Cuff, J. P., Drake, L. E., Tercel, M. P. T. G., Stockdale, J. E., Orozco□ter Wengel, P., Bell, J. R., Vaughan, I. P., Muller, C. T. and Symondson, W. O. C. 2020. Money spider dietary choice in pre□ and post□harvest cereal crops using metabarcoding. *Ecological Entomology.* https://doi.org/10.1111/een.12957

Ecological Entomology

Ecological Entomology (2020), DOI: 10.1111/een.12957

Money spider dietary choice in pre- and post-harvest cereal crops using metabarcoding

JORDAN P. CUFF,[1,2] LORNA E. DRAKE,[1] MAXIMILLIAN P.T.G. TERCEL,[1] JENNIFER E. STOCKDALE,[1,3] PABLO OROZCO-TERWENGEL,[1] JAMES R. BELL,[2] IAN P. VAUGHAN,[1] CARSTEN T. MÜLLER[1] and WILLIAM O.C. SYMONDSON[1] [1]School of Biosciences, Cardiff University, Cardiff, UK, [2]Rothamsted Insect Survey, Rothamsted Research, Harpenden, UK and [3]Faculty of Medicine & Health Sciences, University of Nottingham, University Park, Nottingham, UK

Abstract. 1. Money spiders (Linyphiidae) are an important component of conservation biological control in cereal crops, but they rely on alternative prey when pests are not abundant, such as between cropping cycles. To optimally benefit from these generalist predators, prey choice dynamics must first be understood.

2. Money spiders and their locally available prey were collected from cereal crops 2 weeks pre- and post-harvest. Spider gut DNA was amplified with two novel metabarcoding primer pairs designed for spider dietary analysis, and sequenced.

3. The combined general and spider-exclusion primers successfully identified prey from 15 families in the guts of the 46 linyphiid spiders screened, whilst avoiding amplification of *Erigone* spp. The primers show promise for application to the diets of other spider families such as Agelenidae and Pholcidae.

4. Distinct invertebrate communities were identified pre- and post-harvest, and changes in spider diet and, to a lesser extent, prey choice reflected this. Spiders were found to consume one another more than expected, indicating their propensity towards intraguild predation, but also consumed common pest families.

5. Changes in spider prey choice may redress prey community changes to maintain a consistent dietary intake. Consistent provision of alternative prey via permanent refugia should be considered to sustain effective conservation biocontrol.

Key words. Araneae, biological control, Linyphiidae, predator–prey interactions.

Introduction

Effective pest control requires an integrated, augmentative approach, aimed at maximising the effectiveness of natural enemies (Bale *et al.*, 2008; Peterson *et al.*, 2016). Spiders are polyphagous generalist predators naturally abundant at densities of 200–600 m^{-2} in UK crops (Nyffeler & Sunderland, 2003; Shayler, 2005). As a community, they employ a diversity of foraging techniques (Turnbull, 1973; Riechert & Lockley, 1984) which influence food webs via a range of hunting strategies including passive sit-and-wait predation from webs, and active hunting (Michalko & Pekár, 2016). Acknowledgement that spiders are an effective biocontrol agent has existed for decades (Riechert & Lockley, 1984; Sunderland *et al.*, 1997; Sunderland, 1999) since they regularly consume pests such as aphids (Sunderland *et al.*, 1986; Beck & Toft, 2000; Mayntz & Toft, 2000; Bilde & Soren, 2001; Harwood *et al.*, 2003; Nyffeler & Sunderland, 2003), planthoppers (Wang *et al.*, 2016; Wang *et al.*, 2017), psyllids (Petráková *et al.*, 2016), medflies (Monzó *et al.*, 2010), lepidopterans (Quan *et al.*, 2011; Pérez-Guerrero *et al.*, 2013; Senior *et al.*, 2016), and weevils (Vink & Kean, 2013). Whilst crop rotation disrupts biocontrol by many generalist predators, spider generation times often coincide with crop cycles, with early pest population establishment coinciding with peak spider abundances in Spring, thus facilitating early pest suppression (Riechert & Lockley, 1984; Symondson *et al.*, 2002; Harwood & Obrycki, 2005; Welch *et al.*, 2011).

Harvest, akin to mass deforestation at the scale of a spider, changes the fundamental structure of macro- and micro-habitats,

Correspondence: School of Biosciences, Cardiff University, Museum Avenue, Cardiff, CF10 3AX. cuffjp@cardiff.ac.uk

causing major changes to invertebrate community composition and interactions through immigration/emigration and potential exposure to other predators (Opatovsky & Lubin, 2012; Davey *et al.*, 2013). The large degree of turnover in invertebrate communities following harvest profoundly affects the diet of generalist predators, with the changes in spatial co-occurrence of predator and prey fundamentally influencing predation events (Bell *et al.*, 2010). Given that many spiders over-winter in the field, field margins, and hedgerows (Sunderland & Samu, 2000), the post-harvest provision of prey for these predators will influence their abundance and ability to suppress early pest populations in the subsequent crop cycle (Symondson *et al.*, 2002). To understand more precisely how harvest affects spider behavioural dynamics and how to optimise prey availability to support over-wintering and early-season spiders, the prey choice and dietary dynamics of spiders during this period must first be analysed.

Web-building spiders are effective models of prey choice, with webs providing a proxy for foraging investment (Welch *et al.*, 2016). Spiders are known to forgo abundant prey in favour of less locally abundant taxa (Agustí *et al.*, 2003; Welch *et al.*, 2016). Studies of spider prey choice have mostly consisted of laboratory feeding trials (e.g. Mayntz *et al.*, 2005; Rendon *et al.*, 2019). Given that spiders are fluid feeders, morphological analysis of gut contents is impossible, thus field studies are restricted to direct observation and molecular methods (Symondson, 2002; Harwood *et al.*, 2004; Bell *et al.*, 2010; Pompanon *et al.*, 2012; Chapman *et al.*, 2013; Birkhofer *et al.*, 2017). The parallel identification of many species by variations in short sections of variable genes, termed DNA metabarcoding (Pompanon *et al.*, 2011; Taberlet *et al.*, 2012), is one such molecular method for field-based analysis of spider diet (Piñol *et al.*, 2014; Lafage *et al.*, 2019). A target gene must be selected and a region identified which is variable between target species (the barcode) but flanked by two conserved regions, commonly in the cytochrome c oxidase subunit I (COI) gene for studies of animals due to its mutation rate and the extensive reference libraries available (Hebert *et al.*, 2003; Deagle *et al.*, 2014). These short sections of DNA, typically of 200–400 base pairs, can then be amplified using polymerase chain reactions (PCRs) with primers. Primers are short synthetic oligonucleotides which are complementary to the regions either side of the barcode that are conserved across the target taxa; these catalyse the amplification of the short sections of DNA (Folmer *et al.*, 1994; Piñol *et al.*, 2018). Using high-throughput sequencing (HTS), such PCR-amplified short sections of DNA can be identified in parallel from a single sample, such as the prey in a spider's gut. This ultimately provides an efficient and accurate method for analysing the prey range of predators (Pompanon *et al.*, 2012; Piñol *et al.*, 2018).

Two factors, however, must be overcome if HTS and DNA metabarcoding are to be used for dietary analysis of arthropod-consuming arthropods: (1) if the predator is an arthropod, then there is a high probability that existing general arthropod PCR primers will amplify the predator as well as the prey in its guts; (2) tissue from the spider predator will be undegraded and will hence swamp amplification of the prey (Vestheim & Jarman, 2008; Piñol *et al.*, 2014). Applying universal primers and accepting a loss of data to amplification of the predator is feasible with sufficient read depth and allows analysis of interactions between closely-related species (Piñol *et al.*, 2014). Blocking primers can be used alongside universal primers for the prevention of amplification of the predator whilst still amplifying the prey (Vestheim & Jarman, 2008; Deagle *et al.*, 2009), but these can introduce strong taxonomic biases (Piñol *et al.*, 2015, 2018). Primers can be designed to exclude amplification of the predator whilst still amplifying a broad range of prey species, and such primers have been designed for wolf spiders (Lycosidae; Lafage *et al.*, 2019). However, these primers amplify money spider DNA and are thus not appropriate for analysing linyphiid gut contents. Taxonomically-similar predators and prey, such as intraguild spider-spider interactions, may also be undetected when using primers that exclude amplification of the predator DNA (Vestheim & Jarman, 2008; Piñol *et al.*, 2014).

In this study, we aimed to analyse the diets and prey choices of linyphiid money spiders (Linyphiidae) in cereal crops pre- and post-harvest using DNA metabarcoding. To facilitate this, we developed novel PCR primers for the analysis of spider diet using high-throughput sequencing, with a specific focus on the diet of linyphiids. We hypothesised that linyphiid prey availability and diet would change following harvest, with diet largely reflecting prey assemblage turnover. However, we expected the prevalence of some prey species in the diet of the linyphiid predators to be disproportionate to their availability, reflecting prey choice.

Materials and methods

Primer development and testing

Existing PCR primers were tested and ultimately redesigned to match the target taxa of this study. Two novel primer pairs were used for amplification of DNA for the dietary analysis of spider gut contents to overcome the problems associated with the taxonomic proximity of spiders and their prey (particularly other spider species). Novel PCR primers were adapted for the exclusion of all spider DNA, with a focus on linyphiids (henceforth spider exclusion primers, titled TelperionF-LaurelinR), based upon a primer site slightly 3′ of the general animal barcoding primers LCO1490 (forward primer Folmer *et al.*, 1994), and mICOIintR (Leray *et al.*, 2013). A second primer pair was employed for broad amplification of both spiders and their prey (henceforth general primers, titled BerenF-LuthienR), based upon mICOIintF (Leray *et al.*, 2013) and HCO2198 (Folmer *et al.*, 1994). Both primer pairs were adapted via base changes designed with reference to mass alignments of invertebrate COI sequences and tested *in silico* and *in vitro*. The spider-exclusion primers were designed to overcome the loss of reads to predator DNA, whilst the general primers were designed to avoid the taxonomic biases associated with the exclusion primers.

Mass-alignments of COI sequences were batch-downloaded from GenBank (NCBI) and BOLD (Ratnasingham & Hebert, 2007) using PrimerMiner (Elbrecht & Leese, 2016) in R v.3.3.4 (R Core Team, 2020) to aid visual inspection of existing

Table 1. Primers designed via PrimerMiner within the COI gene.

Primer	Sequence (5′–3′)	Source	3′ Location	Direction	BP
TelperionF (spider excluding)	GGAACWHTATAYTTWATWTTYGG	This study	1535	F	23
LaurelinR (spider excluding)	GGRTAWACWGTTCAWCCWGT	Adapted from mICOIintR (Leray *et al.*, 2013)	1837	R	20
BerenF (general)	CAGGWTGAACWGTWTAYCCYCC	Adapted from mICOIintF (Leray *et al.*, 2013)	1859	F	22
LuthienR (general)	ACTTCWGGRTGWCCAAARAAYCA	Adapted from HCO2198 (Folmer *et al.*, 1994)	2173	R	23

Note: The designed primer pairs, TelperionF–LaurelinR and BerenF–LuthienR, with amplicon sizes of 302 and 314 bp, respectively.

and novel primer sites. PrimerMiner clusters batch downloads into operational taxonomic units (OTUs) based on sequence similarity and visualises mass alignments of sequence data for primer design. By merging overrepresented and duplicate sequences through taxonomy-independent clustering, PrimerMiner accounts for within-species variation and cryptic species whilst ignoring rare haplotypes (Elbrecht & Leese, 2016). Sequences were downloaded for all terrestrial invertebrate orders available, and consensus sequences were created by clustering these into OTUs for each order. The COI sequences were trimmed to include only the Folmer region (Folmer *et al.*, 1994) using Geneious R10 (Kearse *et al.*, 2012) for subsequent use in PrimerMiner. Alignments of prey sequences created via PrimerMiner included cereal crop spiders, in order to find primer sites conserved between a wide range of potential prey, but different for spiders. Where these sites were 100–400 base pairs apart on sequences from one another or from existing primer sites, they were paired, and primers designed (Table 1, Figs. S2 and S3). Existing general invertebrate primer sites were compared against the PrimerMiner alignments to identify any potential improvements to the primers for the amplification of cereal spider prey. The coverage of primers (% amplified) was determined via PrimerMiner using the same mass alignments used for primer design. PrimerMiner uses a taxonomy-independent database and accounts for adjacent base mismatches and the position of each base in the primer. Primers were also analysed using the online ThermoFisher Scientific Multiple Primer Analyzer tool. After the primers were deemed successful *in silico*, they were tested *in vitro*.

The primer pairs were tested *in vitro* against a wide range of extracted invertebrate DNA including spiders, common spider prey, and additional invertebrates. For this, invertebrate samples included those collected from the field site at Burdons Farm, Wenvoe, South Wales, the study site used for subsequent ecological analysis. Invertebrates were found via manual searching, collected via aspirator, and placed in microcentrifuge tubes of 100% ethanol. These were identified at 20-50X magnification using a light stereomicroscope and taxonomic keys (Goulet & Huber, 1993; Roberts, 1993; Unwin, 2001; Ball, 2008; Barber, 2008; Duff, 2012; Dallimore & Shaw, 2013). Additional invertebrates and DNA were taken from existing archived collections within Cardiff University. Extraction of DNA used DNeasy Blood & Tissue Kits (QIAGEN Inc., Chatsworth, CA) following the manufacturer's protocol for animal tissue. For predatory invertebrates, DNA was extracted from the lower legs, excluding the femur, to avoid the inclusion of prey DNA in the gut diverticulae and leg coxae (Macıas-Hernández *et al.*, 2018). To verify successful extraction, the DNA and negative controls were amplified via PCR with the Qiagen PCR Multiplex Kit (Qiagen) with 95 °C for 15 min to activate the HotStarTaq® DNA polymerase, 35 cycles of 95 °C for 30 s, 40 °C for 90 s and 72 °C for 90 s, respectively, followed by a final extension at 72 °C for 10 min using universal invertebrate primers LCO1490 and HCO2198 (Folmer *et al.*, 1994). PCR reactions comprised 25 µl reaction volumes containing 12.5 µl Qiagen PCR Multiplex kit, 0.2 µmol (0.5 µl of 10 µM stock) of each primer, 6.5 µl DNase-free, water and 5 µl template DNA. Amplification was confirmed by gel electrophoresis.

Primers were initially tested against a small selection of spider and non-spider DNA. Temperature-gradient PCRs were used to determine the optimal annealing temperatures for the primer pairs selected, with temperatures between 40 and 60 °C considered and initial tests starting 5 °C lower than the mean melting temperature of both primers. Inclusion of the Q reagent supplied with Multiplex Kits was trialled for each pair to ascertain whether this could improve performance, but was ultimately excluded in all cases. PCR conditions were: 95 °C for 15 min to activate the HotStarTaq® DNA polymerase, 35 cycles of 95 °C, the annealing temperature and 72 °C for 30, 90, and 90 s, respectively, and a final extension at 72 °C for 10 min. PCR reactions comprised 25 µl reaction volumes containing 12.5 µl Qiagen PCR Multiplex kit, 0.2 µmol (0.5 µl of 10 µM stock) of each primer, 6.5 µl DNase-free, water and 5 µl template DNA. Successful amplification was confirmed by gel electrophoresis. Once optimised to amplify a range of non-spider species whilst amplifying few spiders, or amplify a broad range of all species included, primers were further tested on a broader range of DNA (Table S3).

The TelperionF–LaurelinR primer pair has well-conserved sites, facilitating broad coverage with few degenerate bases necessary. The terminal base at the 3′ end of Laurelin, being a thymine base, critically mismatches with the guanine base present for most spider taxa tested; this should theoretically prevent or at least severely reduce amplification of spiders with little cost to amplification breadth otherwise. The BerenF–LuthienR pair similarly makes use of conserved primer sites employed in other studies but adapted for universal amplification of the focal taxa of this study.

Field collection and identification

Linyphiids were visually located on transects through two adjacent spring barley fields at Burdons Farm, Wenvoe in

South Wales (51°26′24.8″N, 3°16′17.9″W), and collected from occupied webs and the ground, in August and September 2017. Each transect comprised 4 m^2 searching areas at least 10 m apart and all observed linyphiids were collected. Spiders were taken from 20 locations along the aforementioned transects, 10 pre-harvest, and 10 post-harvest. Spiders were collected 2 weeks prior (7–13th August) to harvest (~20th August) of the crop and 2 weeks after harvest (4–8th September) in crop stubble and placed in 100% ethanol using an aspirator. Ground-active linyphiid spiders were collected when webs were not abundant. Spiders were taken to Cardiff University, transferred to fresh 100% ethanol, adults identified to species-level and juveniles to genus, and stored at −20 °C until subsequent DNA extraction.

Invertebrate prey communities were collected using a converted McCulloch GBV 325 G-vac leaf blower suction sampler for 1 min over 4 m^2 areas near to those from which spiders for DNA analysis were collected. Samples were taken in transects, with 10 samples each pre- and post-harvest (20 total), split evenly between two adjacent fields. Invertebrate prey community samples were taken approximately 10 m apart, in sites near to those from which spiders were collected, with different sites used pre- and post-harvest. Invertebrates were killed with ethyl acetate and stored in 70% ethanol at −20 °C, as these samples were for measurement of the invertebrate community and not for molecular analysis. All invertebrates were identified to family level under an Olympus SZX7 stereomicroscope using morphological keys, except for springtails of Sminthuroidea (Sminthuridae and Bourletiellidae, which were often indistinguishable following vacuum sampling and preservation due to the fine features necessary to distinguish them) which were left at super-family and mites (many of which were immature or in poor condition), which were identified to order level.

Dietary analysis

Erigone atra and *E. dentipalpis* (Erigoninae), and *Tenuiphantes tenuis* (Blackwall, 1852; Linyphiinae) were the focus of our study, although a few juveniles were included from other genera due to the difficulties associated with morphological identification of linyphiid juveniles; these misidentifications were confirmed in the subsequent metabarcoding. In total, 66 spiders were screened (Table S2), unevenly split across the 20 corresponding prey sampling sites. Spiders were washed in and transferred to fresh 100% ethanol to reduce external contaminants prior to identification using Roberts (1993) morphological key. Abdomens were removed from spiders and washed again in fresh 100% ethanol. Only abdomens were used for molecular analysis of their gut contents given their higher concentration of prey DNA than that of the cephalothorax (Krehenwinkel *et al.*, 2016; Macıas-Hernández *et al.*, 2018). To ascertain optimal extraction technique, samples were split into two groups. From one group, DNA was extracted from the abdomens via Qiagen TissueLyser II as per the DNeasy Blood & Tissue Kit (Qiagen, Hilden, Germany) manufacturer's protocol and abdomens kept in the lysis buffer during incubation. From the other group, DNA was extracted by splitting the abdomen with a sterile micropestle, swilling it in the lysis buffer, and then removing the bulk tissue. Neither method ultimately afforded a significantly greater proportion of prey DNA reads post-amplification (Fig. S1), so were combined for analysis. Post-lysis, all extractions followed the DNeasy Blood & Tissue Kit (Qiagen) manufacturer's protocol but with an extended lysis time of 12 h (recommended: 1–3 h) to account for the complex and branched gut system in spider abdomens (Krehenwinkel *et al.*, 2016). Per 12 spiders, each DNA extraction session included at least one negative control consisting of an empty tube treated identically to the samples.

Primers were labelled with unique 10 bp molecular identifier tags (MID-tags) and samples had a unique pairing of forward and reverse tags for identification of each sample post-sequencing. PCR reactions of 25 µl reaction volumes contained 12.5 µl Qiagen PCR Multiplex kit, 0.2 µmol (2.5 µl of 2 µM) of each primer, 2.5 µl DNase-free, water and 5 µl template DNA. Reactions were carried out in the same Veriti Thermal Cycler (ThermoFisher Scientific, Waltham, USA), with annealing temperatures optimised via temperature gradient PCRs in the same machine. PCRs comprised 15 min at 95 °C, followed by 35 cycles of: 95 °C for 30 s, the primer-specific annealing temperature for 90 s, and 72 °C for 90 s; followed by a final extension at 72 °C for 10 min. The new primers, designated BerenF-LuthienR (universal) and TelperionF-LaurelinR (spider-excluding), used annealing temperatures of 52 and 42 °C, respectively.

Within each PCR 96-well plate, 12 negative (extraction and PCR), and two positive controls were included following Taberlet *et al.* (2018). Negative PCR controls consisted of DNase-free water. Positive controls comprised known-concentration mixtures of the invertebrate DNA used for primer testing, detailed above, quantified using Qubit dsDNA High-sensitivity Assay Kits (ThermoFisher Scientific) to ascertain any effects of primer bias. All concentrations were standardised at 0.1 ng μl^{-1} by diluting the DNA in DNase-free water. Five mixtures of different species richness and proportions were prepared (Table S1). A negative control was present for each MID-tag to identify any contamination of primers. Each plate was pooled according to concentrations determined by Qiaxcel Advanced System (Qiagen). Each pool was cleaned via SPRIselect beads (Beckman Coulter, Brea, USA), with a left-side size selection using a 1:1 ratio (retaining ~300–1000 bp fragments). The concentration of the pooled DNA was determined via Qubit dsDNA High-sensitivity Assay Kits, quality-checked via TapeStation 2200 (Agilent, Santa Clara, USA), and all pools sharing the same primer pair were pooled again into a 'super pool', thus forming one pool per primer pair. Library preparation for Illumina sequencing was carried out on these cleaned 'super pools' via NEXTflex Rapid DNA-Seq Kit (Bioo Scientific, Austin, USA) and samples were sequenced on an Illumina MiSeq via a Nano chip with 2 × 250 bp paired-end reads (expected capacity ≤1 000 000 reads).

Bioinformatic analysis

The Illumina run generated 405 270 and 482 249 reads using BerenF-LuthienR (universal) and TelperionF-LaurelinR

(spider-excluding), respectively. All reads were quality-checked and trimmed in Trimmomatic v0.38 (Bolger *et al.*, 2014) with a minimum quality score and sliding window of 20 and 4 bp, respectively, and a minimum length of 135 bp. The read pairs were aligned via FLASH v1.2.11 (Magoč & Salzberg, 2011) and demultiplexed via Mothur v1.39.5 (Schloss *et al.*, 2009), removing the MID and primer sequences. Replicates were removed, and denoising and clustering to zero-radius OTUs (ZOTUs; clustered without % identity to avoid multiple species represented within a single OTU) completed via Unoise3 in Usearch11 (Edgar, 2010). The resultant sequences were assigned a taxonomic identity from GenBank via BLASTn v2.7.1. (Camacho *et al.*, 2009) using a 97% identity threshold (Alberdi *et al.*, 2017). The BLAST output was analysed in MEGAN v6.15.2 (Huson *et al.*, 2016). Where the top BLAST hit, determined by lowest e-value, was resolved at a higher taxonomic level than species-level, the results were checked by blasting the sequence manually in GenBank and comparing the results; where possibly erroneous entries were preventing species-level assignment (e.g. poorly-resolved identifications on GenBank), finer resolution was considered. Where ZOTUs were assigned the same taxon, these were aggregated. Given the prevalence of family-level assignments (e.g. Chloropidae), the data were eventually converted to family-level, but were retained at their respective output assignments for clean-up.

To clean data prior to statistical analysis, all read counts less than the maximum read count present in blanks (negative controls and unused MID-tag combinations) for its respective ZOTU were removed. Instances of non-positive control taxa present in positive controls were calculated as a percentage of the maximum read count for that taxon. The greatest of these percentages was used to guide a universal percentage of the maximum read within each taxon to be removed. This accounts for tag-jumping and "bleeding" of over-represented taxa into other samples during sequencing. For BerenF-LuthienR, 0.54% was optimal, whilst there were no obvious instances for TelperionF-LaurelinR, so the conservative 0.54% was also applied for that library.

Simultaneously, known lab contaminants (e.g. German cockroach *Blatella germanica* (L.) and various species for which molecular analysis was recently undertaken that could be differentiated from the target taxa in this study, such as tropical species) were identified and the percentage of these occurrences of the total read count for their respective samples was calculated. The highest of these percentages was used as a guide for the universal percentage of each total sample read count to be removed. This accounts for environmental and lab contamination, and artefacts and errors of the sequencing process, which for BerenF-LuthienR and TelperionF-LaurelinR were 0.43% and 0.45%, respectively. The data from the two libraries were then aggregated together, first removing non-target taxa (e.g. fungi) and instances in which predator DNA was amplified (i.e. ZOTUs matching the individual spider's morphological identity). All taxa were converted to family-level to standardise the taxonomic level since many ZOTUs could not be resolved further; this also increases evenness for subsequent analyses. Whilst all conspecific reads were removed to account for predator amplification, interspecific linyphiid interactions were still retained, thus any counts of linyphiids in the diet exclusively represent the consumption of other species. Finally, read counts were converted to presence-absence data.

Statistical analysis

All analyses were conducted in R v.3.3.4 (R Core Team, 2020). Invertebrate communities were compared between pre- and post-harvest using multivariate generalised linear models (MGLMs) via 'manyglm' in the 'mvabund' package (Wang *et al.*, 2012) with a binomial error family and Monte Carlo resampling. Spider diets were compared between pre- and post-harvest, and spider life stage (juvenile, sub-adult, and adult) and sex, using MGLMs that included all two-way interactions between these variables. Whilst the samples selected for sequencing primarily comprised *Tenuiphantes tenuis* and *Erigone* spp. with the intention of comparing these taxa, samples sizes were not sufficient for taxonomic comparisons due to sample drop-out in the sequencing process and the misidentification of some juvenile linyphiids (pre-harvest: *Tenuiphantes* = 19, *Erigone* = 5, other = 3; post-harvest: *Tenuiphantes* = 16, *Erigone* = 1, other = 2). Models were simplified using 'step' from the base R 'stats' package to determine an optimal model based on the lowest AIC value by removing variables. Dietary differences were also visualised by non-metric multidimensional scaling via metaMDS in the 'vegan' package (Oksanen *et al.*, 2016) with Jaccard distance in two dimensions (stress = 0.082). Spider plots were created with nMDS results via 'ordispider' and plotted through 'ggplot2' (Wickham, 2016) to illustrate the distinction in communities between categories. Spider prey choice was analysed using network-based null models in the 'econullnetr' package (Vaughan *et al.*, 2018) with the 'generate_null_net' command and visually represented with the 'plot_preferences' command. This detects whether prey taxa are consumed more or less frequently than expected based on their relative abundance in the community.

Results

Novel primer performance

Both primer pairs amplified a broad range of prey *in silico* and *in vitro* (Fig. S4-5, Table S3). BerenF-LuthienR outperformed both the widely-used animal barcoding primers LCO1490-HCO2198 (Folmer *et al.*, 1994) and ZBJ-ArtF1c-ZBJ-ArtR2c (Zeale *et al.*, 2011) *in silico* for most taxa, with the exceptions of Lepidoptera, Coleoptera, and Thysanoptera (Fig. S4). TelperionF-LaurelinR performed comparably to ZBJ-ArtF1c-ZBJ-ArtR2c but with far greater coverage of several taxa. The only taxa for which TelperionF-LaurelinR did not outperform ZBJ-ArtF1c-ZBJ-ArtR2c were Araneae (intentionally so) and Thysanoptera (Fig. S4). The primers were similarly successful *in vitro*. BerenF-LuthienR amplifyied all but *Psilochorus simoni* (Pholcidae spider) and a nudibranch (Table S3). TelperionF-LaurelinR avoided amplification of spiders in the families Agelenidae, Pholcidae, and Clubionidae, but did

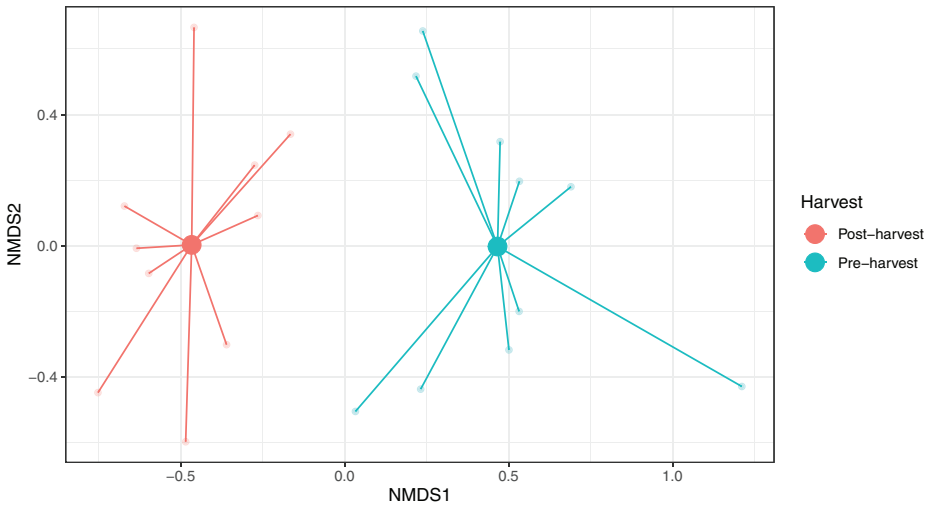

Fig. 1. Spider plot of invertebrate community samples pre- and post-harvest showing distinction between sample taxonomic composition (smaller nodes) and centroids of communities (larger nodes, mean coordinates of the samples for each category) within each category. Species plots align with the distinctions between pre- and post-harvest communities (Fig. S6). [Colour figure can be viewed at wileyonlinelibrary.com].

amplify representatives of Amaurobiidae, Dysderidae, Philodromidae, and Lycosidae, as well as some Linyphiidae, but not the two focal genera of this study. TelperionF-LaurelinR otherwise amplified a very broad range of prey, but exhibited a reduced coverage compared to BerenF-LuthienR, failing to amplify 6/45 of the non-spider taxa tested (Table S3). In the mock community samples, BerenF-LuthienR exhibited some bias towards Lepidoptera and Diptera, whereas TelperionF-LaurelinR exhibited a stronger bias toward Hemiptera, Collembola, Hymenoptera, and Neuroptera (Fig. S5).

Invertebrate community comparison

Identified invertebrates comprised 67 families: 45 pre- and 51 post-harvest (Table S4). Of the 67 families, 29 were recorded in both periods. Distinct invertebrate communities were associated with pre- vs. post-harvest crops (MGLM: LRT = 227.8, P = 0.001; Fig. 1 and S6). Specifically, Ephydridae (shore flies, Diptera; LRT = 34.301, P = 0.001) and Isotomidae (Entomobryomorpha; LRT = 18.761, P = 0.001) were significantly more abundant pre-harvest, whilst Eupelmidae (Hymenoptera; LRT = 11.728, P = 0.014), Microphysidae (Hemiptera; LRT = 12.957, P = 0.006), Parasitiformes (Acari; LRT = 11.879, P = 0.012), and Thripidae (Thysanoptera; LRT = 16.821, P = 0.002) were significantly more abundant post-harvest.

Dietary analysis

The general primers (BerenF-LuthienR) recovered an average of 4204 reads (8.39% reads were prey) and 1.93 prey taxa per spider, and the spider-exclusion primers (TelperionF-LaurelinR) recovered 3530 reads and 0.72 prey taxa per spider from those spiders from which prey were recovered by at least one of the primer pairs. The spider-exclusion primers successfully avoided amplifying any *Erigone* DNA but did amplify the DNA of the other predators (mostly Linyphiinae; in those cases, 2.99% reads were prey).

Of the 66 spiders screened, dietary data were recovered for 46 (Table 2), comprising 15 families (Tables S5 and S6). Several pest taxa were identified from the spider gut DNA, including aphids (*Sitobion avenae* (Fabricius, 1775), true flies (Cecidomyiidae sp., *Oscinella* sp. Becker, 1909), hoppers (*Macrosteles sexnotatus* (Fallén, 1806), *Javasella* sp. Fennah, 1963, *Nothodelphax* sp. Fennah, 1963), and thrips (*Anaphothrips obscurus* (Müller, 1776), *Frankliniella tenuicornis* (Uzel, 1895)). Distinct spider diets were associated with pre- and post-harvest (MGLM: LRT = 27.93, P = 0.027; Figs 2 and S7), but this was affected by the life stage of the spider (MGLM: LRT = 27.43, P = 0.001); however, no specific prey were associated with these differences. Five families were only found in one of the two periods: Aphididae (aphids, Hemiptera), Cecidomyiidae (gall midges, Diptera), and Cicadellidae (leaf hoppers, Hemiptera) were only detected in spider diets pre-harvest, whilst Chironomidae (non-biting midges, Diptera) and Ephydridae (shore flies, Diptera) were only detected post-harvest.

Prey choice analysis

Spiders exhibited prey choice (prey consumed by predators at a higher or lower relative frequency than expected based on availability) in both pre- and post-harvest periods (Figs 3 and S8). Pre-harvest spiders consumed Cicadellidae (leaf hoppers, Hemiptera), Linyphiidae, and Phoridae (humpbacked flies, Diptera) significantly more than expected, and Chloropidae (frit flies, Diptera), Ephydridae (shore flies, Diptera), and Isotomidae (Entomobryomorpha) significantly less than expected. Post-harvest spiders consumed Entomobryidae (Entomobryomorpha), Ephydridae (shore flies, Diptera), and Linyphiidae significantly more than expected, and Thripidae (Thysanoptera)

Table 2. The 46 spiders from which dietary data was recovered.

Pre-harvest	Erigoninae	Adult	Female	3
			Male	2
		Sub-adult	Female	0
			Male	0
		Juvenile		2
	Linyphiinae	Adult	Female	5
			Male	4
		Sub-adult	Female	0
			Male	6
		Juvenile		5
Post-harvest	Erigoninae	Adult	Female	0
			Male	1
		Sub-adult	Female	0
			Male	0
		Juvenile		1
	Linyphiinae	Adult	Female	3
			Male	7
		Sub-adult	Female	0
			Male	7
		Juvenile		0

and Delphacidae (plant hoppers, Hemiptera) significantly less than expected. All other taxa were consistent with the null models and were consumed at the rates expected from their density.

Discussion

Through the application of novel PCR primers, linyphiid diet was shown to differ following harvest, primarily due to changes in prey availability (cf. changing prey choice).

The primers presented in this study are the first designed specifically for gut content analysis of linyphiid spiders and provide an effective means of dietary analysis for a range of species. Both novel primer pairs performed well both *in silico* and *in vitro*, ultimately successfully detecting prey in the guts of field-captured linyphiid spiders with complementary coverage and overlap of detected prey taxa. The broad range of species amplified by the primers suggests applicability to other species and study systems. In the case of TelperionF-LaurelinR, this is particularly true for dietary analysis of spiders in the families Agelenidae, Pholcidae, and Clubionidae at least, as well as some success with ants and other predatory invertebrates given their lack of amplification. Whilst TelperionF-LaurelinR did ultimately amplify the DNA of *Tenuiphantes*, these primers facilitated effective gut content analysis for *Erigone* spiders, likely further Erigoninae species and possibly the other taxa listed above. Whilst no significant difference was found between the two extraction methods trialled, the mean percentage of prey reads per sample was slightly greater when the abdomen was left in the lysis buffer for the longer incubation period, aligning with the findings by Krehenwinkel *et al.* (2016).

This study provides an example of the use of general and predator-exclusion PCR primers together for the same dietary samples. That both primers exhibited taxonomic biases is unsurprising given this widespread phenomenon in metabarcoding (Piñol *et al.*, 2018), but their complementary biases, determined via mock communities, is promising for their combined use, ultimately providing a more complete snapshot of diet. Given the increasing understanding of the effects of primer bias in metabarcoding (Piñol *et al.*, 2018; Braukmann *et al.*, 2019; Elbrecht *et al.*, 2019), further studies should consider employing

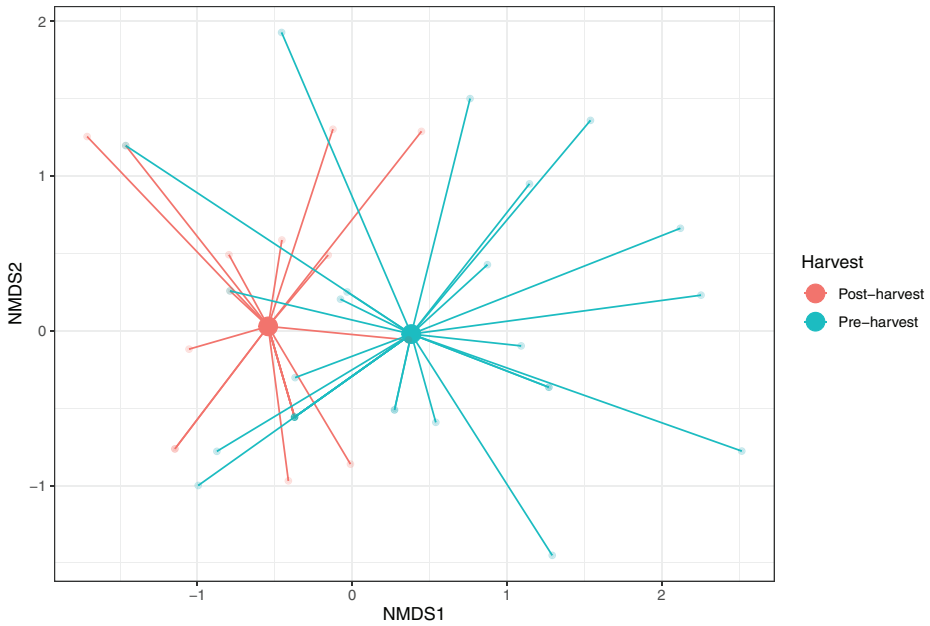

Fig. 2. Spider plot of spider diets pre- and post-harvest showing some distinction with a degree of overlap between the prey families in the diets (smaller nodes) and centroids of diets (larger nodes) within both categories. Species plots align with the distinctions between pre- and post-harvest communities (Fig. S7). [Colour figure can be viewed at wileyonlinelibrary.com].

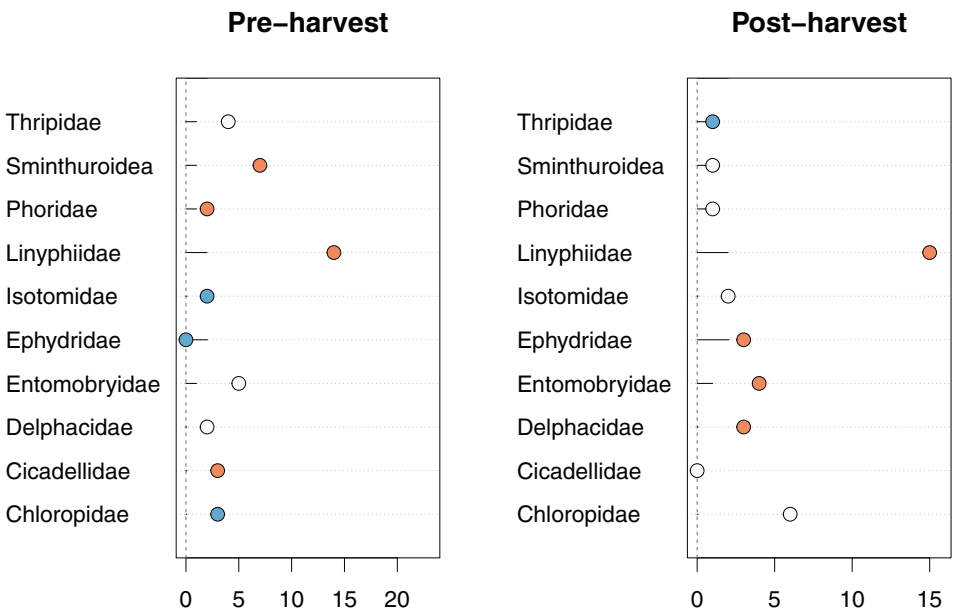

Fig. 3. Significant deviations from expected frequencies of trophic interactions. Blue = lower consumption than expected (avoidance), white = as expected (in proportion to relative abundance), red = higher than expected (consumed more frequently than predicted from relative abundance). Horizontal lines denote 95% confidence limits of the observed frequency of predation. [Colour figure can be viewed at wileyonlinelibrary.com].

this approach to begin mitigating these effects without relying on the heavier biases associated with other methodologies, such as the use of blocking probes (Piñol *et al.*, 2015).

The low annealing temperature used for Telperion-Laurelin (42 °C) was selected due to the exclusion of several prey species at higher temperatures, but could present issues such as dimerisation or non-specific binding, the latter of which can create differently-sized amplicons in some cases, including in another application of these primers (pers. obs. Jordan Cuff). Given the melting temperatures of around 54 °C for these primers, other users could consider higher annealing temperatures when testing against their target organisms to avoid any associated issues, but, as highlighted by the data of this study, the primers will successfully amplify HTS-appropriate DNA from the guts of spiders even at these lower temperatures. Care must be taken, however, to ascertain whether the focal spider species is amplified prior to analysis. Modifications to the primer sequences may increase their specificity for the study of other spider species diets. These spider-exclusion primers, alongside NoSpi2 (Lafage *et al.*, 2019), provide a complementary suite of primers for the metabarcoding of the diets of many spider taxa.

That distinct invertebrate communities were identified pre- and post-harvest indicates a substantial ecological impact of harvest on community composition and dynamics, and thus the provision of alternative prey for generalist predators. The significant decrease in isotomid abundance following harvest is noteworthy given that linyphiids regularly predate springtails such as *Isotoma anglicana* (Schäffer, 1896; Agustí *et al.*, 2003; Harwood *et al.*, 2003; Piñol *et al.*, 2014). The loss of a major prey species could be detrimental to linyphiids immediately prior to winter, especially for those that overwinter as adults. The relative avoidance of isotomids pre-harvest suggests that this may not be the case though, and that reduced abundance of isotomids may still be sufficient for linyphiid populations. Continuous provision of alternative prey not only reduces intraguild predation (Athey *et al.*, 2016), but likely supports over-winter and early-season predation of pests, which is critical in curbing pest populations with short generation times such as aphids (Symondson *et al.*, 2002; Korenko *et al.*, 2010; Pekár *et al.*, 2015). The similarity in linyphiid abundance pre- and post-harvest contrasts findings in a Mediterranean system which evidence declines in agrobiont linyphiid species (Opatovsky & Lubin, 2012); climatic differences may account for this, but the importance of immigrant spiders in rapidly recolonising fields at the start of each cropping cycling and the substantial pest suppression that they provide cannot be overlooked (Opatovsky *et al.*, 2012).

The difference in diet between pre- and post-harvest spiders reflects the community turnover following harvest, with many crop-dependent species such as aphids disappearing from the diet post-harvest. This finding is consistent with Bell *et al.* (2010) in that community turnover, and thus changes in co-occurrences, facilitate dietary changes. Whilst only an overall significant difference was found, not specific taxonomic associations, some taxa were only found in diets in one of the two periods. Regardless of dietary differences, spiders differentially selected from the prey taxa available to them both pre- and post-harvest, inferring differences in prey choice between the two periods. A larger number of significant deviations from expected trophic interaction frequencies were found pre-harvest than post-harvest (7 vs. 4), suggesting that spiders can be more selective regarding their dietary intake, possibly due to greater prey abundance in this period.

Isotomids were predated less frequently than expected pre-harvest when their abundance was greatest, but not post-harvest, following the significant reduction in their abundance. Pre-harvest, spiders may choose alternative prey over isotomids to diversify their dietary intake despite the ease of consuming isotomids, possibly redressing nutritional deficiencies (including accumulation of prey toxins) that would result from consuming only one species (Mayntz *et al.*, 2005). The greater-than-expected frequency of post-harvest consumption of ephydrids and entomobryids suggests that greater provision of these and similar taxa could benefit the health of linyphiid populations post-harvest. Such taxa could be supported by increased habitat complexity (Michalko *et al.*, 2017) and continuous refuges following harvest, particularly grass margins and in-field 'beetle banks' (MacLeod *et al.*, 2004; Mansion-Vaquie *et al.*, 2017). The relative avoidance of thrips post-harvest is likely a consequence of their dominance of the post-harvest invertebrate community, and the need to diversify dietary intake by the spiders. It should be noted that the primers used in this study may exhibit some bias against thrips in PCRs (Fig. S4), thus this result may alternatively be due to disproportional representation of thrips; greater sequencing depth or alternative PCR primers could mitigate this in future studies. Such biases are also observed for most metabarcoding primers, including poor amplification of thrips by the commonly-used primers ZBJ-ArtF1c and ZBJ-ArtR2c (Zeale *et al.*, 2011; Gomez-Polo *et al.*, 2016; Piñol *et al.*, 2018; Silva *et al.*, 2019).

Across both periods, linyphiids were prominent in the diets of the linyphiids screened. Since sequences conspecific with the predator were removed and contamination by tag-jumping accounted for, these were legitimate detections of intra-guild predation between linyphiid species. The inability to detect cannibalism via metabarcoding could mean that these instances of intra-guild predation are in fact an underestimate. That linyphiids are so prominent in their own diet, particularly post-harvest, and were preferentially consumed across both periods, could indicate increased competition for prey resources. Linyphiids will usurp one another for optimal web sites based on prey abundance, sometimes consuming one another in the process (Harwood & Obrycki, 2005), which is likely to increase during periods of prey scarcity. If prey choice is indeed influenced by nutritional requirement, linyphiids may also predate one another as perfect vessels for all of their nutritional needs. This intraguild predation is well-documented in linyphiids (Harwood & Obrycki, 2005; Davey *et al.*, 2013) and is a concern regarding their effectiveness as biological control agents. Intraguild predation may, however, sustain linyphiid populations when suitable prey are absent. Increased habitat complexity could thus reduce intraguild predation through improved prey provisioning (Michalko *et al.*, 2017) and improve the biocontrol capacity of linyphiids. Given that this harvest interface period lies beyond typical crop spraying times, the importance of natural enemies for biocontrol is arguably even greater. Linyphiids are consuming pests in this period (e.g. aphids, as shown in this study), so reductions in intraguild predation by altered management regimes could further enhance this primary control of pests as they begin over wintering, maximally impacting the early-season return of pests.

Acknowledgements

Thanks to Robert Reader for use of fields at Burdons Farm, Wenvoe for sample collection, and to Angela Marchbank and Georgina Smethurst for assistance with high-throughput sequencing. Thanks also to Vasco Elbrecht and Lolita Ammann for advice on PCR primer design. JPC was funded by the Biotechnology and Biological Sciences Research Council through the South West Biosciences Doctoral Training Partnership (grant BB/M009122/1). JRB was funded by BBSRC's Core Capability Grant (BBS/E/C/000 J0200).

Conflict of interest

There are no conflicts of interest to declare and no disputes regarding ownership of the data presented in the manuscript.

Data availability

The data that support the findings of this study are openly available on Dryad at https://doi.org/10.5061/dryad.8cz8w9gng.

Author Contributions

JPC led the study, field work, and preparation of the manuscript. JPC, WOCS, PO-tW, JRB, IPV, and CTM designed and managed the study. JPC, LED, and JES contributed to the design and testing of the primers. JPC and LED processed and analysed the high-throughput sequencing data. JPC and MPTGT identified the invertebrates and processed these data. All authors contributed to the writing and editing of the manuscript.

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Table S1: Mock community DNA mixtures. The volume of each 0.1 ng μl^{-1} sample added to each mock community mixture. Mock communities include some uniformly distributed volumes and others weighted in favour of specific taxa (most often spiders to reflect the abundance of spider DNA in spider diet extracts).

Table S2: The 66 spiders included in dietary screening included 37 pre-harvest, 29 post-harvest, 16 Erigoninae, 50 Linyphiinae, 33 adult, 17 sub-adult, 16 juvenile, 19 female, and 31 male spiders.

Table S3: *In vitro* results for novel primer pairs. A) Arachnids, B) other invertebrates, C) vertebrates and marine. Both primer pairs show a broad amplification of many groups. Beren-Luthien proves to be very general, whilst Telperion-Laurelin avoids amplification of some spiders, but also a few other invertebrate species tested.

Table S4: Invertebrate taxa identified from vacuum samples pre- and post-harvest. Vacuum samples were collected for 30 s over

4 m^2. Each period represents the total from 10 samples evenly split between two adjacent fields.

Table S5: Percentage of spiders in each category (pre- or post-harvest, sex, subfamily, and age class) that had consumed each of the families detected in gut content analysis.

Table S6: Each spider from which a dietary sample was taken is given alongside the raw dietary data and the other information associated with that sample.

Fig. S1 Comparison of % prey reads (of total reads per sample) recovered via the two extractions methods for instances in which the predator was amplified: crush (abdomens disrupted and lysed in solution) and flush (abdomen disrupted and removed before lysis). Left = Beren-Luthien (general primers); right = Telperion-Laurelin (spider-exclusion primers). All samples amplified via Beren-Luthien were compared, but only those in which the linyphiid predator was amplified via Telperion-Laurelin were compared. Neither method recovered a significantly greater % prey reads, although "crushing" recovered a higher mean % prey reads (Beren-Luthien: crush = 10.67%, flush = 6.10%; Telperion-Laurelin: crush = 4.55%, flush = 1.43%).

Fig. S2 PrimerMiner mass alignments of spiders and common cereal crop species for Beren and Luthien. Both primers utilise degenerate bases to increase the amplification of a broad taxonomic range. The conserved primer sites enable broad amplification with few degenerate bases.

Fig. S3 PrimerMiner mass alignments of spiders and common cereal crop species for Telperion and Laurelin. Both primers utilise degenerate bases to increase the amplification of a broad taxonomic range. The terminal base at the 3′ end of Laurelin is a critical mismatch for almost all spiders, making the primer pair efficacious for the amplification of all but spiders (G, rather than A).

Fig. S4 *In silico* analysis of primer bias for novel primer pairs compared against standard animal barcoding primers LCO1490 and HCO2198 (Folmer *et al.*, 1994), and other primers used for linyphiid dietary analysis ZBJ-ArtF1c and ZBJ-ArtR2c (Zeale *et al.*, 2011). The further from the centre that each coloured polygon extends at each anchor point reflects the relative performance of that primer for the respective taxon. Distances are relative, not absolute. Beren-Luthien show almost universally strong amplification potential for most taxa, except for thrips (Thysanoptera). Telperion-Laurelin show the expected bias against spiders (Araneae), but also weaker amplification of mesostigmatic mites and thrips; despite this, the amplification potential for most other groups is stronger than that of ZBJ, which has been used in other studies of linyphiid diet. The reduced amplification potential for thrips by both primer pairs illustrates a potential under-representation of them when paired for dietary analyses.

Fig. S5 Primer bias ascertained using mock community positive control and high-throughput sequencing. The distance of each point radially from the centre is the proportional success in amplifying the mock community component taxa. The "Expected" amplification is based on the proportion of each taxon's DNA in the mock community. Beren-Luthien shows some bias towards Lepidoptera (*Euproctis similis*) and some Diptera (*Melieria crassipennis*), but reduced amplification of linyphiids, springtails, wasps, and aphids. Telperion-Laurelin expectedly avoids amplification of spiders in most cases, although they are still amplified to some extent. The two primer pairs show moderately complementary biases, but both show some bias against linyphiids, and some springtails, flies, and wasps.

Fig. S6 Spider plot of invertebrate communities pre- and post-harvest showing distinction between communities (smaller nodes) and centroids of communities (larger nodes, mean coordinates for each category) within each category. Species present in the communities are plotted according to their alignment with sample dissimilarities.

Fig. S7 Spider plot of spider diets pre- and post-harvest showing some distinction with a degree of overlap between the prey families in the diets (smaller nodes) and centroids of diets (larger nodes) within both categories. Species present in the diets are plotted according to their alignment with sample dissimilarities.

Fig. S8 Full prey choice plot. Blue = lower consumption than expected (avoidance), white = as expected (in proportion to relative abundance), red = higher than expected (consumed more frequently than predicted from relative abundance). Horizontal lines denote 95% confidence limits of the observed frequency of predation.

References

Agustí, N., Shayler, S.P., Harwood, J.D., Vaughan, I.P., Sunderland, K.D. & Symondson, W.O.C. (2003) Collembola as alternative prey sustaining spiders in arable ecosystems: prey detection within predators using molecular markers. *Molecular Ecology*, **12**, 3467–3475.

Alberdi, A., Aizpurua, O., Gilbert, M.T.P. & Bohmann, K. (2017) Scrutinizing key steps for reliable metabarcoding of environmental samples. *Methods in Ecology and Evolution*, **9**, 1–14.

Athey, K.J., Dreyer, J., Kowles, K.A., Penn, H.J., Sitvarin, M.I. & Harwood, J.D. (2016) Spring forward: molecular detection of early season predation in agroecosystems. *Food Webs*, **9**, 25–31. https://doi.org/10.1016/j.fooweb.2016.06.001.

Bale, J.S., van Lenteren, J.C. & Bigler, F. (2008) Biological control and sustainable food production. *Philosophical Transactions of the Royal Society B: Biological Sciences [Online]*, **363**, 761–776. http://rstb.royalsocietypublishing.org/cgi/doi/10.1098/rstb.2007.2182.

Ball, S. (2008) *Introduction to the Families of British Diptera Part 2: Key to Families*. Dipterists Forum, Wardlow, UK.

Barber, A.D. (2008) *Key to the Identification of British Centipedes*. FSC Publications, Shrewsbury, UK.

Beck, J.B. & Toft, S. (2000) Artificial selection for aphid tolerance in the polyphagous predator Lepthyphantes tenuis. *Journal of Applied Ecology*, **37**, 547–556.

Bell, J.R., King, R.A., Bohan, D.A. & Symondson, W.O.C. (2010) Spatial co-occurrence networks predict the feeding histories of polyphagous arthropod predators at field scales. *Ecography*, **33**, 64–72.

Bilde, T. & Soren, T. (2001) The value of three cereal aphid species as food for a generalist predator. *Physiological Entomology*, **26**, 58–68.

Birkhofer, K., Bylund, H., Dalin, P., Ferlian, O., Gagic, V., Hambäck, P.A. *et al.* (2017) Methods to identify the prey of invertebrate predators in terrestrial field studies. *Ecology and Evolution*, **7**, 1–12. http://doi.wiley.com/10.1002/ece3.2791.

Bolger, A.M., Lohse, M. & Usadel, B. (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, **30**, 2114–2120.

Braukmann, T.W.A., Ivanova, N.V., Prosser, S.W.J., Elbrecht, V., Steinke, D., Ratnasingham, S. *et al.* (2019) Metabarcoding a diverse arthropod mock community. *Molecular Ecology Resources*, **19**, 711–727.

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K. *et al.* (2009) BLAST+: architecture and applications. *BMC Bioinformatics*, **10**, 1–9.

Chapman, E.G., Schmidt, J.M., Welch, K.D. & Harwood, J.D. (2013) Molecular evidence for dietary selectivity and pest suppression potential in an epigeal spider community in winter wheat. *Biological Control*, **65**, 72–86. https://doi.org/10.1016/j.biocontrol.2012.08.005.

Dallimore, T. & Shaw, P. (2013) *Illustrated Key to the Families of British Springtails (Collembola)*. FSC Publications, Shrewsbury, UK.

Davey, J.S., Vaughan, I.P., Andrew King, R., Bell, J.R., Bohan, D.A., Bruford, M.W. *et al.* (2013) Intraguild predation in winter wheat: prey choice by a common epigeal carabid consuming spiders. *Journal of Applied Ecology*, **50**, 271–279.

Deagle, B.E., Kirkwood, R. & Jarman, S.N. (2009) Analysis of Australian fur seal diet by pyrosequencing prey DNA in faeces. *Molecular Ecology*, **18**, 2022–2038.

Deagle, B.E., Jarman, S.N., Coissac, E., Taberlet, P. & Deagle, B.E. (2014) DNA metabarcoding and the cytochrome c oxidase subunit I marker: not a perfect match. *Biology Letters*, **10**, 2014–2017.

Duff, A.G. (2012) *Beetles of Britain and Ireland Volume 1: Sphaeriusidae to Silphidae*. A.G. Duff Publishing, West Runton, UK.

Edgar, R.C. (2010) Search and clustering orders of magnitude faster than BLAST. *Bioinformatics*, **26**, 2460–2461.

Elbrecht, V. & Leese, F. (2016) PrimerMiner: an R package for development and in silico validation of DNA metabarcoding primers. *Methods in Ecology and Evolution*, **8**, 622–626.

Elbrecht, V., Braukmann, T.W.A., Ivanova, N.V., Prosser, S.W.J., Hajibabaei, M., Wright, M. *et al.* (2019) Validation of COI metabarcoding primers for terrestrial arthropods. *PeerJ*, **7**, e7745.

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.

Gomez-Polo, P., Alomar, O., Castañé, C., Aznar-Fernández, T., Lundgren, J.G., Piñol, J. *et al.* (2016) Understanding trophic interactions of Orius spp. (Hemiptera: Anthocoridae) in lettuce crops by molecular methods. *Pest Management Science*, **72**, 272–279.

Goulet, H. & Huber, J.T. (1993) *Hymenoptera of the World: An Indentification Guide to Families*. Agriculture Canada, Ottawa, Ontario.

Harwood, J.D. & Obrycki, J.J. (2005) Web-construction behavior of linyphiid spiders (Araneae, Linyphiidae): competition and co-existence within a generalist predator guild. *Journal of Insect Behavior*, **18**, 593–607.

Harwood, J.D., Sunderland, K.D. & Symondson, W.O.C. (2003) Web-location by linyphiid spiders: prey-specific aggregation and foraging strategies. *Journal of Animal Ecology*, **72**, 745–756.

Harwood, J.D., Sunderland, K.D. & Symondson, W.O.C. (2004) Prey selection by linyphiid spiders: molecular tracking of the effects of alternative prey on rates of aphid consumption in the field. *Molecular Ecology*, **13**, 3549–3560.

Hebert, P.D.N., Cywinska, A., Ball, S.L. & deWaard, J.R. (2003) Biological identifications through DNA barcodes. *Proceedings of the Royal Society B: Biological Sciences*, **270**, 313–321. http://rspb.royalsocietypublishing.org/cgi/doi/10.1098/rspb.2002.2218.

Huson, D.H., Beier, S., Flade, I., Górska, A., El-Hadidi, M., Mitra, S. *et al.* (2016) MEGAN Community edition - interactive exploration and analysis of large-scale microbiome sequencing data. *PLoS Computational Biology*, **12**, 1–12.

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S. *et al.* (2012) Geneious basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*, **28**, 1647–1649.

Korenko, S., Pekár, S. & Honěk, A. (2010) Predation activity of two winter-active spiders (Araneae: Anyphaenidae, Philodromidae). *Journal of Thermal Biology*, **35**, 112–116.

Krehenwinkel, H., Kennedy, S., Pekár, S. & Gillespie, R.G. (2016) A cost-efficient and simple protocol to enrich prey DNA from extractions of predatory arthropods for large-scale gut content analysis by Illumina sequencing. *Methods in Ecology and Evolution*, **8**, 126–134.

Lafage, D., Elbrecht, V., Cuff, J.P., Steinke, D., Hambäck, P.A. & Erlandsson, A. (2019) A new primer for metabarcoding of spider gut contents. *Environmental DNA*, **2**, 234–243.

Leray, M., Yang, J.Y., Meyer, C.P., Mills, S.C., Agudelo, N., Ranwez, V. *et al.* (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. *Frontiers in Zoology*, **10**, 1–14.

Macıas-Hernández, N., Athey, K., Tonzo, V., Wangensteen, O.S., Arnedo, M. & Harwood, J.D. (2018) Molecular gut content analysis of different spider body parts. *PLoS One*, **13**, 1–16.

MacLeod, A., Wratten, S.D., Sotherton, N.W. & Thomas, M.B. (2004) 'Beetle banks' as refuges for beneficial arthropods in farmland: long-term changes in predator communities and habitat. *Agricultural and Forest Entomology*, **6**, 147–154.

Magoč, T. & Salzberg, S.L. (2011) FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics*, **27**, 2957–2963.

Mansion-Vaquie, A., Ferrante, M., Cook, S.M., Pell, J.K. & Lovei, G.L. (2017) Manipulating field margins to increase predation intensity in fields of winter wheat (*Triticum aestivum*). *Journal of Applied Entomology*, **141**, 600–611.

Mayntz, D. & Toft, S. (2000) Effect of nutrient balance on tolerance to low quality prey in a wolf spider (Araneae: Lycosidae). *Ekologia Bratislava*, **19**, 153–158.

Mayntz, D., Raubenheimer, D., Salomon, M., Toft, S. & Simpson, S.J. (2005) Nutrient-specific foraging in invertebrate predators. *Science*, **307**, 111–113.

Michalko, R. & Pekár, S. (2016) Different hunting strategies of generalist predators result in functional differences. *Oecologia*, **181**, 1187–1197.

Michalko, R., Petráková, L., Sentenská, L. & Pekár, S. (2017) The effect of increased habitat complexity and density-dependent non-consumptive interference on pest suppression by winter-active spiders. *Agriculture, Ecosystems and Environment*, **242**, 26–33. https://doi.org/10.1016/j.agee.2017.03.025.

Monzó, C., Sabater-Muñoz, B., Urbaneja, A. & Castañera, P. (2010) Tracking medfly predation by the wolf spider, Pardosa cribata Simon, in citrus orchards using PCR-based gut-content analysis. *Bulletin of Entomological Research*, **100**, 145–152. http://www.ncbi.nlm.nih.gov/pubmed/19500431.

Nyffeler, M. & Sunderland, K.D. (2003) Composition, abundance and pest control potential of spider communities in agroecosystems: a comparison of European and US studies. *Agriculture, Ecosystems and Environment*, **95**, 579–612.

Oksanen, J., Blanchet, F.G., Kindt, R., Legendre, P., Minchin, P.R., O'Hara, R.B., Simpson, G.L., *et al.* (2016). vegan: Community Ecology Package. https://cran.r-project.org/package=vegan.

Opatovsky, I. & Lubin, Y. (2012) Coping with abrupt decline in habitat quality: effects of harvest on spider abundance and movement. *Acta Oecologica*, **41**, 14–19.

Opatovsky, I., Chapman, E.G., Weintraub, P.G., Lubin, Y. & Harwood, J.D. (2012) Molecular characterization of the differential role of immigrant and agrobiont generalist predators in pest suppression. *Biological Control [Online]*, **63**, 25–30. https://doi.org/10.1016/j.biocontrol.2012.06.003.

Pekár, S., Michalko, R., Loverre, P., Líznarová, E. & Černecká, Ľ. (2015) Biological control in winter: novel evidence for the importance of generalist predators. *Journal of Applied Ecology*, **52**, 270–279.

Pérez-Guerrero, S., Gelan-begna, A., Tamajón, R., Gelan-begna, A., Tamajo, R. & Vargas-osuna, E. (2013) Potential predation of non-webbuilding spider assemblage on cotton pests Helicoverpa armigera and Spodoptera littoralis (Lepidoptera: Noctuidae). *Biocontrol Science and Technology*, **23**, 335–347.

Peterson, J.A., Ode, P.J., Oliveira-Hofman, C. & Harwood, J.D. (2016) Integration of plant defense traits with biological control of arthropod pests: challenges and opportunities. *Frontiers in Plant Science*, **7**, 1–23.

Petráková, L., Michalko, R., Loverre, P., Sentenská, L., Korenko, S. & Pekár, S. (2016) Intraguild predation among spiders and their effect on the pear psylla during winter. *Agriculture, Ecosystems and Environment*, **233**, 67–74. https://doi.org/10.1016/j.agee.2016.08.008.

Piñol, J., San Andrés, V., Clare, E.L., Mir, G. & Symondson, W.O.C. (2014) A pragmatic approach to the analysis of diets of generalist predators: the use of next-generation sequencing with no blocking probes. *Molecular Ecology Resources*, **14**, 18–26.

Piñol, J., Mir, G., Gomez-Polo, P. & Agusti, N. (2015) Universal and blocking primer mismatches limit the use of high-throughput DNA sequencing for the quantitative metabarcoding of arthropods. *Molecular Ecology Resources*, **15**, 819–830.

Piñol, J., Senar, M.A. & Symondson, W.O.C. (2018) The choice of universal primers and the characteristics of the species mixture determines when DNA metabarcoding can be quantitative. *Molecular Ecology*, **28**, 407–419.

Pompanon, F., Coissac, E. & Taberlet, P. (2011) Metabarcoding, a new way of analysing biodiversity. *Biofutur*, **319**, 30–32.

Pompanon, F., Deagle, B.E., Symondson, W.O.C., Brown, D.S., Jarman, S.N. & Taberlet, P. (2012) Who is eating what: diet assessment using next generation sequencing. *Molecular Ecology*, **21**, 1931–1950.

Quan, X., Wu, L., Zhou, Q., Yun, Y., Peng, Y. & Chen, J. (2011) Identification of predation by spiders on the diamondback moth Plutella xylostella. *Bulletin of Insectology*, **64**, 223–227.

R Core Team. (2020) R: A language and environment for statistical computing. http://www.r-project.org/.

Ratnasingham, S. & Hebert, P.D.N. (2007) Barcoding BOLD: the barcode of life data system (www.barcodinglife.org). *Molecular Ecology Notes*, **7**, 355–364.

Rendon, D., Taylor, P.W., Wilder, S.M. & Whitehouse, M.E.A. (2019) Does prey encounter and nutrient content affect prey selection in wolf spiders. *PLoS One*, **14**, 1–15.

Riechert, S. & Lockley, T. (1984) Spiders as biological control agents. *Annual Review of Entomology*, **29**, 299–320.

Roberts, M.J. (1993) *The Spiders of Great Britain and Ireland (Compact Edition)*, 3rd edn. Harley Books, Colchester, UK.

Schloss, P.D., Westcott, S.L., Ryabin, T., Hall, J.R., Hartmann, M., Hollister, E.B. *et al.* (2009) Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Applied and Environmental Microbiology*, **75**, 7537–7541.

Senior, L.J., Healey, M.A. & Wright, C.L. (2016) The role of spiders as predators of two lepidopteran brassica pests. *Austral Entomology*, **55**, 383–391.

Shayler, S.P. (2005) *Molecular Detection of Predation: The Effects of Detritivore Diversity and Abundance on Pest Control by Generalist Predators*. Cardiff, UK: Cardiff University.

Silva, L.P., Jarman, S.N., Mata, V.A., Lopes, R.J. & Beja, P. (2019) Advancing the integration of multi-marker metabarcoding data in dietary analysis of trophic generalists. *Molecular Ecology Resources*, **19**, 1420–1432.

Sunderland, K. (1999) Mechanisms underlying the effects of spiders on pest populations. *The Journal of Arachnology*, **27**, 308–316.

Sunderland, K. & Samu, F. (2000) Effects of agricultural diversification on the abundance, distribution, and pest control potential of spiders. *Entomologia Experimentalis et Applicata*, **95**, 1–13.

Sunderland, K.D., Fraser, A.M. & Dixon, A.F.G. (1986) Distribution of linyphiid spiders in relation to capture of prey in cereal fields. *Pedobiologia*, **29**, 367–375.

Sunderland, K.D., Axelsen, J.A., Dromph, K., Freier, B., Hemptinne, J.L., Holst, N.H. *et al.* (1997) Pest control by a community of natural enemies Powell, Wilf (ed.). *Acta Jutlandica*, **72**, 271–326.

Symondson, W.O.C. (2002) Molecular identification of prey in predator diets. *Molecular Ecology*, **11**, 627–641.

Symondson, W.O.C., Sunderland, K.D. & Greenstone, M.H. (2002) Can generalist predators be effective biocontrol agents? *Annual Review of Entomology*, **47**, 561–594.

Taberlet, P., Coissac, E., Hajibabaei, M. & Rieseberg, L.H. (2012) Environmental DNA. *Molecular Ecology*, **21**, 1789–1793.

Taberlet, P., Bonin, A., Zinger, L. & Coissac, E. (2018) *Environmental DNA*. Oxford University Press, Oxford.

Turnbull, A.L. (1973) Ecology of the true spiders (Araneomorphae). *Annual Review of Entomology*, **18**, 305–348.

Unwin, D. (2001) *A Key to the Families of British Bugs (Insecta, Hemiptera)*. FSC Publications, Shrewsbury, UK.

Vaughan, I.P., Gotelli, N.J., Memmott, J., Pearson, C.E., Woodward, G. & Symondson, W.O.C. (2018) Econullnetr: an r package using null models to analyse the structure of ecological networks and identify resource selection. *Methods in Ecology and Evolution*, **9**, 728–733.

Vestheim, H. & Jarman, S.N. (2008) Blocking primers to enhance PCR amplification of rare sequences in mixed samples–a case study on prey DNA in Antarctic krill stomachs. *Frontiers in Zoology*, **5**, 12. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2517594&tool=pmcentrez&rendertype=abstract.

Vink, C.J. & Kean, J.M. (2013) PCR gut analysis reveals that *Tenuiphantes tenuis* (Araneae: Linyphiidae) is a potentially significant predator of argentine stem weevil, Listronotus bonariensis (Coleoptera: Curculionidae), in New Zealand pastures. *New Zealand Journal of Zoology*, **40**, 304–313. https://doi.org/10.1080/03014223.2013.794847.

Wang, Y., Naumann, U., Wright, S.T. & Warton, D.I. (2012) Mvabund – an R package for model-based analysis of multivariate abundance data. *Methods in Ecology and Evolution*, **3**, 471–474.

Wang, X., Wang, G., Zhu, Z., Tang, Q., Hu, Y., Qiao, F. *et al.* (2016) Spider (Araneae) predations on white-backed planthopper *Sogatella furcifera* in subtropical rice ecosystems, China. *Pest Management Science*, **73**, 1277–1286.

Wang, B., Li, W. & Yan, H. (2017) Analysis of digestion of rice planthopper by *Pardosa pseudoannulata* based on CO-I gene. *Saudi Journal of Biological Sciences*, **24**, 711–717. https://doi.org/10.1016/j.sjbs.2017.01.046.

Welch, K.D., Crain, P.R. & Harwood, J.D. (2011) Phenological dynamics of web-building spider populations in alfalfa: implications for biological control. *Journal of Arachnology*, **39**, 244–249.

Welch, K.D., Whitney, T.D. & Harwood, J.D. (2016) Non-pest prey do not disrupt aphid predation by a web-building spider. *Bulletin of Entomological Research*, **106**, 91–98. http://www.journals.cambridge.org/abstract_S0007485315000875.

Wickham, H. (2016) *Ggplot2: Elegant Graphics for Data Analysis*. Springer-Verrlag, New York. https://ggplot2.tidyverse.org.

Zeale, M.R.K., Butlin, R.K., Barker, G.L.A., Lees, D.C. & Jones, G. (2011) Taxon-specific PCR for DNA barcoding arthropod prey in bat faeces. *Molecular Ecology Resources*, **11**, 236–244.

Accepted 16 September 2020

Associate Editor: Dirk Sanders