**AGRICULTURAL RESEARCH FOUNDATION**
**FINAL REPORT**
**FUNDING CYCLE 2016 – 2018**

**TITLE:** Molecular diet analysis of insects that vector vegetable pathogens

**RESEARCH LEADER:** Kenneth Frost, Assistant Professor and Plant Pathologist, Department of Botany and Plant Pathology, Hermiston Agricultural Research and Extension Center, Hermiston, OR

**COOPERATORS:** Sandra DeBano, Associate Professor, Department of Fisheries and Wildlife, Hermiston Agricultural Research and Extension Center (HAREC), Hermiston, OR

David Wooster, Sandra DeBano Assistant Professor, Department of Fisheries and Wildlife, Hermiston Agricultural Research and Extension Center (HAREC), Hermiston, OR

**EXECUTIVE SUMMARY:** Each year, vegetable crops are threatened by the occurrence of multiple plant pathogens that are transmitted by insects. The feeding host species used by the insect-vectors prior to arrival in the crop field is rarely known. Information about the identity of specific plant taxa can be can be obtained by amplifying, sequencing, and comparing small fragments of genomic or chloroplast DNA to sequence information from known plant species present in a reference sequence library. The primary objectives of this research was to develop methods to conduct molecular diet analyses on polyphagous plant-feeding insect species with a focus on insect species known to vector vegetable pathogens. The short-term goal of this project was to determine if plant DNA barcoding regions could be detected and amplified from gut extracts of different insect species with differing feeding behaviors. A secondary goal is to determine the length of time that different plant DNA barcoding regions can be detected in insect bodies after transfer to a different feeding host or insect diet. Finally, we would like to use sequence information generated from barcoding to identify plant species that may be present in samples with mixtures of DNA. The long-term goal of this project will be to learn which non-crop plants are being used by insect pests so that eventually growers will be able to better manage non-crop areas to reduce pest pressure in their crops.

**OBJECTIVES:** Our objectives were to 1) develop and refine methods to detect and amplify the chloroplast trnL (UAA) intron present in the bodies of three insect species with different feeding behavior, 2) use no-choice feeding studies to determine the length of time plant DNA remains detectible in the insect body, and 3) identify plant taxa mixtures of plant species by PCR amplification of the chloroplast trnL intron, sequencing the resulting amplicon library, and comparison of sample sequences obtained to a reference database.

**PROCEDURES:**

**2016**

As proof of concept and to test primer sets designed to target chloroplast DNA, a population of thrips were collected from dandelion, *Taraxacum officinale*, flowers around the HAREC in non-crop areas. From the collection of thrips, DNA was extracted using a modified cetyltrimethylammonium bromide (CTAB) methodology. Quantity and quality of the extracted DNA was assessed by scanning 1.5 µl of each sample in a NanoDrop spectrophotometer (Thermo Fisher Scientific, Inc. Waltham, MA). Extracted DNAs

were used as template in polymerase chain reaction that used primer sets that targeting the chloroplast trnL (UAA) intron. Amplicons generated from PCR reactions were cloned into a plasmid (pGEM, Promega Madison, WI) and plasmids were transformed into E. coli. Bacterial colonies from the resulting library were selected and plasmids were purified and inserts were sequenced. The BLAST was used to compare insert sequences to sequence present in GenBank.

**2017**

***Plant species mixtures.*** Five common weedy plant species were used including bull thistle (*Cirsium vulgare* (Savi) Ten.), field bindweed (*Convolvulus arvaensis* L.), Russian olive (*Elaeagnus angustifolia* L.), bittersweet nightshade (*Solanum dulcamrar* L.), and white clover (*Trifolium repens* L.). Samples of each plant species were collected from locations at the HAREC, and several other locations (i.e. Riverside Park and the corner of NE 8$^{th}$ and E Oregon) in Hermiston, OR.

DNA extraction of plant samples was attempted on both fresh and dried plant samples. Nucleic acid extraction of dry plant samples resulted in lower DNA concentrations for both 10 and 100 mg of plant material than fresh plant material. A fresh weight of 50 mg of plant material yielded the highest concentration and quality of DNA so individual plant species and plant species mixtures were created to total 50 mg fresh plant tissue. As a proof of concept and to determine if relative abundance of plant species in a mixture could be inferred from sequence data obtained using the metabarcoding procedures, each plant species (50 mg) and three mixtures of each plant species in equal, highly skewed, and slightly skewed amounts (Table 1) of the five plant species were created based on fresh weight. Nucleic acids from each of the plant samples were then extracted using the Dellaporta method (REF.) and subjected to PCR and metabarcoding.

***Insect Field Collection.*** In June 2017, Lygus bugs and western flower thrips were collected from two field sites in Umatilla County of northeastern Oregon. The insects were collected by sweep net sampling or aspirating from specific plant hosts in both crop and non-crop areas around potato. At least 10 individual thrips and 5 lygus bugs were collected from each plant species at each location.Insects for each insect-plant host-pairing were collected from the same plant species in at least two locations for "proof of concept" metabarcoding experiments. DNA was extracted from groups of 5 Lygus individuals and 10-15 thrips individuals using a modified CTAB extraction protocol.

***Amplicon library preparation and sequencing****.* For all plant and insect samples collected, the chloroplast trnLgene was amplified using PCR with dB49863 and cA49325 (Table 2) primers that included adaptors for Illumina indexing and sequencing. Amplicons from the PCR reaction were visualized in an agarose gel, stained with gel red (Figure 1), and purified using xxx kit. PCR products from each of the five plant species, individually, and the three plant species mixtures were prepared for sequencing. Sample indexing PCR was performed by Oregon State University's Center for Genome Research and Biocomputing OSU-CGRB and sequences were obtained using the MiSeq Protocol resulting in paired 300-bp reads.

***Work-flow for Miseq metabarcode sequences.*** The sequence data were downloaded from the OSU-CGBR server to local machine and fastqc was used to check the quality of the raw reads. Trimmomatic (version 0.36) was used to trim pair-end adapters and reads were organized into folders by samples identification. Paired- reads were assembled with COPE () and cutadapt () was used to remove the primer sequences. Sequences were converted from fastq to fasta and all fasta files from each study were pooled together. Identical sequences were dereplicated, clustered excluding chimearas and Megan6 was used to produce an OTU table. A python script was written to import trnL sequences and retrieve taxonomic information associated with each sequence from GenBank. The trnL sequences and taxonomic information were concatenated together to be used as a reference database. We are

currently working to create a reference database that contains plant species specific to the Columbia Basin. QIIME was used to assign taxonomy to each OTU in the OTU table using the blastn method. Quantification of sequence abundance, data visualization, and statistical analysis were conducted in R.

**The following additional experiments have been conducted** and nucleic acids have been extracted from all of the samples generated from those experiments using a modified CTAB extraction protocol. Also, trnL amplicon libraries have been generated using the dB49863 and cA49325 primers, but none of the trnL libraries have been sent for sequencing at this time.

***Gut clearance time experiments.*** Our preliminary work suggested that plant DNA could be detected from lygus bug (*Lygus Hesperus*), Western flower thrips (*Frankliniella occidentalis*), and beet leafhopper (*Circulifer tenellus*). All three insects have different feeding behaviors which may impact the time plant meals can be detected in the insect gut (Backus 1988). Lygus spp. have a "lacerate and flush" feeding style, where repeated insertion and withdrawal of a stylet into plant cells macerates and discharge of saliva extra-orally digests the plant tissues and cell contents. Saliva is then used to flush the digested cell contents into the sucking mouthpart. Thrips generally have a punch and suck feeding style meaning the insect breaks up plant cells with their mouthparts (i.e., punch) prior to ingesting the plant juices (i.e., suck). Leafhoppers are sheath-feeding insects that generally feed on the vascular tissue of the plant (i.e., phloem) but ingest cell contents from other plant tissues during numerous short duration tests probes prior to settling and feeding for longer duration on the vascular tissue (Backus 1988). To determine the length of time a plant species remains detectable and identifiable in the insect gut, the following gut passage experiment were conducted in 2017:

***Lygus.*** Lygus bugs (*Lygus spp.*) were collected in several locations around the HAREC and were placed in a cage and given access to either green bean plants or water. Lygus bugs were sampled after 0, 12, 24, 72, and 168 hours of access to each diet treatment. This experiment was repeated four times and Lygus bugs were sampled every three days until there were no remaining insects in the cages (13 days for replicates 2 and 3; 43 days for replicate 4, and 28 days for replicate5).

***Leafhopper.*** Similar experiments were conducted using beet leafhoppers (*Circulifer tenellus* Baker) and aster leafhoppers (*Macrosteles* spp.) caught around the HAREC. Leafhoppers were caged on beet (*Beta vulgaris*) plants or water and sampled after 0, 12, 24, 72, and 168 hours of access to their respective diet.

***Thrips.*** Thrips caught from potato, white clover, and puncture vine flowers were caged on green bean, water or a green 5% sucrose solution and sampled after 0, 12, 24, 72, and 168 hours of access to their respective diet treatment.

***Brown marmorated stink bug (BMSB).*** Since BMSB were easy to maintain in culture and produces eggs that are not exposed to plant materials and has a lacerate-and-flush feeding style similar to lygus bugs, we decided to use BMSB for a proof of concept feeding experiment. Following eclosion, 1$^{st}$ instar individuals were given access to choice and no-choice diets consisting of seven treatments, 1) no food, 2) carrot, 3) sugar snap pea, 3) apple, 4) blueberry, 5) blueberry & apple, 6) carrot & sugar snap pea, and 7) all four foods. Currently, a sequential no-choice feeding experiment with three diets is being conducted.

## SIGNIFICANT ACCOMPLISHMENTS:

### 2016

We identified several primer sets that target plant DNA and have had success amplifying target sequences extracted from insects. From thrips collected from dandelion, we detected DNA sequences

that matched sequences from the genera *Taraxicum* (50%), Pstacia (25%), Solanum (6%), and Medicago (6%). Approximately 13% of the sequences did not match known sequences in GenBank.

**2017**

**Plant species mixture.** The proof of concept experiment to examine if changing the relative abundance of a specific plant species based on fresh tissue weight revealed that (Table 4, Figure 1) all species that were put in the mixture were detected, at some level, by sequencing. Other sequences were also present in low abundance, less than 2% of the total sequences, suggesting the existence of some background contamination or sequencing errors. There is not a direct relationship between relative abundance of tissue fresh weight and relative abundance of sequences obtained after sequencing. Using this metabarcoding method, we cannot infer relative abundance of plant materials based on sequence information; only the presence absence of plant is currently all that can be inferred. These results are not surprising given the PCR amplification step is known to introduce bias into the methodology. There are currently some new methods being used to account for bias in the PCR amplification step – we may adopt some of those methods.

***Insect Field Collections*.** We have learned that we can detect plant DNA in plant feeding insects with multiple feeding behaviors. In total, we have detected plant materials in Thrips (punch-and-suck), lygus (lacerate and flush), aphid (phloem-feeding), leafhopper (phloem-feeding) and psyllid (phloem-feeding). Here we only report our work on thrips and lygus.

Plant sequences obtained from nucleic acid extracts of insects collected from a single plant host were often diverse, and often represented multiple plant hosts (Figure 2 & 3). There are multiple reasons that this may have occurred. Multiple insects were sampled from a single host, but we do not know if or for how long each individual may have fed on the host plant; each insect may have had a different feeding history. If the plant hosts on which the insects were collected is not a preferred host, the insects might have been transitioning to a preferred host, test-probing any green tissue they encountered on the way. The gut passage and sequential feeding studies that have been completed, but are not yet sequenced will reveal

Thrips are known to be highly polyphagous. The relative abundance of dandelion and potato sequences amplified was high from Thrips collected from dandelion and potato flowers (Figure 2). This might suggest that the thrips collected from those hosts had been feeding on those hosts for longer periods of time. There was a greater diversity of plant sequences amplified from thrips collected on white clover and onion. Lygus spp. are very mobile and highly polyphagous. There was a high diversity of plant sequences detected in lygus spp. despite being collected on a single host plant (Figure 3). This suggest that the individual lygus bugs may have been moving host-to-host looking for a preferred host on which to settled and feed.

Other accomplishments: We have increased our capacity to study insect diet, both in terms of the molecular methods we used as well as the computational capacity to examine the sequence data generated from these studies. Additional data will be forthcoming as our additional experiments are sequenced.

**BENEFITS & IMPACT:** We have developed methods to conduct molecular diet analysis of small soft-bodied insects including the detection and identification of plant species based fragments of plant DNA ingested by insects. We using these same methods to examine the species composition of pollen loads on native bee species and food web interactions of an invasive crayfish species introduced into the John Day River system. The methods can be very generally applied to many systems. Applied to agricultural problems, have the ability to directly determine insect diet will help to identify plant species in non-crop

habitats that have the greatest importance as food sources of insect pests or as virus inoculum sources. If successful, this project will directly benefit producers vegetable and associated processing industries in the Pacific Northwest by helping to refine pest management strategies for insect-vectored plant pathogens of potato. PI-Frost has delivered presentations at multiple commodity meetings and will extend research results from this project in presentations aimed at target audiences in OR, ID, and WA.

***Meeting Proceedings:***
Frost, K.E. 2016. Exploring the landscape of plant viruses using plant feeding insects. Proceedings of the 2016 Washington-Oregon potato Conference, January 27 & 28, Three Rivers Convention Center, Kennewick, WA.

***Presentations including information from this project:***
Molecular diet analysis and program research update 2018. Washington State Potato Commission. January 17, 2018. Pasco, WA.

Aster Yellows in Vegetable Crops. Pacific Northwest Vegetable Association Conference & Trade Show. November 11, 2016. Kennewick, WA.

Exploring the landscape of plant viruses using plant-feeding insects. Washington Oregon Potato Conference, January 27, 2016. Kennewick, WA.

Insect vectors in vegetables, new approaches. Pacific Northwest Vegetable Association Conference & Trade Show. November 19, 2015. Kennewick, WA.

**ADDITIONAL FUNDING RECEIVED DURING PROJECT TERM:** The Northwest Potato Research Consortium (NPRC) has continued to fund work to conduct diet analysis of Lygus spp. and Thrips spp. affecting potato crops ($20,000 in FY 2016-17; $21,875 in FY 2017-18). This proposal was submitted to the NPRC for a third year of funding ($20,616 for FY 2018-19), but awards have not been announced.

**FUTURE FUNDING POSSIBILITIES:** In 2018, Frost has plans to submit a proposal with Dr. Anders Huseth (North Carolina State University) to the USDA AFRI Foundational *Pests and Beneficial Species in Agricultural Production Systems* program area to study how Lygus spp. use non-crop habitats in potato and cotton agroecosystems.

**TABLES AND FIGURES**

**Table 1**. Number of milligrams of fresh tissue from each plant species in each mixtures based on fresh tissue weight.

| Mixture | Bull thistle | Bindweed | Russian Olive | Nightshade | White Clover |
|---|---|---|---|---|---|
| Bull thistle | 50 | 0 | 0 | 0 | 0 |
| Bindweed | 0 | 50 | 0 | 0 | 0 |
| Russian Olive | 0 | 0 | 50 | 0 | 0 |
| Nightshade | 0 | 0 | 0 | 50 | 0 |
| White clover | 0 | 0 | 0 | 0 | 50 |
| Equal | 10 | 10 | 10 | 10 | 10 |
| Highly Skewed | 46 | 1 | 1 | 1 | 1 |
| Slightly Skewed | 15 | 12.5 | 10 | 7.5 | 5 |

**Table 2.** Primer sequences used for amplification of plant barcode regions

| Primer | Primer Seq. (5'-3') | ExpectedAmplicon Size (bp) |
|---|---|---|
| trnL(c) | TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCGAAATCGGTAGACGCTACG | ~400-600 |
| UAA(h) | GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGCCATTGAGTCTCTGCACCTATC | |
| ITS86 | GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGTGAATCATCGAATCTTTGAA | |
| ITS4 | TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGTCCTCCGCTTATTGATATGC | |

Note: Symbol Key: W = A or T. Full length primer sequences for plant barcode library preparation (i.e. highlighted portion of the sequence) with Illumina adaptor sequences.

**Table 3**. Host plant species from which Lygus and Thrips were collected (i.e. insect-host plant species pairing) prior to nucleic acid extraction, trnL library preparation, and sequencing.

| Insect | # Replicates | # Individuals | Plant Species |
|---|---|---|---|
| Lygus spp. | 3 | 5 | Potato (*Solanum tuberosum* L.) |
| | 2 | 5 | Russian Thistle (*Kali tragus* L.) |
| | 2 | 5 | Pigweed (*Amaranthus spp*.) |
| | 2 | 5 | Lamb's quarter (*Chenopodium album* L.) |
| Thrips spp. | 3 | 10-15 | Potato (*Solanum tuberosum* L.) |
| | 2 | 10-15 | White Clover (*Trifolium repens* L.) |
| | 2 | 10-15 | Common Dandelion (*Taraxacum officinale* L.) |
| | 2 | 10-15 | Onion (*Allium cepa* L,) |

**Table 4**. Average percent (SD) of sequences of each plant species from Illumina MiSeq sequencing of plant species mixtures.

| Mixture | Not mixed | Equal[a] | Highly Skewed | Slightly Skewed |
|---|---|---|---|---|
| Bull thistle | 98.7 | 17.7 (5.1) | 73.4 (1.4) | 31.7 (5.6) |
| Bindweed | 98.2 | 5.5 (2.4) | 2.7 (0.6) | 7.1 (1.6) |
| Russian Olive | 99.4 | 54.4 (12.5) | 12.8 (3.3) | 39.9 (10.1) |
| Nightshade | 99.5 | 17.0 (4.0) | 6.9 (2.3) | 14.6 (2.6) |
| White clover | 99.9 | 1.6 (0.7) | 0.4 (0.2) | 1.0 (0.4) |

[a]Other sequences represented 3.8, 3.8, and 5.7 percent of the equal, high skewed, and slightly skewed plant mixtures, respectively.

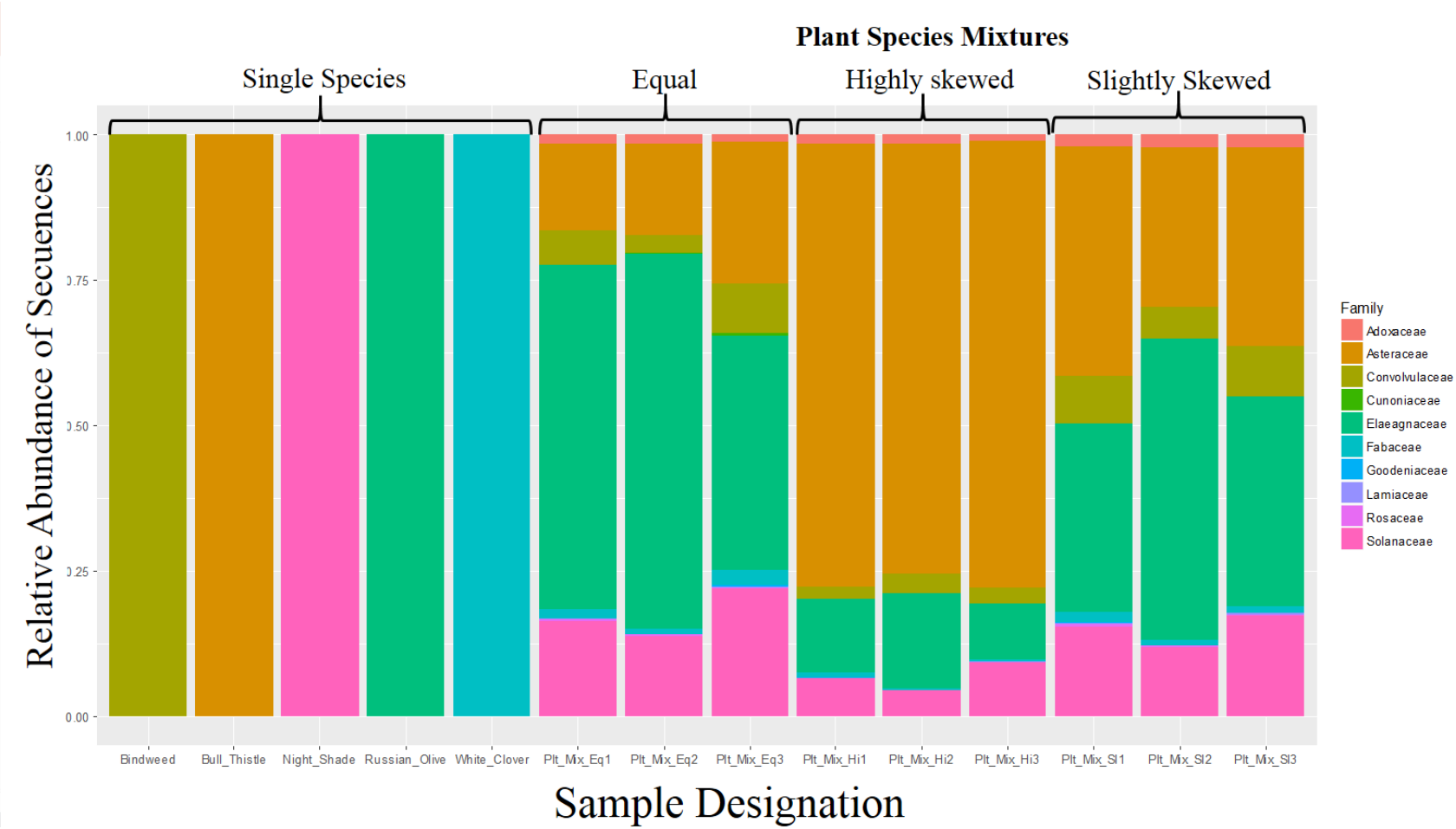


**Figure 1.** Relative abundance of sequences and assigned families for plant species mixture experiment. Only one replicate of each Individual plant species was sequenced and plant species mixtures were run in triplicate.

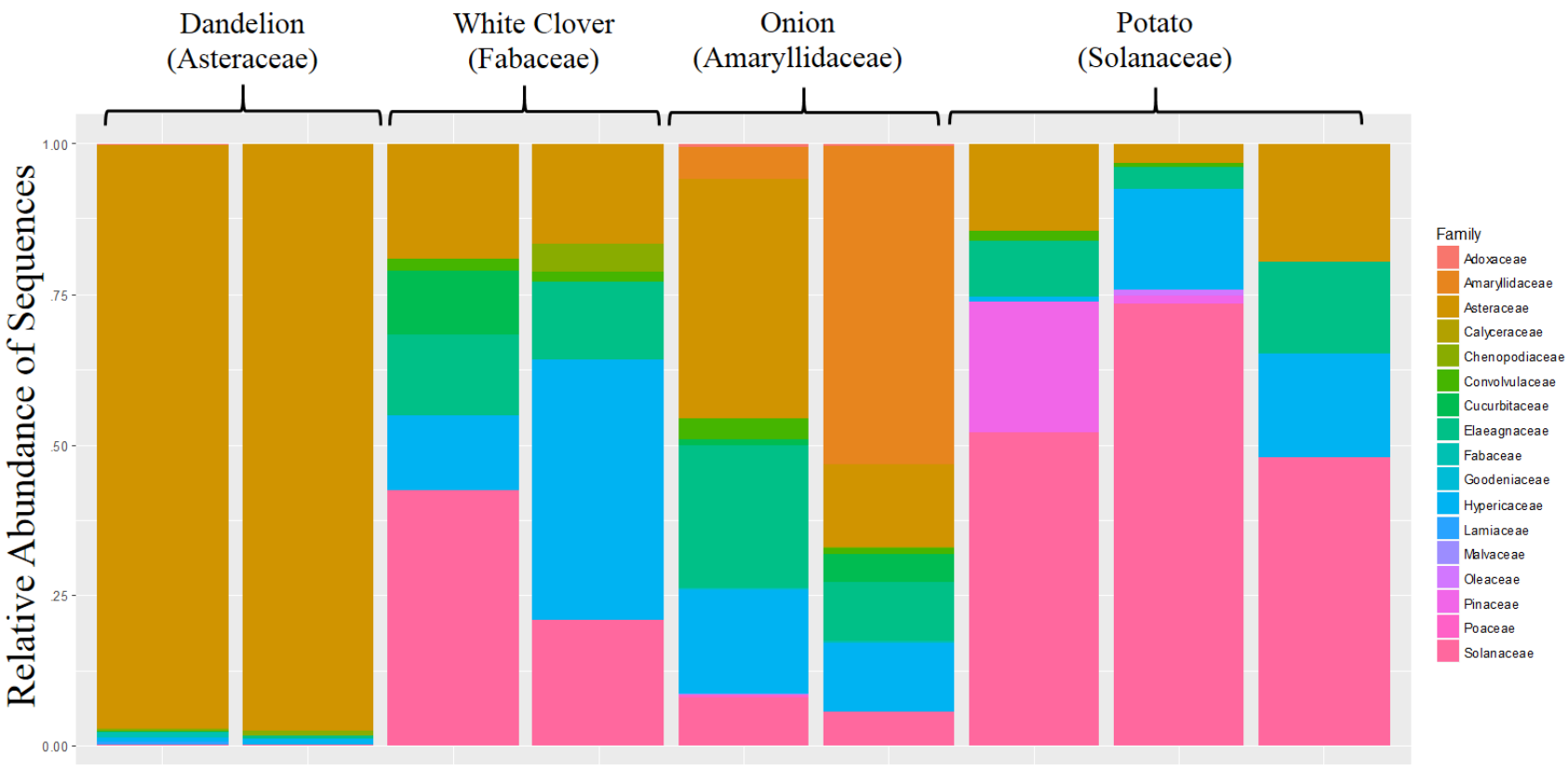


Figure 2. Relative abundance of sequences and assigned families of amplicons from thrips species collected from different plant species, dandelion, white clover, onion, and potato. Multiple thrips individuals (10-15) were in each sample and thrips collected from each plant species were run in duplicate or triplicate.

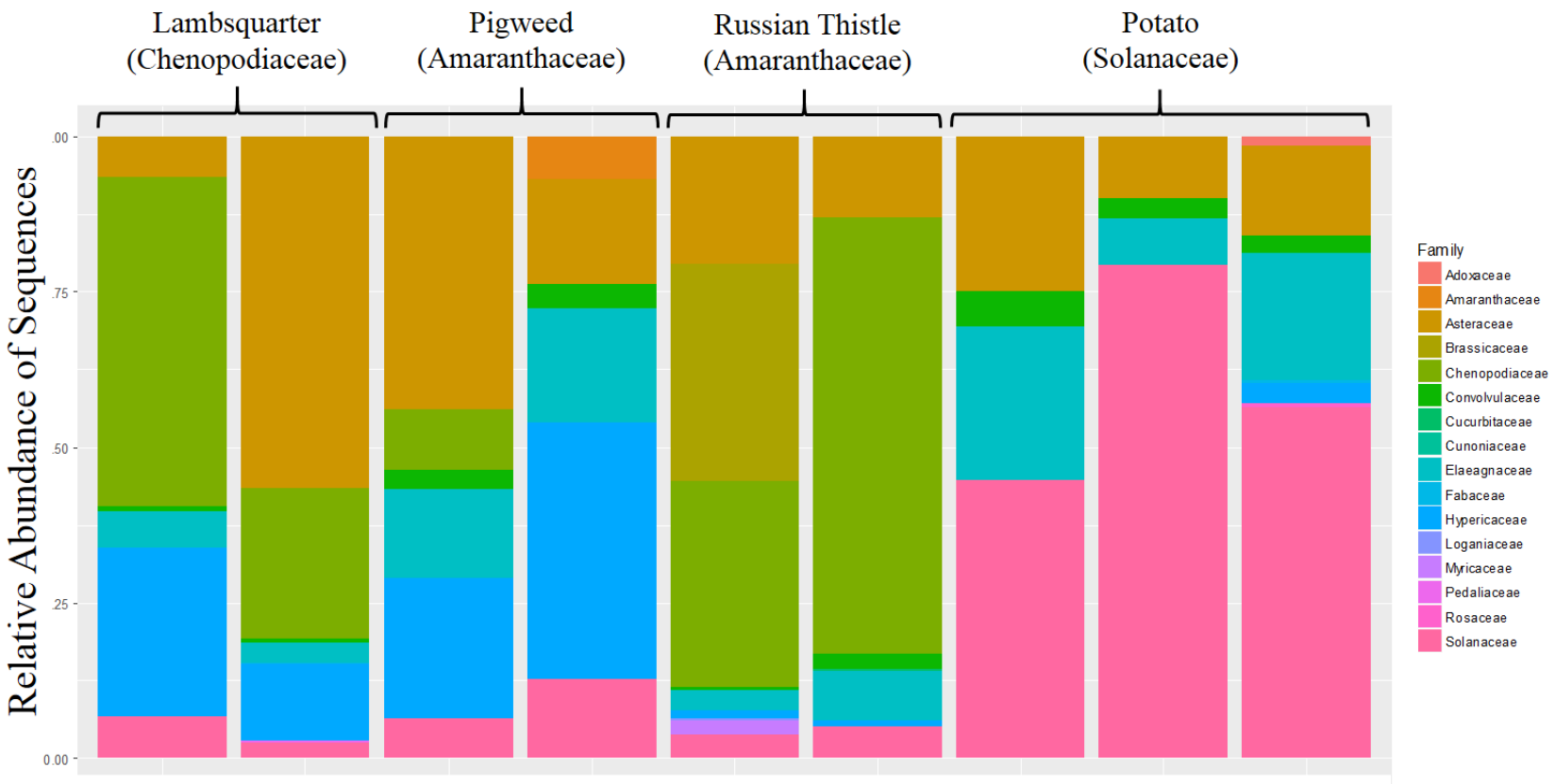


Figure 3. Relative abundance of sequences and assigned families of amplicons from lygus bugs collected from different plant species, lambsquarter, pigweed, Russian thistle, and potato. Multiple thrips individuals (10-15) were in each sample and thrips collected from each plant species were run in duplicate or triplicate.