MOLECULAR ECOLOGY
RESOURCES
Molecular Ecology Resources (2015) doi: 10.1111/1755-0998.12366

Molecular detection of invertebrate prey in vertebrate diets: trophic ecology of Caribbean island lizards

TYLER R. KARTZINEL and ROBERT M. PRINGLE
Department of Ecology & Evolutionary Biology, Princeton University, 106A Guyot Hall, Princeton, NJ 08544, USA

Abstract

Understanding community assembly and population dynamics frequently requires detailed knowledge of food web structure. For many consumers, obtaining precise information about diet composition has traditionally required sacrificing animals or other highly invasive procedures, generating tension between maintaining intact study populations and knowing what they eat. We developed 16S mitochondrial DNA sequencing methods to identify arthropods in the diets of generalist vertebrate predators without requiring a blocking primer. We demonstrate the utility of these methods for a common Caribbean lizard that has been intensively studied in the context of small island food webs: *Anolis sagrei* (a semi-arboreal 'trunk-ground' anole ecomorph). Novel PCR primers were identified in silico and tested in vitro. Illumina sequencing successfully characterized the arthropod component of 168 faecal DNA samples collected during three field trips spanning 12 months, revealing 217 molecular operational taxonomic units (mOTUs) from at least nine arthropod orders (including Araneae, Blattodea, Coleoptera, Hemiptera, Hymenoptera, Isoptera, Lepidoptera and Orthoptera). Three mOTUs (one beetle, one cockroach and one ant) were particularly frequent, occurring in ≥50% of samples, but the majority of mOTUs were infrequent (180, or 83%, occurred in ≤5% of samples). Species accumulation curves showed that dietary richness and composition were similar between size-dimorphic sexes; however, female lizards had greater per-sample dietary richness than males. Overall diet composition (but not richness) was significantly different across seasons, and we found more pronounced interindividual variation in December than in May. These methods will be generally useful in characterizing the diets of diverse insectivorous vertebrates.

Keywords: *Anolis sagrei*, diet analysis, DNA barcodes, metabarcoding, mitochondrial 16S DNA, predator–prey interactions

Received 9 September 2014; revision received 18 December 2014; accepted 19 December 2014

Introduction

Molecular technologies are facilitating incorporation of DNA-based species detection into non-invasive dietary analysis (Valentini *et al.* 2009; Pompanon *et al.* 2012). Developing diet profiles from faecal DNA is particularly advantageous because samples can be obtained in ways that minimize interaction with and harm to animals (Pompanon *et al.* 2012; De Barba *et al.* 2014). This in turn facilitates diet characterization for species that are difficult to observe in the act of eating (e.g. because they are rare, dangerous, shy, nocturnal, etc.: Baamrane *et al.* 2012; Shehzad *et al.* 2012; Hibert *et al.* 2013) or whose prey is difficult to identify visually. The same advantage applies in the context of long-term population studies and field experiments, where investigators have often been forced to choose among a vexing set of inadequate alternatives: forego detailed diet analysis; destructively harvest individuals at risk of confounding or curtailing the experiment; or rely on museum specimens or natural history observations collected long ago, far away and under different environmental conditions.

Caribbean island *Anolis* lizards ('anoles') are a model system for the study of competition, niche partitioning, adaptive radiation, food webs and other ecological and evolutionary phenomena (Schoener 1968, 2011; Williams 1972; Roughgarden 1995; Losos 2009). Much of this research is based on a combination of observational and experimental field studies, and thus, molecular diet analysis has enormous potential as a mechanistic probe to explore observed patterns. For example, interspecific competition between sympatric anole species has repeatedly been demonstrated (Pacala & Roughgarden 1985; Rummel & Roughgarden 1985; Losos & Spiller 1999) and is considered a central driver of both evolutionary

Correspondence: Tyler R. Kartzinel, Fax: +1-609-258-7892; E-mail: tylerk@princeton.edu

diversification and contemporary patterns of species coexistence within the genus (Losos 2009). Yet whereas the partitioning of microhabitat by sympatric anole species is clear, the role of competition for food remains incompletely resolved. Losos (2009, p. 229) wrote, 'Likely, by partitioning space, anoles are partitioning prey', but continued, 'Admittedly, few relevant data are available'. Most of what is known about lizard diets has been learned by sacrificing and dissecting animals (Wolcott 1923; Schoener 1967, 1968; Andrews 1979; Spiller & Schoener 1990; Huang *et al.* 2008), which is problematic in the context of longitudinal field studies. Moreover, analyses of lizard gut contents are often taxonomically imprecise, in that prey species are lumped at the order level (Lister 1976) or by size (Pacala & Roughgarden 1985),which may or may not reveal niche partitioning. In contrast, DNA-based methods can enable species-level (or even subspecific) discrimination, especially when combined with well-developed reference libraries, and prey sequence assignments can be refined in the future as additional sequence data are generated and archived.

The ability to detect niche partitioning would similarly illuminate *intra*specific competition and its consequences. For example, precise diet profiles for sympatric male and female lizards would aid in evaluating the hypothesis that sexual size dimorphism evolved in part to minimize intraspecific resource overlap (Schoener 1967; Shine 1989; Manicom *et al.* 2014). These competition-oriented examples illustrate just a few of the many applications of this technology in testing general principles of community and evolutionary ecology in lizards that have long been workhorses of these disciplines (Losos 2009).

Here, we present DNA-based methods to characterize anole diets, along with field protocols for their application within free-living populations. A broad-spectrum technique was necessary, because anoles tend to be opportunistic and generalized predators of a diverse array of invertebrates, particularly insects and spiders (Losos 2009). In all molecular diet analyses, it is necessary to balance the need for broad taxonomic coverage with the ability to amplify short fragments of degraded DNA that nonetheless contain sufficient information to discriminate among prey taxa (Pompanon *et al.* 2012). In many cases, the hypervariability of the standard DNA barcode region (COI) may reduce its utility for broad-spectrum analyses due to the difficulty of designing internal primers that amplify the majority of potential prey (Clarke *et al.* 2014). Using multiple short markers could improve the taxonomic coverage and/or specificity in generalist diets (Baamrane *et al.* 2012; Pompanon *et al.* 2012; De Barba *et al.* 2014), but this requires considerable labour and expense to establish multilocus protocols specific to the taxa and/or locales being investigated. Thus, we sought a single marker to maximize both coverage and specificity.

We developed novel PCR primers capable of detecting and distinguishing among a large diversity of invertebrates via metabarcoding. We demonstrate the utility of our approach for characterizing dietary richness and composition in the geographically widespread and locally common anole *Anolis sagrei*, which is native to Cuba and invasive in Florida, Hawaii and elsewhere (Kolbe *et al.* 2004). We compare *A. sagrei* dietary richness and composition across multiple time points and between size-dimorphic males and females. Our approach can easily be adapted for use in other vertebrate insectivores.

Methods

Study site and focal species

Protocols were developed and tested using *Anolis sagrei* populations on Staniel Cay, Exuma, Bahamas (24.17°, −76.44°) and nearby small islands. This species is a 'trunk-ground' ecomorph (Williams 1972). Individuals tend to perch head-down on low vegetation, whence they descend to capture prey (Losos 2009); however, more arboreal habits have been recorded for populations in the presence of ground-dwelling predators (Schoener *et al.* 2002). These lizards feed primarily on insects and spiders, occasionally on other invertebrates, and incidentally on other lizards and plant matter (Spiller & Schoener 1990; Losos 2009). Diet composition can vary considerably among populations and locations: Lister (1976) reported that within the Bahamas, rank order of prey by weight on Exuma was lepidopterans > orthopterans > beetles > termites, whereas on nearby Abaco lepidopterans were followed in order by beetles, ants and hemipterans.

Reference DNA collection and sequencing

We collected local arthropods to create a (noncomprehensive) reference library of potential prey DNA. Specimens were captured in bowl traps or nets, sorted into morphospecies, preserved in RNAlater and frozen until DNA extractions from a leg or wing were carried out using a Qiagen DNeasy Blood and Tissue kit or glass fibre DNA extractions (Ivonona *et al.* 2006). Mitochondrial 16S DNA was sequenced using primers 16SAr/16SBr (Palumbi 1996) to enable matching with dietary sequences amplified by the novel primers that we describe below. To assist in verifying taxonomic assignations, we also sequenced the standard COI barcode locus for each specimen, using primers LCO1490/HCO2198 (Folmer *et al.* 1994) and the Barcode of Life Database

('BOLD'; Ratnasingham & Hebert 2007). Bidirectional Sanger sequencing was carried out at GeneWiz (South Plainfield, New Jersey) on an ABI 3730.

Faecal DNA collection

We captured *A. sagrei* individuals from Staniel Cay and 15 nearby small islands in May 2013 (N = 81), December 2013 (N = 230) and May 2014 (N = 128). We measured lizards' snout-vent length (SVL) and sexed adults based on diagnostic visual characteristics (*A. sagrei* are conspicuously sexually dimorphic with respect to body size and dewlap prominence). Males were captured at least twice as frequently as females in all seasons due to their larger size and bolder behaviour (Table 1). Lizards were held individually in previously unused disposable plastic containers in an air-conditioned room. Containers were checked for faecal samples (and lizard condition monitored) several times per day, such that all samples were preserved within 8 h of defecation, and most considerably sooner than that.

We obtained faecal samples from a subset of these lizards (54 in May 2013; 156 in December 2013; 109 in May 2014); we included multiple faecal samples from some individuals to yield a total of 344 samples (May 2013 = 54; December 2013 = 161; May 2014 = 129). Samples were preserved using either RNAlater (all samples from May 2014 and a subset of those from other seasons) or Zymo Xpedition Soil & Fecal mini kits, frozen (−4 °C), and transported to Princeton University, where they were stored (−80 °C) until analysis. We preserved entire faecal samples, which were within the recommended <0.2 g quantity for extraction by Zymo Xpedition kits. For samples preserved in RNAlater, preservative was removed by centrifugation at 3000 ***g*** for 15 min, followed by centrifugation at 300 ***g*** for 5 min, with RNAlater pipetted off following each centrifugation step; the sample was then transferred into a Zymo Xpedition Soil & Fecal mini kit for extraction. This step was omitted for samples preserved directly in Zymo Xpedition kits. Extraction blanks were used to monitor for cross-contamination each time extractions were performed on small batches of 5–23 (typically 15) samples.

Primer design

Primers for broad-spectrum dietary analysis should (i) enable detection of a broad array of potential prey taxa; (ii) amplify short DNA regions to minimize PCR length inhibition of damaged DNA (Deagle *et al.* 2006); and (iii) vary at the species level to enable taxonomic assignation. Mitochondrial DNA (mtDNA, e.g. COI, 16S) is widely preferred because it occurs with multiple copies per cell and is well represented in public databases (Pompanon *et al.* 2012). Ideally, primers will amplify only the DNA of ingested material (Pompanon *et al.* 2012). We sought to develop a set of primers that best adhered to these characteristics for the arthropod prey of lizards.

We identified potentially suitable mtDNA regions in silico (Riaz *et al.* 2011). We downloaded all Arthropoda and Squamata complete mitochondrial genomes from GenBank (1475 Arthropoda and 242 Squamata, as of August 2013). Using *ecoPrimers* (Riaz *et al.* 2011), we searched for primers that met the following criteria: (i)

Table 1 Sample sizes at key stages in the analysis

N overall and by sex	May 2013	December 2013	May 2014	Total
Initial PCR samples sequenced*	30	61	86	177
Male	23	39	64	126
Female	5	20	22	47
High-quality sequence reads	7 242 684	15 835 229	27 598 793	50 676 706
Male	5 704 270	10 346 644	21 597 577	37 648 491
Female	1 076 797	4 662 462	6 001 216	11 740 475
Unique sequences after removing low-frequency noise and sequences outside expected size range	1213	1246	1307	1378
Male	1200	1204	1273	1378
Female	567	972	1176	1316
Final molecular operational taxonomic units after removing putative PCR errors and samples with few reads	93	131	151	217
Male	81	109	126	191
Female	27	66	105	141
Final samples*	30	55	83	168
Male	23	37	61	121
Female	5	16	22	43

*Total sample sizes do not equal the sum of sample sizes by sex due to inclusion of four juveniles with undetermined sex.

18 nt length; (ii) 40- to 300-bp sequence; and (iii) ≤3 total mismatches to the target, with none in the three 3′ bases. These criteria yielded a set of suitable 16S primers for insects, the best-represented arthropod class in the database (N = 1052 accessions), with minimal cross-amplification of Squamata. More suitable primers were not identified using a range of search criteria (e.g. 16- to 20-nt oligos) for all Arthropoda.

Primers were evaluated for taxonomic coverage and specificity. Coverage (i.e. the percent of taxa with an amplifiable sequence) was evaluated using the mitochondrial genome database and *ecoPCR* (Ficetola *et al.* 2010). We searched for sequences of 50–300 bp with a maximum of three primer mismatches. We evaluated overall coverage of arthropod genera and the potential amplification of Squamata under these conditions. To evaluate specificity of the amplified region (i.e. its power to distinguish among taxa), we downloaded all Arthropoda and Squamata 16S sequences from GenBank. The resulting database included 20 662 accessions (19 028 Arthropoda and 1634 Squamata). We performed *ecoPCR*, allowing a maximum of five primer mismatches for amplicons of 50–500 bp, to evaluate specificity of this larger database. We did not evaluate primer coverage for this data set, as some accessions only partially span the target region. Finally, we used the forward and reverse primer sequences for all 'amplified' accessions in this data set to visualize primer mismatches to arthropod orders frequently represented in the diets of these lizards (Araneae, Blattodea, Coleoptera, Diptera, Hymenoptera, and Lepidoptera) using WEBLOGOS (Crooks *et al.* 2004).

Dietary DNA amplification and sequencing

Arthropod DNA in faecal samples was analysed by amplicon sequencing. Arthropod DNA was amplified from faecal samples by PCR using novel primers identified in silico as described above (IN16STK-1F: TGAACTCAGA TCATGTAA and IN16STK-1R: TTAGGGATAACAGC GTAA). Primers were engineered with a unique 8-nt sequence tag on the 5′ end to allow combinatorial demultiplexing. At least 4-nt pairwise differences separated the tags and 96 forward and reverse tag combinations were used. The PCR comprised 20 μL Amplitaq Gold reactions, with 2.5 mM $MgCl_2$, 200 μM each dNTP, 0.1 mg/mL BSA, 4% DMSO. Cycling conditions used initial denaturing at 95 °C for 10 min, followed by 35 cycles of denaturing at 95 °C for 30 s, annealing at 50 °C for 30 s and extension at 72 °C for 30 s, with a final extension at 72 °C for 2 min.

In addition to sequencing faecal samples, we sequenced three 'mock diets'. Such 'mock community' samples are routinely sequenced to gauge error and evaluate reliability of microbial 16S rRNA analyses (Schloss *et al*. 2011; Kozich *et al.* 2013). Our mock diets comprised a mixture of Qiagen-extracted arthropod DNA from five samples, each diluted to 2.5% final volume. Taxa from Table S1 (Supporting information) included in these assays were as follows: the cockroach *Blaberus* sp. 2 (BAHAR168-14), the fly Dolichopodidae 1 (BAHAR162-14), the ant *Camponotus* sp. (BAHAR167-14), the spider *Gasteracantha cancriformis* (BAHAR125-14) and the moth Lepidoptera 3 (BAHAR169-14).

Successful PCRs were normalized using SequalPrep 96-well plates, and up to 96 uniquely tagged PCRs were pooled and concentrated using Zymo clean and concentrator kits. At Princeton University's Lewis Sigler Institute, approximately 100 ng of DNA from each pool of PCR products was used to generate Illumina sequencing libraries using the Apollo 324™ NGS Library Prep System and PrepX DNA library kit (WaferGen Biosystems, Fremont, CA, USA), which included DNA end repairing, A-tailing, adapter ligation and limited amplification. Adapters with different Illumina barcodes were ligated to each PCR pool. The libraries were examined using Agilent Bioanalyzer DNA HS chips (Agilent Technologies, Santa Clara, CA, USA) to assess size distributions and were quantified using a Qubit fluorometer (Invitrogen, Carlsbad, CA, USA). Each set of four libraries was further pooled together in equal amounts and sequenced on an Illumina HiSeq 2500 Rapid Flowcell as single-end 170-nt reads, along with the 7-nt index reads (Illumina Inc., San Diego, CA, USA). Illumina HiSeq Control Software was used to generate pass-filter reads for analysis.

Sequence processing and taxon assignation

For demultiplexing and quality control, we used the software packages *obitools* (http://metabarcoding.org/obitools), MOTHUR (Schloss *et al.* 2009) and R (R Core Development Team 2013). We used obitools' *ngsfilter* to assign each sequence to its sample of origin based on exact matches to both multiplex identifier (MID) tags and allowing up to two primer mismatches. Sequences with ambiguities or mean quality scores <30 (Illumina fastq quality values) across the head, tail or total length of the sequence were discarded. Primers and tags were removed. Identical sequences were merged in obitools using the *obiuniq* command, which tallies per-sample representation. As a conservative measure to eliminate potentially spurious sequences from further analysis, we eliminated sequences that were shorter than the expected length (<107 bp) or sequenced infrequently (<1000 reads, or a proportional representation $<1.8 \times 10^{-5}$ initial sequence reads). We chose these thresholds because they retained the vast majority of the sequence reads while filtering out a large number of low-abundance and putatively erroneous sequences—the resulting set of unique sequences (N = 1378; 0.3% of starting value) included

the vast majority of sequence reads (94%). These remaining sequences were taxonomically classified using obitools' *ecoTag* command, based on a reference library that was constructed using (i) 16S sequences downloaded from GenBank as described above, which we refer to as the '*GB*' database, or (ii) 16S sequences obtained from our Bahamian reference specimen collection, which we refer to as our 'local' database. We refer to the combined database as the *ecoTag* database. When inferring the taxonomy of a sequence, we gave preference to the database with the greater percent identity.

In dietary investigations with relatively comprehensive DNA reference libraries, PCR errors and chimeras can be filtered out by requiring stringent matching with reference sequences (De Barba *et al*. 2014). However, there are relatively few existing publically available reference sequences at this 16S locus for arthropod taxa in our study system, and our local reference collection currently comprises just 62 molecular operational taxonomic units (mOTUs; see Results), a fraction of the hundreds of potential arthropod prey species. Thus, data were imported to MOTHUR to screen for PCR errors, eliminate potential chimeras, evaluate error in samples of known composition and cluster sequences into mOTUs (Kozich *et al*. 2013).

Our strategy using MOTHUR was as follows: Sequences were aligned, trimmed and further dereplicated using a reference alignment comprising local reference sequences (Schloss *et al*. 2009; Kozich *et al*. 2013). Within samples, sequences differing by up to 2 bp from a more abundant sequence were considered possible PCR errors and merged with the more abundant sequence using the function *pre.cluster* (a threshold of 1 bp is recommended for each 100 bp sequenced; Schloss *et al*. 2009). Then, de novo chimera checking was performed using the function *chimera.uchime*. Following these steps to clean the data set, we evaluated errors in the mock diets using the *seq.error* command. This command outputs the number of mismatches between mock diet and reference sequences, from which per-nucleotide error rates can be calculated (i.e. N errant nucleotides read/N nucleotides read) (Schloss *et al*. 2011; Kozich *et al*. 2013).

To mitigate errors that inevitably arise from degraded DNA templates, polymerase or sequencing errors, undetected chimeras, low levels of undetected cross-contamination and/or true intraspecific variation (Pompanon *et al*. 2012), we clustered mOTUs at the 3% level using the functions *dist.seqs* and *cluster* in MOTHUR. We chose the 3% clustering level because: (i) the number of dietary mOTUs approached an asymptote at 3%; (ii) clustering was cut-off above the 4% level (i.e. it was not possible to cluster sequences at levels above 4% and to retain all the data); and (iii) clustering within samples at levels below 3% would be meaningless because sequences differing at this level (~2 bp) were previously merged as putative sequence errors (Schloss *et al*. 2009). Although the threshold that best captures species-level differences can vary among taxa and loci, it is not necessary to determine an exact threshold because results will be correlated across biologically relevant levels (3% falls within the 1–5% range that is considered realistic for insects; Clare *et al*. 2011). To facilitate taxonomic assignments, we selected representative sequences within mOTUs that best matched the *ecoTag* database and removed mOTUs with very poor matches (<80% identity) or that were obvious contaminants (i.e. human DNA detected in two samples). Finally, we removed mOTUs sequenced at low frequencies within samples (<0.001) as well as samples that probably exhibited a low target DNA quality or quantity (final total of <5000 sequence reads; less than the 5% quantile) to diminish the impact of low-level spurious sequences on our results (see Results pertaining to mock diets).

Automatic taxonomic assignments were inspected. Although we present most taxonomic affiliations of mOTUs at the order level, more refined identifications were often possible even in the absence of a comprehensive reference database. We considered species-level assignments reliable if we found an exact match (100% identity) to a single species in the *ecoTag* database. If a close but nonidentical match was made to one or more species (≥95% identity), we used *obitools* to classify the mOTU. Classifications for these well-matched sequences were consistently made at the order level (see Results). For marginal matches (80–95% identity), we confirmed that mOTUs corresponded to phylum Arthropoda and included them in the ecological analyses, but did not attempt finer taxonomic assignation because current algorithms have difficulty-making identifications based on low-identity sequence matches that result from incomplete reference databases (Luo *et al*. 2014). We note, however, that even when species-level identifications were not made, mOTUs can serve as an index of species diversity for use in ecologically meaningful dietary comparisons among groups of consumers.

Comparisons of dietary richness and composition

To illustrate the potential of our methods for testing more conceptual hypotheses, we compared dietary richness and composition among (i) field seasons and (ii) sexes. First, we used ANOVA to test for a significant difference in the mean number of mOTUs per sample between collections from (i) May 2013, December 2013 and May 2014 and (ii) males and females (excluding four samples from juveniles that could not be reliably sexed). We used ANCOVA to evaluate the effect of both sex and body size (SVL) on diet, although there was virtually no overlap in

SVL between sexes. Second, we compared overall mOTU richness using sample-based interpolation (i.e. rarefaction) and extrapolation in ESTIMATES 9.1.0 (Colwell 2013), which generates reliable curves for values up to two- to threefold greater than the actual sample size (Colwell *et al*. 2012). Complete nonoverlap of 95% confidence intervals as the rarefaction curves asymptote is a simple (albeit conservative) indication of significant differences in richness among groups (Colwell *et al*. 2012). Third, we compared dietary similarity of these groups (excluding a single extreme outlier from which only a single infrequent beetle was detected—mOTU193; Table S2, Supporting information) using the *adonis* function with 1000 permutations in VEGAN (Oksanen *et al*. 2013) in R. The *adonis* function performs robust multivariate analyses similar to perMANOVA (Oksanen *et al*. 2013). To enable visualization of similarity comparisons, we present nonmetric multidimensional scaling (NMDS) plots for each comparison. When multiple samples were successfully sequenced per individual, only the first was included in these analyses.

Results

Reference DNA

We analysed DNA from 145 locally collected arthropods representing the following 12 orders (number of specimens in parentheses): Araneae (20), Blattodea (8), Coleoptera (4), Decapoda (1), Diptera (12), Hemiptera (18), Hymenoptera (50), Isoptera (5), Lepidoptera (18), Littorinimorpha (1), Odonata (3) and Orthoptera (5) (see Table S1, Supporting information for complete list). We obtained 16S reference sequences from 113 of these specimens, yielding 62 unique sequences in the local reference database.

Primer design

In silico analysis of complete mitochondrial genomes yielded a suitable primer pair for analysis of coverage and specificity. Taxonomic coverage of these primers based on the compete arthropod mitochondrial genome database was estimated to comprise 87.5% of Arthropoda accessions, but only 19.1% of Squamata accessions. Importantly, this primer set did not amplify *Anolis sagrei*. The amplified region had considerable power to distinguish among taxa (i.e. specificity): 85.1% of families, 74.6% of genera and 50.5% of species were unambiguously discriminated from the global set of arthropod 16S accessions in GenBank. Arthropod orders known to be common in anole diets exhibited infrequent mismatches to the primers, few of which occurred in the final 3 nt (Fig. 1). The general utility of these primers for detecting invertebrate prey in vertebrate diets is indicated by the pattern of primer mismatches to a diverse array of consumers (subphylum Vertebrata; e.g. bats, birds, fish and frogs), in contrast to their better matches with a diverse array of arthropods (subphyla Chelicerata, Crustacea, Hexapoda and Myriapoda; Fig. S1, Supporting information).

Dietary DNA sequencing and identification

We successfully amplified arthropod DNA from ~51% of lizard faecal samples preserved in both RNAlater and Zymo Xpedition lysis buffer ($N = 177$ of 344 tested; Table 1). Sequencing yielded ~50 million reads from *A. sagrei* faecal samples obtained between May 2013 and 2014 (Table 1). After applying initial quality controls in *obitools*, 1378 unique arthropod sequences remained (Table 1). Subsequent removal of PCR or sequencing errors and clustering of mOTUs resulted in a final count of 217 mOTUs (93–151 per season; Table 1).

Error rates evaluated using the three mock diet samples were consistent across replicates (range: 0.012–0.014 errors per nucleotide). Four of five expected sequences were recovered from each mock diet, but one was consistently missing (*Camponotus* sp., BAHAR167-14; Table S1, Supporting information). The number of unique sequences per mock diet ranged from 73 to 110, of which 16–33% were flagged as chimeric and dropped. A bimodal mismatch distribution was apparent for remaining sequences and the cumulative proportion of sequence reads declined rapidly for sequences with ≥1 mismatch (a result of merging putative PCR and sequencing errors of up to 2 bp within samples), whereas sequences with large numbers of mismatches occurred at low frequency (Fig. S2, Supporting information). These low levels of nontarget sequences inflate error rate estimates (i.e. undetected chimeras, cross-contaminants and the DNA of prey consumed by the reference specimens).

The largest numbers of mOTUs with ≥95% identity were assigned to Diptera and Lepidoptera, with at least seven additional orders identified (Fig. 2). Three mOTUs occurred in ≥50% of samples and were assigned to Coleoptera, Blattodea and Hymenoptera (Table S2, Supporting information). However, the majority of mOTUs were infrequent (180, or 83%, occurred in ≤5% of samples; Table S2, Supporting information). In total, 26 perfect matches were made between dietary mOTUs and the *ecoTag* database (Fig. S3; Table S2, Supporting information). Most perfect matches (19 of 26; 70%) were to the local library, and >25% (59 of 217) of all dietary mOTUs were better matched to the local library than to *GB*—a substantial proportion given the limited size of the local database (Table S2, Supporting information). Approximately

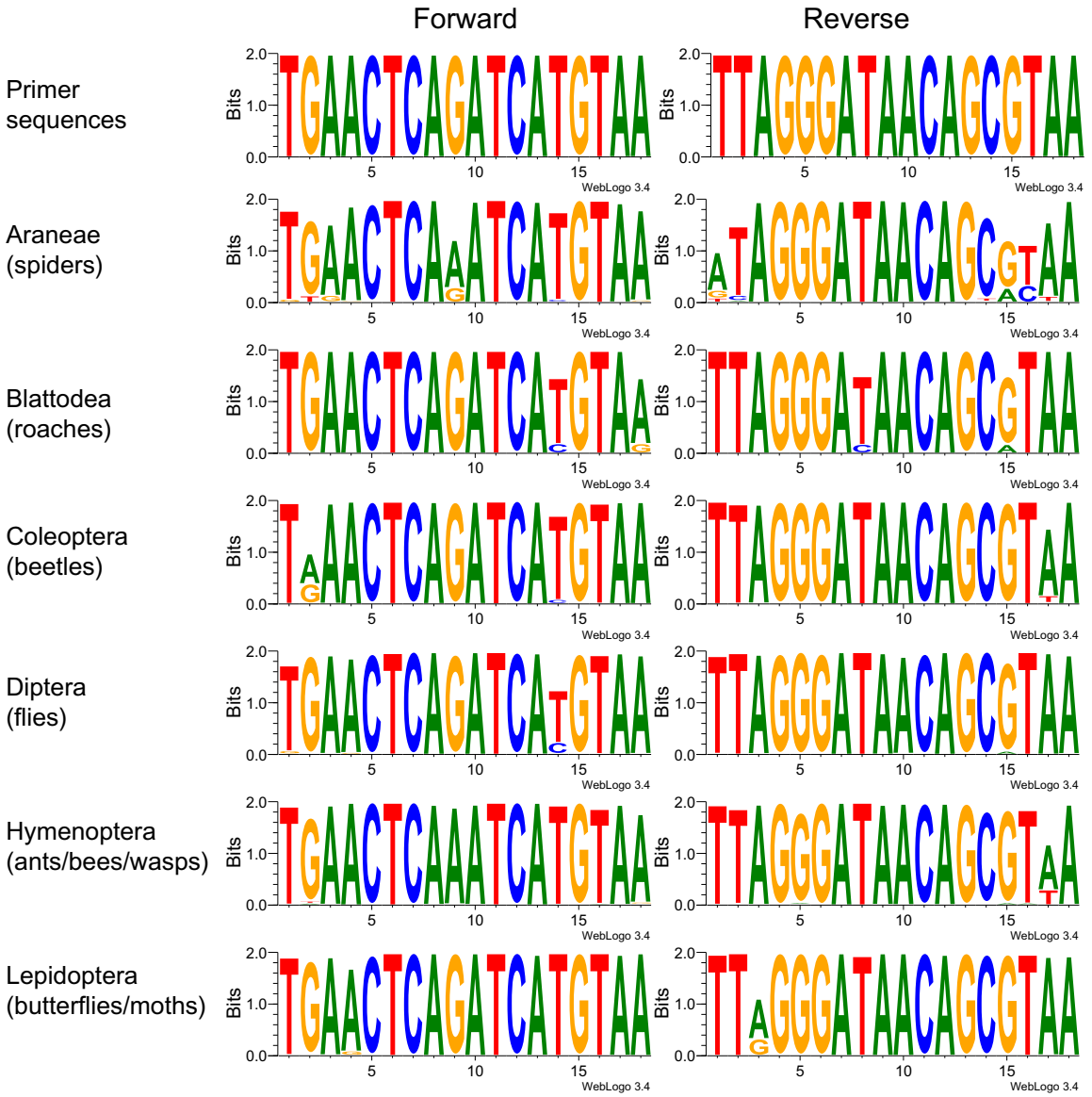

Fig. 1 WEBLOGOS depicting primer matches with major orders of *Anolis sagrei* prey.

66% of mOTUs (146 of 217) were 80–95% identical to a reference sequence (Fig. 2).

Comparisons of dietary richness and composition

There was no significant difference in mean number of mOTUs per sample across seasons ($F_{2,155} = 0.97$, $P = 0.38$; Fig. 3A), but females had a greater mean number of mOTUs per faecal sample than did males ($F_{1,152} = 4.11$, $P = 0.044$; Fig. 3B). Body size was greater for males (mean SVL = 5.56 cm ± 0.03 SE) than females (4.29 cm ± 0.05 SE; ANOVA $F_{1,148} = 365.89$, $P < 0.001$), but the main effect of sex on per-sample richness remained significant when SVL was included as a covariate (ANCOVA: $F_{1,146} = 4.53$, $P = 0.035$), and there was no significant effect of SVL ($F_{1,146} = 0.12$, $P = 0.73$) or sex × SVL interaction ($F_{1,146} = 1.84$, $P = 0.18$). Overall dietary mOTU richness did not differ significantly across seasons (Fig. 3C) or between sexes (Fig. 3D). Multivariate analysis revealed statistically significant dissimilarity of diet

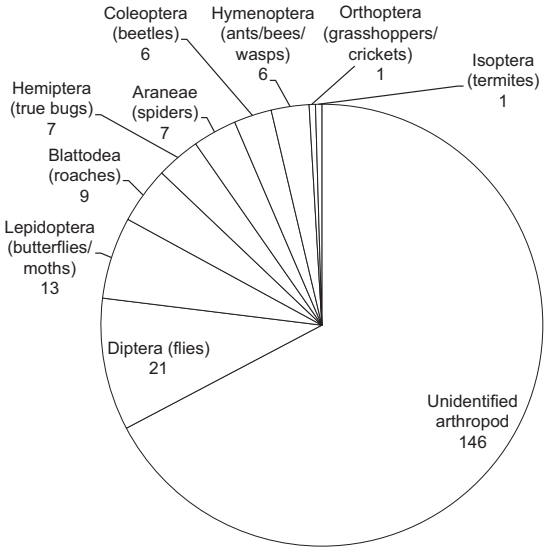

Fig. 2 Arthropod orders identified from *Anolis sagrei* faecal DNA. Pie charts show all 217 molecular operational taxonomic units (mOTUs), including the number of mOTUs in each group. See Tables S1 and S2 and Fig. S3 (Supporting information) for a complete list of reference sequences, arthropod mOTUs and photographs of reference specimens exactly matched to sequences from *Anolis sagrei* faecal samples.

composition across seasons (pseudo-$F_{2,155} = 3.72$, $R^2 = 0.046$, $P < 0.001$; Fig. 3E). Statistically significant relationships that do not explain much clustering (i.e. low R^2) can indicate differences in the degree of dispersion among groups (i.e. differences in the average distance of a sample to the centroid of its group). Indeed, we found significant pairwise differences in dispersion between the December and both of the May field seasons (post hoc test using the *betadisper* function in VEGAN; $P < 0.05$), but not between May 2013 and May 2014. This distance is reflected in the greater spread among samples taken in December versus May (Fig. 3E). In contrast, diet composition did not differ significantly by sex (pseudo-$F_{1,152} = 0.91$, $R^2 = 0.006$, $P = 0.590$, Fig. 3F).

Discussion

The methods described here improve our ability to detect and identify a broad range of arthropods in generalist vertebrate diets, facilitating characterization of diet profiles and their comparison among groups. Illumina sequencing of anole faecal samples yielded 217 mOTUs spanning a broad range of arthropods, fulfilling the intended purpose of maximizing coverage without also amplifying consumer DNA. We further illustrated the use of these data by evaluating several questions about dietary niche breadth and similarity that have general conceptual relevance in ecology and evolution.

The range of taxa sequenced corresponded well with prior compendia of prey identified to order level from Bahamian *Anolis sagrei* gut contents. Prey of populations in Bimini primarily comprised Lepidoptera and Coleoptera (Schoener 1968); in Exuma, Lepidoptera, Orthoptera and Coleoptera (Lister 1976); in Abaco, Lepidoptera, Coleoptera, Hemiptera (Lister 1976); and in our own site of Staniel Cay, Hemiptera, Hymenoptera and Coleoptera (Spiller & Schoener 1990). Even in such a well-studied system, metabarcoding revealed previously unreported predator–prey interactions. For example, the second most common dietary mOTU (occurring in >60% of samples) was a cockroach (*Blaberus* sp.; Fig. 4; Fig. S3; Table S2, Supporting information), yet cockroaches were not identified as frequent diet items in any of these prior studies.

It is noteworthy that despite a potential PCR bias in favour of insects (Fig. 1), we detected at least seven spider mOTUs in *A. sagrei* diets (Fig. 2), providing further evidence that DNA metabarcoding can detect relatively soft-bodied organisms that may be more completely digested and thus under-represented in traditional analyses of gut contents (Bowser *et al*. 2013). One spider mOTU (*Hibana* sp.) occurred in as many as 10% of samples per season (7% overall: Table S2, Supporting information). Our observation of spiders in 10% of *A. sagrei* diets falls within the range of 3.3% by weight (Lister 1976) and 23% by occurrence (Spiller & Schoener 1990) suggested by prior work in this system (Importantly, this latter analysis was of lizards from enclosures in which spider density was exceptionally high.). It remains to be determined whether arthropods that are distantly related to spiders and insects (e.g. subphyla Crustacea and Myriapoda) were (i) not frequently eaten by anoles at our study sites, (ii) not well identified due to the limitations of current reference databases (Fig. 2), or (iii) not detectable due to primer mismatches (Fig. S1, Supporting information).

Relative to conventional approaches, molecular dietary assessments have both strengths and weaknesses. Whereas comparisons of lizard diets based on gut content analyses are usually limited to coarse taxonomic distinctions (Lister 1976), greater detail about prey identity and diversity may ultimately be obtained using molecular approaches—and metabarcoding with a thorough reference library often enables a greater number of prey species to be determined (Cristescu 2014). Although many mOTUs from our analysis remain in need of precise taxonomic identification, our power to identify them will improve as the number of reference sequences increases. Moreover, unlike many gut contents studies in which data are forever limited to whatever level of

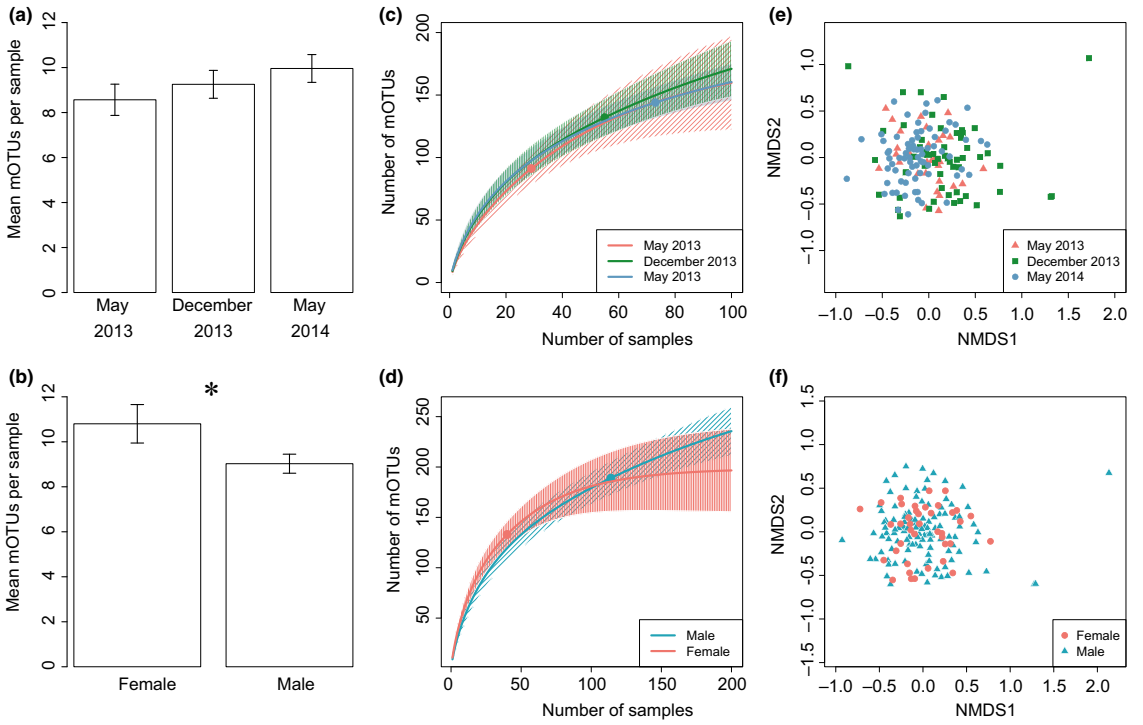

Fig. 3 Dietary richness and overlap. Seasons (top) and sexes (bottom) are compared in plots of (A,B) per-sample molecular operational taxonomic unit (mOTU) richness (±SE; *$P < 0.05$); (C,D) dietary mOTU accumulation curves of (±95% confidence intervals, with observed values plotted as a large circle on each curve); and (E,F) nonmetric multidimensional scaling plots showing two-dimensional representations of dietary similarity among groups.

Fig. 4 Adult male *Anolis sagrei* consuming a *Blaberus* sp. cockroach, the second most frequently detected prey item (molecular operational taxonomic unit 004; Table S2, Supporting information). (Photograph courtesy of Rowan D. H. Barrett.)

taxonomic resolution the original researcher achieved (because the gut contents are discarded), archived molecular data can be updated as reference coverage improves and can be electronically accessed by researchers worldwide. Moreover, for many ecological applications, such as comparisons of niche overlap, mOTUs are sufficient even in the absence of taxonomic identifications. However, compared to traditional gut content analyses, molecular methods are currently limited in their ability to resolve the relative abundance of different prey types in a sample (Spiller & Schoener 1990) without independent assessment of bias (Bowles *et al*. 2011; Thomas *et al*. 2013a) and cannot quantify prey size distributions (Pacala & Roughgarden 1985) or subdivide prey into ontogenetic stages (e.g. larvae vs. adults of holometabolous insects; Lister 1976), all of which have ecological implications.

Our approach to arthropod metabarcoding could be supplemented by additional methods to more completely study trophic interactions. For example, plants may also contribute to anole diets (Losos 2009), and the plant species in their diets can be identified using existing molecular tools (Pompanon *et al*. 2012). Likewise, if predation on a limited set of known species was of interest—for example the potential for intraguild predation by invasive *A. sagrei* on native *A. carolinensis* in Florida

(Campbell & Gerber 1996)—then species-specific DNA-detection approaches could be used (Murray *et al.* 2011; Egeter *et al.* 2014). Stable isotope analyses (or gut contents, if they can be obtained without jeopardizing other study objectives) can enable complementary ecological inferences about the source, quantities or ontogenetic stages of prey (especially when larvae and adults have distinct food sources leading to isotopic differentiation; Carreon-Martinez & Heath 2010). Finally, feeding experiments can help 'calibrate' some molecular dietary analysis, such that the number of sequence reads can reliably be used as a proxy for relative biomass of different prey types (Thomas *et al.* 2013b); this emerging frontier requires further research.

In comparing dietary richness and similarity across seasons and sexes, our aim was primarily to illustrate how dietary mOTU data generated by our approach can be used to address conceptually oriented ecological questions, even in the absence of Latin binomials for prey taxa (although species identifications are always desirable). We found significant variation in diet composition between May and December (a comparatively cooler and drier month). One plausible explanation for this finding is the fact that 'the composition of the diet of many lizards changes as the relative abundances of different types of prey fluctuate with the seasons' (Pianka 1973). Additional factors, such as variation in feeding selectivity or foraging activity, might also have contributed to this result, but evaluating them would require additional data on lizard foraging behaviour and temporal fluctuations in prey populations. We also showed that although per-sample dietary richness was greater for females than males, both consumed a statistically indistinguishable number of arthropod mOTUs overall. Thus, our data set provides no support for the hypothesis that sex-based size differences reduce niche overlap. Previous studies have found evidence for intersexual niche partitioning in some cases (e.g. *Anolis conspersus*: Schoener 1967) and not in others (e.g. *Carlia* skinks: Manicom *et al.* 2014) based on prey size alone. Because our approach cannot evaluate differences in prey size per se, future studies might couple metabarcoding with prey size data to more completely evaluate niche partitioning between sexes.

It is possible that 'secondary predation' (i.e. prey of prey) detected by our metabarcoding methods might have influenced our results (Bowser *et al.* 2013). For example, spiders consumed by *A. sagrei* might contain DNA of insects previously consumed by spiders (Bowser *et al.* 2013; Piñol *et al.* 2014), although this concern does not apply to exclusively herbivorous lizard prey items (e.g. insect genera *Artipus*, *Acanalonia*, *Loxa*, *Nasutitermes*, *Strymon*: Table S2, Fig. S3, Supporting information). In addition to chimeras, some of the spurious sequences in our mock diet samples might have originated from gut contents of arthropod predators in lizard diets. Our protocols could be useful for future studies addressing questions that require molecular characterization of multitrophic interactions (Bowser *et al.* 2013).

Poor taxonomic identification of mOTUs can result from (i) the absence of prey sequences in public or local databases, resulting in low percent identity; and/or (ii) the presence of multiple taxa with identical sequences in public databases, but not the local prey pool (or vice versa), causing precise identifications to be under- (or over-) reported. To account for uncertainty in public data (and even utilize incompletely overlapping accessions), it is possible to apply the BLAST algorithm to a data set like ours and assess how many mOTUs can be assigned to successively higher taxonomic levels (using e.g. Huson *et al.* 2011), as demonstrated in molecular diet analyses of bats (Emrich *et al.* 2014). Nevertheless, a purported advantage of metabarcoding is that exact dietary assignments can be made, freeing investigators from analytical difficulties associated with arbitrary clustering thresholds and potentially high misidentification rates (Valentini *et al.* 2009; De Barba *et al.* 2014). Thus, we emphasize the importance of local reference libraries, which would strengthen future uses of this method by (i) reducing reliance on loosely curated public databases; (ii) increasing the frequency of exact matches; (iii) decreasing spurious matches; and (iv) replacing dependence on clustering algorithms for error-checking and taxonomic assignments with more stringent sequence-matching criteria (Pompanon *et al.* 2012).

Because a substantial global effort has been made to sequence arthropods using the standard COI barcode locus, existing COI-based approaches to identify unknown arthropod sequences (Zeale *et al.* 2011) are poised to more precisely classify sequences using public repositories (Ratnasingham & Hebert 2007). However, given the difficultly in designing broad-coverage primers for environmental DNA analysis within the hypervariable COI region (Deagle *et al.* 2014), we and others having increasingly opted to use short, general 16S primers in lieu of or in conjunction with COI to evaluate insectivore, piscivore and carnivore diets (Boyer *et al.* 2011; Pompanon *et al.* 2012; Bowser *et al.* 2013; Clarke *et al.* 2014; De Barba *et al.* 2014; Waterhouse *et al.* 2014). Although the value of COI as a standard DNA barcode for taxonomic discrimination is well established (Hebert *et al.* 2003), the potential advantages of 16S markers for analyses of environmental DNA (including faecal DNA) suggest the importance of archiving well-annotated reference sequences from both loci whenever possible.

Acknowledgements

We thank R. Barrett, T. Coverdale, J. Daskin, N. Man in 't Veld, T. Palmer, D. Spiller, T. Thurman, and L. Wyman for assistance in the field. P. Chen, M. Kuzmina, B. von Holt, W. Wang and the staff of Princeton's Lewis-Sigler core sequencing facility provided invaluable assistance with laboratory analyses. We thank three anonymous reviewers for constructive comments on the original manuscript. This work was conducted in accordance with applicable permits from the Bahamas Environment, Science & Technology Commission and Princeton University's Institutional Animal Care & Use Committee (permit #1922). Funding was provided by Princeton University.

References

Andrews RM (1979) Evolution of life histories: a comparison of *Anolis* lizards from matched island and mainland habitats. *Breviora*, **454**, 1–51.

Baamrane MAA, Shehzad W, Ouhammou A *et al.* (2012) Assessment of the food habits of the Moroccan Dorcas Gazelle in M'Sabih Talaa, West Central Morocco, using the *trn*L approach. *PLoS One*, **7**, e35643.

Bowles E, Schulte PM, Tollit DJ, Deagle BE, Trites AW (2011) Proportion of prey consumed can be determined from faecal DNA using real-time PCR. *Molecular Ecology Resources*, **11**, 530–540.

Bowser KA, Diamond AW, Addison JA (2013) From Puffins to Plankton: a DNA-based analysis of a seabird food chain in the Northern Gulf of Maine. *PLoS One*, **8**, e83152.

Boyer S, Yeates GW, Wratten SD, Holyoake A, Cruickshank RH (2011) Molecular and morphological analyses of faeces to investigate the diet of earthworm predators: example of a carnivorous land snail endemic to New Zealand. *Pedobiologia*, **54**, S153–S158.

Campbell TS, Gerber GP (1996) *Anolis sagrei*. Saurophagy. *Herpetological Review*, **27**, 200.

Carreon-Martinez L, Heath DD (2010) Revolution in food web analysis and trophic ecology: diet analysis by DNA and stable isotope analysis. *Molecular Ecology*, **19**, 25–27.

Clare EL, Barber BR, Sweeney BW, Hebert PDN, Fenton MB (2011) Eating local: influences of habitat on the diet of little brown bats (*Myotis lucifugus*). *Molecular Ecology*, **20**, 1772–1780.

Clarke LJ, Soubrier J, Weyrich LS, Cooper A (2014) Environmental metabarcodes for insects: in silico PCR reveals potential for taxonomic bias. *Molecular Ecology Resources*, **14**, 1160–1170.

Colwell RK (2013) ESTIMATES: Statistical estimation of species richness and shared species from samples. Version 9. User's Guide and application published at: http://purl.oclc.org/estimates.

Colwell RK, Chao A, Gotelli NJ *et al.* (2012) Models and estimators linking individual-based and sample-based rarefaction, extrapolation, and comparison of assemblages. *Journal of Plant Ecology*, **5**, 3–21.

Cristescu ME (2014) From barcoding single individuals to metabarcoding biological communities: towards an integrative approach to the study of global biodiversity. *Trends in Ecology & Evolution*, **29**, 566–571.

Crooks GE, Hon G, Chandonia J-M, Brenner SE (2004) WebLogo: a sequence logo generator. *Genome Research*, **14**, 1188–1190.

De Barba M, Miquel C, Boyer F *et al.* (2014) DNA metabarcoding multiplexing and validation of data accuracy for diet assessment: application to omnivorous diet. *Molecular Ecology Resources*, **14**, 306–323.

Deagle BE, Eveson JP, Jarman SN (2006) Quantification of damage in DNA recovered from highly degraded samples—a case study on DNA in faeces. *Frontiers in Zoology*, **3**, 11.

Deagle BE, Jarman PJ, Coissac E, Pompanon F, Taberlet P (2014) DNA metabarcoding and the cytochrome *c* oxidase subunit I marker: not a perfect match. *Biology Letters*, **10**, 20140562.

Egeter B, Bishop PJ, Robertson BC (2014) Detecting frogs as prey in the diets of introduced mammals: a comparison between morphological and DNA-based diet analyses. *Molecular Ecology Resources*, doi:10.1111/1755-0998.12309. [Epub ahead of print].

Emrich MA, Clare EL, Symondson WOC, Koenig SE, Fenton MB (2014) Resource partitioning by insectivorous bats in Jamaica. *Molecular Ecology*, **23**, 3648–3656.

Ficetola GF, Coissac E, Zundel S *et al.* (2010) An in silico approach for the evaluation of DNA barcodes. *BMC Genomics*, **11**, 434.

Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek R (1994) DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Molecular Marine Biology and Biotechnology*, **3**, 294–299.

Hebert PDN, Cywinska A, Ball SL, deWaard JR (2003) Biological identifications through DNA barcodes. *Proceedings of the Royal Society of London, Series B: Biological Sciences*, **270**, 313–321.

Hibert F, Taberlet P, Chave J *et al.* (2013) Unveiling the diet of elusive rainforest herbivores in next generation sequencing era? The tapir as a case study. *PLoS One*, **8**, e60799.

Huang S-C, Norval G, Tso I-M (2008) Predation by an exotic lizard, *Anolis sagrei*, alters the ant community structure in betelnut palm plantations in southern Taiwan. *Ecological Entomology*, **33**, 569–576.

Huson DH, Mitra S, Ruscheweyh HJ, Weber N, Schuster SC (2011) Integrative analysis of environmental sequences using MEGAN4. *Genome Research*, **21**, 1552–1560.

Ivonona N, Grainger C, Hajibabaei M (2006) Glass fiber DNA extraction: A new inexpensive method for high throughput DNA isolation. *CCDB Advances*, Methods Release No. 1, November 3rd, 2006.

Kolbe JJ, Glor RE, Rodriguez-Schettino L *et al.* (2004) Genetic variation increases during biological invasion by a Cuban lizard. *Nature*, **431**, 177–181.

Kozich JJ, Westcott SL, Baxter NT, Highlander SK, Schloss PD (2013) Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on MiSeq Illumina sequencing platform. *Applied and Environmental Microbiology*, **79**, 5112–5120.

Lister BC (1976) The nature of niche expansion in West Indian *Anolis* lizards I: ecological consequences of reduced competition. *Evolution*, **30**, 659–676.

Losos JB (2009) *Lizards in an Evolutionary Tree: Ecology and Adaptive Radiation of Anoles*. University of California Press, Berkeley, CA.

Losos JB, Spiller DA (1999) Differential colonization success and asymmetrical interactions between two lizard species. *Ecology*, **80**, 252–258.

Luo C, Rodriguez-R LM, Konstantinidis KT (2014) MyTaxa: an advanced taxonomic classifier for genomic and metagenomic sequences. *Nucleic Acids Research*, **14**, 1–12.

Manicom C, Alford R, Schoener TW, Schwarzkopf L (2014) Mechanisms causing variation in sexual size dimorphism in three sympatric, congeneric lizards. *Ecology*, **85**, 1531–1544.

Murray DC, Bunce M, Cannell BL *et al.* (2011) DNA-Based faecal dietary analysis: a comparison of qPCR and high throughput sequencing approaches. *PLoS One*, **6**, e25776.

Oksanen J, Blanchet FG, Kindt R *et al.* (2013) VEGAN: Community Ecology Package, http://CRAN.R-project.org/package=vegan.

Pacala SW, Roughgarden J (1985) Population experiments with the *Anolis* lizards of St. Maarten and St. Eustatius. *Ecology*, **66**, 129–141.

Palumbi SR (1996) Nucleic acids II: the polymerase chain reaction. In: *Molecular Systematics* (eds Hillis DM, Moritz C & Mable BK), pp. 205–247. Sinauer Associates, Inc., Sunderland, MA.

Pianka ER (1973) The structure of lizard communities. *Annual Review of Ecology and Systematics*, **4**, 53–74.

Piñol J, San Andrès V, Clare EL, Mir G, Symon Dson WOC (2014) A pragmatic approach to the analysis of diets of generalist predators: the use of next-generation sequencing with no blocking primers. *Molecular Ecology Resources*, **14**, 18–26.

Pompanon F, Deagle BE, Symondson WOC *et al.* (2012) Who is eating what: diet assessment using next generation sequencing. *Molecular Ecology*, **21**, 1931–1950.

R Core Development Team (2013) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Ratnasingham S, Hebert PDN (2007) BOLD: the barcode of life datasystem (http://www.barcodinglife.org). *Molecular Ecology Notes* **7**, 355–364.

Riaz T, Shehzad W, Viari A *et al*. (2011) ecoPrimers: inference of new DNA barcode markers from whole genome sequence analysis. *Nucleic Acids Research*, **39**, e145.

Roughgarden J (1995) *Anolis Lizards of the Caribbean: Ecology, Evolution, and Plate Tectonics (Oxford Series in Ecology & Evolution)*. Oxford University Press, New York, NY.

Rummel JD, Roughgarden J (1985) Effects of reduced perch-height separation on competition between two *Anolis* lizards. *Ecology*, **66**, 430–444.

Schloss PD, Westcott SL, Ryabin T *et al*. (2009) Introducing MOTHUR: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Applied and Environmental Microbiology*, **75**, 7531–7541.

Schloss PD, Gevers D, Westcott SL (2011) Reducing the effects of PCR amplification and sequencing artifacts on 16S rRNA-based studies. *PLoS One*, **6**, e27310.

Schoener TW (1967) The ecological significance of sexual dimorphism in size in the lizard *Anolis conspersus*. *Science*, **155**, 474–477.

Schoener TW (1968) The *Anolis* lizards of Bimini: resource partitioning in a complex fauna. *Ecology*, **49**, 704–726.

Schoener TW (2011) The newest synthesis: understanding the interplay of evolutionary and ecological dynamics. *Science*, **331**, 426–429.

Schoener TW, Spiller DA, Losos JB (2002) Predation on common *Anolis* lizard: can the food-web effects of a devastating predator be reversed. *Ecological Monographs*, **72**, 383–407.

Shehzad W, Riaz T, Nawaz MA *et al*. (2012) Carnivore diet analysis based on next-generation sequencing: application to the leopard cat (*Prionailurus bengalensis*) in Pakistan. *Molecular Ecology*, **21**, 1951–1965.

Shine R (1989) Ecological causes for the evolution of sexual dimorphism: a review of the evidence. *Quarterly Review of Biology*, **64**, 419–461.

Spiller DA, Schoener TW (1990) Lizards reduce food consumption by spiders: mechanisms and consequences. *Oecologia*, **83**, 150–161.

Thomas AC, Jarman SN, Haman KH, Trites AW, Deagle BE (2013a) Improving accuracy of DNA diet estimates using food tissue control materials and an evaluation of proxies for digestion bias. *Molecular Ecology*, **23**, 3706–3718.

Thomas AC, Jarman SN, Haman KH, Trites AW, Deagle BE (2013b) Improving accuracy of DNA diet estimates using food tissue control materials and an evaluation of proxies for digestion bias. *Molecular Ecology*, **23**, 3706–3718.

Valentini A, Pompanon F, Taberlet P (2009) DNA barcoding for ecologists. *Trends in Ecology & Evolution*, **24**, 110–117.

Waterhouse BR, Boyer S, Wratten SD (2014) Pyrosequencing of prey DNA in faeces of carnivorous land snails to facilitate ecological restoration and relocation programs. *Oecologia*, **175**, 737–746.

Williams EE (1972) The origin of faunas. Evolution of island congeners in a complex island fauna: a trial analysis. *Evolutionary Biology*, **94**, 2886–2897.

Wolcott GN (1923) The food of Porto Rican lizards. *Journal of the Department of Agriculture of Porto Rico*, **7**, 5–43.

Zeale MR, Butlin RK, Barker GA, Lees DC, Jones G (2011) Taxon-specific PCR for DNA barcoding arthropod prey in bat faeces. *Molecular Ecology Resources*, **11**, 236–244.

T.R.K. and R.M.P. conceived the study, designed the study, collected samples, and wrote the manuscript. T.R.K. developed the laboratory protocols and analysed the molecular data.

Data accessibility

Reference DNA sequences: GenBank accessions KP253753–KP253873 (COI) and KP253874–KP253981 (16S).

Reference DNA descriptors and photographs: Online Supporting information.

Dietary mOTU sequences: DRYAD entry doi: 10.5061/dryad.f18k5.

Metabarcode sequence data (fastq files and corresponding MID tags): DRYAD entry doi: 10.5061/dryad.f18k5.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Table S1 Specimens in our local Bahamian arthropod reference library.

Table S2 Arthropod mOTUs detected in *Anolis sagrei* diets, listed in decreasing order of frequency of occurrence.

Fig. S1 Primer mismatches with sequences of representative potential vertebrate consumers and arthropod subphyla from GenBank accessions.

Fig. S2 Mock diet error distributions.

Fig. S3 Photographs of reference DNA specimens exactly matched to dietary mOTUs (Tables S1 and S2).