

Genetic Structure of *Liriomyza trifolii* (Diptera: Agromyzidae) Associated With Host Plants From Southeastern Mexico

Jessica Pérez-Alquicira,^{1,10} Esther van der Knaap,^{2,3} Kristin Mercer,⁴ Leah McHale,⁴ Thomas Mitchell,⁵ Jose Luna-Ruiz,⁶ Esther Texocotitla-Vázquez,⁷ Ofelia Vargas-Ponce,⁸ and Andy Michel⁹

¹Departamento de Botánica y Zoología, CONACYT – Laboratorio Nacional de Identificación y Caracterización Vegetal, Centro Universitario de Ciencias Biológicas y Agropecuarias, Universidad de Guadalajara, Zapopan, Mexico, ²Department of Horticulture, University of Georgia, Athens, GA, ³Department of Horticulture and Crop Sciences, Ohio State University, Wooster, OH, ⁴Department of Horticulture and Crop Science, Ohio State University, Columbus, OH, ⁵Department of Plant Pathology, Ohio State University, Columbus, OH, ⁶Centro de Ciencias Agropecuarias, Universidad Autónoma de Aguascalientes, Aguascalientes, México, ⁷Facultad de Ciencias, Universidad Nacional Autónoma de México, México, ⁸Departamento de Botánica y Zoología, CUCBA, Universidad de Guadalajara, Zapopan, Jalisco, México, ⁹Department of Entomology, Ohio State University, Wooster, OH, and ¹⁰Corresponding author, e-mail: perezalquicira@gmail.com

Subject Editor: Nick Miller

Received 13 July 2018; Editorial decision 28 November 2018

Abstract

Host-associated differentiation (HAD) has played a major role in insect diversification at both macroevolutionary and microevolutionary scales. This evolutionary process has been reported in insects associated with wild and domesticated plant species. In particular, domesticated species harbor large genetic and phenotypic diversity associated with traits of human interest, including variation in nutrition, phenology, fruit, and leaf shape. This diversity may alter selection regimes affecting insect evolution and host specialization. The genus *Liriomyza* includes highly polyphagous species that are characterized for living and feeding inside plant leaves. Ecological and genetic data suggest the presence of cryptic species within this genus. Moreover, there is evidence of HAD in a group of populations of *Liriomyza trifolii* (Burgess) associated with *Capsicum annum* L. (Solanaceae). In this work, we explored HAD in *L. trifolii* populations from southeastern Mexico, and inquire into differentiation specific to peppers based on cytochrome oxidase I. We also evaluated the relationship between the genetic structure of leafminers and the different types of *C. annum*. Our main results did not support previous findings of specialization of *L. trifolii* on *C. annum*. Nevertheless, we found a divergent group of haplotypes associated to *Allium cepa* (Asparagales: Amaryllidaceae) in sympatric condition to *Physalis philadelphica* Lam. (Solanales: Solanaceae) and *C. annum*, suggesting the presence of HAD, as well as significant genetic differentiation of *L. trifolii* associated to peppers from Oaxaca and Yucatán.

Key words: host-associated differentiation, genetic diversity, domestication, phytophagous

Ecological divergence has been a major driver of species diversification (Rundle and Nosil 2005). Previous research has shown that among polyphagous insects, specialization and ecological divergence occur as a consequence of the selection pressures imposed by host traits (Antwi et al. 2015). This process has been called host-associated differentiation (HAD) (Bush 1969, Abrahamson et al. 2001), and it has been demonstrated in multiple insect species feeding on wild and cultivated host plant species (Guttman et al. 1981, Via 1991, DeBarro et al. 1995, Nason et al. 2002, Ruiz-Montoya et al. 2003, Conord et al. 2006, Alvarez et al. 2007, Barman et al. 2012). Adaptation to these host plants could lead to HAD (Abrahamson et al. 2001, Medina et al. 2012, Antwi et al. 2015, Forbes et al. 2017, Ramírez-Romero et al. 2017).

One of the processes that has also influenced the patterns of genetic diversity of insects is crop domestication. During domestication, plant traits, including morphology, phenology, defense, and nutrition, have been selected to satisfy human needs, producing a larger diversity within these plant traits than is found in crop-progenitor wild populations (Lindig-Cisneros et al. 1997, Benrey et al. 1998, Gols et al. 2008, Meyer and Purugganan 2013, Chen et al. 2015). This novel phenotypic diversity, along with common agricultural practices, may impose different and unique selective pressures on the biotic (e.g., insect) community in agroecosystems, influencing the ecology, genetic structure, and evolution of individual species (Lindig-Cisneros et al. 1997, Benrey et al. 1998, Gols et al. 2008, Chen et al. 2015).

Moreover, the geographical and ecological contexts in which the host crop plants grow may play an important role in shaping the genetic structure and evolutionary dynamics of insects (Reitz and Trumble 2002). For example, Barman et al. (2012) found that insects use different host plants across their geographical distribution as a consequence of ecological components including variation in the abundance and availability of host plants. Therefore, multiple human-mediated and natural factors can influence the adaptive potential and population genetic structure of insects in agroecosystems.

The leafminer *Liriomyza trifolii* (Burgess) is a very important insect pest around the world, infesting >120 species in 21 families (Minkenberg and Lenteren 1986). *Liriomyza trifolii* and two other species, *Liriomyza sativae* Blanchard and *Liriomyza huidobrensis* (Blanchard), are the most damaging species in the genus (Capinera 2001). The three species originated in the Americas and subsequently have spread to different regions of the globe (Minkenberg 1988, Scheffer 2000, Scheffer and Lewis 2006). Although *L. trifolii* is a polyphagous species, previous research suggested that this species is in the process of specialization on *Capsicum annuum* (chile pepper; Solanaceae) through incipient speciation. Within California, Reitz and Trumble (2002) found that *L. trifolii* groups are adapted to different host plants depending on geographical region. While *L. trifolii* populations from southern California successfully fed, oviposited, and reproduced on all the associated hosts, central Californian populations fed significantly less on all hosts and only reproduced on pepper. These results suggested the presence of differentiated groups of *L. trifolii* in California (Reitz and Trumble 2002). Additionally, research on mitochondrial genetic markers (*cytochrome c oxidase subunit I* [COI]) of *L. trifolii* indicated clade divergence associated with different plant hosts (Scheffer and Lewis 2006). This evidence supports the main results found by Reitz and Trumble (2002) of incipient speciation of *L. trifolii* in peppers and suggest that HAD is in an early, but ongoing, process within this species. However, more genetic information and broader host sampling is needed to support this hypothesis, including from samples infesting the diverse phenotypic variation of *C. annuum* in Mexico.

Our goal was to determine if *L. trifolii* from Mexican agricultural fields supports HAD and if additional divergence is occurring within a pepper-specific clade. Mexico is an interesting location for this work, because it harbors the centers of domestication and diversity of *C. annuum*. In particular, two states in southern Mexico boast a large diversity of peppers landraces (local cultivars improved by traditional agriculture): Oaxaca and Yucatán. Since pre-Columbian times, indigenous groups have selected fruit traits of agronomic importance, including the shape, size, taste, and degree of pungency, producing a great diversity of pepper types. This diversity might have influenced the microevolutionary dynamics of *L. trifolii* and the emergence of the ‘pepper clade’ (Scheffer and Lewis 2006). Along with *C. annuum*, southern Mexico produces many other crops used by leafminers as hosts. Indeed, *L. trifolii* from onion (*Allium cepa* (Asparagales: Amaryllidaceae)), tomatoes (*Solanum lycopersicum* L. (Solanales: Solanaceae)), and beans (*Phaseolus vulgaris* L. (Fabales: Fabaceae)) clustered separately from the pepper clade in a previous study (Scheffer and Lewis 2006). We also included leafminers from tomatillo (*Physalis philadelphica* Lam. (Solanales: Solanaceae)) as this is a commonly grown crop, but yet genetic information of leafminers is lacking. Our specific aims were to 1) analyze the genetic structure of *L. trifolii* from different hosts to investigate the HAD hypothesis; 2) explore whether a specific clade of *L. trifolii* is associated to *C. annuum* in Yucatán and Oaxaca; 3) analyze if the diversity of pepper landraces in Oaxaca and Yucatán is related to the genetic structure of *L. trifolii*; and 4) evaluate whether geographical isolation influences the genetic structure of *L. trifolii*.

Materials and Methods

Sampling

In 2012 and 2013, we collected *L. trifolii* samples from *C. annuum* landraces (chile de agua and chile costeño) at six locations in Oaxaca and one location from Yucatán (chile dulce and chile xcatik) (Fig. 1, Table 1; the spanish names are in italics). Additionally, we collected samples of leafminers from *C. chinense* (chile habanero) from backyards in two towns in Yucatán state as well as wild peppers (*C. annuum* var. *glabriusculum* (chile maax'ik), Table 1). Chile maax'ik was identified as *C. annuum* var. *glabriusculum* (wild pepper) based on morphology; however, recent genetic research has showed that chile maax'ik is grouped in a clade with chile dulce and not within the *glabriusculum* clade (Taitano et al. 2018). Although the taxonomic status of chile maax'ik is unclear, we will use the chile maax'ik nomenclature in this paper, with genotypic information from this collection present in Taitano et al. (2018). Moreover, this pepper is not intensively cultivated as the remaining peppers included in this study. We surveyed tropical deciduous forest along the Oaxacan coast, where *C. annuum* var. *glabriusculum* grows; however, we did not find any leafminers on this small fruited pepper. In total, we collected 66 *L. trifolii* samples on four *C. annuum* landraces (chile de agua, costeño, xcatik, dulce), 12 samples from chile maax'ik, and 6 samples from *Capsicum chinense* (chile habanero). Because our aims were to explore whether a specific clade of *L. trifolii* is associated with *C. annuum* and also investigate HAD, we collected leafminers from other hosts growing in the same agricultural fields with *C. annuum*. We collected 13 leafminers on tomatillo, 7 on onion, and 12 on tomato in Oaxaca, and 3 leafminers on beans and 2 on tomato in Yucatán (Table 1). In the field, we inspected plants for the presence of larval mines on leaves. When found, we removed the mined leaf from the plant and placed it in a plastic bag for transport back to the laboratory. Leafminers were either dissected from the leaf (larval collections), allowed to pupate (pupal collections), or reared to adults (adult collections). For all life stages, leafminers were preserved inside 1.5-ml tubes with silica gel and a piece of cotton.

DNA Extraction and Sequencing

Total leafminer DNA was extracted using the Omega Bio-Tek E.Z.N.A. tissue DNA kit (Norcross, GA). We amplified COI, because this DNA barcode has been considered a valuable tool for the identification of leafminers (Scheffer et al. 2006). The amplification DNA protocol was followed as described by Scheffer and Lewis (2006). The sequencing reactions were performed by Functional Biosciences DNA sequencing service (WI, functionalbio.com).

Data Analyses

Sequences were edited with CodonCode aligner (version 7.1.1; CodonCode Corporation, www.codoncode.com). These were compared with sequences in GenBank using BLAST (Altschul et al. 1990) to corroborate that we had the correct gene fragment (COI). As we included larval and pupal collections, we confirmed that the COI sequences were Dipteran and not from Hymenoptera or any other potential parasitoid. Alignment was conducted using Muscle (Edgar 2004). We calculated genetic diversity parameters, including haplotype number (H), haplotypic diversity (*h*), and nucleotide diversity (*π*) using DnaSP software version 5 (Librado and Rozas 2009). Deviation from neutral equilibrium was evaluated by Tajima's *D* (Tajima 1989) and Fu's *F_s* (Fu 1997) test using Arlequin software version 3.5.2.2 (Excoffier and Lischer 2010). We also generated a haplotype list combining our data with COI sequences produced by Scheffer and Lewis (2005, 2006) (GenBank AY697731–AY697843

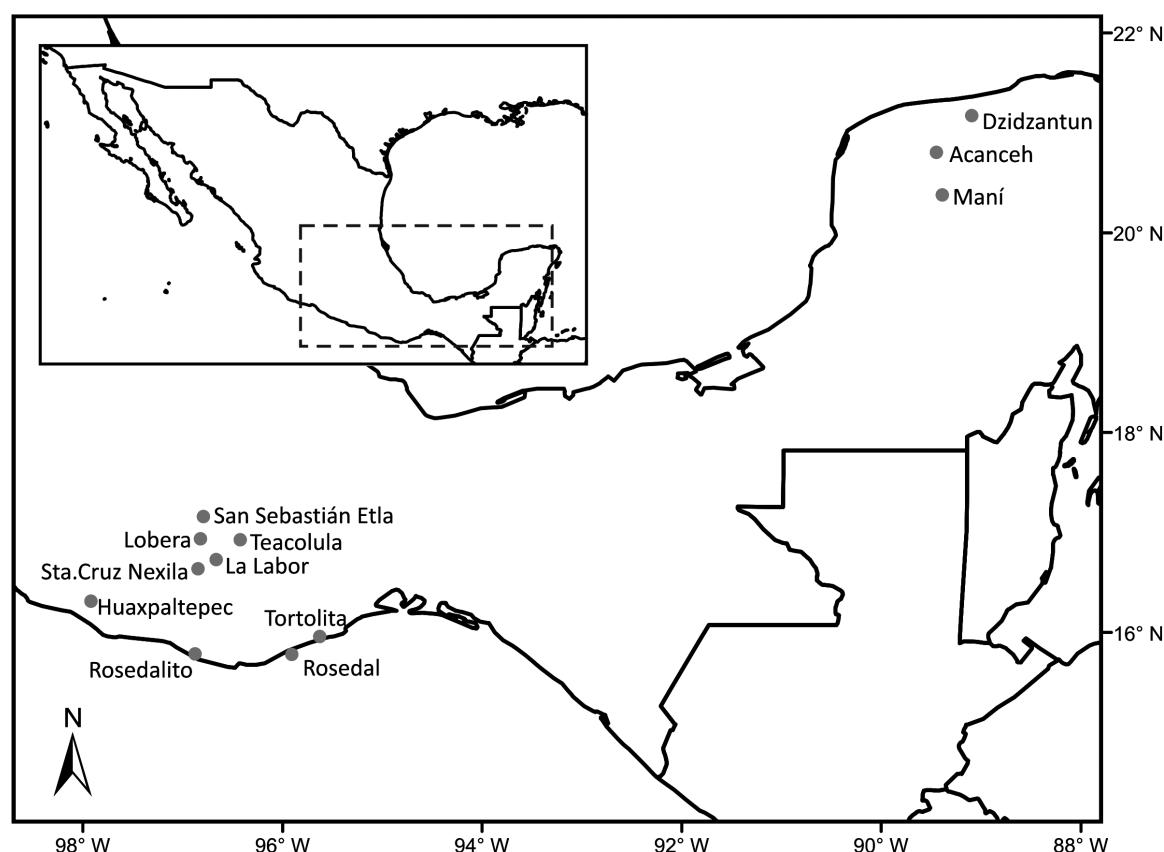


Fig. 1. Map of sampling locations of *L. trifolii* and *L. sativae* in Oaxaca and Yucatán, Mexico.

and DQ516539–DQ516678, respectively). This database will allow us to compare our results with those generated by Scheffer and Lewis (2005, 2006). We obtained a final database with 492 base pairs and we will refer it as combined database.

We calculated a pairwise genetic distance matrix for *L. trifolii* based on F_{ST} , estimated using allele frequencies (Wright 1951) and N_{ST} , which takes into account genetic similarities among haplotypes (Pons and Petit 1996). We used DnaSp version 5 (Librado and Rozas 2009) for these estimations with 1,000 permutations to assess significance. We implemented an analysis of molecular variance (AMOVA) using Arlequin version 3.5.2.2 (Excoffier and Lischer 2010) with 1,000 permutations to estimate hierarchical genetic structure within and among hosts. We included insects collected on three crop species in this analysis: peppers (including all the *C. annuum* varieties, also chile *maax'ik* and *C. chinense*), tomatillo, and onion. To understand how the diversity in peppers related to the genetic structure of *L. trifolii*, we performed a second AMOVA for *L. trifolii* collected only from peppers. Finally, we performed a third AMOVA using *L. trifolii* collected from peppers and separating them by geographic regions, Oaxaca and Yucatán, to test if there are genetic differences between peppers from both geographical regions.

To assess the relationships between geographical and genetic distances, we performed a Mantel test (Mantel 1967) for *L. trifolii* samples with 10,000 permutations using the package ADE4 in R (Dray and Dufour 2007). Then, we performed a second Mantel test for *L. trifolii* associated with peppers. We obtained the geographical distance matrix with the Geographic Distance Matrix Generator version 1.2.3 (Ersts, Internet 2018) using the F_{ST} pairwise distance matrix.

To determine the presence of genetic clusters among all leafminer samples (including non-*L. trifolii* species), we used the Bayesian Analysis of Population Structure (BAPS) version 5.2 (Corander et al.

2006, 2008) with $K = 10$ and 20 repetitions using the method of clustering for linked loci. A second analysis using the same parameters was performed to test for the presence of HAD within *L. trifolii*.

Genetic Relationships Among Haplotypes

We determined the genetic relationships among insect haplotypes using a median joining network implemented in the PopART software (Leigh and Bryant 2015) and we performed Bayesian approach for phylogenetic inference using Beast version 2.3.1 (Bouckaert et al. 2014). We used the substitution model (HKY) determined by the Akaike information criterion in jModelTest version 2.1.10 (Guindon and Gascuel 2003, Darriba et al. 2012).

Results

We obtained high-quality sequence results for 121 individuals and 1,361 base pairs (GenBank MK111649–MK111752 and MK168162–MK168299; Table 1). According to the results from BLAST, we had collected 104 *L. trifolii* samples and 17 *L. sativae* (Table 1). *Liriomyza trifolii* was found in *C. annuum*, *C. chinense*, *P. philadelphica*, and *A. cepa*, whereas *L. sativae* was associated with *S. lycopersicum* and *P. vulgaris* (Table 1). Both *L. trifolii* and *L. sativae* were collected in the same geographical location (i.e., farm).

Genetic Diversity of *L. trifolii* and *L. sativae*

We detected a total of 32 variable sites and 14 haplotypes for *L. trifolii*, and 35 variable sites, and 7 haplotypes for *L. sativae*. The genetic diversity for *L. trifolii* based on haplotype diversity (h) was 0.70; and the nucleotide diversity (Π) was 0.004. The highest levels of diversity were found in the pepper landrace chile *de agua* and tomatillo for h

Table 1. Host plant and sampling locations for *L. trifolii* and *L. sativae* in Oaxaca and Yucatán, Mexico

Host	Common Name	N	Location	Management	Geographical coordinate	State	Species	H
<i>Capsicum annum</i>	Chile de agua	16	San Sebastián Etla	Landrace	N17.16510 W096.79075	Oaxaca	<i>Liriomyza trifolii</i>	4,5, 6, 7
<i>Capsicum annum</i>	Chile de agua	3	Sta Cruz Nexila	Landrace	N16.640231 W96.846306	Oaxaca	<i>Liriomyza trifolii</i>	7,12,13
<i>Capsicum annum</i>	Chile de agua	6	La Labor	Landrace	N16.73175 W96.664806	Oaxaca	<i>Liriomyza trifolii</i>	5,6,11
<i>Capsicum annum</i>	Chile de agua	1	Lobera	Landrace	N16.941222 W96.819972	Oaxaca	<i>Liriomyza trifolii</i>	11
<i>Capsicum annum</i>	Chile de agua	1	Teacolula (Paraje de Pedirillo)	Landrace	N16.92840 W096.42232	Oaxaca	<i>Liriomyza trifolii</i>	5
<i>Capsicum annum</i>	Chile de costeño	2	Rosedalito	Landrace	N15.788917 W96.876472	Oaxaca	<i>Liriomyza trifolii</i>	5,11
<i>Capsicum annum</i> var. <i>Glabriusculum</i> *	Chile maax'ik	2	Acanceh	Backyard	N20.812809 W89.4469	Yucatán	<i>Liriomyza trifolii</i>	6, 8
<i>Capsicum annum</i> var. <i>Glabriusculum</i> *	Chile maax'ik	10	Maní	Backyard	N20.38226 W89.38948	Yucatán	<i>Liriomyza trifolii</i>	9
<i>Capsicum chinense</i>	Chile habanero	6	Maní	Backyard	N20.38226 W89.38948	Yucatán	<i>Liriomyza trifolii</i>	6
<i>Capsicum annum</i>	Chile dulce	10	Dzidzantun	Landrace	N21.18117 W89.0929	Yucatán	<i>Liriomyza trifolii</i>	1,6,9
<i>Capsicum annum</i>	Chile xcatik	27	Dzidzantun	Landrace	N21.18117 W89.0929	Yucatán	<i>Liriomyza trifolii</i>	6,9
<i>Physalis philadelphica</i>	Tomatillo	3	Lobera	Landrace	N16.941222 W96.819972	Oaxaca	<i>Liriomyza trifolii</i>	5,6
<i>Physalis philadelphica</i>	Tomatillo	5	La Labor	Landrace	N16.73175 W96.664806	Oaxaca	<i>Liriomyza trifolii</i>	5,6,10, 11
<i>Physalis philadelphica</i>	Tomatillo	5	Teacolula (Paraje de Pedirillo)	Landrace	N16.92840 W096.42232	Oaxaca	<i>Liriomyza trifolii</i>	2,5,6
<i>Allium cepa</i>	Onion	7	Teacolula (Paraje de Pedirillo)	Landrace	N16.92840 W096.42232	Oaxaca	<i>Liriomyza trifolii</i>	3,14
<i>Solanum lycopersicum</i>	Tomato	2	Tortolita	Landrace	N15.965972 W95.625583	Oaxaca	<i>Liriomyza sativae</i>	20
<i>Solanum lycopersicum</i>	Tomato	4	Huaxpaltepec	Landrace	N16.316889 W97.921556	Oaxaca	<i>Liriomyza sativae</i>	15,16, 18
<i>Solanum lycopersicum</i>	Tomato	6	Rosedal	backyard	N15.783056 W95.906389	Oaxaca	<i>Liriomyza sativae</i>	15,17, 19
<i>Solanum lycopersicum</i>	Tomato	2	Dzidzantun	Landrace	N21.18117 W89.0929	Yucatán	<i>Liriomyza sativae</i>	20,21
<i>Phaseolus vulgaris</i>	Bean	3	Dzidzantun	Landrace	N21.18117 W89.0929	Yucatán	<i>Liriomyza sativae</i>	21

N, sample size, H, Haplotype number based on the genetic network and phylogenetic tree (Figs. 2 and 3).

*Chile maax'ik corresponds to *C. annum* var. *glabriusculum* based on morphology (see further explanation in Materials and Methods).

and II, considering only those with sample size above 10. The lowest levels of genetic diversity were found in pepper landraces chile *xcatik* and chile *maax'ik*. However, these measurements might be biased due to an uneven number of collecting sites per host (Table 1). We did not compare genetic diversity for *L. sativae*, because we obtained only 17 samples for this species. Tajima's D had significantly negative values for chile *dulce* and tomatillo, whereas *Fu's* had significant negative value for chile *de agua*, indicating departure from mutation-drift equilibrium in the direction of population expansion (Table 2).

Haplotype Network of *L. trifolii* and *L. sativae*

The median joining network for leafminers exhibited two main groups that are separated by 37 mutational steps (if we consider the most related haplotype between both groups) and correspond to 1) *L. trifolii* W (pepper-tomatillo group) and A group and 2) *L. sativae* W and A group (note: we used A and W groups to maintain the same genetic group names used by Scheffer and Lewis (2005, 2006).

The groups *L. trifolii*-W and *L. trifolii*-A were separated by 18 mutational steps (Fig. 2). The W group included 14 haplotypes collected from only two hosts: various chile pepper landraces (chile *agua*, chile *costeño*, chile *dulce*, chile *xcatik*, chile *habanero*, chile *maax'ik*), and tomatillo. The haplotype with the highest frequency was H6 and occurred in all pepper types and tomatillo. The H9 haplotype included samples from Yucatán (chile *maax'ik*, chile *dulce* and chile *xcatik*), whereas H5 and H11 included samples associated to *C. annum* and tomatillo, both from Oaxaca state. The H2 haplotype corresponded to two samples of chile *de agua*. The A group of *L. trifolii* included 2 haplotypes (H3 and H14) occurring in samples only from onion.

Separation within *L. sativae* was more complex among geography and hosts. Two main groups were detected and were separated by 29 mutational steps. The group W included samples associated with tomatoes from the central and western coast of Oaxaca (Huaxpaltepec and Rosedal locations, respectively), and

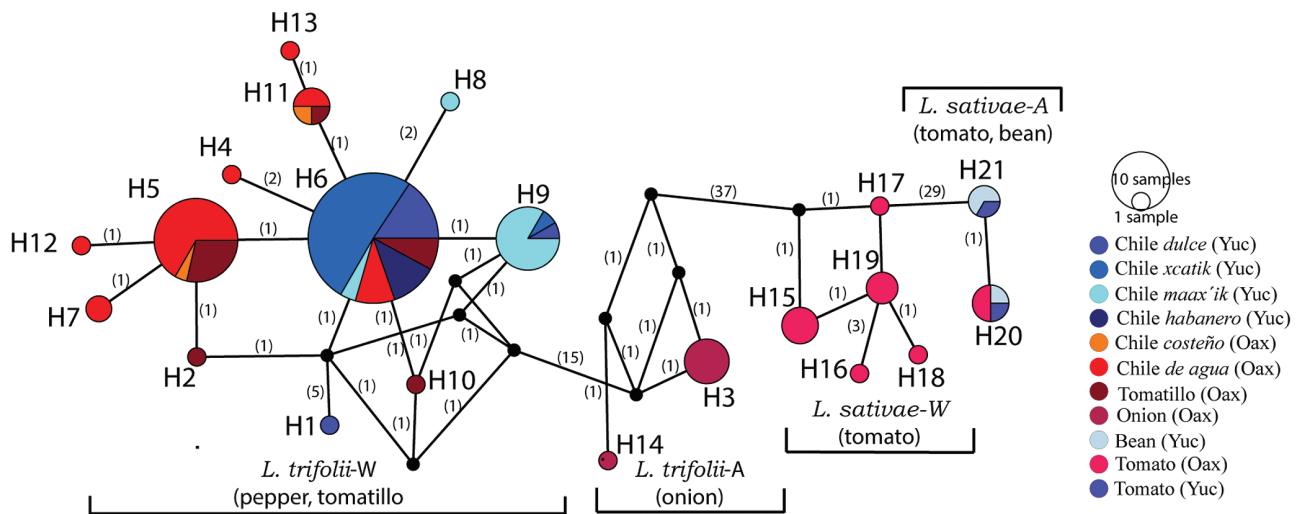


Fig. 2. Haplotype network for *Liriomyza trifolii* (group W and A) and *L. sativae* (group W and A) based on mitochondrial cytochrome oxidase subunit I. The size of the circles corresponds to haplotype frequency, black dots represent missing haplotypes, and the numbers above lines connecting haplotypes correspond to the number of mutations.

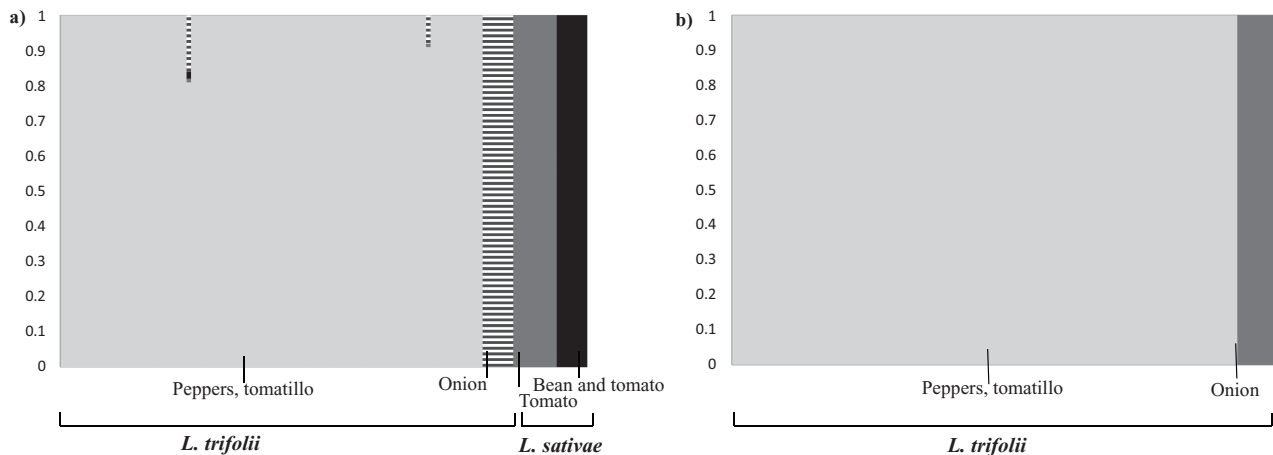


Fig. 3. Bayesian clustering analysis of mtDNA sequences (a) *L. trifolii* and *L. sativae* samples produced $K = 4$ and (b) *L. trifolii* samples produced $K = 2$.

the group A included samples associated with tomatoes from the most eastern coast of Oaxaca (Tortolita location), as well as with tomatoes and beans from Yucatán (Dzidzantun location) (Fig. 2).

Bayesian Analyses of *L. trifolii* and *L. sativae*

The Bayesian genetic assignment performed with BAPS for the whole data set (121 sequences) produced four genetic groups (Fig. 3a), similar to the haplotype network (Fig. 2). We also ran a second BAPS analysis including only *L. trifolii* samples, which recovered the W pepper-tomatillo and A groups (Fig. 3b). The Bayesian tree for 21 haplotypes recovered each species as monophyletic group, *L. trifolii* and *L. sativae* (Fig. 4). Moreover, within *L. trifolii*, the two main genetic lineages, the W and A groups, were also detected; however, the genetic relatedness of samples within group W was unresolved. For *L. sativae*, we recovered the A group, but the W group was unresolved and not well supported.

Genetic Structure of *L. trifolii*

The F_{ST} pairwise comparisons among *L. trifolii* collected on onion, tomatillo and peppers ranged from 0 to 0.96. Most of the values

were statistically significant (Table 3). The lowest values of genetic difference were found between leafminers collected on chile peppers: 1) between chile *costeño* and most other chile landraces (except samples collected from chile *maax'ik*); and 2) among samples from some Yucatan peppers (chiles *xcatik*, *habanero*, *dulce*). There was also little differentiation between samples from tomatillo and some peppers: chile *de agua* and chile *costeño* (both from Oaxaca) and chile *dulce* (from Yucatan). The highest levels of genetic differences (F_{ST} above 0.9) were found between *L. trifolii* from onion and all the remaining hosts (all peppers and tomatillos). The pairwise genetic differences between *L. trifolii* from Oaxaca and Yucatán were high and statistically significant, and the range of F_{ST} varied from 0.16 to 0.55. Interestingly, leafminers associated to chile *maax'ik* exhibited high levels of genetic differences with the rest of the hosts, ranging from 0.31 to 0.95. Large genetic differences were also detected between chile *de agua* and chile *habanero* (0.41). The N_{ST} values were very similar to F_{ST} (Table 3).

The hierarchical analysis and AMOVA for *L. trifolii* reflected that 89% of the variation was explained by differences among host species, whereas 11% was explained by variation within host. A hierarchical analysis of landrace peppers (*C. annuum* and

Table 3. Pairwise F_{ST} below diagonal and N_{ST} above diagonal for cytochrome oxidase I between all pairs of *L. trifolii* samples associated to onion, tomatillo, and peppers.

Host	1	2	3	4	5	6	7	8
1. Onion	—	0.94	0.94	0.93	0.96	0.96	0.95	0.92
2. Chile de agua	0.94	—	0	0.22	0.39	0.41	0.55	0
3. Tomatillo	0.94	0	—	0.16	0.33	0.34	0.53	0
4. Chile dulce	0.93	0.22	0.16	—	0	0	0.31	0
5. Chile xcatik	0.96	0.39	0.33	0	—	0	0.63	0
6. Chile habanero	0.96	0.41	0.35	0	0	—	0.68	0
7. Chile maax'ik	0.95	0.55	0.53	0.31	0.63	0.68	—	0.34
8. Chile costeño	0.92	0	0	0	0	0	0.34	—

The values in bold represent $P < 0.05$.

Table 4. AMOVA results for cytochrome oxidase subunit I for *L. trifolii* samples (whole samples, host pepper-plant and locations)

Source of variation	df	Sum of squares	Variance components	Percentage of variation
Whole samples				
a) Among host species	2	301.02	8.77	89.44
b) Within host species	101	105.24	1.04	10.56
Host Pepper-plant				
c) Among pepper host	5	41.55	0.59	44.06
d) Within pepper host	78	61.77	0.79	55.94
Locations (Oax and Yuc)				
e) Among regions	1	20.40	0.24	25.68
f) Among locations within region	4	21.15	0.44	24.12
g) Within locations	78	61.77	0.79	50.20

The values in bold represent $P < 0.05$.

Table 5. List of haplotypes shared by Scheffer and Lewis (2005 and 2006), and the present study using a database of 492 base pairs

Haplotypes reported by Scheffer and Lewis (2005, 2006)	Host	Location (State)	Haplotype	Host	Location (State)
T4 (Trifolii-A)	Bean, melon, onion	California, Arizona, New York, USA	H3	Onion	Oax. Mex
T7 (Trifolii-W Pepper)	Pepper	Tampico, Méx	H11	Chile de agua, costeño y tomatillo	Oax. Mex
T15 (Trifolii-W Pepper)	Pepper	Tampico, Méx	H5	Chile de agua, costeño y tomatillo	Oax. Mex
S1 (Sativae-A)	Swept, Bean, Tomato, Cucumber	Florida USA, Guatemala, Honduras	H20, H21	Tomato, Bean	Oax. Mex Yuc. Mex
SX (Sativae-W) (S4,S5,S7,S8,S9, S11,S20,S27,S28)	Bean, melon, Tomato, Chrysanthemum among others	Asia, Africa, North and South America	H15,H17, H18,H19	Tomato	Oax. Mex

from Florida, Guatemala, and Honduras; and our samples included tomato and bean host from Oaxaca and Yucatán. This haplotype corresponds to the clade *L. sativa*-A. The SX (*L. sativae*-W) haplotype included samples from melon, bean, tomato, and among other hosts (see Scheffer and Lewis 2005), and were collected worldwide (Asia, Africa, North and South America; Table 5). We named SX haplotype because included 9 haplotypes from Scheffer and Lewis (2005). The SX haplotype also included our leafminer samples collected on tomato from west and central coast of Oaxaca.

Discussion

In our surveys for leafminers in Oaxaca and Yucatan, we collected two main species, *L. trifolii* and *L. sativae*. The former was found on peppers, tomatillo and onion, whereas *L. sativae* was found

on beans and tomatoes. The phylogenetic tree recovered each species of *Liriomyza* as a monophyletic group. Within *L. trifolii*, we did not find a specific clade associated with *C. annuum*, as Scheffer and Lewis (2006) reported. Instead, we recovered two clades: W, which was restricted to peppers and tomatillos, and A, which was restricted to onion. Strong HAD existed among the groups despite that some leafminer collections from different hosts (onion, tomatillo, and some of the pepper including chile de agua) occurred from the same farm (Teacolula, Paraje de Pedirillo; Table 1). Previous studies did not include tomatillo, probably because this species mainly grows in Mexico and Central America, although *L. trifolii* is known as an occasional pest of tomatillo in California. Both peppers and tomatillos are close relatives within the *Solanaceae* family and may share chemical and nutritional compounds attractive to *L. trifolii* and they usually occur in the

same agricultural fields. Interestingly, tomatoes also belong to *Solanaceae*; however, this species is phylogenetically further from tomatillo and peppers, and thus, the chemical and nutrient content of tomatoes is likely to be more different from peppers and tomatillo. This may explain why different species of *Liriomyza* attack specific *Solanaceae* species.

Our results based on the phylogenetic tree and haplotype network recovered the A clade for *L. trifolii*, which is 18 mutational steps away from the W group. These data suggest the presence of cryptic species and a process of specialization in different hosts (i.e., HAD) under sympatric conditions, since samples from the W and A clades were collected in the same agricultural fields. Scheffer and Lewis (2006) also detected two divergent groups within *L. trifolii* species. According to our combined database, we recovered three main haplotypes T4, T7, and T15. The haplotypes T7 and T15 included pepper samples and tomatillo: all of them were collected in Oaxaca and the pepper samples from Scheffer and Lewis (2006) were collected in Tampico, Mexico. Our samples from Yucatán corresponded to a haplotype that included peppers as well as nonpeppers (celery, potato, zucchini, *Coreopsis* among others) from Scheffer and Lewis (2006). However, this grouping may be artificial as the combined datasets included fewer base pairs than the dataset from previous studies. Nonetheless, it is still interesting that *L. trifolii* from peppers in Oaxaca and Tampico are part of the same haplotype and different than *L. trifolii* infesting peppers in Yucatán. We also recovered the haplotype T4 that matched our *L. trifolii* collected from onion with samples from Scheffer and Lewis (2006) collected from onion, melon, and bean. This haplotype corresponded to *L. trifolii*-A clade. Interestingly, in our study, beans were attacked only by *L. sativae*. Furthermore, we only found leafminers on onion host that correspond to the clade *L. trifolii*-A, in contrast to Scheffer and Lewis (2006) where haplotypes from onion were grouped in the clades *L. trifolii*-W and *L. trifolii*-A. This result suggests geographical variation in host preferences among leafminers, which may also impact HAD in this species. Multiple studies have shown that the traits important to species interactions may differ geographically among groups (Sword and Dopman 1999; Althoff and Thompson 2001).

The AMOVA for *L. trifolii* exhibited large genetic differentiation among host species (almost 90% of the variation was explained by differences among host species). This result suggests divergence by host use in sympatric conditions, supporting HAD between our two haplotype groups (A+W). The second AMOVA that included only leafminers associated with pepper also exhibited large levels of genetic differences among individual pepper landraces (44%). Additionally, the hierarchical AMOVA indicated that almost 26% of the variation was explained by differences between regions, but these differences were not statistically significant. However, the haplotype network indicated geographical structure of the haplotype distribution since only one haplotype was shared between Oaxaca and Yucatán, whereas eight haplotypes were private to Oaxaca and three haplotypes were private to Yucatán. These results suggest that isolation by distance may have influenced the genetic divergence found in leafminers infesting pepper varieties. However, our data likely do not have enough sampling to detect a statistically significant Mantel test. The importance of geographical isolation on genetic divergence has been demonstrated in *L. sativae* from populations collected in China (Du et al. 2016) and Brazil (Parish et al. 2017).

Our data also indicated that the pairwise genetic differences within each geographical region were mostly nonsignificant, suggesting that phenotypic divergence in pepper varieties (including

C. annuum and *C. chinense* landraces) themselves are not influencing the genetic structure of *L. trifolii*. However, samples from chile *maax'ik* were the most divergent from all other hosts, including other pepper varieties. The highest F_{ST} values were detected between chile *habanero* and chile *maax'ik*, collected from different backyards within the same locality. Moreover, there were also large and statistically significant genetic differences between chile *maax'ik* and chile *xcatik* leafminers, both from Yucatán with a distance of ~160 km between the two sites. The high levels of genetic divergence of chile *maax'ik* from other hosts appears to be due to their low frequencies of the H6 haplotype (the most common haplotype among the other hosts), as well as a high frequency of the H9 haplotype, which was relatively rare among the other peppers (chile *xcatik*). Also, the H8 haplotype was found only in leafminers from chile *maax'ik*. The high genetic divergence of leafminers collected from chile *maax'ik* might have been driven by host traits since chile *maax'ik* is likely a wild pepper not grown in the same frequency or intensity as other varieties. Therefore, the content of volatiles, nutrients, or secondary metabolites might be different in *maax'ik* compared with peppers landraces as consequence of artificial selection driven by humans. This result may suggest that domestication processes have impacted the genetic structure of *L. trifolii* with the possibility of adaptive divergence between leafminers associated with peppers. However, additional sampling is needed among other, more ancestral and wild *C. annuum* varieties. We also detected large genetic differences between chile *agua* (Oaxaca) and chile *habanero* (Yucatán), but the influence of geography and pepper species (*C. annuum* vs *C. chinense*, respectively) complicates inferences in our study.

The leafminers collected on tomatoes and beans corresponded to *L. sativae* species. Similar to *L. trifolii*, we detected two divergent groups (W and A) for *L. sativae*, based on the haplotype network. Group W included samples from tomatoes located on the central and western coast of Oaxaca (Huaxpaltepec and Rosedal locations, respectively). The A group included samples associated with beans from Yucatán (Dzidzantun location) and tomatoes from the easternmost coast of Oaxaca (Tortolita location). Unexpectedly, leafminers collected on tomatoes from Rosedal (group W) and Tortolita (group A) are separated by 29 mutational steps, despite a geographical distance of only 50 km between these locations and sharing a host. The deep genetic divergence between the W and A groups may suggest the presence of cryptic species within *L. sativae*. Similar to *L. trifolii*, our analyses based on the combined database recovered the SX haplotype for *L. sativae* (W clade) (Scheffer and Lewis 2005). We named this haplotype SX because it included samples from multiple haplotypes found by Scheffer and Lewis (2006). The main hosts were bean, melon, tomato, and *Chrysanthemum* and were collected worldwide. Our samples included in this haplotype were from tomatoes collected in Oaxaca west and central coast. We also recovered the S1 haplotype that is part of the divergent clade *L. sativae*-A. This haplotype was found on bean, tomato, and cucumber and were collected from Guatemala and Honduras (Scheffer and Lewis 2005). Interestingly, S1 was also included in our samples from bean and tomato located in Yucatán and the east coast of Oaxaca. These results indicate that haplotype S1 has a distribution mainly in Central America. Therefore, it is possible that SX has a worldwide distribution, whereas S1 exhibits a narrower distribution in south-eastern Mexico and Central America, and both haplotypes belong to two divergent genetic clades. Interestingly, even though these clades are extremely divergent, they shared the same plant host (tomato). Further studies are necessary to test the importance of geographical and ecological context on host preference divergence.

Other studies have also detected divergent genetic clades within *L. sativae*. For example, Parish et al. (2017) found a genetic clade of *L. sativae* in Brazil, which is separated by 36 mutational steps from the clade *L. sativae*-W. Within the *Liriomyza* genus, other species also include different evolutionary lineages. For example, for *L. huidobrensis*, two distinct evolutionary groups were detected, one group included flies from California and Hawaii and the second group included flies from South and Central America (Scheffer 2000; Scheffer and Lewis 2001). The name *L. langei* Frick was reinstated for the populations from North America (California and Hawaii) and the name *L. huidobrensis* was restricted to the South America clade (Scheffer and Lewis 2001). All these results, including this study, suggest that the presence of cryptic species is common within *Liriomyza*, driven mainly by geographical isolation and host specialization.

Our results did not show a clade specific to *C. annuum*, as other studies have proposed (Scheffer and Lewis 2006). Instead, leafminers from peppers and tomatillos were genetically undifferentiated. We did not detect significant genetic differences among leafminers associated with different types of *C. annuum* within the same geographic region (Oaxaca or Yucatán), but large and significant differences were found between the two regions, indicating that isolation by distance has played an important role in this divergence. Our results and the comparison with previous work suggest that host plant use, as well as geographical and ecological divergence, have influenced the genetic structure of *L. trifolii*. The divergent genetic lineages within *L. trifolii* and *L. sativae* support previous evidence for the presence of cryptic species within these polyphagous insects (Scheffer 2000, Scheffer and Lewis 2001, 2006). Our study has produced information on the geographical distribution of genetic variation of *L. trifolii* and *L. sativae* on different hosts that may help in management programs of these important pests. Our study also provides information on the identity of leafminer species attacking major crops in southeastern Mexico. Interestingly, there is no overlap on the species attacking the same host.

Acknowledgments

We extend thanks to Lev Jardón Barbolla, Ernesto González Gaona, Esaú Ruiz Sánchez, Luis Latournerie Moreno, Salvador Montes Hernandez, Jose Carrillo, Araceli Aguilar Melendez, Brian Pace, Rachel Capouya, and Nathan Taitano for providing information and assistance in collecting leafminer samples. J.P.-A. acknowledges the postdoctoral fellowship from Ohio State University Center of Applied Plant Sciences and for providing financial support to the project.

References Cited

- Abrahamson, W. G., M. D. Eubanks, C. P. Blair, and A. V. Whipple. 2001. Gall flies, inquilines, and goldenrods: A model for host-race formation and sympatric speciation. *Am. Zool.* 41: 928–938.
- Althoff, D. M., and J. N. Thompson. 2001. Geographic structure in the searching behaviour of a specialist parasitoid: Combining molecular and behavioural approaches. *J. Evol. Biol.* 14: 406–417.
- Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman. 1990. Basic local alignment search tool. *J. Mol. Biol.* 215: 403–410.
- Alvarez, N., M. Hossaert-McKey, G. Restoux, A. Delgado-Salinas, and B. Benrey. 2007. Anthropogenic effects on population genetics of phytophagous insects associated with domesticated plants. *Evolution*. 61: 2986–2996.
- Antwi, J. B., G. A. Sword, and R. F. Medina. 2015. Host-associated differentiation in a highly polyphagous, sexually reproducing insect herbivore. *Ecol. Evol.* 5: 2533–2543.
- Barman, A. K., M. N. Parajulee, C. G. Sansone, C. P. C. Suh, and R. F. Medina. 2012. Geographic pattern of host-associated differentiation in the cotton fleahopper, *Pseudatomoscelis seriatus*. *Entomol. Exp. Appl.* 143: 31–41.
- Benrey, B., A. Callejas, L. Rios, K. Oyama, and R. F. Denno. 1998. The effects of domestication of *Brassica* and *Phaseolus* on the interaction between phytophagous insects and parasitoids. *Biol. Control*. 11: 130–140.
- Bouckaert, R., J. Heled, D. Kühnert, T. Vaughan, C. H. Wu, D. Xie, M. A. Suchard, A. Rambaut, and A. J. Drummond. 2014. BEAST 2: A software platform for Bayesian evolutionary analysis. *PLoS Comput. Biol.* 10: e1003537.
- Bush, G. L. 1969. Sympatric host race formation and speciation in frugivorous flies of the genus *Rhagoletis* (Diptera, Tephritidae). *Evolution*. 23: 237–251.
- Capinera, J. L. 2001. American serpentine leafminer, *Liriomyza trifolii* (Burgess) (Insecta: Diptera: Agromyzidae). IFAS Extension. Institute of Food and Agricultural Sciences, University of Florida, pp. 1–5.
- Chen, Y. H., R. Gols, and B. Benrey. 2015. Crop domestication and its impact on naturally selected trophic interactions. *Annu. Rev. Entomol.* 60: 35–58.
- Conord, C., G. Lempérière, P. Taberlet, and L. Després. 2006. Genetic structure of the forest pest *Hylobius abietis* on conifer plantations at different spatial scales in Europe. *Heredity* (Edinb.). 97: 46–55.
- Corander, J., P. Marttinen, and S. Mäntyniemi. 2006. Bayesian identification of stock mixtures from molecular marker data. *Fish. Bull.* 104: 550–558.
- Corander, J., P. Marttinen, J. Sirén, and J. Tang. 2008. Enhanced Bayesian modelling in BAPS software for learning genetic structures of populations. *BMC Bioinformatics*. 9: 539.
- Darriba, D., G. L. Taboada, R. Doallo, and D. Posada. 2012. jModelTest 2: More models, new heuristics and parallel computing. *Nat. Methods*. 9: 772.
- Dray, S., and A. B. Dufour. 2007. The ade4 package: Implementing the duality diagram for ecologists. *J. Stat. Softw.* 224: 1–20.
- Du, Y. Z., X. T. Tang, L. P. Wang, Y. Shen, and Y. W. Chang. 2016. Genetic differentiation of geographical populations of *Liriomyza sativae* (Diptera: Agromyzidae) in China based on mitochondrial COI gene sequences. *Mitochondrial DNA A. DNA Mapp. Seq. Anal.* 27: 3936–3940.
- Edgar, R. C. 2004. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32: 1792–1797.
- Ersts, P. J. 2018. [Internet] Geographic distance matrix generator (version 1.2.3). American Museum of Natural History, Center for Biodiversity and Conservation. Retrieved from http://biodiversityinformatics.amnh.org/open_source/gdmg
- Excoffier, L., and H. E. Lischer. 2010. Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Mol. Ecol. Resour.* 10: 564–567.
- Forbes, A. A., S. N. Devine, A. C. Hippee, E. S. Tvedte, A. K. G. Ward, H. A. Widmayer, and C. J. Wilson. 2017. Revisiting the particular role of host shifts in initiating insect speciation. *Evolution*. 71: 1126–1137.
- Fu, Y.-X. 1997. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics*. 147: 915–925.
- Gols, R., T. Bukovinsky, N. M. van Dam, M. Dicke, J. M. Bullock, and J. A. Harvey. 2008. Performance of generalist and specialist herbivores and their endoparasitoids differs on cultivated and wild *Brassica* populations. *J. Chem. Ecol.* 34: 132–143.
- Guindon, S., and O. Gascuel. 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst. Biol.* 52: 696–704.
- Guttman, S. I., T. K. Wood, and A. A. Karlin. 1981. Genetic differentiation along host plant lines in the sympatric *Enchenopa binotata* Say complex (Homoptera: Membracidae). *Evolution*. 35: 205–217.
- Leigh, J. W., and D. Bryant. 2015. PopART: Full-feature software for haplotype network construction. *Methods Ecol. Evol.* 6: 1110–1116.
- Librado, P., and J. Rozas. 2009. DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*. 25: 1451–1452.
- Lindig-Cisneros, R., B. Benrey, and F. J. Espinosa-García. 1997. Phytoalexins, resistance traits, and domestication status in *Phaseolus coccineus* and *Phaseolus lunatus*. *J. Chem. Ecol.* 23: 1997–2011.
- Mantel, N. 1967. The detection of disease clustering and a generalized regression approach. *Cancer Res.* 27: 209–220.

- Medina, R. F., S. M. Reyna, and J. S. Bernal. 2012. Population genetic structure of a specialist leafhopper on *Zea*: Likely anthropogenic and ecological determinants of gene flow. *Entomol. Exp. Appl.* 142: 223–235.
- Meyer, R. S., and M. D. Purugganan. 2013. Evolution of crop species: Genetics of domestication and diversification. *Nat. Rev. Genet.* 14: 840–852.
- Minkenberg, O. P. J. M. 1988. Dispersal of *Liriomyza trifolii*. *Bull. OEPP/EPPO*. 18: 173–182.
- Minkenberg, O. P. J. M., and J. C. Van Lenteren. 1986. The leafminers *Liriomyza bryoniae* and *L. trifolii*, their parasites and host plants: a review. Agricultural University Wageningen, Netherlands. Papers no. 86-2.
- Nason, J. D., S. B. Heard, and F. R. Williams. 2002. Host-associated genetic differentiation in the goldenrod elliptical-gall moth, *Gnorimoschema gallaesolidaginis* (Lepidoptera: Gelechiidae). *Evolution*. 56: 1475–1488.
- Parish, J. B., G. A. Carvalho, R. S. Ramos, E. A. Queiroz, M. C. Picanço, R. N. C. Guedes, and A. S. Corrêa. 2017. Host range and genetic strains of leafminer flies (Diptera: Agromyzidae) in eastern Brazil reveal a new divergent clade of *Liriomyza sativae*. *Agric. For. Entomol.* 19: 235–244.
- Pons, O., and R. J. Petit. 1996. Measuring and testing genetic differentiation with ordered versus unordered alleles. *Genetics*. 144: 1237–1245.
- Ramirez-Romero, R., D. Garibay-Benítez, O. Vargas-Ponce, A. Joyce, and J. S. Bernal. 2017. Do assortative mating and immigrant inviability help maintain population genetic structuring of an herbivore on a crop and a wild relative? *J. Insect. Sci.* 1–14.
- Reitz, S. R., and J. T. Trumble. 2002. Interspecific and intraspecific differences in two *Liriomyza* leafminer species in California. *Entomol. Exp. Appl.* 102: 101–113.
- Ruiz-Montoya, L., J. Núñez-Farfán, and J. Vargas. 2003. Host-associated genetic structure of Mexican populations of the cabbage aphid *Brevicoryne brassicae* L. (Homoptera: Aphididae). *Heredity (Edinb.)*. 91: 415–421.
- Rundle, H. D., and P. Nosil. 2005. Ecological speciation. *Ecol. Lett.* 8: 336–352.
- Scheffer, S. J. 2000. Molecular evidence of cryptic species within the *Liriomyza huidobrensis* (Diptera: Agromyzidae). *J. Econ. Entomol.* 93: 1146–1151.
- Scheffer, S. J., and M. L. Lewis. 2001. Two nuclear genes confirm mitochondrial evidence of cryptic species within *Liriomyza huidobrensis* (Diptera: Agromyzidae). *Ann. Entomol. Soc. Am.* 94: 648–653.
- Scheffer, S. J., and M. L. Lewis. 2005. Mitochondrial phylogeography of vegetable pest *Liriomyza sativae* (Diptera: Agromyzidae): Divergent clades and invasive populations. *Ann. Entomol. Soc. Am.* 98: 181–186.
- Scheffer, S. J., and M. L. Lewis. 2006. Mitochondrial phylogeography of the vegetable pest *Liriomyza trifolii* (Diptera: Agromyzidae): Diverged clades and invasive populations. *Ann. Entomol. Soc. Am.* 99: 991–998.
- Scheffer, S. J., M. L. Lewis, and R. C. Joshi. 2006. DNA barcoding applied to invasive leafminers (Diptera: Agromyzidae) in the Philippines. *Ann. Entomol. Soc. Am.* 99: 204–210.
- Sword, G. A., and E. B. Dopman. 1999. Developmental specialization and geographic structure of host plant use in a polyphagous grasshopper, *Schistocerca emarginata* (=Lineata) (Orthoptera: Acrididae). *Oecologia*. 120: 437–445.
- Taitano, N., V. Bernau, L. Jardón-Barbolla, B. Leckie, M. Mazourek, K. Mercer, L. McHale, A. Michel, D. Baumler, M. Kantar, et al. 2018. Genome-wide genotyping of a novel Mexican Chile Pepper collection illuminates the history of landrace differentiation after *Capsicum annuum* L. domestication. *Evol. Appl.* 1–15.
- Tajima, F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*. 123: 585–595.
- Via, S. 1991. The genetic structure of host plant adaptation in a spatial patchwork: Demographic variability among reciprocally transplanted pea aphid clones. *Evolution*. 45: 827–852.
- Wright, S. 1951. The genetical structure of populations. *Ann. Eugen.* 15: 323–354.