A detection assay to identify alternative food sources of the two-spotted stink bug, *Bathycoelia distincta* (Hemiptera: Pentatomidae)

Arista Fourie[a], Stephanus N. Venter[a], Bernard Slippers[a], Gerda Fourie*[a]

[a]Department of Biochemistry, Genetics and Microbiology, Forestry and Agricultural Biotechnology Institute (FABI), University of Pretoria, Pretoria, 0002, South Africa

*Correspondence to:

Dr. Gerda Fourie
Department of Biochemistry, Genetics and Microbiology
Forestry and Agricultural Biotechnology Institute
University of Pretoria
Lunnon street
Pretoria, South Africa
0002
Phone: +2712 420 3854
E-mail: gerda1.fourie@fabi.up.ac.za

Abstract

The two-spotted stink bug, *Bathycoelia distincta* Distant, is a serious pest in South African macadamia orchards. This pest is predominantly controlled using insecticides, thus alternative control methods are essential. The stink bugs arrive as adults in the orchards, during the early nut set season, but little is known about their alternative plant hosts prior to their arrival. The aim of this study was to develop a PCR-based metabarcoding assay to identify plant material in the gut of *B. distincta.* Thereafter, the persistence of plant DNA in the gut, after switching food sources, was determined by rearing the stink bugs on *Zea mays*, transferring them to *Macadamia* sp. and then collecting insects at different time points. As a proof of concept, the assay was tested on insects collected from commercial macadamia orchards to determine if it can identify alternative food sources. The chloroplast gene markers, *trnL* and *trnF,* were most successful for plant DNA amplification. The time trial suggested that plant material can be detected 24 hours after switching to the alternate food source and one of the samples still contained *Z. mays* DNA after five days. Various plant species were detected from the orchard collected samples, including known food sources of other stink bugs, such as tea plants (*Camellia sinensis*) and sunflowers (*Helianthus annuus*). This study provides the first indication of potential alternative food sources of *B. distincta*. The assay developed in this study can now be implemented for large-scale field surveys to contribute to future integrated pest management strategies.

Keywords: metabarcoding, stink bug, food assay, Hemiptera, Pentatomidae

Introduction

Macadamia nuts are a highly profitable crop (van Wyk 2018) and the industry is well established in sub-tropical countries such as Australia (its native range), South Africa, Kenya, USA (Hawaii), China, and at least 13 other countries (Topp et al. 2019). South Africa is currently the largest producer of macadamia nuts, predominantly grown in three provinces, of which Limpopo has the oldest industry, Mpumalanga is the largest producer and KwaZulu-Natal has a newer but rapidly growing industry (South African Macadamia Association, SAMAC, Loss Factor Benchmark 2021). Several pests contribute to annual macadamia nut losses but the two spotted stink bug, *Bathycoelia distincta* Distant (Hemiptera: Pentatomidae), is the most serious and contributes to 15.23 million USD of losses annually (Taylor et al. 2018).

Stink bugs have very adaptive feeding strategies and can feed on either the xylem/phloem sap or mesophyll/parenchyma cells of seeds, fruit and new plant shoots, using their stylets for sheath feeding or cell rupture feeding (Serteyn et al. 2020). This can induce serious damage in agricultural crops (McPherson et al. 2018) or even introduce pathogens (Medrano et al. 2007). During a recent survey in South Africa, 20 stink bug species were detected in macadamia orchards, including *Parachinavia prunasis* Dallas, *Nezara viridula* L., *Chinavia pallidoconspersa* Stål, *Boerias* species and *Piezodorus* species, but *B. distincta* was the dominant species (Sonnekus 2021). *C. pallidoconspersa* and *B. distincta* are also considered as the most damaging species of mature macadamia nuts (Schoeman 2019). Although *B. distincta* is considered a native species, little is currently known of where this species overwinters or what alternative food sources it feeds on. Scouting surveys have confirmed that these insects arrive in the macadamia orchards as adults, early in the nut season, and must therefore be residing on other plants when not in the orchards. Later in the season egg patches and nymphal stages are detected during scouting, when the stink bugs establish in the orchards, although more adults may also enter from surrounding areas throughout the nut season.

Stink bug species in South Africa are currently controlled using insecticides, especially pyrethroids (Schoeman 2014). The spraying of chemicals is, however, becoming more problematic; as the trees become older they are too tall to effectively reach the insects at the top of the canopy. In addition, sub-lethal exposure to the chemicals leave the pending threat of the development of insecticide resistance (Schoeman 2019). An integrated pest management system is, therefore, imperative to prevent insecticide resistance and reduce the detrimental impact on the environment.

Biological control strategies explored for stink bugs include the use of insectivorous bats (Weier et al. 2019), entomopathogenic fungi (Erper et al. 2016, Dalla Nora et al. 2021) and egg parasitoids (Buffington et al. 2018, Stahl et al. 2019). An alternative approach might be to identify the food sources of stink bugs when they are not in the macadamia orchards, since these plants could serve as breeding hosts where the insects accumulate before and during the nut set season. Knowledge of other host plants of *B. distincta* could identify other crops to target for early pest control. Alternatively, the plant species could be planted as a trap crop to reduce the insect numbers entering the orchard or for an "attract and kill" approach, where more localised control is implemented (Shelton and Badenes-Perez 2006, Gordon et al. 2017).

Assays to detect plant material in the insect gut have been developed for various insects (Navarro et al. 2010, García-Robledo et al. 2013, Avanesyan et al. 2021). Phloem-feeding insects, however, present a greater challenge, since plant DNA are less likely to be present or are at very low levels in the sap these insects feed on. Nonetheless, effective assays have been developed for psyllid species (Cooper et al. 2016, Cooper et al. 2019) and leafhoppers (Avanesyan et al. 2020). Stink bugs likely feed on both xylem/phloem sap and mesophyll cell contents (Serteyn et al. 2020) so plant DNA detection might be less of a restriction for this insect family. The first assay for a stink bug species was recently developed for *Halyomorpha halys* (Hepler et al. 2020), where a feeding time trial suggested the initial plant food source can be detected between 3-14 days after the food were switched to another plant species. The

longevity of plant material could vary largely between the aforementioned insects, due to the very different gut structures, and thus food digestion, between heteroptera (a straight gut) and homoptera insects (a midgut loop and filter chamber) (Terra and Ferreira 2003).

The aim of this study was to develop an effective PCR-based plant detection assay for the two-spotted stink bug, *B. distincta*, and to use this assay to determine the lifetime of plant material in the gut once feeding was switched to another food source. For this purpose, several plant barcode markers were evaluated and the two best-performing markers were optimised for metabarcoding with Illumina MiSeq technology. As a proof of concept, this assay was further tested in a small number of macadamia orchard-collected stink bugs, obtained during different times of the nut set season, to evaluate its efficiency in detecting alternative plant material in their gut.

Materials and methods

Time trial experiment

A lab-reared population of *B. distincta* was maintained at the Forestry and Agricultural Biotechnology Institute Biological Control Centre of the University of Pretoria. The insects were kept in modified plastic containers, with sieve-covered openings in the lids for aeration, at controlled temperature and light conditions of 25-27 °C and a 16:8 h day/night cycle. The insects were fed with *Zea mays* (maize) kernels on the cob.

A feeding time trial was performed, to determine how long one plant species remain detectable in the gut once meals were switched to a different plant, by using adult insects reared on *Z. mays* and transferring them to nuts and branches of a *Macadamia integrifolia* x *M. tetraphylla* hybrid cultivar (Beaumont). Five adult insects were used per time point. Insects were collected at 6 hours, 1 day, 3

days, 5 days and 7 days after transfer to macadamia and immediately frozen at -20˚C for later dissection and DNA extraction. The adults included a mixture of both males and females.

Macadamia orchard collected samples

A total of 32 insects were collected in 2020 and 2021 from two macadamia farms in the Limpopo province (Limpopo 1 and 2) and two farms in Mpumalanga (Mpumalanga 1 and 2) (Table 1). Collections were made from Limpopo 1 and Mpumalanga 1 farms during early nut set stages, when nuts were between match head and pea sized (October - November), and from Limpopo 2 and Mpumalanga 2 farms at the end of the nut set season when nuts were fully developed (February - March). Insects were collected as part of routine scouting surveys on the farms, where different blocks were selected each week and knock-down sprays were applied to 10 random trees to represent trees from the edge as well as deeper inside the orchard. The insects were frozen at -20˚C and stored for later processing.

Table 1. Sample collection and PCR amplification of *B. distincta* samples, obtained from macadamia orchards.

Province and Farm	**Sample name**	**No. of insects collected[a]**	**Samples with successful *trnL* amplification**	**Samples with successful *trnF* amplification**	**Date collected**
Limpopo 1	LP1.Oct20 (early season)	4	EF2, EF3, EF4	EF2, EF3, EF4	October 2020
Limpopo 2	LP2.Feb20 (late season)	7	Ar4, Ar5, Ar8, Ar10	Ar2, Ar4, Ar5, Ar8, Ar10	February 2020
Mpumalanga 1	MP1.Nov20 (early season)	4	DB1, DB4, DB5	DB1, DB4, DB5	November 2020
Mpumalanga 1	MP1.Feb21 (late season)	4	DB14, DB17	DB10, DB11, DB14, DB17	February 2021
Mpumalanga 1	MP1.Mar21 (late season)	4	DB18, DB21, DB23	DB18, DB21, DB23	March 2021
Mpumalanga 2	MP2.Feb21 (late season)	9	DK2, DK3, DK9	DK3	February 2021

[a]DNA extraction and PCR amplifications were not successful for all samples collected. The third and fourth column indicates the final samples used for Illumina MiSeq sequencing.

The farms in the Mpumalanga province are located in a bushveld/forest biome where various native forest trees occur as well as planted forests and other nut, fruit and vegetable crops such as citrus, avocado, mango and banana. The farms in the Limpopo province are located in a bushveld/savanna grassland biome but are also surrounded by forest plantations and fruit orchards.

Insect dissections and DNA extraction

The insects collected from the time trial and from the macadamia farms were surface sterilised by washing for 2 minutes in 100% ethanol, twice in 70% ethanol and then rinsed in Sabax water. The insects were then dissected in phosphate buffered saline solution under a stereomicroscope, using insect dissection tools, to remove the entire gut from the insect. The last section of the midgut (M4 section) contains a high abundance of bacterial symbiont colonies and, based on a pilot trial, if this section was removed the downstream PCR amplification was more effective. Hence, the M4 section was removed from the gut and the rest of the gut was used for DNA extraction. The gut was rinsed in 70% ethanol and water and then stored at -20 ˚C for later processing.

DNA was extracted from the gut samples using a commercial kit. The first samples were processed with the Macherey-Nagel Nucleospin Tissue XS kit (Düren, Germany), but this proved less effective for downstream plant DNA amplification. The rest of the samples were processed using the QIAamp PowerFecal DNA Kit (Qiagen Sciences, Germantown, MD, USA). DNA concentrations and quality were determined using a Thermo Scientific Nanodrop ND_100 spectrophotometer (Wilmington, DE, USA) and samples with high concentrations were diluted to a standard of 50 ng/µl.

PCR optimisation and sequencing of plant barcode markers

One nuclear plant DNA barcode marker, internal transcribed spacer region 2 (ITS2), and four chloroplast markers, *rbcl*, *trnH-psbA*, *trnL* and the intergenic spacer between *trnL* and *trnF* (referred

to here as *trnF*) were tested for PCR amplification success. The universal plant-specific primers used for ITS2 were ITS-p3 and ITS-u4 (Cheng et al. 2016), the primers for *rbcL* were rbcL-a_f and rbcL-a_r (Kress and Erickson 2007), as well as an alternative reverse primer rbcL_Rev (Kress et al. 2009), those for *trnH-psbA* were trnH-psbA_f and trnH-psbA_r (Kress and Erickson 2007), for *trnL* were trnL_c (B49317) and trnL_d (A49855) and for trnF were trnL_e (B49873) and trnF_f (A50272) (Taberlet et al. 1991), respectively.

Initial PCR optimisation of the five different barcode markers were tested with *Z. mays* DNA and the DNA from gut samples of two insects fed on *Z. mays*. Despite various optimisations, PCR amplification for the *rbcL* and *trnH-psbA* gene regions either resulted in no amplification or multiple fragments. A single PCR amplicon fragment could be obtained for the ITS2 region in an insect sample, but a pre-screening with Illumina MiSeq (data not shown) indicated only bacterial DNA contamination and no *Z. mays* DNA. Optimisations included different Taq DNA polymerases (MyTaq (Bioline Ltd., UK) and FastStart Taq (Roche Applied Science, South Africa), annealing temperatures ranging from 48˚C to 53˚C, addition of bovine serum albumin and alternative amplification programs such as touchdown PCR (-1˚C annealing temperature each cycle) and expand PCR (5s increase of elongation step each cycle).

The two barcode markers with the most successful amplification were *trnL* and *trnF*. The primers for these barcode markers were modified with Illumina Nextera adapter sequences, required for Illumina MiSeq library preparation. PCR amplification reactions were performed in a total volume of 25 µl and consisted of 50-100 ng of DNA, 5 µl of 5 x Q5 Reaction buffer, 0.02 U/µl Q5 High-Fidelity DNA Polymerase (New England Biolabs, Ipswich, MA, USA), 200 µM of all dNTPs combined, 1 µl of the forward and reverse *trnF* primers each (10 µM stock concentration, 0.4 µM final) or 0.7 µl of each *trnL* primer (10 µM stock concentration, 0.28 µM final). The PCR cycler program was set at 98˚C for 30 sec, 35 cycles of 98˚C for 10 sec, 53-55˚C (ranged for optimisation) for 30 sec, 72˚C for 30 sec, and a final

extension step at 72˚C for 2 min. PCR amplicons were visualised by combining the DNA with 1.5 µl Novel Juice (Sigma-Aldrich, St. Louis, Missouri, United States), performing agarose gel electrophoresis (AGE) on a 2% gel and visualising it under UV illumination.

PCR amplifications of the orchard collected samples were performed separately for each sample, but all the successful amplicons were combined for each orchard and collection timepoint (two to five amplicons combined) for Illumina MiSeq sequencing (Table 1). This resulted in six bulked samples for the orchard collections. The PCR amplicons for all the time points of the feeding time trial were sequenced separately for each sample. The PCR amplicons of the *trnL* and *trnF* gene regions were purified with Agencourt AMPure XP beads (Beckman Coulter, CA, USA) and, to reduce sequencing costs, were then combined in equimolar concentrations for each sample for further library preparation and sequencing. Barcodes where then added to each sample through another PCR step and samples were combined for multiplex sequencing with Illumina MiSeq paired-end sequencing. Library preparation and sequencing was performed at the Bioinformatics division, Agricultural Research Council, Pretoria, South Africa.

Illumina MiSeq sequence data analysis

The quality of the Illumina MiSeq reads was determined using FastQC v0.11.5 (Andrews 2010). QC summary data indicated the presence of short fragments (<70 bp) for which the sequencer sequenced the entire short PCR fragment, as well as the Nextera Transposase Adapter at the 3' end of the read. All reads were trimmed for any residual Nextera Transposase Adapter sequences on the 3' end, using cutadapt v3.2 (Martin 2011). The *trnL* and *trnF* sequence reads were then separated for each sample, with cutadapt (Martin 2011), by trimming the primer sequences at the 5' end and splitting the read output based on which primer was present in the read. All reads that were untrimmed and those shorter than 100 bp were discarded.

The processed sequence reads for *trnL* and *trnF*, respectively, were further analysed with DADA2 (Callahan et al. 2016) using R v. 4.0.3 (R Core Team 2020) in RStudio (RStudio Team 2019) to obtain amplicon sequence variants (ASVs). For the *trnF* sequence reads, the reads were filtered and trimmed based on a minimum read length of 100 bp, maximum number of N's = 0, maximum number of errors = 3 (forward) and 5 (reverse) and truncQ = 2. Sequences were then de-replicated, denoised and merged after which chimeras were removed to produce the final ASVs. The *trnL* region can have a large variation in size between plant species and, for some amplicons, the forward and reverse reads could not be assembled, hence, the forward and reverse reads were analysed separately. All analyses in DADA2 were performed the same as for trnF, except reads were not merged. The obtained ASVs were separated and analysed separately for the orchard collected samples and the time trial.

The time trial samples' analyses were only performed with ASVs identified as *Macadamia* sp. or *Z. mays* (based on BLASTn to NBCI nucleotide database; Sayers et al. 2019) to compare the abundance of the plant material, since reads with any other identity are most likely sequence errors, chimeras or contamination. The ASV taxonomy and abundance data were analysed with Phyloseq (McMurdie and Holmes 2013) using R v. 4.0.3 (R Core Team 2020) in RStudio (RStudio Team 2019) for relative abundance analyses. Data were rarefied based on the number of reads present in the smallest sample size, then transformed to relative abundance and plotted using ggplot (Wickham 2016) in RStudio.

The identity of all *trnL* and *trnF* ASVs of the orchard collected samples were determined by performing a BLASTn analysis to the NCBI GenBank nucleotide database (Sayers et al. 2019). If discrepancies were found in the taxonomic classification of the forward and reverse reads of *trnL*, both sequences were aligned to the full length *trnL-trnF* region of the top five most likely hits and analysed in combination. The orchard collected samples were each analysed individually, data was thus not rarefied for comparison between samples, but sequence read abundance was converted to relative abundance

for each sample. All ASVs with a relative abundance above 1 x 10^{-2} were considered for downstream analyses and a relative abundance bar plot was generated in Excel.

Results

Plant barcode markers PCR optimisation and metabarcoding

The ITS2, *rbcL* and *trnH-psbA* markers could not be optimised for PCR amplification of plant material from the insect gut. The *trnF* primers produced a ±450 bp fragment and the *trnL* primers a ±530 bp fragment, based on AGE, but many of the samples resulted in a second small fragment of ±150 bp for the *trnL* region. This small fragment was removed from the sequence data during sequence trimming and filtering. Not all samples for which DNA was extracted could be amplified. The average PCR success rate for the time trial samples were 100% for *trnL* and *trnF*, although three samples produced faint bands or double bands for *trnL*. The orchard collected samples had a lower PCR success rate with *trnL* 56% successful and *trnF* 59%. The samples from the Mpumalanga 2 farm were specifically difficult to amplify, possibly due to the DNA quality, and when these samples are excluded, *trnL* had a 65% success rate and *trnF* 78%.

The number of sequence reads per sample ranged from 110 000 to 690 000 (Table S1). After reads were trimmed and split between the two markers, the *trnF* reads ranged from 39 500 to 350 000 and *trnL* from 5 250 to 360 000 per sample. The low number of reads retained for the *trnL* marker for some samples was due to many reads assigned to the short additional amplicon that was removed during sequence trimming. Due to the difference in quality between forward and reverse reads and the read length retained after quality trimming, only the forward reads were used for downstream analysis of the *trnL* marker. Sequence data were submitted to the NCBI SRA database (BioProject ID PRJNA769672).

Time trial experiment

DADA2 analyses of the time trial samples resulted in 40 ASVs for the *trnF* region (Table S2) and 14 ASVs for the *trnL* region (Table S3). In the *trnF* data, the majority of the reads were assigned to five abundant ASVs, identified as *Macadamia* sp. (1 ASV) and *Z. mays* (4 ASVs) and 30 of the ASVs were low abundance sequences that either had not BLAST hit (likely chimeras) or matched to bacterial sequences. Four samples (7Days.5, 6Hours.1, 1Day.1 and 1Day.5) had apparent cross-contamination, as they also contained ASVs from plant species present in the orchard collected samples, which could have occurred during amplicon processing or sequencing. These contaminants were removed before analyses.

The *trnL* data consisted of four abundant ASVs, identified as *Macadamia* sp. (3 ASVs) and *Z. mays* (1 ASV) and 9 ASVs that either had no BLAST hit or were identified as bacterial or stink bug DNA. Five samples (6Hours.2, 6Hours.3, 1Day.4, 5Days.5, 7Days.4) had a low level of other plant DNA contamination. Interestingly, the two samples that had a high contamination of alternative plant species in the *trnF* region (1Day.1 and 1Day.5) did not contain any contaminating plant species sequences in the *trnL* region. This further confirms that there likely was a cross contamination from orchard collected samples in the *trnF* marker. Only the *Z. mays* and *Macadamia* sp. sequences were considered for analysing the relative abundance of plant material for the time trial.

The summarised data of the relative abundance of *Z. mays* and *Macadamia* sp. in the feeding time trial indicated a high presence of *Z. mays* in all five the samples collected at six hours and in four of the five samples collected one day after transfer (Fig. 1). Very little *Z. mays* material was detected later than one day. One sample, however, still contained low levels of *Z. mays* DNA 5 days after transfer (based on the *trnF* marker) (Fig. 1).

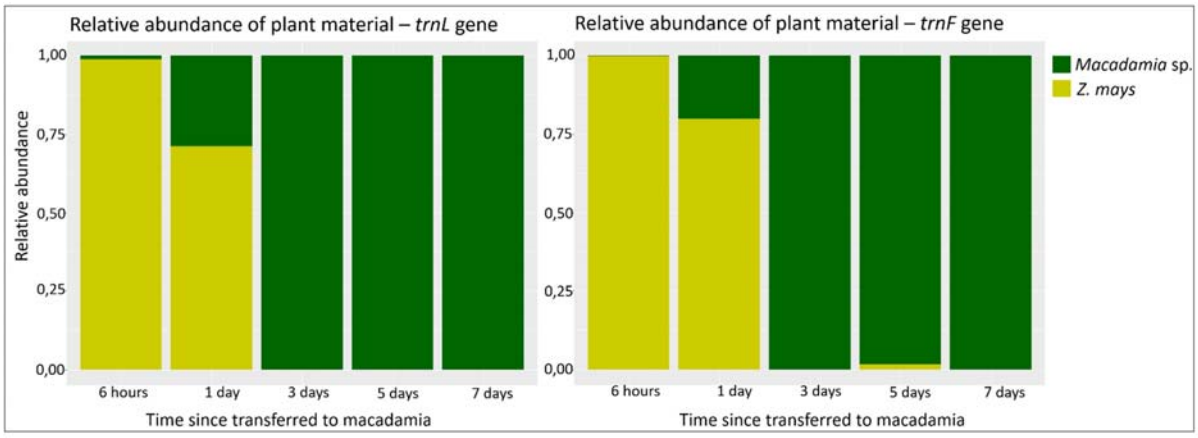

Figure 1. Relative abundance of *Z. mays* and *Macadamia* sp. ASVs in the stink bug guts from the feeding time trial, based on the *trnL* and *trnF* gene marker data. This data represents a summary of five replicates per timepoint.

Macadamia orchard collected samples

In the six orchard collected samples 25 ASVs were identified with the *trnF* region (Table S2) and 16 ASVs with the *trnL* region (Table S3). Ten of the *trnF* ASVs, present at low frequencies, either had no BLAST hit or were bacterial contamination. The most abundant *trnF* ASV was a *Macadamia* sp. sequence, that was present in all samples, and ten ASVs representing other plant species present at a relative abundance above 1.5% in the samples (Fig. 2). The late nut set season collections, samples LP2.Feb20 and MP1.Mar21, only contained *Macadamia* sp. sequences, the insects collected at Mpumalanga 1 farm in Nov ’20 (MP1.Nov20) contained *Camellia sinensis* (tea plant) and those from February ’21 (MP1.Feb21) contained low levels of *Lauraceae* family and high levels of *Prunus dulcis* (almond) sequences. The MP2.Feb21 sample contained *Daucus syrticus*/*D. carota* (carrot) and the LP1.Oct20 sample contained *Lauraceae* family, *Helianthus annuus* (sunflower), *Celtis africana*, *Xylopia* sp., *Bridelia* sp. and very low levels (1.5 – 2%) of *Pinus roxburghii* and *Flindersia* sp.

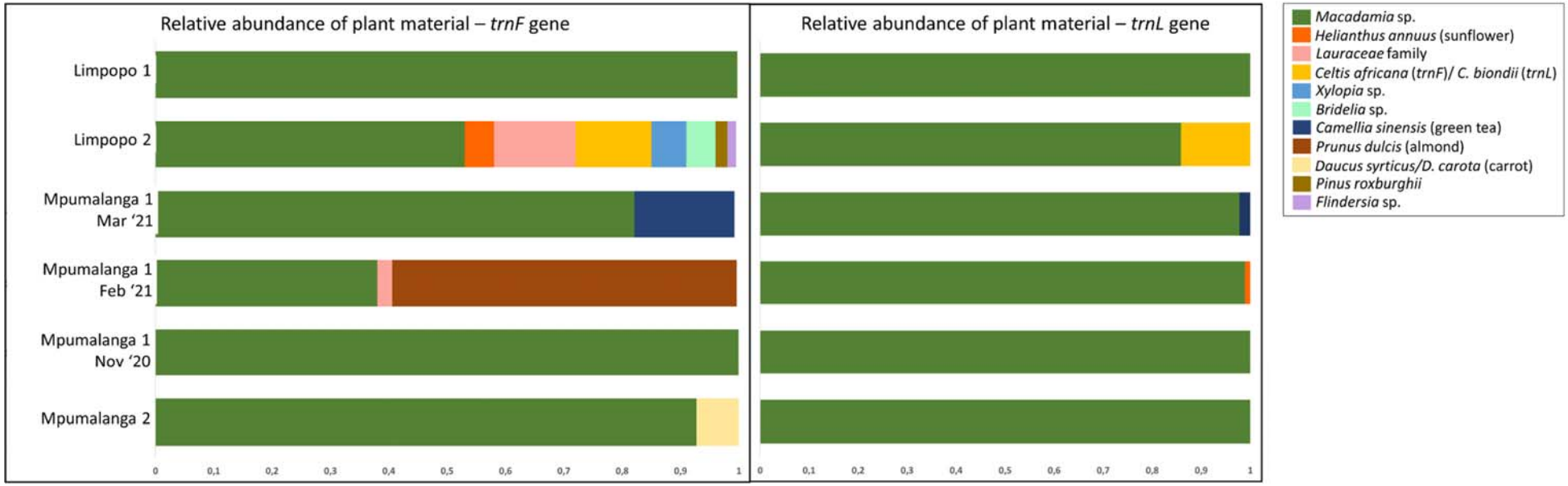

Figure 2. Relative abundance of different plant material found in the gut of stink bugs collected from macadamia farms. Data is based on the *trnF* and *trnL* gene regions' ASV sequence data.

The three most abundant ASVs in the *trnL* region were all assigned to *Macadamia* sp. (Table S3). Nine of the 16 ASVs either had no BLAST hit or were identified as stink bug or bacterial contamination. The *trnL* data confirmed a *Celtis* species at the Limpopo 1 farm but this gene region classified it as *Celtis biondii* instead of *C. africana*, yet *C. africana* is native to South Africa and is most likely the correct species identification. The *trnL* region also confirmed the presence of *Camellia sinensis* in the MP1.Nov20 sample. In addition, *Helianthus annuus* was detected in the MP1.Feb21 sample (Fig. 2).

The combined data obtained from the two gene regions thus indicated the following alternative plant food sources for *B. distincta* in the Limpopo province: *H. annuus*, *C. africana, Lauraceae* family, *Xylopia* sp. and *Bridelia* sp. early in the nut set season. However, only macadamia was detected late in the season. In Mpumalanga, *C. sinensis* and *H. annuus* were prominent food sources early in the nut set season whereas the *Lauraceae* family, *P. dulcis* and *D. syrticus* were detected in the late nut set season. *H. annuus* was the only plant detected in both provinces.

Discussion

In this study we developed and optimised a plant barcoding assay that can detect food content in the gut of the two-spotted stink bug, *B. distincta*. A metabarcoding analysis of the two optimal plant barcode markers (*trnL* and *trnF*) could determine the lifetime of plant material in the gut and was also effective in detecting various plant species from stink bugs collected from macadamia orchards. This study provides the first indication of alternative food sources of *B. distincta,* migrating into macadamia orchards, which can guide integrated pest management strategies for this insect.

Several genes have been used effectively for plant meal detection in insects (Staudacher et al. 2011, García-Robledo et al. 2013, Cooper et al. 2016), but many of these could not be optimised for *B. distincta*. The majority of insect plant meal studies have used *trnL*, suggesting this is an effective barcode marker in insects (Avanesyan et al. 2021). However, from the current study, the *trnF* region

had a higher PCR success rate and is suggested as the primary marker for plant screening in *B. distincta*. The *trnF* marker has also been used for plant detection in psyllids (Cooper et al. 2019) and the *H. halys* stink bug (Hepler et al. 2020) supporting this as a good marker for sap-feeding insects. The combination of *trnL* and *trnF,* however, provided more extensive plant identifications and better resolution in this study and it is suggested to combine these two markers for future plant meal detections.

The time trial data suggested that plant material can still be detected in the gut at least 24 hours after feeding on another plant source. No *Z. mays* DNA could be detected at three days, but it is likely that some would still have been present after two days. This provides a good indication for future surveys; that insects should be collected as soon as possible, between one and two days, after arrival in the macadamia orchards. The presence of *Z. mays* DNA after five days in one of the samples, however, suggests that samples collected later may still provide information of alternative food sources. This timeline differs from what was found for *H. halys*, fed on Lima beans and transferred to carrots, where the bean DNA could still be detected in low amounts between seven and 14 days after transfer (Hepler et al. 2020). This could be due to the different plant species used for feeding, physiological differences between the two stink bug species or the use of different sequencing technologies (PacBio vs MiSeq). Additional studies could also consider tracking the longevity of plant material when no alternative food source is provided to determine the lifetime of the plant material without feeding.

The metabarcoding assay of the orchard collected samples proved the effectiveness of this method to identify alternative plant material from the gut of stink bugs arriving in macadamia orchards. The time of season seems to influence the efficacy of detecting alternative food sources, as the samples collected early in the nut set season contained a higher abundance and diversity of plant species. The insects collected later in the season predominantly contained macadamia DNA and were likely locally breeding populations in the orchard.

The data obtained from this study provided the first indication to alternative food sources of *B. distincta*. One plant species that could be linked to known surrounding vegetation was *Camellia sinensis* (tea plant). It was detected in insects from the Mpumalanga 1 farm, where a green tea farm is located 26 km from this macadamia farm. Although this is not the average flight distance for stinkbugs, long distance flight capabilities of 70-100 km have been recorded for species such as *H. halys* (Lee and Leskey 2015; Wiman et al. 2015).

Some of the plants species detected are known food sources of other stink bug species, such as almond (Rijal et al. 2021) and sunflower (Lee et al. 2013), or carrot, which is used as a diet for lab reared *H. halys* (Medal et al. 2012). Many of the plants detected could also be confirmed from the surrounding vegetation. In Mpumalanga, almond trees occur on farms in the area, although no commercial plantations are established. Sunflower is planted commercially in both provinces, and the one farm in Limpopo also used it as a cover crop to attract pollinators. The Lauraceae family was detected in both provinces (classification not confirmed to genus level) and various plants from this family occur in both regions, such as avocado, Cape Laurel and camphor trees. *Celtis africana* are common trees in the savanna areas of Limpopo and two native *Bridelia* tree species (*B. micrantha* and *B. mollis*) occur in this province. The *Pinus* sequences detected at low levels in Limpopo could also be supported, since various pine plantations occur in the same area.

Sunflower, one of the prominent plants detected in *B. distincta*, has been considered as part of a polyculture trap crop where it was planted around vegetable plots, to reduce the number of *H. halys* insects entering and damaging the crops (Gordon et al. 2017, Mathews et al. 2017). The sunflowers could also serve to attract the insects where they can be killed more locally, reducing the amount of insecticide used. Although the development of a trap crop strategy needs significant optimization, it might also hold potential for *B. distincta*.

The assay developed in this study can be used in further screening surveys, including more insect samples and collections over a longer timeframe, to identify the most prominent food sources of *B. distincta.* Such a host range database could aid in integrated pest management in various ways. For example, the tea farm was identified as a potential food source and this farm could serve as a site for early pest control. Alternatively, the prominent plant species can be tested for efficacy as a trap crop. In addition, as seen for psyllid insects, this information can provide a better understanding of the landscape-level movements of this insect (Cooper et al. 2019) and of its biology and ecology.

Acknowledgements

This study was made possible by the Forestry and Agricultural Biotechnology Institute at the University of Pretoria, who provided the facilities and resources for this research, as well as the South African Macadamia Association (SAMAC) who provided a postdoctoral research grant and additional research funds for this project. The stink bugs collected from macadamia orchards were all provided by local macadamia growers and we thank each of the growers that contributed to the sample collections.

References

Andrews, S. 2010. FastQC: A quality control tool for high throughput sequence data. Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/.

Avanesyan, A., N. Illahi, and W. O. Lamp. 2020. Detecting ingested host plant DNA in potato leafhopper (Hemiptera: Cicadellidae): Potential use of molecular markers for gut content analysis. J. Econ. Entomol. 114: 472-475.

Avanesyan, A., H. Sutton, and W. O. Lamp. 2021. Choosing an effective PCR-based approach for diet analysis of insect herbivores: A systematic review. J. Econ. Entomol. 114: 1035-1046.

Buffington, M. L., E. J. Talamas, and K. A. Hoelmer. 2018. Team Trissolcus: Integrating taxonomy and biological control to combat the brown marmorated stink bug. Am. Entomol. 64: 224-232.

Callahan, B. J., P. J. McMurdie, M. J. Rosen, A. W. Han, A. J. A. Johnson, and S. P. Holmes. 2016. DADA2: High-resolution sample inference from Illumina amplicon data. Nat. Methods 13: 581-583.

Cheng, T., C. Xu, L. Lei, C. Li, Y. Zhang, and S. Zhou. 2016. Barcoding the kingdom Plantae: new PCR primers for ITS regions of plants with improved universality and specificity. Mol. Ecol. Res. 16: 138-149.

Cooper, W. R., D. R. Horton, T. R. Unruh, and S. F. Garczynski. 2016. Gut content analysis of a phloem-feeding insect, *Bactericera cockerelli* (Hemiptera: Triozidae). Environ. Entomol. 45: 938-944.

Cooper, W. R., D. R. Horton, M. R. Wildung, A. S. Jensen, J. Thinakaran, D. Rendon, L. B. Nottingham, E. H. Beers, C. H. Wohleb, D. G. Hall, and L. L. Stelinski. 2019. Host and non-host 'whistle stops' for psyllids: Molecular gut content analysis by high-throughput sequencing reveals landscape-level movements of Psylloidea (Hemiptera). Environ. Entomol. 48: 554-566.

Dalla Nora, D., B. C. Piovesan, C. Bellé, R. S. Stacke, R. R. Balardin, J. V. C. Guedes, J. P. Michaud, and R. J. S. Jacques. 2021. Isolation and evaluation of entomopathogenic fungi against the neotropical brown stink bug *Euschistus heros* (F.) (Hemiptera: Pentatomidae) under laboratory conditions. Biocontrol Sci. Technol. 31: 22-34.

Erper, I., I. Saruhan, I. Akca, H. M. Aksoy, and C. Tuncer. 2016. Evaluation of some entomopathogenic fungi for controlling the green shield bug, *Palomena prasina* L. (Heteroptera: Pentatomidae). Egypt. J. Biol. Pest Control 26: 573-578.

García-Robledo, C., D. L. Erickson, C. L. Staines, T. L. Erwin, and W. J. Kress. 2013. Tropical plant–herbivore networks: Reconstructing species interactions using DNA barcodes. PLOS ONE 8: e52967.

Gordon, T. L., M. Haseeb, L. H. B. Kanga, and J. C. Legaspi. 2017. Potential of three trap crops in managing *Nezara viridula* (Hemiptera: Pentatomidae) on tomatoes in Florida. J. Econ. Entomol. 110: 2478-2482.

Hepler, J., R. Cooper, and E. Beers. 2020. Host plant signal persistence in the gut of the brown marmorated stink bug (Hemiptera: Pentatomidae). Environ. Entomol. 50: 202-207.

Kress, W. J., and D. L. Erickson. 2007. A two-locus global DNA barcode for land plants: The coding *rbcL* gene complements the non-coding trnH-psbA spacer region. PLOS ONE 2: e508.

Kress, W. J., D. L. Erickson, F. A. Jones, N. G. Swenson, R. Perez, O. Sanjur, and E. Bermingham. 2009. Plant DNA barcodes and a community phylogeny of a tropical forest dynamics plot in Panama. Proc. Natl. Acad. Sci. 106: 18621-18626.

Lee, D. H., B. D. Short, S. V. Joseph, J. C. Bergh, and T. C. Leskey. 2013. Review of the biology, ecology, and management of *Halyomorpha halys* (Hemiptera: Pentatomidae) in China, Japan, and the Republic of Korea. Environ. Entomol. 42: 627-641.

Lee, D., and Leskey, T. 2015. Flight behavior of foraging and overwintering brown marmorated stink bug, *Halyomorpha halys* (Hemiptera: Pentatomidae). Bull. Entomol. Res., 105: 566-573.

Martin, M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal 17: 10-12.

Mathews, C. R., B. Blaauw, G. Dively, J. Kotcon, J. Moore, E. Ogburn, D. G. Pfeiffer, T. Trope, J. F. Walgenbach, C. Welty, G. Zinati, and A. L. Nielsen. 2017. Evaluating a polyculture trap crop

for organic management of *Halyomorpha halys* and native stink bugs in peppers. J. Pest Sci. 90: 1245-1255.

McMurdie, P. J., and S. Holmes. 2013. phyloseq: An R package for reproducible interactive analysis and graphics of microbiome census data. PLOS ONE 8: e61217.

McPherson, J. E., C. S. Bundy, and A. G. Wheeler. 2018. Overview of the superfamily Pentatomoidea, pp. 3-17. In J. E. McPherson (ed.), Invasive stink bugs and related species (Pentatomoidea): biology, higher systematics, semiochemistry, and management. CRC Press, Boca Raton, FL.

Medal, J., T. Smith, A. Fox, A. S. Cruz, A. Poplin, and A. Hodges. 2012. Rearing the brown marmorated stink bug *Halyomorpha halys* (Heteroptera: Pentatomidae). Fla. Entomol. 95: 800-802, 803.

Medrano, E. G., J. F. Esquivel, and A. A. Bell. 2007. Transmission of cotton seed and boll rotting bacteria by the southern green stink bug (*Nezara viridula* L.). J. Appl. Microbiol. 103: 436-444.

Navarro, S. P., J. A. Jurado-Rivera, J. Gómez-Zurita, C. H. C. Lyal, and A. P. Vogler. 2010. DNA profiling of host–herbivore interactions in tropical forests. Ecol. Entomol. 35: 18-32.

R Core Team, 2020. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available online at: https://www.R-project.org/.

Rijal, J. P., A. L. Joyce, and S. Gyawaly. 2021. Biology, ecology, and management of hemipteran pests in almond orchards in the United States. J. Integr. Pest Manag. 12.

RStudio Team, 2019. RStudio: Integrated Development for R. RStudio, Inc., Boston, MA. Available online at: http://www.rstudio.com/.

SAMAC (South African Macadamia Association). 2021. Industry statistics on the South African Macadamia Industry 2021. Available online at: https://www.samac.org.za/industry-statistics/

Sayers, E. W., M. Cavanaugh, K. Clark, J. Ostell, K. D. Pruitt, and I. Karsch-Mizrachi. 2019. GenBank. Nucleic Acids Res. 48: D84-D86.

Schoeman, P. 2014. Stink bug IPM on macadamias in South Africa: current status and the road ahead. Trends Entomol. 10: 87-95.

Schoeman, P. S. 2019. Stink bug resistance – Causes and action plans, SAMAC pest management technical information sheets. https://www.samac.org.za/.

Schoeman, P. S. 2019. Damage potential of indigenous Heteroptera species occurring on Macadamia nuts (Macadamia integrifolia Maiden & Betche & Macadamia tetraphylla L. Johnson) in South Africa during the early and late season. Int. J. Trop. Insect Sci. 40: 217-219.

Serteyn, L., L. Ponnet, E.A. Backus, and F. Francis. 2020. Characterization of electropenetrography waveforms for the invasive heteropteran pest, *Halyomorpha halys*, on *Vicia faba* leaves. Arthropod Plant Interact. 14: 113–126.

Shelton, A. M., and F. R. Badenes-Perez. 2006. Concepts and applications of trap cropping in pest management. Annu. Rev. Entomol. 51: 285-308.

Sonnekus, B. 2021. Diversity of stink bugs (Pentatomidae) and associated egg parasitoids in macadamia orchards in South Africa. MSc. dissertation, University of Pretoria, South Africa.

Stahl, J. M., D. Babendreier, C. Marazzi, S. Caruso, E. Costi, L. Maistrello, and T. Haye. 2019. Can *Anastatus bifasciatus* be used for augmentative biological control of the brown marmorated stink bug in fruit orchards? Insects 10: 108.

Staudacher, K., C. Wallinger, N. Schallhart, and M. Traugott. 2011. Detecting ingested plant DNA in soil-living insect larvae. Soil Biol. Biochem. 43: 346-350.

Taberlet, P., L. Gielly, G. Pautou, and J. Bouvet. 1991. Universal primers for amplification of three non-coding regions of chloroplast DNA. Plant Mol. Biol. 17: 1105-1109.

Taylor, J. P., I. Grass, A. J. Alberts, E. Joubert, and T. Tscharntke. 2018. Economic value of bat predation services – A review and new estimates from macadamia orchards. Ecosyst. Serv. 30: 372-381.

Terra, W.R., and C. Ferreira. 2003. Digestive System, pp. 273-280. In Resh, V.H., and C.T. Ring (eds.), Encyclopedia of insects. Academic Press, Amsterdam.

Topp, B. L., C. J. Nock, C. M. Hardner, M. Alam, and K. M. O'Connor. 2019. Macadamia (*Macadamia* spp.) breeding, pp. 221-251. In J. M. Al-Khayri, S. M. Jain and D. V. Johnson (eds.), Advances in

plant breeding strategies: Nut and beverage crops, vol. 4. Springer International Publishing, Cham.

van Wyk, A. 2018. Earning you almost R375,000 per hectare, this must be SA's most profitable crop, Business Insider SA. Business Insider Inc. Accessed online: https://www.businessinsider.co.za/sas-most-profitable-crop-2018-3, Business Insider South Africa.

Weier, S. M., Y. Moodley, M. F. Fraser, V. M. G. Linden, I. Grass, T. Tscharntke, and P. J. Taylor. 2019. Insect pest consumption by bats in macadamia orchards established by molecular diet analyses. Glob. Ecol. Conserv. 18: e00626.

Wickham, H. 2016. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York. ISBN 978-3-319-24277-4, https://ggplot2.tidyverse.org.

Wiman, N.G., V.M. Walton, P.W. Shearer, S.I. Rondon, and J.C. Lee. 2015. Factors affecting flight capacity of brown marmorated stink bug, *Halyomorpha halys* (Hemiptera: Pentatomidae). J. Pest Sci. 88: 37–47.