Ecological Entomology

Ecological Entomology (2021), DOI: 10.1111/een.13101

Cuticular hydrocarbon variation among *Rhagoletis* fruit flies (Diptera: Tephritidae): implications for premating reproductive isolation and ecological speciation

GLEN RAY HOOD,[1] JACKSON H. JENNINGS,[2] DANIEL J. BRUZZESE,[3] MELANIE BEEHLER,[2] THOMAS SCHMITT,[4] JEFFREY L. FEDER[3] and WILLIAM J. ETGES[2] [1]Department of Biological Sciences, Wayne State University, Detroit, Michigan, U.S.A, [2]Ecology, Evolution and Organismal Biology, Department of Biological Sciences, University of Arkansas, Fayetteville, Arkansas, U.S.A, [3]Department of Biological Sciences, University of Notre Dame, Notre Dame, Indiana, U.S.A. and [4]Department of Animal Ecology and Tropical Biology, Biocentre, University of Würzburg, Würzburg, Germany

Abstract. 1. Fruit flies in the genus *Rhagoletis* are a model for ecological speciation via sympatric host plant shifting. *Rhagoletis* mate on or near the fruit of their respective host plants, generating premating reproductive isolation between taxa specialised on different hosts. However, non-host-related premating isolation has been observed between some *Rhagoletis* species associated with morphological differences in body colour and wing patterns.

2. Here, the extent of epicuticular hydrocarbon (CHC) variation as a possible, additional determinant of mate choice in adults of six *Rhagoletis* taxa, including the apple and hawthorn-infesting host races of *R. pomonella* is investigated.

3. Gas-chromatography-mass-spectrometry revealed 36 repeatable and quantifiable hydrocarbon components present on the epicuticles of all six fly taxa, comprised of at least 53 different CHC compounds, with chain lengths varying from 27 to 34 carbon atoms, consisting of n-alkanes, mono-, dimethyl-, and trimethyl-alkanes, alkenes, and alkadienes. There were significant differences in the relative proportions of CHCs between adult *R. cingulata*, *R. cornivora*, *R. zephyria*, *R. mendax*, and *R. pomonella*, as well as between the apple- and hawthorn-infesting host races of *R. pomonella*. Furthermore, within the *R. pomonella* host races and *R. mendax*, significant CHC differences were observed between the sexes and across collecting sites.

4. The results are consistent with variation in CHCs potentially contributing to observed patterns of premating isolation between *Rhagoletis* taxa, possibly due to combination of sexual and host-related selection, which will necessitate further, in depth chemical analyses and future mating trials to substantiate.

Key words. Apple maggot fly, courtship, host race, mate choice, pheromone.

Introduction

During the early stages of divergence, local populations can evolve differences in mate recognition phenotypes and courtship behaviours, leading to increased levels of assortative mating and reproductive isolation, potentially contributing to speciation (Sobel *et al.*, 2010). In many cases, variation in mate recognition is multi-modal, including behavioural, chemical, auditory, visual, and time series-dependent signalling between the sexes that results in acceptance/rejection of prospective mates during courtship (Gerhardt & Huber, 2002; Greenfield, 2002; Taylor *et al.*, 2011; Gleason *et al.*, 2012; Giglio & Dyer, 2013). In insects that mate on or near their oviposition sites, host-finding or 'settling' behaviour (Jaenike, 1986) can also result in differences in encountering prospective mates

Correspondence: Glen Ray Hood, Department of Biological Sciences, Wayne State University, Detroit, MI 48202, U.S.A. E-mail: glenrayhood@wayne.edu

and can influence premating isolation and drive reproductive divergence (Feder *et al.*, 1994; Feder, 1998; Berlocher & Feder, 2002; Dres & Mallet, 2002). Thus, deciphering the roles and relative importance that different components of the mating system have for sexual isolation and speciation requires an integrative understanding of not only the nature of behaviours contributing to successful mating, but the biological and ecological contexts of where, when, and how they are expressed (Etges, 2002, 2014; Mullen & Shaw, 2014; Hood *et al.*, 2019; Zhang *et al.*, 2020).

Here, we explore variation in chemical signalling, an overlooked and potentially important component of the mating system, which may contribute to premating isolation in fruit flies in the *Rhagoletis pomonella* sibling species complex (Diptera: Tephritidae), a model for studying contemporary evolution via host plant shifting in phytophagous insects (Bush, 1969, 1993; Feder *et al.*, 2012). In particular, the recent shift of *R. pomonella* from its native host plant, downy hawthorn (*Crataegus* spp.), to introduced domesticated apple (*Malus domestica*) sometime in the mid-1800s is an example of host race formation in action (Funk *et al.*, 2002; Jiggins & Birdle, 2004), a hypothesised early stage of ecological speciation (Berlocher & Feder, 2002; Dres & Mallet, 2002). The apple- and hawthorn-infesting host races of *R. pomonella* are members of a closely related group of sibling species that each infest the fruit of different host plant species: *R. zephyria* (host: snowberry, *Symphoricarpos* spp.), *R. mendax* (host: blueberry, *Vaccinium* spp.), *R. cornivora* (host: silky dogwood, *Cornus amomum*), and the undescribed flowering dogwood fly (host: *Cornus florida*) (Berlocher, 2002; Powell *et al.*, 2013). In addition, several other taxa infesting the fruit of non-overlapping host plants such as the eastern cherry fruit fly, *R. cingulata* (host: black cherry; *Prunus serotina*), belong to different species groups in the genus and are sympatric with *R. pomonella* group flies in North America. These distinct host plant affiliations and broadly overlapping ranges led Bush (1966) to propose that many flies in the genus *Rhagoletis* speciated ecologically by sympatric host plant shifting.

Previous studies have shown that *Rhagoletis* mate only on or near the fruit of their respective host plants (Prokopy *et al.*, 1971, 1972). Thus, differences in host preference among flies generates strong premating isolation, facilitating their ecological divergence in sympatry (Bush, 1969; Feder *et al.*, 1994; Linn *et al.*, 2003, 2012; Forbes *et al.*, 2005; Powell *et al.*, 2012; Sim *et al.*, 2012). Additionally, following adult emergence and after flies alight on fruit, *Rhagoletis* can engage in courtship displays and prolonged tactile interactions prior to and during copulation. These displays and interactions suggest that, following host choice, additional mate choice behaviours may act to increase premating isolation between taxa (Bush, 1966; Alonso-Pimentel *et al.*, 2000). In this regard, mate choice in *Rhagoletis* has been shown to correlate with the degree to which flies differ in their body coloration, striping, and wing patterns, with strong isolation observed between morphologically distinct species and lesser degrees between morphologically similar taxa (Smith, 1986; Hood *et al.*, 2012). For example, complete premating isolation was observed between seven pairings of morphologically distinct *Rhagoletis* species in laboratory mate choice tests in cages devoid of host plant cues. However, premating isolation was not complete between the morphologically similar sibling species *R. mendax* and *R. zephyria* (Schwarz & McPheron, 2007) or between western (*R. indifferens*) and eastern (*R. cingulata*) cherry-infesting fruit flies (Hood *et al.*, 2012). While Hood *et al.* (2012) conclude that the observed differences in premating isolation among taxa may be due to differences in chemical signalling, their experiments were not designed to explicitly test that hypothesis.

Compared to their well-characterised morphology (Bush, 1966; Smith & Bush, 2001), little is known about whether *Rhagoletis* vary chemically in their epicuticular hydrocarbon (CHC) profiles, and if variation does exist, whether these differences might affect sexual selection via mate choice. Here, we assess CHC variation among *Rhagoletis* taxa with the aim of eventually determining the role that chemical communication may play between these flies during their courtship and mating. Epicuticular hydrocarbons, composed of alkanes, methyl-branched alkanes, alkenes, alkadienes, and alkatrienes, are known to serve as contact pheromones in many species of arthropods, including mites (Margolies & Collins, 1994), and several species of insects (reviewed in Blomquist & Bagnères, 2010; Howard & Blomquist, 2005), such as walking sticks (Riesch *et al.*, 2017), house flies (Reed *et al.*, 1996), ants (Sharma *et al.*, 2015), crickets (Mullen *et al.*, 2007), seaweed flies (Berdan *et al.*, 2019), and *Drosophila* (Blomquist *et al.*, 1985; Jackson & Bartelt, 1986; Jallon & David, 1987; Toolson *et al.*, 1990; Etges & Jackson, 2001; Blows & Higgie, 2002; Ferveur, 2005; Yew *et al.*, 2011; Jennings *et al.*, 2014). We therefore contend that CHCs may also serve as contact pheromones and contribute to courtship and copulation success in *Rhagoletis*, potentially accentuating premating isolation between taxa generated by host specificity and differences in morphology. Thus, the goal of this study was to characterise CHC variation and determine if differences exist among several species of North American *Rhagoletis*, including the recently formed apple- and hawthorn-infesting host races of *R. pomonella*, using gas chromatography–mass spectrometry (GC–MS). The discovery of CHC differences among these taxa and/or between sexes would springboard studies to test for a role of epicuticular hydrocarbons in mate choice and species recognition, and determine whether the variation observed is genetically based and/or environmentally influenced.

Methods

Taxonomy of flies

Five different species of *Rhagoletis* were analysed for CHCs (Table 1). Four of the species (*R. pomonella*, *R. mendax*, *R. zephyria*, and *R. cornivora*) are members of the *R. pomonella* sibling species group that each infests the fruit of a different host plant species (Bush, 1966; Berlocher *et al.*, 1993). Within the species *R. pomonella*, we also analysed the ancestral hawthorn- (*Crataegus mollis*) and recently sympatrically derived apple-infesting host races of the fly (Feder *et al.*, 1988; McPheron *et al.*, 1988). *Rhagoletis pomonella* group flies, including the hawthorn and apple host races, are all morphologically similar, although *R. pomonella* is generally larger in

Table 1. Numbers of adult *Rhagoletis* of each taxon, population and host-association that were assayed for CHC variation. Numbers in parentheses following sites in the 'Sample locations' column correspond to site numbers shown in Fig. 1, while those following semicolons reflect the numbers of female and male flies analysed at each site. The general sampling dates of each taxa at each sampling location are also given. Due to the patchiness of suitable host plants and low infestation rates, sample sizes for *R. cornivora* are relatively low.

Species/host race	Host plant species	Common host plant name	Females	Males	Sample locations
R. cingulata	*Prunus serotina*	Black cherry	15	15	Urbana, IL: early-Aug (1; 2, 4); Granger, IN: late Aug (3; 10, 10); Fennville, MI: late Aug (6; 3, 1)
R. cornivora	*Cornus amomum*	Silky dogwood	1	4	East Lansing, MI: late July (8; 1, 4)
R. mendax	*Vaccinium corymbosum*	Blueberry	33	28	Sawyer, MI: late July (4; 10, 10); Fennville, MI: late July (6; 14, 12); Otis Lake, MI: late July (7; 9, 6)
R. pomonella	*Malus domestica*	Apple	71	35	South Bend, IN: mid-Aug (2; 7, 4); Fennville, MI: mid-Aug (6; 16, 4); East Lansing, MI: mid-Aug (8; 22, 25); Grant, MI: mid-Aug (9; 26, 2)
R. pomonella	*Crataegus mollis*	Hawthorn	67	52	South Bend, IN: early-Sept (2; 42, 34); Cassopolis, MI: early-Sept (5; 20, 15); Fennville, MI: early-Sept (6; 0, 1); Grant, MI: early-Sept (9; 5, 2)
R. zephyria	*Symphoricarpos occidentalis*	Snowberry	9	8	E. Lansing, MI: early-Aug (8; 6, 8); Mt. Pleasant, MI: late-Aug (10; 3, 0)

body size than the other species (Yee *et al.*, 2009). Sexual dimorphism is also generally limited for *R. pomonella* group flies, with two notable exceptions: adult females are typically ∼10% larger than adult males, and females possess four white pollinose bands on their black abdomens, one each along the posterior margins of tergites II–V, while males have three bands, one each on the posterior margins of tergites II–IV (Bush, 1966).

The fifth species in the study, *R. cingulata*, belongs to the *R. cingulata* sibling species group of cherry- and olive-infesting flies (Bush, 1966; Doellman *et al.*, 2019, 2020). The *R. cingulata* group is not the immediate sister group to *R. pomonella* (Hulbert *et al.*, 2018). As a result, *R. cingulata* is morphologically distinct from the other *R. pomonella* group flies analysed in the study, differing in body coloration, striping, and wing banding pattern. Moreover, while *R. cingulata* is sympatric with *R. pomonella* group flies, it displays complete premating isolation from *R. pomonella* in mating trials (Hood *et al.*, 2012).

Sources of flies

Larval-infested fruits were collected from 10 sites in the Midwestern United States (Fig. 1; Table 1), following the methods described in Hood *et al.* (2015). Fruits were sampled directly from plants or the ground below plants in the late summer and early fall of 2012 to 2015 (see Table 1 for collection dates), transported to a greenhouse at the University of Notre Dame, Notre Dame, Indiana, U.S.A., and placed on wire mesh racks in plastic collecting trays. The trays were monitored twice daily for newly emerging larvae. All larvae were allowed to pupariate and were then placed in Petri dishes filled with moist vermiculite and exposed to a pre-winter period of 7–10 days at 21 °C and a 14:10 L:D photoperiod. After this time, Petri dishes were placed in a refrigerator at 4 °C for 4 months to simulate winter conditions. Following the overwintering period, Petri dishes were removed from the refrigerator and placed in 15 cm^3 Plexiglass cages stocked with a mixture of honey and brewer's yeast as a standard food source and water. Cages were kept inside a 24 °C, 14:10 L:D incubator. Upon eclosion, adults were aged to sexual maturity (7–10 days; Prokopy *et al.*, 1972), and then placed in micro-centrifuge tubes and preserved at −80 °C. By using sexually mature adults of similar ages, we attempted to minimise potential age-related variation in the quantities and/or composition of CHCs in the study. All samples were shipped on dry ice to the University of Arkansas, Fayetteville, Arkansas, U.S.A., for CHC analysis (see below). Details on the numbers of adult females and males analysed for CHC variation from each site are listed in Table 1.

Cuticular hydrocarbon analysis

For each host-associated population, we characterised variation in the CHC profiles of *Rhagoletis*, consisting of 36 peaks identified by gas chromatography (Table 2; Fig. 2). Extracts of adult CHCs were obtained by immersing each individual fly in HPLC grade hexane for 10 min (mixed by vortexing once after 5 min) in a 300 µl glass vial insert (Microliter Analytical Supplies, Suwanee, Georgia), and subsequently evaporating all hexane at room temperature. All extracts not used immediately for analysis were frozen at −20 °C. Each CHC extract was then re-dissolved in 5 µl of heptane containing 342 ng of C_{22} (docosane) as an internal standard. The use of this internal standard allowed us to check sample quality during each run and calculate total amounts of CHC under each peak analysed in the final extract of each sample, which were then used for statistical analysis (see below). Individual CHC extracts of one 1 µl injections were analysed by capillary gas–liquid chromatography, using an automated Shimadzu GC-17A (Shimadzu Scientific Instruments, Columbia, Maryland) fitted with a flame ion detector and a 15 m (ID = 0.22 mm) Rtx-5 fused-silica column (Restek Corporation, Bellefonte, Pennsylvania). Injector and

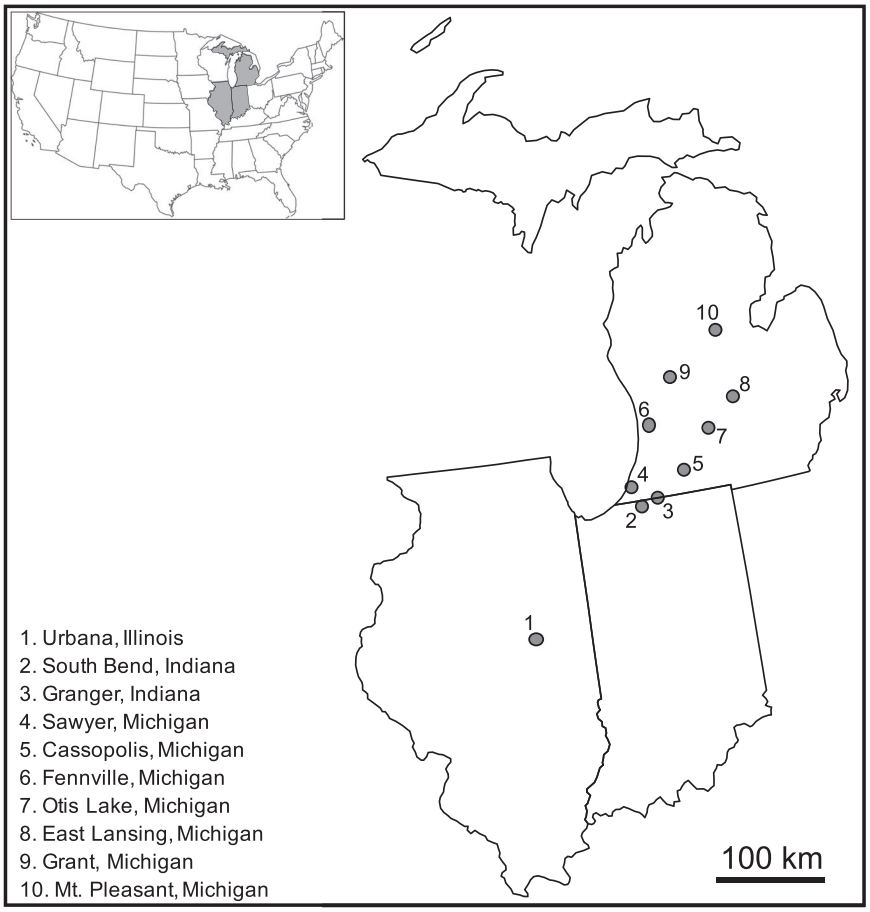

Fig 1. Map of the 10 collecting sites for *Rhagoletis* fruit flies sampled in the Midwestern states of Illinois, Indiana, and Michigan, U.S.A., analysed in the study. See Table 1 for information on taxa collected and sample sizes, including the numbers of each sex analysed at each site.

Table 2. Cuticular hydrocarbons observed for the six taxa of *Rhagoletis* analysed in the study, including the apple- and hawthorn-infesting host races of *R. pomonella*. All peaks were present and quantifiable in all samples, although each peak does not necessarily represent only one compound (e.g., peak eight contain multiple compounds). Peak number refers to the labelled peaks in Fig. 2. Equivalent chain lengths (ECL) were calculated as in Stennett and Etges (1997) and unidentified/unknown chemical structures are labelled with a 'C-' followed by their ECL value.

Peak number	ECL	Hydrocarbon(s)	Peak number	ECL	Hydrocarbon(s)
1	26.62	2-MeC26	19	30.00	C30 (*n*-triacontane)
2	27.00	C27 (*n*-heptacosane)	20	30.35	12-MeC30
3	27.34	13-;11-; 9-C27	21	30.63	C31dien
4	27.51	5-MeC27	22	30.66	2-MeC30
5	27.61	11,13-diMeC27	23	30.84	C31en
6	27.74	3-MeC27; 5,x-diMeC27	24	31.00	C31 (*n*-hentriacontane)
7	28.00	C28 (*n*-octacosane); 3,13-diMeC27	25	31.05	3,x-diMeC31
8	28.33	14-; 13-; 12-; 10-; 8-MeC28	26	31.31	15-; 13-; 11-; 9-MeC31
9	28.50	6-MeC28	27	31.54	13,17-diMeC31
10	28.82	2-MeC28	28	31.59	11,19-diMeC31
11	28.87	C29en; C29en	29	31.69	C-31.69
12	29.00	C29 (*n*-nonacosane)	30	32.33	C-32.33
13	29.35	15-; 13-; 11-; 9-; 7-MeC29	31	32.54	C33dien
14	29.53	11,15-; 11,17-diMeC29	32	32.58	C-32.58
15	29.58	9,13-diMeC29	33	32.73	C-32.73
16	29.63	9,19-diMeC29	34	33.31	C-33.31
17	29.75	3-MeC29	35	33.46	C-33.46
18	29.81	5,x-diMeC29	36	33.59	C-33.59

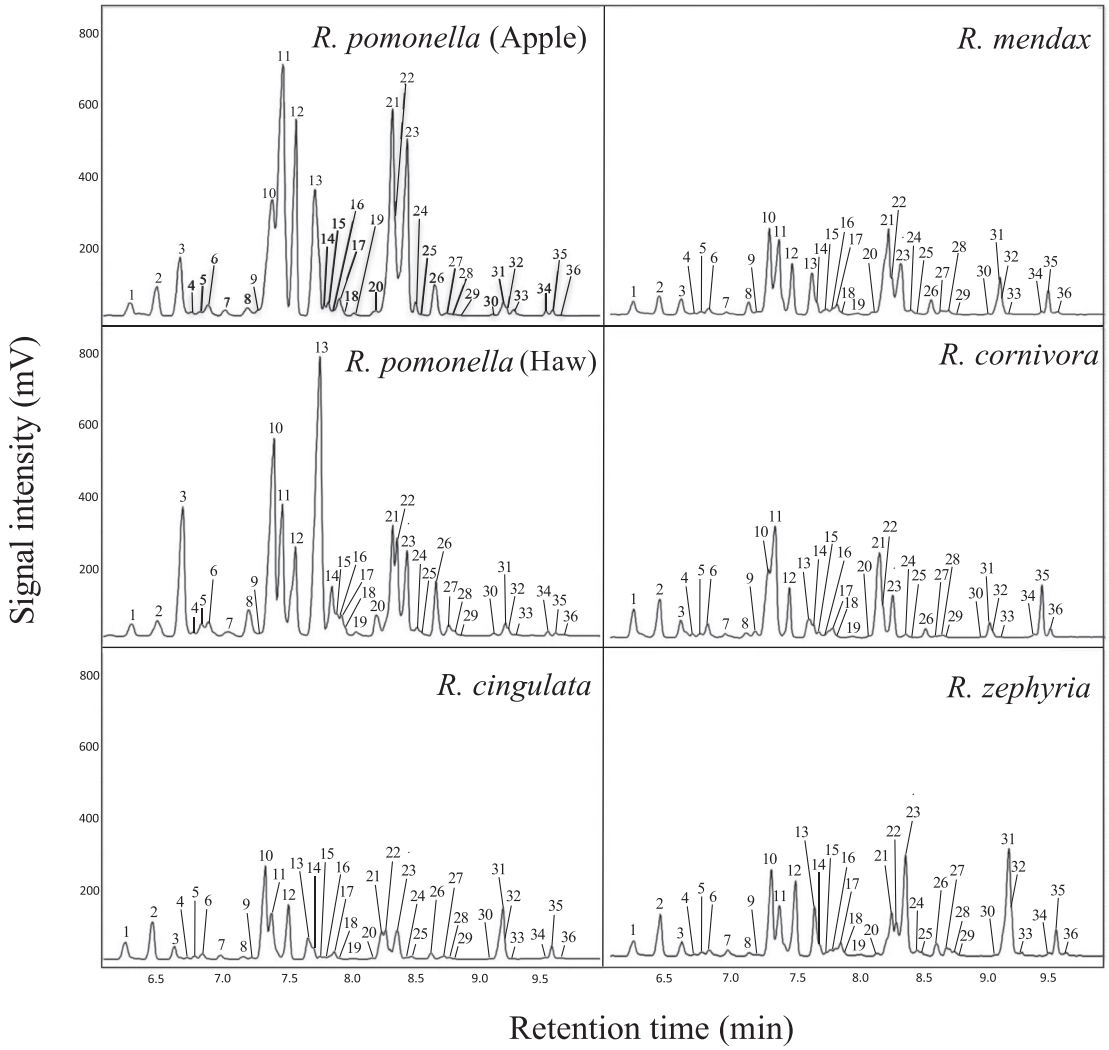

Fig 2. Chromatograms showing representative CHC profiles for the six taxa of *Rhagoletis* analysed in the study, including the apple- and hawthorn-infesting host races of *R. pomonella*. Numbered peaks in each panel correspond to peak numbers and CHC components in Table 2.

detector temperatures were set at 290 °C and 345 °C, respectively, with the injector port in split mode (3:1 ratio), with the column heated from 200 °C to 345 °C at 15 °C min^{-1} holding at 345 °C for 4 min.

Equivalent chain lengths (ECLs) of all observed CHC peaks were determined by sample co-injection with gulf wax and known internal standards including C_{30}, and C_{38} as described in Stennett and Etges (1997). We quantified the amount of CHCs under each of the 36 peaks in all flies by analysis of peak integrations using Class VP 4.2 software provided by Shimadzu, with calculated C_{22} amounts as an internal standard and expressed as nanograms/fly.

While the GC analysis described above allowed for the quantification of 36 measurable and repeatable CHC peak components in all flies, this resolution did not allow for chemical identification of CHCs or the determination of whether multiple compounds might constitute individual peaks. We therefore carried out structural characterisation and identification of CHC components from pooled adult *R. pomonella* samples using GC–MS on an Agilent 7820A GC system equipped with a 5975 Mass Selective Detector (Agilent Technologies, Inc., Santa Clara, California) and a HP-5 ms column [(5%-Phenyl)-methylpolysiloxane, 30 m length, 250 μm ID, 0.25 μm film thickness; Agilent Technologies, Inc.]. Electron ionisation (EI) energy was set at 70 eV. One microliter of the sample was injected in splitless mode and analysed with constant helium flow at 1 ml min^{-1}. The column was set at 40 °C for 3 min, increased to 200 °C at a rate of 35 °C min^{-1}, then increased to 345 °C at a rate of 15 °C min^{-1} for 4 min. The MS was set to detect from *m/z* 33 to 500. Chromatograms and spectra were analysed using MSD

ChemStation (Agilent Technologies, Inc.). Each CHC component was identified on the basis of retention time and EI fragmentation patterns.

Statistical analyses

As is common in studies of CHC variation (Schwander *et al.*, 2013), we used proportional abundances of CHCs rather than absolute values for statistical analyses, thereby reducing experimental error among individuals and standardising for differences stemming from variation in insect body size (Blows & Allan, 1998; Rundle *et al.*, 2005). We calculated CHC proportions by dividing the amount of each CHC peak (i.e., nanograms per sample, based on the C_{22} internal standard) in a single sample by the sum of all peaks for that same sample. These CHC proportions were then log-contrast transformed to remove non-independence among variables (Blows & Allan, 1998). To obtain these values, we first divided the relative values for each of the 36 peaks individually by that of peak 6 (Table 2) and then $\log_{10}$ transformed these new variables. We selected peak 6 because a principal components analysis (PCA) of absolute CHC abundances for all six taxa indicated that this component had a high loading on PC1, which explained almost 65% of the variation, but relatively low loadings compared to the other 35 peaks on additional PCs 2–8, which collectively accounted for ∼27% of the variation. All CHC components (peaks) had appreciable positive loadings on PC1 in a PCA of absolute levels (see Table S1), implying that PC1 largely reflected differences in adult body size, not differences due to host-association. Additionally, the statistical results obtained from log-contrast values derived by dividing relative abundances by other peak values were quantitatively similar to those for peak 6 (see Table S2) and, thus, we restricted the presentation of our findings to those generated using peak 6 as a denominator. These and all other statistical analyses were performed in R ver. 3.3.2 (R Core Team, 2019).

Tests for significant variation in CHC values were conducted by implementing a series of multivariate analyses of variance (MANOVA). In each MANOVA, to further reduce data dimensionality and to account for multicollinearity of CHC peaks, we used a set of the principal components of a PCA on log-contrast transformed relative CHC abundances (Mullen *et al.*, 2007; Stamps & Shaw, 2019; Butterworth *et al.*, 2020; Davis *et al.*, 2021) that collectively explained 90% of the total variation or used one less PC than the number of samples in the dataset (see Table S3 for the number of and percent variation explained by the PCs used in each MANOVA). In each of these analyses described below, the factors 'species/host', 'sex', and 'collecting site' were designated as fixed effects. In addition, we also calculated Euclidean distances between centroids from the MDS output described below, focusing here on comparisons involving differences between species/host, and sex (Table S4).

We performed MANOVA at three different levels of population organisation. At the first level of population organisation, we tested for significant CHC variation among the five described species *R. pomonella*, *R. mendax*, *R. zephyria*, *R. cornivora*, and *R. cingulata* (Bush, 1966), pooling the apple and hawthorn host races together to constitute the *R. pomonella* sample, to determine whether the five species could be distinguished as groupings. In an additional MANOVA, we excluded *R. cornivora* due to a lack of males in the sample to determine the degree to which species level divergence was related to any sex-associated differences. In these analyses, collecting site was excluded as a factor as not all taxa were collected from multiple sites. At the second level of population organisation, sex-related differentiation and geographic divergence across collecting sites were investigated in for the *R. pomonella* host races only. Using the full *R. pomonella* dataset, pooling individuals across sites, we investigated the effects of host and sex on CHC profiles. To explore the effects of collecting site and sex, we performed two additional MANOVA, one each for hawthorn- and apple-infesting *R. pomonella*, with reduced datasets including only those locations where both sexes were appropriately sampled (Table 1). At the third level of population organisation, we explored intraspecific variation (i.e., between sexes and when sample size permits, among sites) in CHC profiles within each of three remaining fly taxa including the factors sex and collecting site for *R. mendax*, and collecting site only for *R. zephyria*, and *R. cingulata*. We adopted this multi-level strategy due to unevenness in the sampling scheme in which representatives of both sexes from each host association were not collected at all sites (see Table 1; Fig. 3).

Using a similar approach to a number of recent CHC studies (e.g., Moore *et al.*, 2021), we implemented a machine learning technique known as a neural net classifier (NNC) to further explore differences between species and among sexes for the above-mentioned levels of biological organisation. In these analyses, our objective was to determine if CHC profiles could accurately predict species and sex identity. To accomplish this goal, we first split our PCA dataset described in the MANOVA into a training dataset (80%) and a testing dataset (20%), oversampling using the *Tidymodels* package (Kuhn & Wickham, 2020) within SMOTE (Chawla *et al.*, 2002) in R when classes were imbalanced. To include *R. cornivora* in the analysis, we up-sampled this taxon due to low sample size. As a result, we classified CHC differences at the species level in two separate analyses, once including *R. cornivora* and one removing the species from the dataset. We then fit a neural network with one hidden layer to the training dataset using the *nnet* (Venables & Ripley, 2002) and *caret* (Kuhn *et al.*, 2020) packages in R with 10-fold cross validation repeated 10 times similar to the methods of Moore *et al.* (2021), to 'classify' each taxa and/or sex based on their CHC profiles and determine statistical significance. The best-fit model, which determined the accuracy at which variation in CHCs distinguished among *Rhagoletis* taxa and between males and females, was then evaluated with a confusion matrix, using a one-sided *t*-test to statistically compare the testing dataset generated in the cross-validation process to the 'no information rate' training dataset.

Finally, we further tested for CHC differentiation between sexes and among taxa via a multidimensional scaling (MDS) approach, similar to that described in Adams and Collyer (2009). First, using MDS, we reduced CHC variation among flies to two dimensions (axes). We then determined the angle of

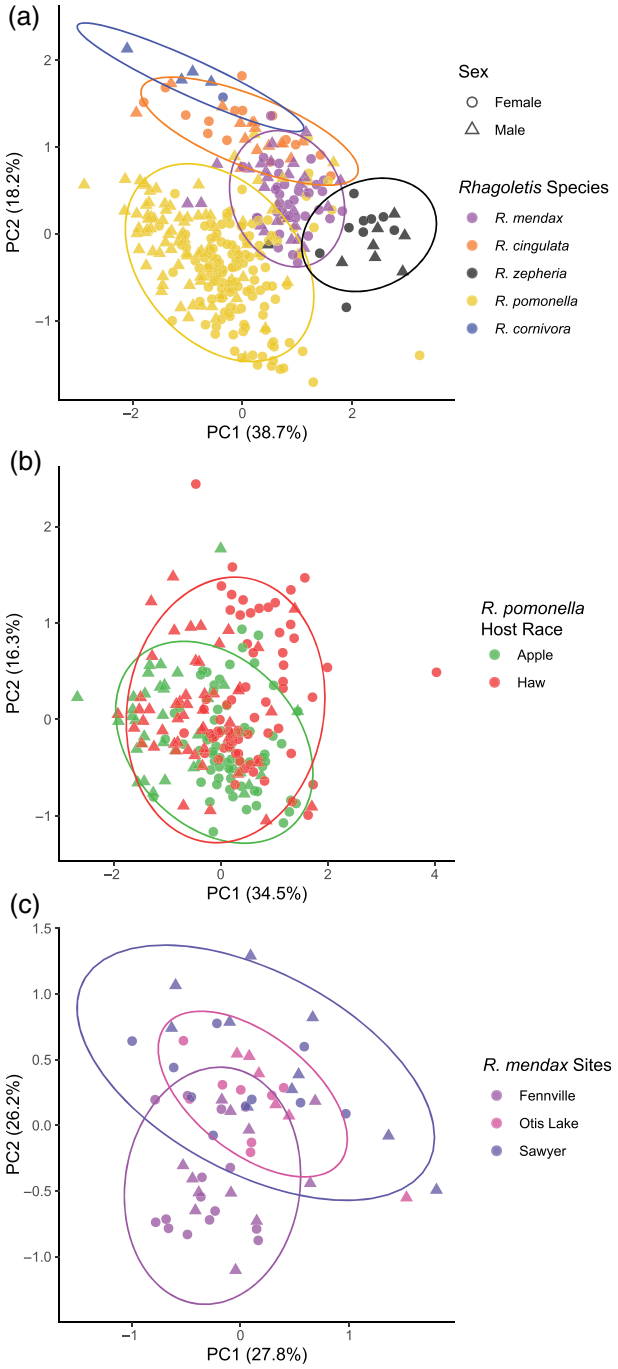

Fig 3. Principal component analysis (PCA) of CHC profiles for (a) the five *Rhagoletis* species, where the apple- and hawthorn-infesting host races are pooled to constitute the *Rhagoletis pomonella* sample; (b) the apple- and hawthorn-infesting host races of *R. pomonella*, assessed by sex and host plant; and (c) *Rhagoletis mendax*, examined by sex and site. Shown are plots of CHC scores for individual female (triangles) and male (circle) flies based on the first two PC axes, with the percentage of the total variation explained given for each axis in parentheses.

intersection of the line connecting the centroids for females and males within each taxon to the horizontal *x*-axis in the two-dimensional MDS plot. Each angle was calculated by taking the arctangent of the slope of the line connecting the centroids, with taxon having a positive slope assigned a value from 0° to 90° and negative slopes assigned a value between 0° and −90°. By determining the angles in this manner, we did not assign a directionality to the lines, as we would, for example, if we designated the centroid for females as the origin of the line. Instead, we focused on how well the lines connecting females and males, in general, aligned across taxa. In other words, this analysis determined whether there was a tendency for the sexes to vary along similar axes of CHC variation among different taxa. Thus, we considered the difference between taxa in their angles of intersection to the *x*-axis to reflect the degree to which their CHC profiles aligned between the sexes, with smaller differences reflecting greater similarity. Consequently, when the *x*-axis angles for the two taxa being compared were both positive or both negative, we used the absolute value of the difference between these angles as our metric of similarity. Alternatively, when one *x*-axis angle was positive and the other negative, we estimated an additional angle for the line having the negative slope as the absolute value of its angle of intersection with the *x*-axis subtracted from 180°. Next, we calculated the absolute values of both this second angle and the original angle subtracted from the angle for the taxon having the positive slope, taking the minimum of the two as our measure of similarity (see Fig. 4). To determine statistical significance, we first generated expected null distributions though non-parametric Monte Carlo computer simulations by randomly generating angles of intersection between the sexes for the six taxa and then calculating the mean and standard deviation for each of the 10 possible pairwise comparisons in 1 000 000 replicates. We then assessed whether the observed values resided in the lower 5% of the simulated null distribution to determine statistical significance for the differences between sexes across taxa being more similar than expected by chance.

We used a similar MDS approach to investigate divergence between taxa with two differences. First, we used the mid-points between the centroids between females and males within taxon to determine the angle at which the lines connecting species/races intersected the *x*-axis. Second, we did not consider all possible pairwise comparisons between taxa to calculate the mean and standard deviation of angle differences. Rather, we focused on four specific taxonomic comparisons: (1) *R. cornivora* to the hawthorn-infesting *R. pomonella*, and hawthorn-infesting *R. pomonella* to (2) *R. zephyria*, (3) *R. mendax*, and (4) apple-infesting *R. pomonella*. These four comparisons represent the evolutionary sequence in which these five *R. pomonella* group taxa analysed in the study are hypothesised to have descended (were derived) from each other by phylogenetic inference (Berlocher, 2002; Hulbert *et al.*, 2018). As discussed above, the *R. cingulata* group is not immediately sister to the *R. pomonella* group. Consequently, we could not interpret the trajectory of the CHC vector between *R. cingulata* and any of the *R. pomonella* group flies as reflecting evolution change occurring directly between ancestral and derived taxa. We therefore did not include *R. cingulata* in the analysis. Thus, the MDS

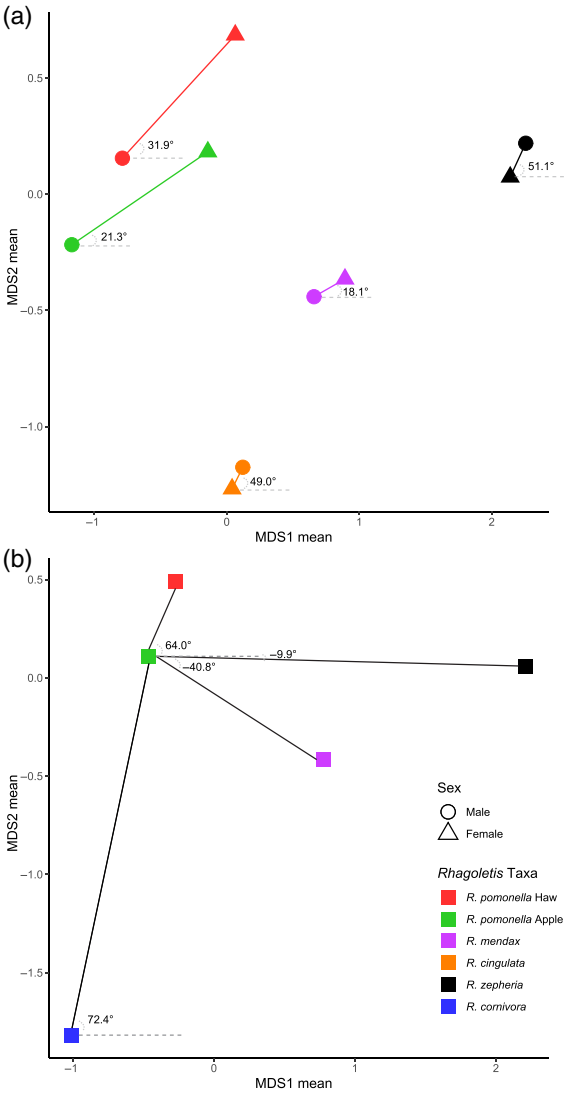

Fig 4. Plots of first two dimensions of multidimensional scaling analysis (MDS) of *Rhagoletis* CHC profiles for (a) sex-related variation within the five taxa (excluding *R. cornivora*); and (b) divergence between the five taxa (excluding *R. cingulata*). In panel (a), lines connecting female (triangles) and male (circles) centroids within each taxon are shown along with the angles that they intersect the *x*-axis, depicted by grey stippled lines. Non-parametric Monte Carlo computer simulations indicated that the mean difference in angles among taxa of 18.78° was significantly more similar than the null expectation of 45°. In panel (b), lines connecting centroids between pairs of taxa that represent the evolutionary sequence of divergence between populations are shown in black along with the angles that they intersect the *x*-axis, depicted as grey stippled lines.

analysis of species/host races explicitly focused on investigating the degree to which the five taxa in the *R. pomonella* group followed similar or different lines of CHC differentiation as they diverged. We evaluated the statistical significance of the mean and standard deviation of angle differences for the four taxonomic comparisons by non-parametric Monte Carlo computer simulation, as described above.

Results

Characterisation of CHCs for Rhagoletis

Based on GC analysis, we quantified 36 different but repeatable CHC peaks based on retention times for flies of all six taxa of *Rhagoletis*, including the *R. pomonella* host races (Fig. 2). However, MS analysis showed that these 36 components are comprised of at least 53 different CHC compounds, with chain lengths varying from 27 to 34 carbons and consisting of n-alkanes, methyl-branched alkanes, and several alkenes, indicating that there was not a one-to-one ratio of peaks to compounds (Table 2). Some compounds in *R. pomonella* and other species remain to be identified because longer ECL compounds had poor quality EI spectra due to lower signal intensities. Heavier compounds require more energy for ionisation and individual peaks were poorly resolved towards the end of the GC run. As a result, there was little interpretable information in the EI spectra and the mass to charge ratios (*m*/*z*) of intact molecules were not reliably detected for the unknown compounds. Based on near-identical retention times, all peaks were assumed to be homologous across species and between apple- and hawthorn-infesting host races of *R. pomonella*, as differences among samples appeared quantitative and not qualitative, that is, there were no observable peaks present in some samples, but not others (Fig. 2).

Variation among species

The full species-level MANOVA revealed significant differences in the relative abundance of CHCs among the five *Rhagoletis* species (Fig. 3a; Table 3). In addition, the species-level MANOVA with *R. cornivora* removed showed a significant interaction between sex and species (Table 3). The CHC variation between males and females was generally (but not always) lower in magnitude than those between species, as indicated by Euclidean distances between the centroids of CHC profiles generated in the MDS between different taxa and sexes of flies (Table S4). For example, the mean distance between the sexes, pooled across sites, among the five species was 0.532 ± 0.214 SE. In comparison, the mean Euclidean distance between females and males, pooled across sites, among species was 1.597 ± 0.191 and 1.884 ± 0.285, respectively, while the mean Euclidean distance between different species, pooling across sex, was 1.659 ± 0.227.

The NNC quantified the degree to which CHC variation distinguished the five *Rhagoletis* species and reinforced the pattern of interspecific CHC variation subsuming sex and site-related differences documented by the full-species level MANOVA. The overall accuracy of correct species assignment of fly taxa in the NNC based on log-transformed PCA values was 87.9% [confidence interval (CI) = 77.5–94.6%; $P < 7.23 \times 10^{-5}$], with all *R. cornivora* and *R. zephyria* being correctly identified, one of the *R. cingulata* individuals mis-classified as *R. pomonella*, 2 of the 61 *R. mendax* mis-assigned as *R. pomonella* and *R. zephyria*, respectively, and one, and four

Table 3. MANOVA results for CHC variation at the species level of population organisation, assessing insert differences among the full dataset including all five taxa (*R. cingulata*, *R. zephyria*, *R. mendax*, *R. cornivora*, and *R. pomonella*), and excluding *R. cornivora*. The apple- and hawthorn-infesting host races were pooled to constitute the *R. pomonella* sample. 'Sex' refers to adult females and males collected in the field as larvae and aged to sexual maturity in the laboratory.

MANOVA description	Source of Variation	d.f.	Wilk's λ	F	P
Full dataset (all taxa)	Species	4	0.197	44.10	2.32×10^{-104}
	Residuals	332			
Excluding *R. cornivora*	Species	3	0.0798	61.40	2.48×10^{-158}
	Sex	1	0.72	17.00	4.60×10^{-19}
	Species × sex	3	0.846	2.61	1.05×10^{-04}
	Residuals	324			

Table 4. MANOVA results for CHC variation at the host race level of population organisation, assessing differences between the apple- and hawthorn-infesting populations of *R. pomonella*. To explore the effect of site, we analysed subsets of apple- and hawthorn-infesting populations of *R. pomonella* separately. 'Sex' refers to adult females and males collected in the field as larvae and aged to sexual maturity in the laboratory. 'Site' refers to geographic locations at which flies were sampled.

MANOVA description	Source of variation	d.f.	Wilk's λ	F	P
Full *R. pomonella* dataset	Host race	1	0.764	5.88	2.58×10^{-8}
	Sex	1	0.467	21.8	2.85×10^{-29}
	Host race × sex	1	0.909	1.91	0.04
	Residual	220			
Apple-infesting *R. pomonella*	Sex	1	0.497	23.66	1.11×10^{-10}
	Site	2	0.829	2.29	0.04
	Sex × site	2	0.879	1.56	0.16
	Residuals	72			
Hawthorn-infesting *R. pomonella*	Sex	1	0.382	15.70	2.29×10^{-16}
	Site	1	0.201	38.60	1.84×10^{-29}
	Sex × site	1	0.709	3.99	1.38×10^{-4}
	Residuals	106			

of the *R. pomonella* mis-classified as *R. cingulata*, and *R. mendax*, respectively. With *R. cornivora* removed from the species-level dataset, the accuracy of the correct species assignment improved to 95.4% (CI = 87.1–99.0%, $P = 5.08 \times 10^{-8}$) with fewer *R. mendax* and *R. pomonella* flies misidentified. Given that variation in CHCs is generally lower in magnitude between males and females than among taxa, it is not surprising that our accuracy in classifying sexes is relatively poor at 65.1% (CI = 52.0–76.7%), albeit statistically significant ($P = 0.0003$).

Variation between the Rhagoletis pomonella *host races*

The full *R. pomonella* dataset MANOVA revealed significant variation in CHC profiles as indicated by a two-way interaction between host race and sex. Similarly, in the reduced dataset for hawthorn-infesting host race, we detected a significant two-way interaction between sex and site, while in the apple-infesting host race we detected significant sex and site terms with the interaction between the two being non-significant (Fig. 3b; Table 4). The Euclidean distances between the centroids of CHC profiles generated in the MDS, pooled across sites, between the host races (0.436) were less than that between the sexes within apple- (1.102) and hawthorn-infesting fly populations (1.003). Overall, the NNC assigned *R. pomonella* to the correct combined host race with an accuracy of 72.7% (CI = 57.2–85.0%, $P = 0.005$). However, when tasked to predict both host race and sex, the accuracy of the model fell to 61.4% (CI = 57.2–75.6%, $P = 5.25 \times 10^{-5}$). The first PC axis primarily differentiated the sexes and explained 34.5% of the overall CHC variation in *R. pomonella*, while the second axis distinguishing the host races explained 16.3% of the variation (Fig. 3b).

Sex and geographic variation within other species

In addition to *R. pomonella*, the blueberry-infesting species *R. mendax* also displayed both significant sex and site-related CHC variations, as determined by MANOVA (Fig. 3c; Table 5). Across the three populations analysed, NNC assigned *R. mendax*-origin flies to the correct sex with an accuracy of 45.5% (CI = 16.8–76.6%, $P = 0.82$) and the correct combined site and sex with a reduced accuracy of 40.0% (CI = 12.2–73.8%, $P = 0.12$). For *R. mendax*, both the first and second PC axes, explaining 27.8% and 26.2% of the variation, respectively, primarily distinguished flies at the Fennville, Michigan site from those elsewhere, with flies sampled from Otis Lake nested within the Sawyer, Michigan population (Fig. 3c). Finally, MANOVA did not reveal a significant sex effect for *R. cingulata* or *R. zephyria* (Table 5).

Table 5. MANOVA results for CHC variation at intraspecific level of population organisation, assessing CHC differences within *R. mendax*, *R. cingulata*, and *R. zephyria*. Due to limitations in sampling, *R. cornivora* could not be analysed, and the interaction of 'Sex × site' could not be analysed for *R. zephyria*. 'Sex' refers to adult females and males collected in the field as larvae and aged to sexual maturity in the laboratory. 'Site' refers to geographic locations at which flies were sampled.

MANOVA description	Source of variation	d.f.	Wilk's λ	F	P
R. mendax	Sex	1	0.619	6.29	0.0001
	Site	2	0.122	19.00	3.07×10^{-19}
	Sex × site	2	0.822	1.05	0.041
	Residuals	55			
R. cingulata	Sex	1	0.747	7.45	0.606
	Residuals	28			
R. zephyria	Sex	1	0.816	1.41	0.260
	Residuals	15			

Parallelism/convergence in patterns of CHC variation

MDS analysis implied that male and female flies show similarities in the axes along which their CHC profiles vary among the different *Rhagoletis* taxa (Fig. 4a). The mean (18.78) and standard deviation (10.81) of the 10 pairwise differences in angles of vectors connecting CHC centroids between sexes among taxa were significantly lower than expected when compared to their computer-simulated null distributions (expected null mean = 45.00, $P = 0.01$; expected null standard deviation = 24.25, $P = 0.008$). While orientation of the vectors between sexes differed among taxa, they were opposite in direction between females and males for *R. cingulata* and *R. zephyria* compared to the apple- and hawthorn-infesting host races and *R. mendax* (Fig. 4a). This difference can help account for the significant two-way interaction between sex and species based on the MANOVA with *R. cornivora* removed conducted at the species level of population organisation discussed above (see Table 3). However, the magnitude of the difference and/or sample sizes for the comparisons between females and males for *R. cingulata* and *R. zephyria* were not sufficient to detect significant sex-related differences for these two species. Thus, a degree of caution is urged in interpreting the results from the MDS analysis. Nevertheless, the pattern is suggestive that the sexes differ along similar CHC axes, although not in a similar direction or magnitude, among fly taxa.

In contrast to sex, vectors describing CHC differentiation among *R. cornivora*, *R. zephyria*, *R. mendax*, and the apple- and hawthorn-infesting host races of *R. pomonella* were not more similar to one another than expected by chance (Fig. 4b). The mean (56.26) and standard deviation (27.06) of the four pairwise differences in angles of vectors connecting CHC centroids between these five taxa were not significantly different than expected compared to their computer-simulated null distributions (expected null mean = 45.00, $P = 0.080$; expected null standard deviation = 23.01, $P = 0.231$). Thus, the differences between the five fly taxa do not appear to reflect evolution along similar CHC axes. Rather, the results suggest that the CHC profiles of taxa tend to be over dispersed as opposed to randomly orientated, although not statistically so.

Discussion

Our results revealed significant quantitative variation in adult CHC composition among the six *Rhagoletis* taxa analysed in this study (Table 3; Figs 3a and 4b). In addition, significant sex- and site-related variation in CHC profiles is present within the apple- and hawthorn-infesting host races of *R. pomonella*, and in *R. mendax* (Tables 4 and 5; Figs 3b,c and 4a). Although statistically significant sex differences appear lacking for *R. cingulata* and *R. zephyria* (Table 5), the limited numbers of these flies assayed across populations do not rule out the possibility that such effects may be detected with increased sampling. Indeed, the CHC profiles of females and males in all six taxa tested, including *R. cingulata*, *R. zephyria*, and *R. cornivora*, appear to vary in magnitude (although not in direction) along similar axes (Fig. 4a). In contrast, as new *Rhagoletis* species and host races are hypothesised to have originated via sympatric host plant shifting (Bush, 1966; Feder *et al.*, 1988), they appear to have diverged along different and independent paths in their CHC profiles (Fig. 4b). Thus, sex and species differences in CHC profiles do not seem to be related in *Rhagoletis*.

Heritability of CHC profiles

Given our results, a key question concerns whether the observed sex and species differences in CHC profiles are heritable. All flies analysed in the study were collected as larvae in fruit in nature, and matured to adulthood in the laboratory under controlled conditions. Thus, the variation we detected in CHCs profiles among taxa and between sexes may have been influenced by differences in the environment experienced by flies prior to being transported into the laboratory (e.g., fruit condition, temperature; Etges *et al.*, 2009; Merli *et al.*, 2018), as the various taxa infest host fruit at different times of the season (Dambroski & Feder, 2007). However, the consistency in the axes along which the CHC profiles vary between females and males across taxa suggests that at least for sex, the differences are perhaps genetically based, related to some developmental difference between females and males that vary in direction and magnitude among taxa. Moreover, if the host fruit environment is a major determinant of CHC profiles, then one might expect that the phylogenetic relatedness of host plants, which may

reflect a degree of similarity in fruit chemistry, would be a significant predictor of CHC divergence. In this regard, apples and hawthorns are both Rosaceous hosts of *R. pomonella* (Bush, 1966) and the apple- and hawthorn-infesting host races have the most similar CHC profiles of any of the six taxa analysed in the study (Fig. 4b; mean Euclidean distance for CHCs between apple and hawthorn flies of both sexes at sympatric sites is 0.436). However, this relationship is also expected given the recent origin of the apple race (∼170 years ago) and estimated levels of ongoing gene flow between the host races of 4–6% per generation (Feder *et al.*, 1994). Also, *R. pomonella* is more similar in its CHC profile to *R. mendax* than to the other three species (Figs 3a and 4b). However, *R. mendax* infests hosts in the family Ericaceae, which is in a different order (Ericales) than apples and hawthorns (Rosales). In addition, the different cherry hosts of *R. cingulata*, like those of *R. pomonella*, are also Rosaceous (Bush, 1966). However, *R. cingulata* has a distinct CHC profile from apple- and hawthorn-infesting flies (Figs 3a and 4b). Thus, there is no obvious relationship between host phylogeny and CHC divergence for fly taxa included in our study.

Aside from the *R. pomonella* host races, there was also no obvious relationship between the CHC profiles of flies and their evolutionary relatedness. Euclidean distances among the taxa for CHCs were not significantly correlated with mean genetic distances between taxa (Mantel test: $r = 0.089$, $P = 0.795$, 10 d.f.) based on the mtDNA sequence variation reported in Hulbert *et al.* (2018). Thus, there was no general trend for more closely related taxa to be more similar in their CHC profiles than more evolutionarily distant species. Moreover, the trajectories of CHC change between ancestral and derived taxa was not statistically different from random expectation and, if anything, are perhaps over dispersed in the MDS analysis of the *R. pomonella* group (Fig. 4b). Consequently, there was no clear phylogenetic signal in the magnitude or direction of CHC change related to the history of taxonomic divergence of the flies in our study. Thus, while the sexes may be constrained to vary along similar axes in their CHC profiles, the same was not true between different *Rhagoletis* taxa. Indeed, there was a non-significant trend for taxa to be more different from one another than expected by chance, which may, if statistically verified by increased sampling of additional taxa, imply a possible role for CHCs in species discrimination. Regardless, the pattern of CHC differentiation among species provides little insight concerning their genetic basis, as might be the case if more closely related taxa sharing a more recent and common genetic ancestry displayed greater CHC similarity. Thus, while there is no current evidence that the observed CHC differences among flies are environmentally induced, we cannot rule out this possibility. Thus, future reciprocal host transplant experiments (rearing flies in common garden studies), crosses, and genome-wide-association studies are needed to confirm that the observed CHC differences are heritable.

The role of CHCs in sexual selection and species recognition

In addition to completing the identification of the CHCs in this system and understanding their genetic basis, more work is needed to resolve the biological significance of the observed CHC variation. If these differences are heritable, then CHC variation may play an important role in sexual selection via mate choice, and/or host race and species recognition, contributing to premating isolation and population divergence. Several aspects of the mating behaviour of *Rhagoletis* are well-characterised. For *R. pomonella*, as well as in other *Rhagoletis* species (Smith & Prokopy, 1982; Smith, 1984, 1985), males have been observed moving from fruit to fruit in search of females, and encounters between adults typically involve posturing and wing waving by both sexes. Male approach towards prospective mates is visual, based on size, shape, colour, and movement. Wing waving and vibration has been recorded in a number of *Rhagoletis* species (Bush, 1966; Alonso-Pimentel *et al.*, 2000) and other tephritids (Sivinski & Burk, 1989) but no courtship songs are known for *R. pomonella*. Males orient around females both from the front and rear, but short-range stimuli that elicit courtship attempts have not been determined. Presence of a male sex pheromone was suggested in *R. pomonella* (Prokopy, 1975) but no analysis of these chemicals has been reported (Sarles *et al.*, 2015). More copulations were recorded when males approached females from behind where they appear to forcibly mate with females and 'jump' onto females from ∼1 to 3 cm away. Receptive females then extend their ovipositors towards the male's aedeagus, otherwise she inhibits the male's advances by moving away. Successful males then clasp their foretarsi on the female's anterior abdomen and use their other legs to grasp terminal segments of her abdomen. This clasping behaviour has been hypothesised to stimulate ovipositor extension and facilitate insertion of the aedeagus into her ovipositor (Prokopy & Bush, 1973). Such close physical contact, somewhat similar to the latter stages of successful courtship and copulation in *Drosophila* species (Spieth, 1974; Hall, 1994), suggests a potential role for CHCs as contact pheromones, although volatile pheromones cannot currently be ruled out. Moreover, it is not uncommon for males to attempt to mount and copulate with other males when they approach from behind (J. L. Feder, pers. observ.), an activity which may be foreshortened when sexes differ in their CHC profiles. Thus, the finding that CHCs vary in a similar manner between males and females is consistent with sex-related differences being used as a means for assessing the quality or identity of conspecific mates, with the differences between fly taxa serving as cues for species and/or host race discrimination. More detailed analyses of courtship behaviour and mate choice will be needed to confirm whether CHC signal and preference are under selection in this system.

Features of the natural history of *Rhagoletis* may imply a lesser role for CHCs in sexual selection and species recognition. For example, in previous mate choice experiments in *Rhagoletis*, strong premating isolation was observed mainly between species with distinct morphological differences related to body coloration, body striping, and wing pattern. In this regard, the cherry-infesting *R. cingulata* is visually distinct from the other taxa we analysed in the current study and displays complete premating isolation from the apple-infesting host race in mating trials (Hood *et al.*, 2012). In comparison, the *R. pomonella* group flies *R. pomonella*, *R. mendax*, *R. zephyria*, and *R. cornivora* are all morphologically similar sibling species (Bush, 1966;

Jiggins & Birdle, 2004). Despite this, moderate premating isolation was observed between *R. mendax* and *R. zephyria* (Schwarz & McPheron, 2007). However, Smith (1986) found no evidence of assortative mating between the apple-infesting host race of *R. pomonella* and its undescribed sister taxon infesting flowering dogwood (Berlocher, 2001), which remains to be characterised for CHCs. In addition, no premating isolation was reported between the apple- and hawthorn-infesting host races of *R. pomonella* (Reissig & Smith, 1978). Moreover, in *R. mendax*, males appear to be unable to discriminate between sexes either visually on blueberries or by physical contact (Smith & Prokopy, 1982). In many systems, sex-specific differences in the presence versus absence of specific CHC compounds among distinct populations or taxa can drive sexual isolation (Davis *et al.*, 2021). Thus, the lack of qualitative differences in CHC profiles among *Rhagoletis* taxa or between males and females may suggest a lesser role of CHCs in sexual isolation. More extensive crossing and manipulative studies are therefore needed to establish the effects that CHCs may have for species recognition, mate choice, and premating isolation in *Rhagoletis*.

Additional roles of variation in CHCs

Epicuticular hydrocarbons can play other important functional roles in insects in addition to affecting mate choice, including serving as hydrophobic barriers to water loss (Gibbs & Rajpurohit, 2010). Insects are vulnerable to desiccation because of their large surface area to volume ratio and must minimise transcuticular water loss. *Rhagoletis* is univoltine and adults have a limited life span of about 1 month in nature (Dean & Chapman, 1973), therefore adults must live long enough to locate host fruits, mate, and reproduce. Different host plants attacked by *Rhagoletis* fruit at different times of the season, thus, the life cycles of different fly species and races are offset to coincide with the phenology of their respective hosts (Dambroski *et al.*, 2005). As a result, not only are adults of different *Rhagoletis* species and races active at different times of the season when rainfall and humidity levels can vary significantly, but pupae must resist desiccation in the soil for differing amounts of time before the onset of winter (Feder *et al.*, 1997; Egan *et al.*, 2015). Furthermore, *Rhagoletis* males, generally being smaller in size and eclosing as adults later than females (Bush, 1966; Feder *et al.*, 1993, 1994), could experience different selection pressures for desiccation and these effects could vary between taxa and locally among sites. Consequently, desiccation rather than mate choice and sexual selection could potentially explain the pattern of CHC variation we observed in the study. Further work is therefore needed to resolve the roles that CHCs play as potential contact pheromones and/or in desiccation resistance in *Rhagoletis* (Chung & Carroll, 2015).

Our results add to a growing list of studies in which CHC variation has been described generally in arthropods and more specifically for tephritid fruit flies. Species specific cuticular alkanes and alkenes have been characterised in *Anastrepha ludens*, *A. suspensa*, *Ceratitis capitata*, *C. rosa*, *Dacus cucurbitae*, and *D. dorsalis*, with several alkadienes observed in *Anastrepha* and *Ceratitis* but not in *Dacus* (Carlson & Yocom, 1986). In invasive Hawaiian populations of *C. capitata*, CHCs composed of alkanes, alkenes, and alkadienes ranged in size from C_{29} to C_{37}, with a large unsaturated C_{35} component (W. J. Etges and Shelley, unpubl. data). In the South American fruit fly, *Anastrepha fraterculus*, both males and females were characterised by CHC profiles including n-alkanes, methyl-branched alkanes, and alkadienes that differed between sexes at maturation (Vaníčková *et al.*, 2012). In this study, the authors identified 66 peaks corresponding to chain-lengths ranging from 13 to 37 carbons including 14 n-alkanes, 20 monomethyl alkanes, 2 dimethyl alkanes, 10 alkenes, 12 alkadienes, and 9 unidentified compounds. Analysis of adult *A. ludens* males revealed nine CHCs ranging from 28 to 31 carbons with 2-methyloctacosane, n-nonacosene, 2-methyltriacontane, and n-hentriacontene as the most abundant components (Bosa *et al.*, 2018). The degree to which *Rhagoletis* CHCs share chemical similarities with these other species will require further sampling and analyses of phylogenetic and host influences on CHC composition (Symonds & Elgar, 2008; Oliveira *et al.*, 2011). Nevertheless, *Rhagoletis* provides another example demonstrating a complex array of CHCs on the surface of an adult tephritids potentially under natural and/or sexual selection.

Conclusion

Understanding the speciation process for sexually reproducing taxa requires discerning how inherent genetic differences arise between formerly interbreeding populations that cause them to become reproductively isolated. Barriers to gene flow acting earlier compared to later in the life cycle are considered to have relatively greater effect on reproductive isolation (Coyne & Orr, 2004). Hence, when compared to post-mating isolation, those traits causing premating isolation can often be among the most significant factors restricting gene flow and facilitating speciation (Coyne & Orr, 1989, 1997; Mendelson, 2003). Fruit flies in the genus *Rhagoletis* are considered a model system for ecological speciation triggered by divergent adaptation to different host plants (Berlocher & Feder, 2002; Funk *et al.*, 2002; Jiggins & Birdle, 2004). While previous studies established that differential host choice is a key behaviour generating premating isolation between *Rhagoletis* (Feder *et al.*, 1994), we have documented host race, species, sex, and site-related differences in CHC profiles, which suggest a possible role for these compounds as contact pheromones. Additional premating isolation may therefore exist after host choice occurs in *Rhagoletis* based on chemical variation in their epicuticular hydrocarbons, as has been found for other insects (Howard & Blomquist, 2005; Blomquist & Bagnères, 2010). Whether CHCs are involved in courtship and premating isolation in *Rhagoletis* and/or affect desiccation resistance remains to be determined by future experiments. If CHCs affect desiccation resistance but not mate choice, then they may still play an important role in speciation for *Rhagoletis*. However, rather than contributing to premating isolation, CHCs may reflect host-related ecological adaptations restricting gene flow between taxa specialised on different plants, thereby potentially generating post-mating reproductive isolation. Indeed, it is even possible that CHCs both

ecologically adapt flies to their respective hosts while also serving as key species and sex recognition cues. Thus, *Rhagoletis* holds promise for clarifying in a model system the roles that different components of host choice and courtship behaviour, together with ecological specialisation, play in generating reproductive isolation, and the order and relative importance of these factors in the speciation process.

Acknowledgements

G. R. Hood thanks P. Morton for support. Funding for this study was provided by National Science Foundation grant DEB-0211125, and a grant from the Center on the Economics and Demography of Aging (CEDA)–UC Berkeley. M. Beehler was supported by the Arkansas INBRE program from NIH grant NIGMS P20 GM103429. J. L. Feder and G. R. Hood were supported by grants from the NSF and USDA. The authors declare no conflict of interest.

Author contributions

The study was designed by GRH, JHJ, WJE, and JLF. Data were collected by GRH, JHJ, MB, TS, JLF, and WJE. Statistical analyses were conducted by GRH, JHJ, DJB, JLF and WJE. The manuscript was written by GRH, JHJ, JLF, and WJE. All authors contributed to revisions and approved the final manuscript.

Data availability statement

The raw data is available upon request from the corresponding author.

Supporting Information

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

Table S1 Loadings for each of the 36 CHC components on the first eight Principal Components among the six taxa of *Rhagoletis*. Peak numbers (1–36) correspond to the labeled peaks in Fig. 2 and compounds in Table 2.

Table S2 Results for MANOVA analysis of CHC variation at three levels of population organization (species, host races of *R. pomonella*, and intraspecific variation in *R. mendax*, *R. cingulata*, and *R. zephyria*). Numbers in column headers represent the CHC used to calculated log-contrasts for the MANOVA (see Table 2 for list of the numerical designations of the compounds). We focused on discussing the results when relative CHC abundances were log-contrast transformed for CHC peak 6. Light grey shaded boxes represent significant effects in the MANOVA, with significant results for CHC peak 6 shown are in dark grey for comparison.

Table S3 The number of principal components (PCs) used in the MANOVA summarised in Tables 3, 4, and 5, and the percent variation in CHC profiles of *Rhagoletis* explained by those PCs.

Table S4 Euclidean distance between *Rhagoletis* taxa of different sexes infesting different host plants (upper portion of table) pooled across sex (lower table).

References

Adams, D.C. & Collyer, M.L. (2009) A general framework for the analysis of phenotypic trajectories in evolutionary studies. *Evolution*, **63**, 1143–1154.

Alonso-Pimentel, H., Spangler, H.G., Rogers, R. & Papaj, D.R. (2000) Acoustic component and social context of the wing display of the walnut fly *Rhagoletis juglandis*. *Journal of Insect Behaviour*, **13**, 511–524.

Berdan, E., Enge, S., Nylund, G.M., Wellenreuther, M., Martens, G.A. & Pavia, H. (2019) Genetic divergence and phenotypic plasticity contribute to variation in cuticular hydrocarbons in the seaweed fly *Coelopa frigida*. *Trends in Ecology and Evolution*, **9**, 12156–12170.

Berlocher, S.H., McPheron, B.A., Feder, J.L. & Bush, G.L. (1993) A revised phylogeny of the *Rhagoletis pomonella* (Diptera: Tephritidae) sibling species group. *Annals of the Entomological Society of America*, **86**, 716–727.

Berlocher, S.H. (2001) Host race or species? Allozyme characterization of the 'flowering dogwood fly', a member of the *Rhagoletis pomonella* complex. *Heredity*, **83**, 652–662.

Berlocher, S.H. (2002) Radiation and divergence in the *Rhagoletis pomonella* species group: inferences from allozymes. *Evolution*, **54**, 543–557.

Berlocher, S.H. & Feder, J.L. (2002) Sympatric speciation in phytophagous insects: moving beyond controversy? *Annual Review of Entomology*, **47**, 773–815.

Blomquist, G.J., Toolson, E.C. & Nelson, D.R. (1985) Epicuticular hydrocarbons of *Drosophila pseudoobscura* (Diptera: Drosophilidae): identification of unusual alkadiene and alkatriene positional isomers. *Insect Biochemistry*, **15**, 25–34.

Blomquist, G.J. & Bagnères, A.-G. (2010) *Insect Hydrocarbons: Biology, Biochemistry, and Chemical Ecology*. New York, NY, USA: Cambridge University Press.

Blows, M.W. & Allan, R.A. (1998) Levels of mate recognition within and between two *Drosophila* species and their hybrids. *American Naturalist*, **152**, 826–837.

Blows, M.W. & Higgie, M. (2002) Evolutionary experiments on mate recognition in the *Drosophila serrata* species complex. *Genetics of Mate Choice: From Sexual Selection to Sexual Isolation* (ed. by W. J. Etges and M. A. F. Noor), pp. 239–250. Kluwer, New York, New York.

Bosa, C.-F., Cruz-López, L., Guillén-Navarro, K., Zepeda-Cisneros, C.S. & Liedo, P. (2018) Variation in the cuticular hydrocarbons of the Mexican fruit fly *Anastrepha ludens* males between strains and age classes. *Insect Biochemistry and Physiology*, **99**, e21513.

Bush, G.L. (1966) The taxonomy, cytology, and evolution of the genus *Rhagoletis* in North America (Diptera: Tephritidae). *Bulletin of the Museumof Comparative Zoology*, **134**, 431–562.

Bush, G.L. (1969) Mating behavior, host specificity, and the ecological significance of sibling species in frugivorous flies of the genus *Rhagoletis* (Diptera: Tephritidae). *American Naturalist*, **103**, 669–672.

Bush, G.L. (1993) Host race formation and sympatric speciation in *Rhagoletis* fruit flies (Diptera: Tephritidae). *Psyche*, **99**, 335–357.

Butterworth, N.J., Wallman, J.F., Drijfhout, F.P., Johnston, N.P., Keller, P.A. & Byrne, P.G. (2020) The evolution of sexually dimorphic cuticular hydrocarbons in blowflies (Diptera: Calliphoridae). *Journal of Evolutionary Biology*, **10**, 1468–1486.

Carlson, D.A. & Yocom, S.R. (1986) Cuticular hydrocarbons from six species of tephritid fruit flies. *Archives of Insect Biochemistry Physiology*, **3**, 397–412.

Chawla, N.V., Bowyer, K.W., Hall, L.O. & Kegelmeyer, W.P. (2002) SMOTE: synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research*, **1**, 321–357.

Chung, H. & Carroll, S.B. (2015) Wax, sex and the origin of species: dual roles of insect cuticular hydrocarbons in adaptation and mating. *BioEssays*, **37**, 822–830.

Coyne, J.A. & Orr, H.A. (2004) *Speciation*. Sinauer Press, Sunderland, Massachusetts.

Coyne, J.A. & Orr, H.A. (1989) Patterns of speciation in *Drosophila*. *Evolution*, **43**, 362–381.

Coyne, J.A. & Orr, H.A. (1997) "Patterns of speciation in *Drosophila*" revisited. *Evolution*, **51**, 295–303.

Dambroski, H.R., Linn, C., Berlocher, S.H., Forbes, A.A., Roelofs, W. & Feder, J.L. (2005) The genetic basis for fruit odor discrimination in *Rhagoletis* flies and its significance for sympatric host shifts. *Evolution*, **59**, 1953–1964.

Dambroski, H.R. & Feder, J.L. (2007) Host plant and latitude-related diapause variation in *Rhagoletis pomonella*: a test for multifaceted life history adaptation on different stages of diapause development. *Journal of Evolutionary Biology*, **20**, 2101–2112.

Davis, J.S., Pearcy, M.J., Yew, J.Y. & Moyle, L.C. (2021) A shift to shorter cuticular hydrocarbons accompanies sexual isolation among *Drosophila americana* group populations. *Evolution Letters*. **5**, 521–540.

Dean, R.W. & Chapman, P.J. (1973) Bionomics of the apple maggot in eastern New York. *Search Agriculture and Entomology Geneva*, **3**, 1–64.

Doellman, M.M., Saint Jean, G., Egan, S.P., Powell, T.H.Q., Hood, G.R., Schuler, H. *et al.* (2020) Evidence for spatial clines and mixed geographic modes of speciation for North America cherry-infesting *Rhagoletis* (Diptera: Tephritidae) flies. *Ecology and Evolution*, **10**, 12727–12744.

Doellman, M.M., Schuler, H., Saint Jean, G., Hood, G.R., Egan, S.P., Powell, T.H.Q. *et al.* (2019) Geographic and ecological dimensions of host plant-associated genetic differentaition and speciation in the *Rhagoletis cingulata* (Diptera: Tephritidae) sibling specis group. *Insects*, **10**, 275.

Dres, M. & Mallet, J. (2002) Host races in plant-feeding insects and their importance in sympatric speciation. *Philosophical Transactions of the Royal Society London B: Biological Sciences*, **357**, 471–592.

Egan, S.P., Ragland, G.J., Assour, L., Powell, T.H.Q., Hood, G.R., Emrich, S. *et al.* (2015) Experimental evidence of genome-wide impact of ecological selection during early stages ofspeciation-with-gene-flow. *Ecology Letters*, **18**, 817–825.

Etges, W.J. & Jackson, L.L. (2001) Premating isolation is determined by larval rearing substrates in cactophilic *Drosophila mojavensis*. VI. Epicuticular hydrocarbon variation in *Drosophila mojavensis* cluster species. *Journal of Chemical Ecology*, **27**, 2125–2149.

Etges, W.J. (2002) Divergence in mate choice systems: does evolution play by rules? *Genetica*, **116**, 151–166.

Etges, W.J., de Oliveira, C.C., Ritchie, M.G. & Noor, M.A.F. (2009) Genetics of incipient speciation in *Drosophila mojavensis*. II. Host plants and mating status influence cuticular hydrocarbon QTL expression and G x E interactions. *Evolution*, **63**, 1712–1730.

Etges, W.J. (2014) No boundaries: genomes, organisms, and ecological interactions responsible for divergence and reproductive isolation. *Journal of Heredity*, **105**, 756–770.

Feder, J.L., Chilcote, C.A. & Bush, G.L. (1988) Genetic differentiation between sympatric host races of *Rhagoletis pomonella*. *Nature*, **336**, 61–64.

Feder, J.L., Hunt, T.A. & Bush, G.L. (1993) The effects of climate, host plant phenology and host fidelity on the genetics of apple and hawthorn infesting races of *Rhagoletis pomonella*. *Entomologia Experimentalis et Applicata*, **69**, 117–135.

Feder, J.L., Opp, S.B., Wlazlo, B., Reynolds, K., Go, W. & Spisak, S. (1994) Host fidelity is an effective pre-mating barrier between sympatric races of the apple maggot fly. *Proceedings of the National Academy of Science United States of America*, **91**, 7990–7994.

Feder, J.L., Roethele, J.B., Wlazlo, B. & Berlocher, S.H. (1997) Selective maintenance of allozyme differences between sympatric host races of the apple maggot fly. *Proceedings of the National Academy of Science United States of America*, **94**, 11417–11421.

Feder, J.L. (1998) The apple maggot fly, *Rhagoletis pomonella*: Flies in the face of conventional wisdom about speciation? *Endless Forms: Species and Speciation* (ed. by D. J. Howard and S. H. Berlocher), pp. 130–144. Oxford University Press, New York, New York.

Feder, J.L., Egan, S.P. & Nosil, P. (2012) The genomics of speciation-with-gene-flow. *Trends in Genetics*, **28**, 342–350.

Ferveur, J.F. (2005) Cuticular hydrocarbons: their evolution and roles in *Drosophila* pheromonal communication. *Behavioral Genetics*, **35**, 279–295.

Forbes, A.A., Fisher, J. & Feder, J.L. (2005) Habitat avoidance: overlooking an important aspect of host specific mating and sympatric speciation? *Evolution*, **59**, 1552–5159.

Funk, D.J., Filchak, K.E. & Feder, J.L. (2002) Herbivorous insects: model systems for the comparative study of speciation ecology. *Genetica*, **116**, 251–267.

Gerhardt, H.C. & Huber, F. (2002) *Acoustic Communication in Insects and Anurans*. University Chicago Press, Chicago, Illinois.

Gibbs, A.G. & Rajpurohit, S. (2010) Water-proofing properties of cuticular lipids. *Insect Hydrocarbons; Biology, Biochemistry and Chemical Ecology* (ed. by G. J. Blomquist and A. G. Bagneres), pp. 100–120. Cambridge University Press, New York, New York.

Giglio, E.M. & Dyer, K.A. (2013) Divergence of premating behaviors in the closely related species *Drosophila subquinaria* and *D. recens*. *Ecology and Evolution*, **3**, 365–374.

Gleason, J.M., Pierce, A.A., Vezeau, A.L. & Goodman, S.F. (2012) Different sensory modalities are required for successful courtship in two species of the *Drosophila willistoni* group. *Animal Behaviour*, **83**, 217–227.

Greenfield, M.D. (2002) *Signalers and Receivers: Mechanisms and Evolution of Arthropod Communication*. Oxford University Press, New York, New York.

Hall, J.C. (1994) The mating of a fly. *Science*, **264**, 1702–1714.

Hood, G.R., Egan, S.P. & Feder, J.L. (2012) Evidence for sexual isolation as a prezygotic barrier to gene flow between morphologically divergent species of *Rhagoletis* fruit flies. *Ecological Entomology*, **37**, 521–528.

Hood, G.R., Forbes, A.A., Powell, T.H.Q., Egan, S.P., Hamerlinck, G., Smith, J.J. *et al.* (2015) Sequential divergence and the multiplicative origin of community diversity. *Proceedings of the National Academy of Sciences United States of America*, **112**, E5980–E5998.

Hood, G.R., Zhang, L., Hu, E.G., Ott, J.R. & Egan, S.P. (2019) Cascading reproductive isolation: plant phenology drives temporal isolation among populations of a host-specific herbivore. *Evolution*, **73**, 554–568.

Howard, R.W. & Blomquist, G.J. (2005) Ecological, behavioral, and biochemical aspects of insect hydrocarbons. *Annual Review of Entomology*, **50**, 371–393.

Hulbert, D., Jackson, M.D., Hood, G.R. & Smith, J.J. (2018) Description of a new *Rhagoletis* (Diptera: Tephritidae) species in the *tabellaria* species group. *Insect Systematics and Diversity*, **2**, 1.

Jackson, L.L. & Bartelt, R.J. (1986) Cuticular hydrocarbons of *Drosophila virilis*. Comparison by sex and age. *Insect Biochemistry*, **16**, 433–439.

Jaenike, J. (1986) Genetic complexity of host-selection behavior in *Drosophila*. *Proceedings of the National Academy of Science United States of America*, **83**, 2148–2151.

Jallon, J.-M. & David, J.R. (1987) Variations in cuticular hydrocarbons among the eight species of the *Drosophila melanogaster* subgroup. *Evolution*, **41**, 294–302.

Jennings, J.H., Etges, W.J., Schmitt, T. & Hoikkala, A. (2014) Cuticular hydrocarbons of *Drosophila montana*: geographic variation, sexual dimorphism and potential roles as pheromones. *Journal of Insect Physiology*, **61**, 16–24.

Jiggins, C.D. & Birdle, J.R. (2004) Speciation in the apple maggot fly: a blend of vintages? *Trends in Ecology and Evolution*, **19**, 111–114.

Kuhn, M. & Wickham, H. (2020) Tidymodels: a collection of packages for modeling and machine learning using tidyverse principles. URL https://www.tidymodels.org. [accessed on 1 August 2021].

Kuhn, M., Wing, J., Weston, S., Williams, A., Keefer, C., Engelhardt, A. *et al.* (2020) Package 'caret'. *The R Journal*, **20**, 223.

Linn, C.E., Feder, J.L., Nojima, S., Dambroski, H.R., Berlocher, S.H. & Roelofs, W. (2003) Fruit odor discrimination and sympatric host race formation in *Rhagoletis*. *Proceedings of the National Academy of Sciences United States of America*, **100**, 11490–11493.

Linn, C.E., Yee, W.L., Sim, S.B., Cha, D.H., Powell, T.H.Q., Goughnour, R.B. *et al.* (2012) Behavioral evidence for fruit odor discrimination and sympatric host races of *Rhagoletis pomonella* flies in the western United States. *Evolution*, **66**, 3632–3641.

Margolies, D.C. & Collins, R.D. (1994) Chemically-mediated pre-mating behavior in two tetranychid species. *Experimental and Applied Acarology*, **18**, 493–501.

McPheron, B.A., Smith, D.C. & Berlocher, S.H. (1988) Genetic differences between host races of *Rhagoletis pomonella*. *Nature*, **336**, 64–66.

Mendelson, T.C. (2003) Sexual isolation evolves faster than hybrid inviability in a diverse and sexually dimorphic genus of fish (Percidae: Etheostoma). *Evolution*, **57**, 317–327.

Merli, D., Mannucci, B., Bassetti, F., Corana, F., Falchetto, M., Malacrida, A.R. *et al.* (2018) Larval diet affects male pheromone blend in a laboratory strain of the medfly, *Ceratitis capitata* (Diptera: Tephritidae). *Journal of Chemical Ecology*, **44**, 339–353.

Moore, H.E., Hall, M.J., Drijfhout, F.P., Cody, R.B. & Whitmore, D. (2021) Cuticular hydrocarbons for identifying Sarcophagidae (Diptera). *Scientific Reports*, **11**, 7732.

Mullen, S.P., Mendelson, T.C., Schal, C. & Shaw, K.L. (2007) Rapid evolution of cuticular hydrocarbons in a species radiation of acoustically diverse Hawaiian crickets (Gryllidae: Trigonidiinae: *Laupala*). *Evolution*, **61**, 223–231.

Mullen, S.P. & Shaw, K.L. (2014) Insect speciation rules: unifying concepts in speciation research. *Annual Review of Entomology*, **59**, 339–361.

Oliveira, C., Manfrin, M., Sene, F., Jackson, L. & Etges, W.J. (2011) Variations on a theme: diversification of cuticular hydrocarbons in a clade of cactophilic *Drosophila*. *BMC Evolutionary Biology*, **11**, 179.

Powell, T.H.Q., Cha, D., Linn, C.E. & Feder, J.L. (2012) On the scent of standing variation for speciation: behavioral evidence for native sympatric host races of *Rhagoletis pomonella* (Diptera: Tephritidae) in the southern United States. *Evolution*, **66**, 2739–2756.

Powell, T.H.Q., Hood, G.R., Murphy, M.O., Heilveil, J.S., Berlocher, S.H., Nosil, P. *et al.* (2013) Genetic divergence across the speciation continuum: the transition from host race to species in *Rhagoletis*. *Evolution*, **67**, 2561–2576.

Prokopy, R.J., Bennett, E.W. & Bush, G.L. (1971) Mating behavior in *Rhagoletis pomonella* (Diptera: Tephritidae): I. Site of assembly. *Canadian Entomology*, **103**, 1405–1409.

Prokopy, R.J., Bennett, E.W. & Bush, G.L. (1972) Mating behavior in *Rhagoletis pomonella* (Diptera: Tephritidae): II. Temporal organization. *Canadian Entomology*, **104**, 97–104.

Prokopy, R.J. & Bush, G.L. (1973) Mating behavior in *Rhagoletis pomonella* (Diptera: Tephritidae): IV. Courtship. *Canadian Entomologist*, **105**, 873–891.

Prokopy, R.J. (1975) Mating behavior in *Rhagoletis pomonella* (Diptera: Tephritidae): V. Virgin female attraction to male odor. *Canadian Entomologist*, **107**, 905–908.

R Core Team (2019) *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/. [accessed on 1 April 2019].

Reed, J.R., Hernandez, P., Blomquist, G.J., Feyereisen, R. & Reitz, R.C. (1996) Hydrocarbon biosynthesis in the house fly, *Musca domestica*: substrate specificity and cofactor requirement of *P*450hyd. *Insect Biochemistry and Molecular Biology*, **26**, 267–276.

Reissig, W.H. & Smith, D.C. (1978) Bionomics of *Rhagoletis pomonella* in *Crataegus*. *Annals of the Entomological Society of America*, **71**, 155–159.

Riesch, R., Muschick, M., Lindtke, D., Villoutreix, R., Comeault, A.A., Farkas, T.E. *et al.* (2017) Transitions between phases of genomic differentiation during stick-insect speciation. *Natire Ecology and Evolution*, **1**, 0082.

Rundle, H.D., Chenoweth, S.F., Doughty, P. & Blows, M.W. (2005) Divergent selection and the evolution of signal traits and mating preferences. *PLoS Biology*, **12**, 1988–1995.

Sarles, L., Verhaeghe, A., Francis, F. & Verheggen, F.J. (2015) Semiochemicals of *Rhagoletis* fruit flies: potential for integrated pest management. *Crop Protection*, **78**, 114–118.

Schwander, T., Arbuthnott, D., Gries, R.G., Nosil, P. & Crespi, B.J. (2013) Hydrocarbon divergence and reproductive isolation in *Timema* stick insects. *BMC Evolutionary Biology*, **13**, 15.

Schwarz, D. & McPheron, B.A. (2007) When ecological isolation breaks down: sexual isolation is an incomplete barrier to hybridization between *Rhagoletis* species. *Evolutionary and Ecological Research*, **9**, 829–841.

Sharma, K.R., Enzmann, B.L., Schmidt, Y., Moore, D., Jones, G.R., Parker, J. *et al.* (2015) Cuticular hydrocarbon pheromones for social behavior and their coding in the ant antenna. *Cell Reports*, **12**, 1261–1271.

Sim, S.B., Mattsson, M., Feder, J.L., Cha, D.H., Yee, W.L., Goughnour, R.B. *et al.* (2012) A field test for host fruit odour discrimination and avoidance behaviour for *Rhagoletis pomonella* flies in the western United States. *Journal of Evolutionary Biology*, **25**, 961–971.

Sivinski, J. & Burk, T. (1989) Reproductive and mating behaviour. *Fruitflies, Their Biology, Natural Enemies and Control*, Vol. **3A** (ed. by A. S. Robinson and G. Hooper), pp. 343–351. Elsevier, Amsterdam, The Netherlands.

Smith, D.C. & Prokopy, R.J. (1982) Mating behavior of *Rhagoletis mendax* (Diptera: Tephritidae) flies in nature. *Annals of the Entomological Society of America*, **75**, 388–392.

Smith, D.C. (1984) Feeding, mating, and oviposition by *Rhagoletis cingulata* (Diptera: Tephritidae) flies in nature. *Annals of the Entomological Society of America*, **77**, 702–704.

Smith, D.C. (1985) General activity and reproductive behavior of *Rhagoletis cornivora* (Diptera: Tephritidae) flies in nature. *Journal of the New York Entological Society*, **93**, 1052–1056.

Smith, D.C. (1986) Genetics and reproductive isolation of *Rhagoletis* flies. PhD dissertation, University of Illinois at Urbana-Champaign.

Smith, J.J. & Bush, G.L. (2001) Phylogeny of the subtribe Carpomyina (Trypetinae), emphasizing relationships of the genus *Rhagoletis*. *Fruit*

Flies (Tephritidae): Phylogeny and Evolution of Behavior (ed. by M. Aluja and A. L. Norrbom), pp. 187–218. CRC Press, Boca Raton, Florida.

Sobel, J.M., Chen, G.F., Watt, L.R. & Schemske, D.W. (2010) The biology of speciation. *Evolution*, **64**, 295–315.

Spieth, H.T. (1974) Courtship behavior in *Drosophila*. *Annual Review of Entomology*, **19**, 385–405.

Stamps, G.F. & Shaw, K.L. (2019) Male use of chemical signals in sex discrimination of Hawaiian swordtail crickets (genus *Laupala*). *Animal Behaviour*, **156**, 111–120.

Stennett, M.D. & Etges, W.J. (1997) Pre-mating isolation is determined by larval rearing substrates in cactophilic *Drosophila mojavensis*. III. Epicuticular hydrocarbon variation is determined by use of different host plants in *Drosophila mojavensis* and *Drosophila arizonae*. *Journal of Chemical Ecology*, **23**, 2803–2824.

Symonds, M.R.E. & Elgar, M.A. (2008) The evolution of pheromone diversity. *Trends in Ecology and Evolution*, **23**, 220–228.

Taylor, R.C., Klein, B.A., Stein, J. & Ryan, M.J. (2011) Multimodal signal variation in space and time: how important is matching a signal with its signaler? *Journal of Experimental Biology*, **214**, 815–820.

Toolson, E.C., Markow, T.A., Jackson, L.L. & Howard, R.W. (1990) Epicuticular hydrocarbon composition of wild and laboratory-reared *Drosophila mojavensis* Patterson and Crow (Diptera: Drosophilidae). *Annals of the Entomological Society of America*, **83**, 1165–1176.

Vaníčková, L., Svatos, A., Kroiss, J., Kaltenpoth, M., Nascimento, R.R.D., Hoskovec, M. *et al.* (2012) Cuticular hydrocarbons of the South American fruit fly *Anastrepha fraterculus*: variability with sex and age. *Journal of Chemical Ecology*, **38**, 1133–1142.

Venables, W.N. & Ripley, B.D. (2002) *Modern Applied Statistics with S-PLUS*, 4th edn. Springer, New York, New York.

Yee, W.L., Chapman, P.S., Sheets, H.D. & Unruh, T.R. (2009) Analysis of body measurements and wing shape to discriminate *Rhagoletis pomonella* and *Rhagoletis zephyria* (Diptera: Tephritidae) in Washington state. *Annals of the Entomological Society of America*, **102**, 1013–1028.

Yew, J.Y., Dreisewerd, K., de Oliveira, C.C. & Etges, W.J. (2011) Male-specific transfer and fine scale spatial differences of newly identified cuticular hydrocarbons and triacylglycerides in a *Drosophila* species pair. *PLoS One*, **6**, e16898.

Zhang, L., Hood, G.R., Roush, A.M., Shzu, S.A., Comerford, M.S., Ott, J.R. *et al.* (2020) Asymmetric, but opposing reductionsin immigrant viability and fecundity promote reproductive isolation among host-associated populations of an insect herbibore. *Evolution*, **75**, 476–489.

Accepted 14 October 2021

Associate Editor: Rob Knell