately 650 nucleotides of the cytochrome *c* oxidase subunit 1 proximal portion (the 5′ end) were chosen as the standard DNA barcoding fragment used for species identification [50,51]. However, the terminal region (i.e., the 3′ end) also has been (and still is) widely used as a genetic marker for insect phylogenies. In dung flies (family Scathophagidae), for example, most studies involving the cytochrome *c* oxidase subunit 1 were based on 810 nucleotides of the 3′ end [52–54], with, only 1 study including sequences from the cytochrome *c* oxidase subunit 1 proximal region [55]. Therefore, to recover most information available in data banks such as GenBank and BOLD, we decided to sequence the first 250 nucleotides of the cytochrome *c* oxidase subunit 1 proximal portion (the 5′ end) plus approximately 250 nucleotides of the terminal region (the 3′ end; Supplemental Data, Figure S1). The 2 sequenced fragments are not overlapping.

DNA extraction methods

The DNA extraction was performed on insect pools of various compositions in 2-mL Eppendorf tubes. Ethanol used for conservation and storage of the samples was removed, and the remaining liquid was evaporated at 56 °C for 30 min. One milliliter of lysis buffer containing 1 mg/mL proteinase K was added to each pool sample. Digestion was performed overnight at 56 °C with end-over-end mixing. After digestions, 150 μL of water, 600 μL of buffer AL, and 600 μL of 100% ethanol were added to each sample. After end-over-end mixing for 5 min, 700 μL of the digestion mixture was loaded on a spin column, followed by 2 washing steps with buffers AW1 and AW2

Table 1. Composition of the 8 samples of dung insects with essential summary statistics[a]

Taxon	Dry body mass (mg)	Sample 0	Sample 1	Sample 2	Sample 3	Sample 4	Sample 5	Sample 6	Sample 7	Mean	SD	*CV*
Beetles (Coleoptera)												
Aphodius inmaturus	13.11								2 (3.10); <98%			
Hydrophilidae (*Cercyon marinus/melanocephalus*)	1.99				1 (2.30); <98%							
Onthophagus taurus	12.43						1 (0); ND	1 (0); ND	1 (3.06); 100%			
Philonthus marginatus	10.72							1 (5.45); 100%				
Scarabaeidae long	13.10							1 (0); ND				
Scarabaeidae small	3.13						2 (0); ND					
Sphaeridium scarabaeoides	20.18							1 (2.53); 99%				
Staphylinidae (*Aleochara* sp.?)	0.55				1 (0); ND							
Flies (Diptera)												
Anthomyiidae	0.47					2 (2.95); <98%						
Drosophila melanogaster	0.21	1 (3.20); 100%	5 (2.61); 100%	16 (3.18); 100%	5 (2.70); 100%	4 (3.85); 100%	3 (2.99); 100%					
Mesembrina meridiana	33.94							1 (5.67); 100%				
Musca domestica	4.21						2 (5.05); 100%	1 (5.44); 100%	4 (5.79); 100%			
Neomyia cornicina	4.82							1 (3.92); 99%				
Polietes lardarius	12.39						1 (4.70); 100%					
Scatophaga stercoraria	3.46						2 (5.04); 100%	1 (5.45); 100%	4 (5.89); 100%			
Scatophaga lutaria	1.78							1 (1.78); 99%				
Sciaridae	0.19					2 (4.39); <98%						
Sepsis cynipsea	0.67	1 (5.75); 100%			2 (5.15); 100%							
Sepsis fulgens	0.28	1 (3.79); 99%	5 (2.88); 98%	32 (3.32); 98%	10 (4.17); 99%		2 (2.80); 99%					
Sepsis neocynipsea	0.69	1 (5.25); 99%	5 (3.48); 99%	16 (3.79); 99%		1 (0); ND						
Sepsis punctum	0.72	1 (2.67); 99%	5 (4.76); 99%	8 (4.68); 100%	1 (4.19); 99%							
Sepsis thoracica	0.81	1 (5.14); 99%	5 (4.24); 100%	4 (4.65); 99%	6 (4.54); 99%	3 (4.43); 99%	3 (3.67); 99%					
Sepsis violacea	0.61	1 (4.51); 99%	5 (3.17); 100%	2 (2.65); 99%		1 (4.25); 100%						
Sphaeroceridae (*Spelobia tenebrarum*)	0.10					3 (3.22); <98%						
Themira lucida	0.65	1 (4.06); 100%	5 (3.18); 100%	1 (2.37); 99%	4 (3.31); 100%							
Wasps (Hymenoptera)												
Braconidae (*Kleidotoma* sp.?)	0.06					3 (2.56); <98%						
Species richness present		8	7	7	8	8+	8	9	4	7.29	1.60	0.22
Species barcoded (extraneous/parasite)		16 (0/1)	15 (0/1)	15 (1/1)	15 (0/1)	16 (1/2)	12 (2/1)	11 (2/0)	7 (2/1)	13.38	3.16	0.24
Species not detected		0	0	0	1	1 (3)	2	2	0			
Species diversity Hill $^1D = \exp(\text{Shannon's H})$		8	7	4.68	5.91	7.29	7.49	9	3.54	6.61	1.80	0.27
Barcode diversity (raw read numbers)		4.79	5.05	4.14	6.3	6.38	4.05	4.02	2.05	4.60	1.40	0.30
Barcode diversity (log10 read numbers)		15.51	14.44	14.34	14.66	15.43	11.51	10.04	6.58	12.81	3.18	0.25

[a]Given is the number of specimens of each species group, with the corresponding log10 (number of detected sequences) in parentheses followed by the percentage of sequence matching.
CV = coefficient of variation; SD = standard deviation; ND = taxa not detected.

according to the manufacturer's instructions. The DNA was eluted with 50 μL buffer AE and used nondiluted for the following polymerase chain reaction (PCR) amplification steps.

Amplification by PCR was performed in 3 individual steps to amplify the respective cytochrome *c* oxidase subunit 1 fragments, followed by addition of sequence tails for Illumina sequencing. All PCR steps were performed individually on each of the 8 DNA extracts from the different species pools, using all primers. The first round of PCR was performed in a 10-μL reaction volume containing 2 μL of genomic DNA, 5 μL HotStarTaq Master Mix (Qiagen; catalog no. 203445), double-distilled water, and 0.3 μM of forward and reverse primers each. The following cycling protocol was used on a TC-412 Programmable Thermal Controller: 5 cycles with 94 °C for 30 s, 45 °C for 90 s, and 72 °C for 90 s, followed by 35 cycles with 94 °C for 30 s, 56 °C for 90 s, and 72 °C for 60 s. Before the first cycle, a prolonged denaturation step (95 °C for 15 min) was included, and the last cycle was followed by a 5-min extension at 72 °C. The first round of PCR was used undiluted for the second round of PCR, adding Illumina sequencing tails specific to the PCR amplificate, followed by a third round of PCR amplification, adding species pool-specific MID tags. The setup of the PCR systems followed recommendations by Illumina for amplicon sequencing. Primer sequences for all PCR steps and the PCR programs are given in Supplemental Data, Table S1.

Sequencing data analysis

After the third round of amplification, the resulting PCRs were pooled and sequenced on a MiSeq sequencer using the 2×250–bp v2 chip with the bidirectional sequencing format, resulting in a total of 17 737 149 reads ranging from 416 298 to 2 062 861 bidirectional reads per sample after demultiplexing. Subsequently the data were assembled with the MIRA4 assembler using the est/denovo/accurate switches. The resulting sample-specific contigs were matched against GenBank using stand-alone BLAST from the National Center for Biotechnolgoy Information (nonredundant nucleotide collection, set at BlastN and word size 7). Rather than using specialized available bioinformatic tools such as OBItools, MEGAN, or EcoPCR to filter out the most likely candidate variants, we kept this procedure simple by utilizing merely Excel spreadsheets and standard statistical programs (SPSS) so as to make it palatable to the practitioner. We accepted only matches with at least 98% identity, no gaps in the sequence, and BLAST hit lengths >170 bp for the known species, largely following Porazinska et al. [17,18]. In the end, we obtained roughly 200 to 500 hits per sample, most of the different sequence matches likely being sequencing errors and only few reflecting real sequence differences. This included occasional identical hits for 2 different species (e.g., *Sepsis cynipsea* and *Sepsis neocynipsea*) but typically not. For all used sepsid specimens (Diptera: Sepsidae) the latter are very unlikely because we used multiple individuals of the same iso-female lines, which essentially should be genetically identical. For final analysis, we summed all sequences for a given species match from the data bank within a sample, equivalent to OTU picking [18,33], deliberately not using their number as a discrimination criterion so as not to miss small and/or rare species.

RESULTS AND DISCUSSION

Qualitative results

All specimens in our samples identified to the species level were detected in at least 1 of the samples, even when they were small and only 1 specimen was present (Table 1). The sequence identities of the majority of these known species were matched 100% with the corresponding data bank sequences (Table 1), as in the case of our control *D. melanogaster* (which is not a dung organism). However, some of the sepsids known to be in the data banks were only matched to an accuracy of 99% or even 98% (*Sepsis fulgens*), depending on whether the match was obtained reading forward versus backward—that is, depending on the sequenced part of the gene and/or the primer pair that detected the match (Table 1; Supplemental Data, Table S1). Obviously, had we only accepted 100% matches, some of the species contained in the sample would have remained undetected, increasing the type I error.

A few specimens known only to the family level remained undetected in mixtures 3 through 6 (Table 1). This includes a small staphylinid parasitoid, presumably of the *Aleochara* group, and 2 larger scarabaeid dung beetles, all of which probably are not in the data banks. In addition, 2 species (*O. taurus* and *S. neocynipsea*) correctly detected in some mixtures remained undetected in others (type I error). *Sepsis neocynipsea* is known to share its cytochrome *c* oxidase subunit 1 barcode with its closely related sister species *S. cynipsea*; however, nondetection of *O. taurus* in 2 of 3 samples must be a true reading or matching error. Our unknown specimens belonging to the common fly families Sciaridae and Anthomiidae were detected at only the family level, whereas our sphaerocerid specimen was matched as *Spelobia tenebrarum*, though all at <98% identity and hence with considerable insecurity as to their species identity (Table 1; Supplemental Data, Table S2).

In general, given our rather strict inclusion criteria (see *Materials and Methods*), approximately twice as many species were detected as were in the sample (high type II error or false positives; Table 1; Supplemental Data, Table S2). Understandably, a large majority of those erroneously matched species were closely related congeners with similar or perhaps even identical barcodes. Notably, their abundances in terms of number of reads on average did not differ from those of the species present in the sample, based on *t* tests ($p > 0.1$ for all sample comparisons). *Wolbachia* bacteria were detected in several samples. Because *Wolbachia* are common endosymbionts of insects and because they have genes homologous to cytochrome *c* oxidase subunit 1 [43,56], their occurrence in the samples was not unexpected. Only few false positives were completely extraneous (Supplemental Data, Table S2), including *Homo sapiens* (human) and *Bos taurus* (cattle). The latter could reveal real contamination of the sample by human skin or cattle dung.

Using 3 primer pairs reading backward and forward—that is, from both ends of the cytochrome *c* oxidase subunit 1 gene (Supplemental Data, Figure S1)—definitely increased the likelihood of detecting the correct species in the data banks, thus minimizing type I error. More than 50% of the specimens were matched congruently by both forward and backward sequencing, whereas approximately 30% of the specimens were only detected by forward and the remaining 20% only by backward sequencing. Using several primer pairs and forward and backward sequencing thus adds precision of identification but also cost and probably more false positives (see below). Note that instead of the 650-bp standard [52,53] we sequenced 2×250 bp only, in practice not more than 212 (initial base-pair range of hits 27–212 [mean 194, standard deviation 28]; after implementing our 170-bp threshold [see *Materials and Methods*] mean 206, standard deviation 5.6).

Overall, these results indicate high sensitivity of our methods in detecting species in the sample (low type I error, or few undetected positives) but at the same time high type II error (or many detected negatives). Every statistician knows that one cannot minimize both type I and type II errors for a given sample. We minimized the former, as is standard. Similar studies on other taxa in other contexts employed various bioinformatic tools, some of them specialized for metabarcoding [18,57], to filter PCR, reading, or assignment errors and minimize false positives [40–43], which are generally frequent in such studies [15]. Others, often systematists, focus on a strategy of barcoding every single specimen inexpensively, for example, by skipping various expensive technical steps and avoiding PCR [16,19,36,39,43]. We deliberately used simple filtering criteria that do not require genomic expertise or specialized software. Nevertheless, we still obtained many false positives, mostly closely related species with similar barcodes, resulting in the species richness detected being approximately twice that present (Table 1). The latter definitely inflates the biodiversity of any sample identified by barcoding. However, the species richness present and that barcoded do correlate positively (Table 1; $r = 0.63$) and have similar coefficients of variation, so the type II error was roughly similar in all samples. Because biodiversity assessments in the present study and other contexts typically compare various samples across populations or treatments (in the present special section livestock medication levels [22,,23,30–32,58]), detection of differences between them might therefore not be strongly affected by inflated barcoding richness (assuming no systematic biases). On the other hand, the barcoding diversity index was actually lower than the actual species diversity of the samples when based on the relative abundances of raw reads for all hits, but higher when based on the decadic logarithm of these relative abundances (Table 1). This occurred because the Shannon-Weaver diversity index H and the resulting Hill number $^{1}D = \exp(H)$, (= evenness [9]), in addition to the number of species (richness), are particularly sensitive to skew in species abundances, which was much greater for the number of barcode reads per detected species than for the actual specimen numbers. (Note that this relative importance of skew in the number of reads was vastly reduced when taking their logarithm, resulting in higher diversity indices in Table 1). Nevertheless, the 2 diversity indices still correlate positively ($r = 0.42$) across samples and have similar coefficients of variation (Table 1), so detection of differences among samples might still be feasible and meaningful; but this remains to be specifically tested or simulated. What remains is that the methods and criteria for detecting species and/or OTUs in a sample crucially quantitatively affect the results obtained and should therefore be chosen with care [15,33].

Quantitative results

As could be expected, specimens of larger-bodied species with more mitochondrial DNA material to be amplified generated more reads (e.g., Amend et al. [38]). Figure 1 shows a positive correlation, for all samples combined, between the total dry body mass of the various species contained in a given sample (means ranging 0.06–34 mg; see Table 1) and the number of corresponding reads generated by PCR and detected by the BLAST search. This positive correlation, often high, was even detectable within most of the individual samples (see correlation coefficients in insert of Figure 1). Even the smallest species used, the parasitoid wasp *Kleidotoma* (Hymenoptera: Cynipoidea; dry body mass = 0.06 mg), was detected, again proving the high sensitivity in principle of genetic barcoding, requiring little genetic material to start with.

On the other hand, we obtained no good correlation between the number of reads and the number of specimens per species across samples (Figure 1). The expected positive relationship was only obtained for the larger flies (*Musca domestica* and *S. stercoraria*; both ~5 mg dry wt) but not for the smaller-bodied sepsid flies (<1 mg), even though the number of sepsid specimens in a sample varied between 1 and maximally 32 (Table 1). This implies that read numbers within samples reflect well the relative proportions and sizes of particular barcoded species, whereas read numbers for the same species across samples are not comparable, presumably because the technical properties of sequencing runs tend to be too variable and idiosyncratic [18,38]. This generally calls into question the calculation of diversity indices, requiring information about species identity and specimen numbers, when using identification by barcoding. However, our summary statistics in Table 1 indicate that although the absolute numbers of richness and diversity differ considerably, their variation among samples is largely proportional, thus potentially still permitting their use when comparing samples.

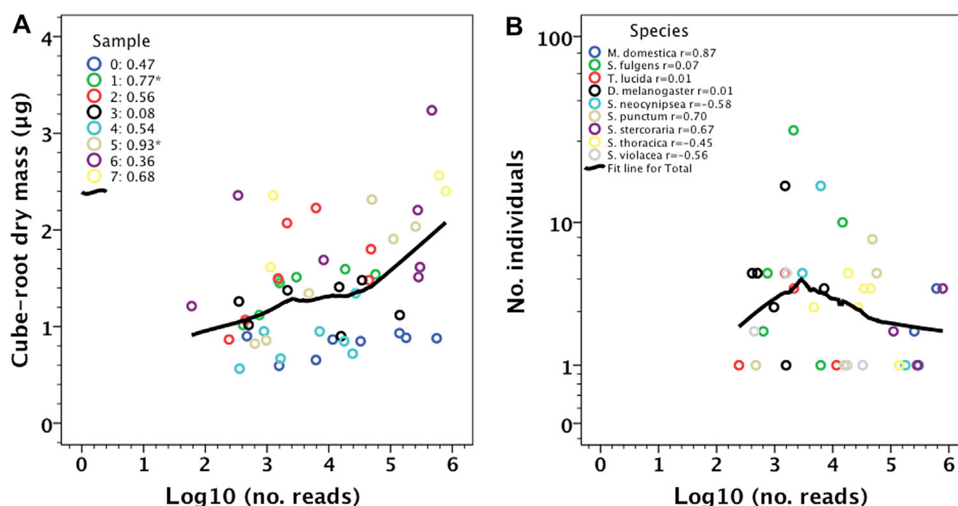

Figure 1. Relationship between the (decadic logarithm of the) number of reads and (**A**) the combined specimen size and number within samples (correlations given in the insert) and overall, and (**B**) the number of specimens per species detected across samples.

Is identification by barcoding useful in dung community studies?

We used largely freshly killed specimens, which usually amplify well, although our field samples older than 1 yr did not. We therefore stress that the age, and hence the storage quality, of the specimens remains an important determinant, potentially limiting sequencing quality and success. Immediate storage of emerged dung insects in pure ethanol in a freezer and swift processing after capture for DNA barcoding are recommended to generate reliable results. Various technical problems with the methods of specimen storage and processing, DNA sequencing, and data bank searching remain, which likely can and will be overcome with time [15,47–51,56]. Crucially, the method of identification by barcoding hinges on a good reference database [19] that does not yet exist for dung organisms and will take considerable time to be assembled. If it is to be applied widely, metabarcoding would benefit from the creation of specific data banks for particular organism groups or habitats, such as the dung organisms tested in the present study, because more pointed and efficient species matching will then be possible [59]. For calibration of quantitative responses, information on body size and corresponding amplification capacity of species should be included in these data banks, along with other useful and relevant morphological and ecological information. Specialized data banks, of course, can and will be cross-referenced with and/or derived partly from existing data banks containing data for all kinds of species.

In the end, especially for practical applications as treated in the present study, costs and main objectives will determine the method of choice. Comparing individual Sanger barcoding, the gold standard for assigning species by barcoding, metabarcoding, and traditional morphological species assignment, Stein et al. [59] concluded recently that the promise that molecular taxonomic species identification will become less expensive than traditional morphological identification is "not yet realized." This agrees with our present study. Morphological identification is largely determined by the availability of taxonomic experts and the funds to hire them. In contrast, the cost of identification by barcoding depends primarily on technological advances but is predicted to become ever less expensive and more efficient with time. However, use of specialized bioinformatics tools will add to these costs because extra expertise and software are required. Furthermore, salaries ultimately also determine the price of barcoding efforts because sample collection, sorting, and classification of insects with complex body structures cannot be automatized effectively [19,59]. The future will tell whether in the end metabarcoding [40–43] or efficient individual Sanger-like species identification [19] will be better and less expensive.

CONCLUSIONS

To summarize, metabarcoding of large numbers of specimens derived from biodiversity experiments seems principally feasible using next-generation sequencing, the cost of which is expected to drop further with technological advances. This was already multiply tried and shown with other taxa in other contexts [17,18,34–38]. We wanted to establish and embed this approach within a very concrete and practical ecotoxicological application in the context of the other studies reported in the present special section. We even obtained reasonable quantitative responses reflecting the combination of specimen numbers and body sizes of species in a sample, and hence their total DNA content, although across samples the number of specimens was not well quantified. The latter is problematic when calculating the Shannon-Weaver and Gini-Simpson diversity indices [9], although the analysis of species richness (presence/absence) will remain possible as long as most, if not all, species in a sample are detected. Irrespective of whether identification by barcoding ultimately overestimates biodiversity or not, meaningful comparisons between sites or treatment groups nevertheless appear possible as long as any overestimation is proportionately similar in all. Other barcoding genes common in data banks—16S, 18S, and the like—could be added to increase detection of more species, especially because cytochrome *c* oxidase subunit 1 may not be the best gene for metabarcoding [40–42], which, however, would again augment effort and cost.

We conclude that barcoding is helpful in higher-tier field tests of the dung fauna as documented in the present special section but, at this time, might still augment rather than reduce costs. This is because classical taxonomic expertise will remain essential for any functional or mechanistic analysis of the dung community until a complete inventory of dung organisms has been archived and possibly beyond. And this will take a very long time.

Supplemental Data—The Supplemental Data are available on the Wiley Online Library at DOI: 10.1002/etc.3275.

Acknowledgment—We thank SwissBOL for funding the present study and the German Umweltbundesamt (UBA) and the University of Zurich for additional funding and continuing support.

Data availability—All data available are presented within the present study. All barcodes are already in existing data banks. We are happy to supply any files on request (wolf.blanckenhorn@ieu.uzh.ch).

REFERENCES

1. Connell JH, Orias E. 1964. The ecological regulation of species diversity. *Am Nat* 98:399–414.
2. Margalef R. 1963. On certain unifying principles in ecology. *Am Nat* 97:357–374.
3. MacArthur RH. 1965. Patterns of species diversity. *Biol Rev* 40:510–533.
4. Pielou EC. 1974. *Population and Community Ecology: Principles and Methods*. Gordon & Breach Science, New York, NY, USA.
5. Balmford A, Green MJB, Murray MG. 1996. Using higher-taxon richness as a surrogate for species richness. 1. Regional tests. *Proc R Soc Lond B* 263:1267–1274.
6. Puniamoorthy N, Schäfer MA, Römbke J, Meier R, Blanckenhorn WU. 2014. Ivermectin sensitivity is an ancient trait affecting all ecdysozoa but shows phylogenetic clustering among sepsid flies. *Evol Appl* 7:548–554.
7. Hooper DU, Adair EC, Cardinale BJ, Byrnes JEK, Hungate BA, Matulich KL, Gonzalez A, Duffy JE, Gamfeldt L, O'Connor MI. 2012. A global synthesis reveals biodiversity loss as a major driver of ecosystem change. *Nature* 486:105–108.
8. Collen B, Nicholson E. 2014. Taking the measure of change. *Science* 346:166–167.
9. Gotelli NJ, Chao A. 2013. Measuring and estimating species richness, species diversity, and biotic similarity from sampling data. In Levin S, ed, *Encyclopedia of Biodiversity*, 2nd ed, Vol 5. Academic, Waltham, MA, USA, pp 195–211.
10. Krishnamurchy PK, Francis RA. 2012. A critical review of the utility of DNA barcoding in biodiversity and conservation. *Biodivers Conserv* 21:1901–1908.
11. Pei N, Chen B. 2013. DNA barcoding of life: A classification of uses according to function and scale after ten years of development. *Biodiversity Science* 21:616–627.
12. Schao K, Lee H, Lin Y, Lai K. 2014. A review of marine biodiversity information resources. *Biodiversity Science* 22:253–263.
13. Shkoralla S, Spall JL, Gibson JF, Hajibabaei M. 2012. Next-generation sequencing technologies for environmental research. *Mol Ecol* 21:1794–1805.
14. Meier R, Zhang G, Ali F. 2008. The use of mean instead of smallest interspecific distances exaggerates the size of the "barcoding gap" and leads to misidentification. *Syst Biol* 57:809–813.

15. Taberlet P, Coissac E, Pompanon F, Brochmann C, Willerslev E. 2012. Towards next-generation biodiversity assessment using DNA metabarcoding. *Mol Ecol* 21:2045–2050.
16. Yu DW, Ji Y, Emerson BC, Wang X, Ye C, Yang C, Ding Z. 2012. Biodiversity soup: Metabarcoding of arthropods for rapid biodiversity assessment and biomonitoring. *Methods Ecol Evol* 3:613–623.
17. Porazinska DL, Giblin-Davis RM, Faller L, Farmerie W, Kanzaki N, Morris K, Powers TO, Tucker AE, Sung W, Thomas WK. 2009. Evaluating high-throughput sequencing as a method for metagenomic analysis of nematode diversity. *Mol Ecol Resour* 9:1439–1450.
18. Porazinska DL, Sung W, Giblin-Davis RM, Thomas WK. 2010. Reproducibility of read numbers in high-throughput sequencing analysis of nematode community composition and structure. *Mol Ecol Resour* 10:666–676.
19. Meier R, Wong W, Srivathsan A, Foo M. 2015. $1 DNA barcodes for reconstructing complex phenomes and finding rare species in specimen-rich samples. *Cladistics* 32:100–110.
20. Floate KD, Wardhaugh KG, Boxall ABA, Sherratt TN. 2005. Faecal residues of veterinary pharmaceuticals: Non-target effects in the pasture environment. *Annu Rev Entomol* 50:153–179.
21. Jochmann R, Blanckenhorn WU, Bussière L, Eirkson CE, Jensen J, Kryger U, Lahr J, Lumaret J-P, Römbke J, Wardhaugh KG, Floate KD. 2011. How to test nontarget effects of veterinary pharmaceutical residues in livestock dung in the field. *Integr Environ Assess Manage* 7:287–296.
22. Floate KD, Düring R-A, Hanafi J, Jud P, Lahr J, Lumaret J-P, Scheffczyk A, Tixier T, Wohde M, Römbke J, Sautot L, Blanckenhorn WU. 2016. Validation of a standard field test method in four countries to assess the toxicity of residues in dung of cattle treated with vetinary medical products. *Environ Toxicol Chem* 35:1934–1946 (*this issue*).
23. Jochman R, Lipkow E, Blanckenhorn WU. 2016. A field test of the effect of spiked ivermectin concentrations on the biodiversity of coprophagous dung insects in Switzerland. *Environ Toxicol Chem* 35:1947–1952 (*this issue*).
24. International Cooperation on Harmonisation of Technical Requirements for Registration of Veterinary Medicinal Products. 2000. Environmental impact assessment (EIAs) for veterinary medicinal products (VMPs)—Phase I. Ecotoxicity. London, UK.
25. International Cooperation on Harmonisation of Technical Requirements for Registration of Veterinary Medicinal Products. 2004. Environmental impact assessment for veterinary medicinal products—Phase II. Guidance. London, UK.
26. Hempel H, Scheffczyk A, Schallna H-J, Lumaret J-P, Alvinerie M, Rombke J. 2006. Toxicity of four veterinary parasiticides on larvae of the dung beetle *Aphodius constans* in the laboratory. *Environ Toxicol Chem* 25:3155–3163.
27. Römbke J, Floate KD, Jochmann R, Schäfer MA, Puniamoorthy N, Knäbe S, Lehmhus J, Rosenkranz B, Scheffczyk A, Schmidt T, Sharples A, Blanckenhorn WU. 2009. Lethal and sublethal toxic effects of a test chemical (ivermectin) on the yellow dung fly (*Scathophaga stercoraria*) based on a standardized international ring test. *Environ Toxicol Chem* 28:2117–2124.
28. Römbke J, Barrett K, Blanckenhorn WU, Hargreaves T, Kadiri N, Knäbe S, Lehmhus J, Lumaret J-P, Rosenkranz B, Scheffczyk A, Sekine T. 2010. Results of an international ring test with the dung fly *Musca autumnalis* in support of a new OECD test guideline. *Sci Total Environ* 408:4102–4106.
29. Blanckenhorn WU, Puniamoorthy N, Schäfer MA, Scheffczyk A, Römbke J. 2013. Standardized laboratory tests with 21 species of temperate and tropical sepsid flies confirm their suitability as bioassays of pharmaceutical residues (ivermectin) in cattle dung. *Ecotoxicol Environ Saf* 89:21–28.
30. Scheffczyk A, Floate KD, Blanckenhorn WU, Düring R-A, Klockner A, Lahr J, Lumaret J-P, Salamon J-A, Tixier T, Wohde M, Römbke J. 2016. Nontarget effects of ivermectin residues on earthworms and springtails dwelling beneath dung of treated cattle in four countries. *Environ Toxicol Chem* 35:1959–1969 (*this issue*).
31. Tixier T, Blanckenhorn WU, Lahr J, Floate K, Scheffczyk A, Düring R-A, Wohde M, Römbke J, Lumaret J-P. 2016. A four-country ring test of nontarget effects of ivermectin residues on the function of coprophilous communities of arthropods in breaking down livestock dung. *Environ Toxicol Chem* 35:1953–1958 (*this issue*).
32. Wohde M, Blanckenhorn WU, Floate KD, Lahr J, Lumaret J-P, Römbke J, Scheffczyk A, Tixier T, Düring R-A. 2016. Analysis and dissipation of the antiparasitic agent ivermectin in cattle dung and soil at four field sites with different climate. *Environ Toxicol Chem* 35:1924–1933 (*this issue*).
33. Coissac E, Riaz T, Puillandre N. 2012. Bioinformatic challenges for DNA metabarcoding of plants and animals. *Mol Ecol* 21:1834–1847.
34. Hajibabaei M, Shokralla S, Zhou X, Singer GAC, Baird DJ. 2011. Environmental barcoding: A next-generation sequencing approach for biomonitoring applications using river benthos. *PLoS One* 6:e17497.
35. Kermarrec L, Franc A, Rimet F, Chaumeil P, Humbert JF, Bouchez A. 2013. Next-generation sequencing to inventory taxonomic diversity in eukaryotic communities: A test for freshwater diatoms. *Mol Ecol Resour* 13:607–619.
36. Zhou X, Li Y, Liu S, Yang Q, Su X, Zhou L, Tang M, Fu R, Li J, Huang Q. 2013. Ultra-deep sequencing enables high-fidelity recovery of biodiversity for bulk arthropod samples without PCR amplification. *GigaScience* 2:4.
37. Carew ME, Pettigrove VJ, Metzeling L, Hoffmann AA. 2013. Environmental monitoring using next generation sequencing: Rapid identification of macroinvertebrate bioindicator species. *Front Zool* 10:45.
38. Amend AS, Seifert KA, Bruns TD. 2010. Quantifying microbial communities with 454 pyrosequencing: Does read abundance count? *Mol Ecol* 19:5555–5565.
39. Zhan A, Hulák M, Sylvester F, Huang X, Adebayo AA, Abbott CL, Adamowicz SJ, Heath DD, Cristescu ME, MacIsaac HJ. 2013. High sensitivity of 454 pyrosequencing for detection of rare species in aquatic communities. *Methods Ecol Evol* 4:558–565.
40. Leray M, Yang JY, Meyer CP, Mills SC, Agudelo N, Ranwez V, Boehm JT, Machida RJ. 2013. A new versatile primer set targeting a short fragment of the mitochondrial COI region for metabarcoding metazoan diversity: Application for characterizing coral reef fish gut contents. *Front Zool* 10:34.
41. Clarke LJ, Soubrier J, Weyrich LS, Cooper A. 2014. Environmental metabarcodes for insects: In silico PCR reveals potential for taxonomic bias. *Mol Ecol Resour* 114:1160–1170.
42. Deagle BE, Jarman SN, Coissac E, Pompanon F, Taberlet P. 2014. DNA metabarcoding and the cytochrome *c* oxidase subunit I marker: Not a perfect match. *Biol Lett* 10:20140562.
43. Shokralla S, Gibson JF, Bakht HN, Janzen DH, Hallwachs W, Hajibabaei M. 2014. Next-generation DNA barcoding: Using next-generation sequencing to enhance and accelerate DNA barcode capture from single specimens. *Mol Ecol Resour* 14:892–901.
44. Clary DO, Wolstenholme DR. 1985. The mitochondrial DNA molecule of *Drosophila yakuba*: Nucleotide sequence, gene organisation, and genetic code. *J Mol Evol* 22:252–271.
45. Lunt DH, Zhang D-X, Szymura JM, Hewitt GM. 1996. The insect cytochrome oxidase I gene: Evolutionary patterns and conserved primers for phylogenetic studies. *Insect Mol Biol* 5:153–165.
46. Bernasconi MV, Valsangiacomo C, Piffaretti J-C, Ward PI. 2000. Phylogenetic relationships among Muscoidea (Diptera, Calyptratae) based on mitochondrial DNA sequences. *Insect Mol Biol* 9:67–74.
47. Caterino MS, Cho S, Sperling FA. 2000. The current state of insect molecular systematics: A thriving Tower of Babel. *Annu Rev Entomol* 45:1–54.
48. Simon C, Buckley TR, Frati F, Stewart JB, Beckenbach AT. 2006. Incorporating molecular evolution into phylogenetic analysis, and a new compilation of conserved polymerase chain reaction primers for animal mitochondrial DNA. *Annu Rev Ecol Evol Syst* 37:545–579.
49. Simon C, Frati F, Beckenbach A, Crespi B, Liu H, Flook P. 1994. Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. *Ann Entomol Soc Am* 87:651–701.
50. Hebert PDN, Cywinska A, Ball SL, De Waard JR. 2003. Biological identifications through DNA barcodes. *Proc R Soc Lond B* 270:313–321.
51. Hebert PDN, Ratnasingham S, De Waard JR. 2003. Barcoding animal life: Cytochrome *c* oxidase subunit 1 divergences among closely related species. *Proc R Soc Lond B* 270:S96– S99.
52. Bernasconi MV, Pawlowski J, Valsangiacomo C, Piffaretti J-C, Ward PI. 2000. Phylogeny of the Scathophagidae (Diptera, Calyptratae) based on mitochondrial DNA sequences. *Mol Phylogenet Evol* 16:308–315.
53. Bernasconi MV, Pawlowski J, Valsangiacomo C, Piffaretti J-C, Ward PI. 2001. Phylogeny of the genus *Scathophaga* (Diptera: Scathophagidae) inferred from mitochondrial DNA sequences. *Can J Zool* 79:517–524.
54. Bernasconi MV, Berger D, Blanckenhorn WU. 2010. Systematic ambiguity in the well-established model system insect *Scathophaga stercoraria* (Diptera: Scathophagidae): Sister species *S. soror* revealed by molecular evidence. *Zootaxa* 2441:27–40.

55. Kutty SN, Bernasconi MV, Šifner F, Meier R. 2007. Sensitivity analysis, molecular systematics and natural history evolution of Scathophagidae (Diptera: Cyclorrhapha: Calyptratae). *Cladistics* 23:64–83.
56. Smith MA, Bertrand C, Crosby K, Eveleigh ES, Fernandez-Triana J, Fisher BL, Gibbs J, Hajibabaei M, Hallwachs W, Hind K, Hrcek J, Huang DW, Janda M, Janzen DH, Li Y, Miller SE, Packer L, Quicke D, Ratnasingham S, Rodriguez J, Rougerie R, Shaw MR, Sheffield C, Stahlhut JK, Steinke D, Whitfield J, Wood M, Zhou X. 2012. *Wolbachia* and DNA barcoding insects: Patterns, potential, and problems. *PLoS One* 7:e36514.
57. Zhang J, Kapli P, Pavlidis P, Stamatakis A. 2013. A general species delimitation method with applications to phylogenetic placements. *Bioinformatics* 29:2869–2876.
58. Adler N, Bachman J, Blankenhorn WU, Floate KD, Jensen J, Römbke J. 2016. Effects of ivermectin application on the diversity and function of dung and soil fauna: Regulatory and scientific background information. *Environ Toxicol Chem* 35:1914–1923 (*this issue*).
59. Stein ED, Martinez MC, Stiles S, Miller PE, Zakharov EV. 2014. Is DNA barcoding actually cheaper and faster than traditional morphological methods: Results from a survey of freshwater bioassessment efforts in the United States? *PLoS One* 9:e95525.