SPECIAL ISSUE: HOST RACE FORMATION AND SPECIATION

A test of adaptive divergence in a newly discovered host association of the soapberry bug *Jadera haematoloma* on Mexican buckeye, *Ungnadia speciosa*

Mattheau S. Comerford[1]* iD, Scott P. Carroll[2,3] iD & Scott P. Egan[1] iD

[1]Department of BioSciences, Rice University, Houston, TX, 77005, USA, [2]Institute for Contemporary Evolution, Davis, CA, 95616, USA, and [3]Department of Entomology & Nematology, University of California, Davis, CA, 95616, USA

Abstract

Host races represent an important step in the speciation process of phytophagous insects as they reflect the maintenance of genetically divergent host-associated populations in the face of appreciable gene flow. The red-shouldered soapberry bug, *Jadera haematoloma* (Herrich-Schäffer) (Hemiptera: Rhopalidae), is an oligophagous seed predator with a history of host race evolution on plant associations in the (soapberry) family Sapindaceae. Soapberry bugs are a model group for understanding rapid ecological adaptation to their hosts, and hence good candidates for investigating evolutionary divergence in host associations over short timescales. Here, we describe the recent discovery of Mexican buckeye, *Ungnadia speciosa* Endl., as a host of *J. haematoloma* in a region of the Chihuahuan desert including west Texas and southeastern New Mexico, USA. This host differs from *J. haematoloma*'s previously recorded hosts in the Sapindaceae in seed chemistry, ecology, and phylogeny. The tendency toward rapid, host-associated adaptations by populations of *J. haematoloma* and the unique biology of the newly discovered *Ungnadia* host create the opportunity for potential host race formation, as it overlaps geographically with two previously recorded host plants in this region – the native western soapberry tree, *Sapindus saponaria* var. *drummondii* (Hook & Arn.), and the non-native goldenrain tree, *Koelreuteria paniculata* Laxm. We explore the possibility of host race formation on *Ungnadia*-associated insects by testing for host-associated differentiation in morphology and feeding behaviors. We find evidence of differentiation in the length of the mouth parts, which is an ecologically relevant feeding trait between host plant species with larger or smaller seed capsules. This divergence is maintained in the face of potential gene flow by reproductive isolation in the form of habitat isolation, which we detect in host preference trials. Together, our results demonstrate that soapberry bugs associated with this newly discovered host exhibit morphological and behavioral traits consistent with host race formation, but additional work is required to confirm its state along the speciation continuum.

Introduction

The study of speciation in herbivorous insects focuses on the role of the host plant environment in generating a source of differentiating selection between populations and its potential role in promoting reproductive isolation

(Funk et al., 2002). The study of 'host races', which were defined by Drès & Mallet (2002) as genetically differentiated, sympatric populations of parasites using different hosts with measurable gene flow between them, has played an important role in speciation research (Coyne & Orr, 2004; Funk, 2012). Host races are commonly treated as evidence of 'speciation-with-gene-flow' as they demonstrate that host-associated differentiation can promote reproductive isolation in the face of the homogenizing effects of gene flow (Berlocher & Feder, 2002; Coyne & Orr, 2004; Medina, 2017). What makes the study of host

*Correspondence: Mattheau S. Comerford, Department of BioSciences, Anderson Biological Laboratories, Rice University, 6100 Main Street, Houston, TX 77005, USA. E-mail: msc10@rice.edu

race formation particularly appealing is that it represents a potential intermediate stage between biotypes and sister species in the speciation continuum, allowing us to capture evolution in progress (Berlocher & Feder, 2002; Drès & Mallet, 2002; Funk, 2012). Examples of well-documented host race formation by dietary specialists include the *Rhagoletis* apple maggot on apple and hawthorn trees (Feder et al., 1988), *Eurosta* gall flies on various species of goldenrod (Craig et al., 1993), *Belonocnema* gall wasps on sister species of live oaks (Egan et al., 2012a; Hood et al., 2019; Zhang et al., 2021), and *Acyrthosiphon* pea aphids on alfalfa and clover (Via, 1999). Each of these systems has made important contributions to our understanding of how new host plant-associated insect species evolve.

However, host races are just one 'stage' in the evolutionary process for host-associated insects along the speciation continuum. The study of herbivore populations on distinct host plant species, especially when those host plants are distantly related, can help in uncovering the mechanisms initiating the speciation process (Funk et al., 2002). Thus, alternative host plants provide the platform for divergent selection and non-random gene flow (Edelaar & Bolnick, 2012). This includes the study of specific reproductive barriers associated with host use, such as sexual isolation (Funk, 1998; Nosil et al., 2002; Egan et al., 2012a), allochronic isolation (Wood, 1993; Wood et al., 1999; Hood et al., 2019), and immigrant inviability (Nosil et al., 2005; Zhang et al., 2021).

Capturing novel host shifts, host-associated divergence, and subsequent formation of host races can be challenging (Berlocher & Feder, 2002; Drès & Mallet, 2002). This is partially because it is difficult to disentangle the past biogeographical history of both the host plant and their associated herbivores (Gildenhuys et al., 2013; Wisz et al., 2013). For this reason, non-native or invasive plants that accumulate native herbivores have become valuable models in the study of host-associated divergence and host race formation. Invading plant species provide colonization opportunities for native insects attracted to a potential new resource. The discovery of previously undocumented, but older host associations may likewise be valuable as they may represent reservoirs of standing genetic variation within the insect species that inhabit them. Moreover, when these newly discovered hosts undergo range expansions, they present a unique paradigm for evolutionary inquiry as they provide pairwise comparisons between both the ancestral hosts in their native distribution, and unrealized or newly derived hosts in their expanded range (Xie et al., 2007). Cataloging host-associated insect differentiation along this gradient allows us to observe the very adaptations that either directly or indirectly lead to reproductive isolation between host-associated populations.

Host-associated populations of the red-shouldered soapberry bug, *Jadera haematoloma* (Herrich-Schäffer) (Hemiptera: Rhopalidae, Serinethinae), have become a leading example of rapid contemporary adaptive diversification (Carroll et al., 2001). Historically, *J. haematoloma* fed on seeds of native plants in the soapberry family (Sapindaceae) across the southern USA, but has expanded its host range to include related ornamental Asian *Koelreuteria* trees that were introduced to North America in the mid-20th century. Owing to differences in seedpod size between native and non-native (*Koelreuteria*) hosts, the insect's mouth parts evolved in length accordingly, along with several other host-associated traits (Carroll & Boyd, 1992; Carroll & Fox, 2007). Most of the studies on this topic in *J. haematoloma* come from work in Florida, USA, although the bugs are distributed across the entire southern USA.

In the southcentral and southwestern USA, including Texas and New Mexico, *J. haematoloma* has been documented to feed on native sapindaceous host plants from the infrafamily plant group Paullinia (Buerki et al., 2009), including: (1) the western soapberry tree, *Sapindus saponaria* var. *drummondii* (Hook & Arn.), distributed across most of Texas, (2) the littlefruit slipplejack, *Serjania brachycarpa* (A. Gray), distributed in a restricted area in south Texas along the Rio Grande River (Carroll & Boyd, 1992; Carroll & Fox, 2007), and (3) native balloon-vines in the genus *Cardiospermum*. The historical biogeography of *Cardiospermum* is unclear as changes in its distribution in Texas and Oklahoma predate the arrival of European settlers (Gildenhuys et al., 2013). Regardless, *J. haematoloma* has had the opportunity to interact with *Cardiospermum* for more than a few hundred years in this region. In addition to native soapberry plants, non-native tree species in the genus *Koelreuteria* (infrafamily plant group Koelreuteria; Buerki et al., 2009) have been introduced to the region and *J. haematoloma* populations associated with this recent arrival (1950–1960) in Texas exhibit a change in beak length that mirrors the well-studied contemporary host shift in Florida (Carroll & Boyd, 1992).

Interestingly, there is an additional native sapind that is quite abundant in Texas: Mexican buckeye, *Ungnadia speciosa* Endl. (infrafamily plant group Delavaya; Buerki et al., 2009). Long suspected as a potential host to *J. haematoloma*, it was surveyed repeatedly without success since 1990 (SP Carroll, unpubl.). However, *U. speciosa* spans the region from the Gulf Coast to the Trans-Pecos (over 1200 km) and had not been intensively sampled across its entire range. Here, we document the discovery that *U. speciosa* is a host of *J. haematoloma* (Figure 1) and test for host-associated differentiation of the insects found on this plant, including morphological and behavioral

Figure 1 The exposed seed of a damaged *Ungnadia speciosa* seedpod is being fed on by a group of *Jadera haematoloma* in the Davis Mountains in west Texas, USA.

differences consistent with local adaptation to *U. speciosa*. We address these topics in the context of understanding divergent host use, potential host race formation, and incipient speciation.

Material and methods

Natural history of *Jadera haematoloma*

The red-shouldered soapberry bug is a dietary specialist granivore whose North American diet breadth is limited to the seeds of host plants within the Sapindaceae (Carroll & Loye, 2012). As a dietary specialist, *J. haematoloma* populations have evolved intimate relationships with their hosts, on and around which all growth and reproduction occur. Until recently, the breeding range of *J. haematoloma* in the USA was limited to the southcentral USA (mainly Texas, New Mexico, and Oklahoma) along with a disjunct population in south Florida (Carroll, 1988; Hoffman & Steiner, 2005). This southerly range limit was dictated by the distribution of its primary host plants *S. saponaria* var. *drummondii* in the southwest and *Cardiospermum corindum* (L.) in south Florida. However, the movement or introduction of exotic species in the soapberry family and subsequent host shifts by *J. haematoloma* have resulted in a

series of range expansions in recent history (Carroll, 2008). Specifically, *J. haematoloma* moved onto the non-native *Koelreuteria elegans* (Seem.) and the more cold-tolerant species *Koelreuteria paniculata* Laxm. and *Koelreuteria bipinnata* Franch., east Asian trees that were widely introduced from southeast Asia beginning in the middle of the 20th century (Carroll & Boyd, 1992). It was this series of contemporary host shifts on to *Koelreuteria* spp. that has resulted in an explosive expansion of the insect's home range, which now reflects the distribution of its derived non-native hosts spanning most of the USA from New York to California (Hoffman & Steiner, 2005). With each new host shift, host-associated divergence of *J. haematoloma* populations soon followed. Novel host-associated populations of the insect exhibited morphological differentiation corresponding with their associated host plants' seedpod characteristics. Specifically, differences have been observed in the length of the insect's straw-like labium beak, which serves as a feeding apparatus used to penetrate and feed on the seeds deep within their host plant's seedpods (Carroll & Loye, 1987; Carroll et al., 2001).

Natural history of *Ungnadia speciosa*

The suspected host plant *U. speciosa* is the only member of this monotypic genus and is endemic to the southern USA and northern Mexico (Vines, 1984), including the Edwards Plateau in central Texas, the Trans-Pecos in west Texas, and more generally to the Chihuahuan desert in west Texas, southern New Mexico, and north-central Mexico. However, starting in the 1950s and growing in popularity at the end of the century, due to its attractive fragrant purple flowers and tolerance to drought, *U. speciosa* has become a common ornamental plant throughout the states of Texas, Oklahoma, and New Mexico (Gilman & Watson, 2011). It is a shrubby deciduous tree that can grow to 10 m high, but rarely reaches 5 m in the wild. It produces three-lobed fruit pods, where each pod contains 1–3 seeds that reach 15 mm in diameter and persist on the tree through winter. The seeds of *U. speciosa* contain toxic cyanolipids, a compound unique to the Sapindaceae (Nahrstedt, 1985). When damaged, *U. speciosa* plant tissues also exude toxic cyanolipid-laden latexes that have likely restricted the herbivore community that attacks the plant and its seeds (Lara-Villalon et al., 2017).

Field sampling

Between 2018 and 2020, we searched 73 *U. speciosa* sites across their geographic distribution for the presence of *J. haematoloma* (Figure 2A; Table S1). When insects were detected, both adults and nymphs were collected when possible and returned to the laboratory where they were placed in shoebox-sized plastic rearing bins. Laboratory insects were
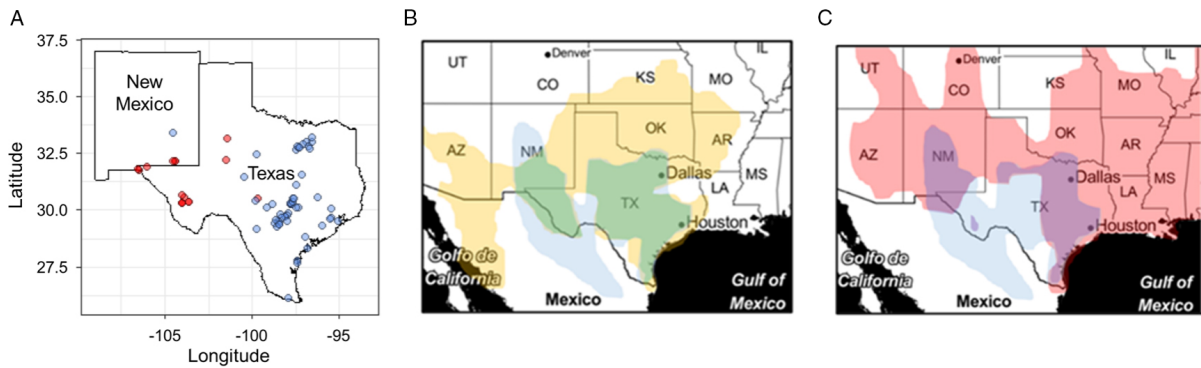
Figure 2 Ranges of *Jadera haematoloma* and its host plants (*Ungnadia speciosa*, *Sapindus saponaria* var. *drummondii*, and *Koelreuteria paniculata*). (A) Sampling locations of *U. speciosa* plants across Texas and New Mexico, USA, where *J. haematoloma* was found (red dots) or not (blue dots) on the plant. (B) Range map of the native host plant *S. saponaria* var. *drummondii* (yellow) and novel host *U. speciosa* (blue) and their overlap (green). (C) Range map of the non-native host plant *K. paniculata* (red) and novel host *U. speciosa* (blue) and their overlap (purple).

maintained on ad libitum seed diets of their native host's seeds. Cardboard egg cartons were included in each rearing container to provide shelter habitat for the insects. As we explored and collected *J. haematoloma* on *U. speciosa*, we also collected *J. haematoloma* on other soapberry family host plants in the areas surveyed (*S. saponaria* var. *drummondii*: n = 9 sites; *K. paniculata*: n = 4 sites).

Herbarium records of regional host plant ranges

To assess the historical distributions of *J. haematoloma*'s host plants in this region, we compiled herbarium records of *U. speciosa*, *S. saponaria* var. *drummondii*, and *K. paniculata*, from 33 regional herbaria (Table S2). Density plot maps along with overlapping host range maps were created using the geographical coordinates of herbarium records and the R package 'leaflet'. To assess current and past ranges of the native hosts we compiled the lists of the earliest USA herbarium record for each of the plants at the county level (*S. saponaria* var. *drummondii*: n = 106; *U. speciosa*: n = 76). These records were used to develop a color-by-date map of Texas and New Mexico (Figure S1A and B).

Testing for physiological differences between host plant seedpods

Beyond chemical defenses, sapindaceous plants regularly employ air-filled seedpods and hard seedcoats as physical barriers to protect their seeds from predators. We assessed the depths in which the beak of a feeding *J. haematoloma* would be required to penetrate to gain access to the embryo of its two native hosts (*S. saponaria* var. *drummondii* and *U. speciosa*). This assessment was made by using the depth gauge of a IP54 digital caliper (iGaging, San Clemente, CA, USA) inserted through the seedpod until contact with the seed was made. The seed was then removed and manually cracked in half and the seed coat

was measured with a Leica M125 microscope at 10× magnification and LAS v.4.4 software (Leica, Wetzlar, Germany). The distances from the surface of the seedpod to the seed and the seed coat were combined as a measure of total 'feeding depth'.

Testing for host plant-associated differences in *Jadera haematoloma* morphology

Past studies have demonstrated the ecological significance of insect proboscis length in *Jadera* species, as it directly reflects their ability to access the well-protected seed located deep inside the pods and fruits of their hosts (Carroll & Loye, 1987). Here, we measured beak length of wild-type adults of both sexes collected from all geographically overlapping host plant associations [*U. speciosa*: n = 194 (95 males, 99 females), *S. saponaria* var. *drummondii*: n = 218 (110 males, 108 females), *K. paniculata*: n = 96 (62 males, 34 females)]. In addition, pronotum widths and hind tibia lengths were measured as a control for overall insect body size. We used a Leica M125 microscope at 8× magnification and LAS v.4.4 software (Leica) to digitally measure each insect morphological trait. For analysis, insects were grouped by their sex and by an important flight polymorphism, where some individuals are macropterous with fully developed wings and others are brachypterous with reduced wings and are incapable of flight (Carroll et al., 2003). This flight polymorphism is an important consideration when measuring body size, as it correlates with changes to overall body size and other dimensions, including beak length.

Cross-rearing study to test the evolution of trait differences

Past studies have demonstrated that beak length differences between *J. haematoloma* host races have a genetic

basis (narrow-sense heritability: $h^2 = 0.50$–0.87; Carroll et al., 2001; Dingle et al., 2009). However, maladaptive plasticity has also been implicated in masking the divergence in beak length when insects consume the non-native host (Carroll, 2008; Cenzer, 2017). Here, we test for the evolution of feeding trait differences between host-associated insects via a cross-rearing experiment that compares beak length differences of insects that were raised on alternative host plant seeds in controlled laboratory conditions. In doing so, we tested the role of diet on observed host-associated differences among *J. haematoloma* populations.

To accomplish this, unmated wild-type penultimate nymphs were collected from *U. speciosa* and *K. paniculata* and mate-paired at maturation with individuals from the same host plant. The mated gravid females were allowed to oviposit into a Petri dish full of sterile sand. Eggs were provided 3 weeks to develop and hatch, with newly hatched individuals removed daily to prevent egg cannibalism. Once collected, neonatal siblings were divided evenly between rearing cages on the two hosts (*U. speciosa* and *K. paniculata*). Each cage was stocked with water and seeds from one of the two species. Dietary treatments consisted of ad libitum access to *K. paniculata* or *U. speciosa* seeds, which were mechanically cracked in half, removing seed coat as a potential source of reduced survival and selection on beak length. Rearing was conducted synchronously in ambient light conditions under a sunshade in the greenhouse, at a constant temperature of 25 °C. Rearing cages were rotated daily to help homogenize external effects that might affect insect growth. Once nymphs reached adulthood, macropterous females were collected and their beak lengths were measured. Dietary contributions to insect beak length were assessed by comparing mean beak lengths between insects raised on the seeds of their natal vs. alternative hosts.

Testing *Jadera haematoloma* host seed preference

To test for host seed feeding preference in *J. haematoloma*, we conducted a four-way choice experiment utilizing 10-cm-diameter Petri dishes for preference arenas lined with sterile sand. Adult insects were individually placed in arenas and simultaneously presented with four seed options. Three of the seeds (*U. speciosa*, *S. saponaria* var. *drummondii*, and *K. paniculata*) are ecologically relevant to the insects as they are viable host options in the region of this study. The fourth (*Cardiospermum halicacabum* L.) is within the general host range of the insect but does not naturally occur in our study region of west Texas or southern New Mexico (Figure 2B and C; Figure S2). Seeds were evenly placed 5 cm apart so that their bisecting axis would form a cross. Sixteen preference arenas arranged in 4 × 4

columns and rows were included in each trial. To account for potential external stimuli, the spatial orientation of seed type within each subsequent Petri dish in the trial was rotated 90° between assays. Prior to the experiment, seeds were tested for viability via water submersion. Seeds that floated were discarded as they lack intact embryos.

Preference tests were conducted with wild-type adults collected from all three hosts that inhabit the study area. Prior to the experiments, subjects were placed on a 48-h water-only diet to encourage active feeding behavior. At the onset of each trial, individual insects were placed at the center of the Petri dish and allowed a 15-min acclimation period, followed by a 60-min assessment period (Figure S3). During assessment, insect feeding behavior was recorded every 5 min for a total 12 observations. Insects that failed to move during the experiment were excluded from analysis. Each subject was assessed under two preference metrics: did individuals from each host plant differ in (1) their propensity to visit the *Ungnadia* seed, and (2) the time (no. of observations, ranging from 0 to 12) they spent on each seed?

Statistical analysis

To assess whether the current home range of the native plants has experienced a contemporary eastward expansion, we performed a linear regression in which we test whether the 'longitude' of herbarium record predicted the 'year' of the plant's earliest herbarium occurrence. All statistical analyses and data illustrations were conducted with R v.3.6.1 (R Core Team, 2019). The linear regression was performed using the 'lme4' package in R (Bates et al., 2015). Additionally, we conducted χ^2 tests of independence between the presence and absence of insects at host plant sample sites located in central/east Texas vs. west Texas, using the Pecos River (longitude: -101.3507), the traditional demarcation of the Trans-Pecos ecoregion, as our dividing line (Hill, 1887).

Next, we tested for differences in morphology and host feeding preference among host-associated populations of *J. haematoloma* on *U. speciosa*, *S. saponaria* var. *drummondii*, and *K. paniculata*. To assess how host plant association influences morphological traits of beak length and body size (thorax width and tibial length), we used generalized linear mixed models (GLMM) with a normal error distribution. All morphological traits (beak length, thorax width, and tibia length) were included as response variables for their own model, with sex, wing morphology, and host association as predictor variables while allowing random intercepts by sample site. All GLMMs were performed using the 'lme4' package in R with P-values calculated via the 'lmerTest' package (Kuznetsova et al., 2017). Post-hoc pairwise Tukey's tests were implemented to

assess differences between predictor levels through the 'emmeans' package (Length, 2021).

We also evaluated the physical differences between host plant seed defense using a GLMM. We assessed if host plant species differed in the distance an insect beak must penetrate to feed on the seed. In this model, the response variable was the total 'feeding depth' (seed depth + seed coat) and the predictor variable was the 'host' plant species (*U. speciosa* vs. *S. saponaria* var. *drummondii*). We included a random intercept of individual plant.

To analyze the effect of diet on *J. haematoloma* trait differentiation between host-associated populations, we built a GLMM to test for beak length variation between population and rearing diet assuming a normal distribution. Again, beak length was included as the response variable with host association (*Ungnadia* vs. *Koelreuteria*) and insect 'diet' (*U. speciosa* vs. *K. paniculata* seeds) as predictor variables. We included collection site as a random intercept in the model. Sex and wing morphology were not included as factors in this analysis, as only macropterous females were included in this comparison. Post-hoc Tukey's tests were performed on pairwise comparisons between predictor levels.

Lastly, we assessed how host association influences insect seed preference using two approaches. The first was a test of each host-associated populations propensity to visit the newly discovered host *U. speciosa*. To accomplish this, we employed a χ^2 test comparing the number of insects from each population that visited *U. speciosa* seeds vs. those that did not in our four-way choice assay. Second, we assessed if host association influences the amount of time insects spent on their natal seed (the seeds from which the insect was collected) vs. the alternative three seed types. To accomplish this, we used a cumulative link mixed model (CLMM) with a Poisson distribution, which is an extension of a logistic regression that accounts for differences between levels. This statistical technique has proven effective in testing food-patch preferences in *Drosophila melanogaster* Meigen (Geiger & Saltz, 2019) and is a common approach utilized for ordinal based psychological studies (Christensen & Brockhoff, 2013). Our model is based on a fixed-time incrementally assessed multi-option preference test. We modeled 'time spent feeding' as our response variable ranging from 0 to 12, with each value representing the number of observations in which an insect fed on the relevant seed. Each individual was characterized by four measurements of 'time spent feeding', one for each host seed, with the variable 'seed type' describing which observation corresponded to each of the four seed options (*K. paniculata*, *S. saponaria* var. *drummondii*, *U. speciosa*, and *C. halicacabum*). Our predictor variables included 'seed type' and 'host association',

allowing 'collection site' and 'individual insect' as random intercept effects that account for population level variation and the non-independence of the four measures of seed type collected from each individual. Initially, insect wing morphology and sex were considered as potential variables, but were found to have a negligible contribution both as a main effect and as an interaction, thus they were ultimately dropped from the final model. We performed CLMM using the 'ordinal' package in R (Christensen, 2019) while estimating the significance of the fixed effects and interactions using the joint_test function χ^2 test of the 'emmeans' package.

Results

Historical biogeography of native host plants

In this study, we document our discovery of *J. haematoloma* using *U. speciosa* as a host for feeding and reproduction at 19 localities across multiple years (Figures 1 and 2A). Those occurrences span an area of about 10000 km^2 of west Texas and southern New Mexico in the Chihuahuan desert region. However, this newly discovered association does not appear to be constant across the plant's range, which may explain why *J. haematoloma* had not been found on *U. speciosa* prior to this study. Whereas 80% (16/20) of *U. speciosa* sites sampled in the western portion of the range (the Trans-Pecos region) hosted *J. haematoloma*, occurrence dropped to a mere 5.6% (3/53) in the eastern part of the plant's range (Edwards Plateau, Panhandle, and coastal Texas) ($\chi^2 = 41.68$, d.f. = 1, P<0.001).

Herbarium records indicate the contemporary range of *U. speciosa* includes most of Texas, as well as parts of adjacent New Mexico in the USA and the Mexican states of Chihuahua and Coahuila. Outside of cultivation and botanical gardens, the approximate western limit of *U. speciosa* is indicated by a 1932 collection of a specimen in Dona Ana County, New Mexico (32.352648, −106.832721) (New York Botanical Garden Steere Herbarium, New York, NY, USA; collection ID 2491284). The northernmost specimen was collected in 1995 in Cooke County, Texas (33.6392, −97.212626) (BRIT Philecology Herbarium, Fort Worth, TX, USA; collection ID BRIT48030). The easternmost specimen was observed in Harris County, Texas, in 2019 (29.79373, −94.9548; iNaturalist; collection ID 21647640). The earliest herbarium record of *U. speciosa* in the USA is an 1895 specimen that matches the western limit in Dona Ana County, NM (Indiana University Deam Herbarium, Bloomington, IN, USA; collection ID IND-0079465).

Early records of *U. speciosa* were predominantly found in western Texas and southern New Mexico; however, it

was not until after 1965 that the plant appears in herbarium records to be common throughout central Texas, and to grow sporadically through the western portions of the Edwards Plateau (Figure S1B). In fact, 81.4% of the 1778 herbarium records we reviewed were collected along the USA Interstate Highway corridor connecting the centrally located cities of San Antonio, Austin, and Dallas/Fort Worth. Yet only 13 (ca. 0.9%) of these centrally located plant records occurred prior to the year 2000.

In contrast, *S. saponaria* var. *drummondii* herbarium collections show the plant was commonly and broadly observed across Texas and New Mexico throughout the 19th century (Figure S1A). These records show that the *S. saponaria* var. *drummondii* range far exceeds *U. speciosa* as it is common in Arizona, New Mexico, Texas, Oklahoma, and Kansas, USA (Figure S2). The first recorded occurrence of *S. saponaria* var. *drummondii* in the USA was in the east Texas town of Goliad in 1834 (Harvard University Herbarium, Cambridge, MA, USA; collection ID 1870069) with 18 other county records sporadically occurring east of the Pecos River prior to the turn of the century. Moreover,

S. saponaria var. *drummondii*'s earliest occurrences were not limited to only central and east Texas, as pre-1900 collections of the plant were also common in west Texas and central New Mexico.

Using these historical herbarium records, we tested for patterns of contemporary range expansions in both native host-associated plants (Figure S1C and D). In doing so, we found a strong association between longitude and date of first collection for *U. speciosa*, consistent with our prediction of eastern expansion (r = 0.75, P<0.001; Figure S1D). In contrast, this was not the case for *S. saponaria* var. *drummondii*, where no spatial-temporal pattern was observed in the earliest county records by longitude (r = 0.06, P = 0.55; Figure S1C).

Morphological differentiation in *Jadera haematoloma*

Host plant association had a marked effect on *J. haematoloma* beak length differentiation on all three host plants. Beak lengths of *Ungnadia* bugs in nature were larger than beaks on native *Sapindus* and non-native *Koelreuteria* (Figure 3A, Table 1). Sex and wing polymorphism were also
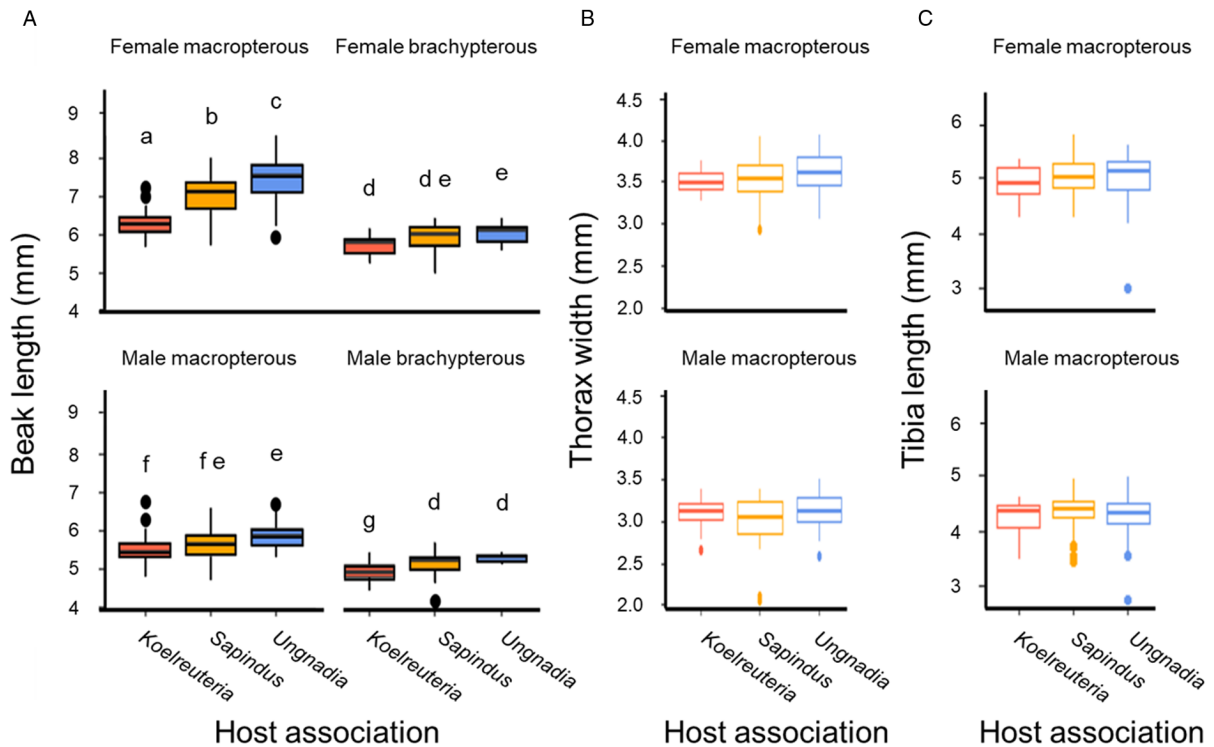
Figure 3 Box plots of morphological trait differentiation observed for both sexes of *Jadera haematoloma* by host plant association in nature: (A) beak length (mm) for macropterous and brachypterous insects, and macropterous insect (B) thorax width (mm), and (C) tibia length (mm). Colors represent host plant association: *Koelreuteria paniculate* (red), *Sapindus saponaria* var. *drummondii* (yellow), and *Ungnadia speciosa* (blue). The upper and lower boxes indicate the first and third quartile, the thicker line in between indicates the median value, and the whiskers show 1.5× the interquartile range. The dots are outliers. Means capped with different letters are significantly different (Tukey's test: P<0.05).

Table 1 Morphological divergence in *Jadera haematoloma* beak length: GLMM results of phenotypic differentiation in beak length between the fixed effects host association, sex, and wing polymorphism as predictors and collection site as a random effect

Predictor	SS	d.f.	MS	F	P
(Intercept)	718.74	1	718.74	4086.75	<0.001
Host association (H)	22.87	2	11.44	65.03	<0.001
Sex (S)	8.34	1	8.34	47.45	<0.001
Wing polymorphism (W)	3.25	1	3.25	18.48	<0.001
H*S	5.86	2	2.93	16.67	<0.001
H*W	2.63	2	1.31	7.48	0.001
S*W	0.01	1	0.01	0.08	0.78
H*S*W	1.64	2	0.82	4.66	0.010
Error	87.23	496	0.18		

strong predictors of insect beak length, as females generally had 17% longer beaks than males (least square mean ± SE = 6.7 ± 0.04 vs. 5.6 ± 0.03 mm), and macropterous morphs had 12.7% longer beaks than their brachypterous counterparts (6.3 ± 0.02 vs. 5.5 ± 0.04 mm). Overall, this resulted in *Ungnadia*-associated *J. haematoloma* females exhibiting 13% larger beaks than *Sapindus*-associated females and 17% larger beaks than *Koelreuteria*-associated females. Similarly, *Ungnadia*-associated *J. haematoloma* males exhibited 6% larger beaks than *Sapindus*-associated males and 9% larger beaks than *Koelreuteria*-associated males.

Interestingly, the interaction host*wing*sex was significant ($F_{2,497} = 4.46$, P = 0.01; Figure 3A, Table 1), which is due to the pattern of host-associated differentiation in beak length being the most pronounced in the macropterous females compared to brachypterous females or males of either wing morph. Pairwise comparisons (that include *Ungnadia*) of macropterous insect beak length indicated that each host association was different (Tukey's test: P<0.05). In contrast, there were no significant differences between host plant association and beak length in brachypterous insects, apart from *Koelreuteria* and *Ungnadia*-associated male insects (Tukey's test: P = 0.03).

Overall comparisons of insect body size included measures of the width of the pronotum, which covers the dorsal surface of the thorax, and tibial length. Both measures exhibited significant differentiation between the sexes, with females having a larger thorax than males by approximately 11% (mean ± SE = 3.26 ± 0.02 vs. 2.90 ± 0.02mm). Wing polymorphism was also a significant

predictor of thorax width when controlled for sex (macropterous: 3.31 ± 0.02, brachypterous: 2.84 ± 0.03 mm). Testing the effect of host plant on thorax width was complicated by the comparatively low numbers of brachypterous individuals collected on *U. speciosa* and, thus, we concentrated the analysis on macropterous females (excluding brachypterous females and all males). There was no significant effect of host on thorax width for macropterous females ($F_{2,450} = 0.87$, P = 0.42; Figure 3B, Table S3).

Tibial length was approximately 13.5% larger in female than in male insects ($F_{1,416} = 34.25$, P<0.001; Table S4). Wing polymorphism was also a good predictor of tibial length ($F_{1,416} = 9.50$, P<0.002), although neither the interaction between the two ($F_{1,416} = 1.07$, P = 0.30), nor the three-way interaction of host*wing*sex were significant ($F_{1,416} = 0.68$, P = 0.51). Tibial length of macropterous females was not affected by host plant association (*U. speciosa*: 5.01 ± 0.05, *S. saponaria* var. *drummondii*: 5.06 ± 0.05, *K. paniculata*: 4.94 ± 0.09 mm; Figure 3C).

Testing for physical differences between host plant seedpods

Whereas *U. speciosa* had a thinner seed coat than *S. saponaria* var. *drummondii* (mean ± SE = 0.308 ±0.0195 vs. 1.367 ±0.0269 mm; $\beta_{Host} = 1.058$, P<0.001), their seedpod depth was greater (3.87 ± 0.176 vs. 1.41 ±0.340 mm; $\beta_{Host} = 2.452$, P<0.001). Combined, the total 'feeding depth' (seedcoat + pod depth) an insect must penetrate with its proboscis to access seed endosperm was greatest when feeding on *U. speciosa* (4.17 ± 0.176 vs. 2.78 ± 0.337 mm; $\beta_{Host} = 1.394$, P<0.01; Figure S4), which is consistent with divergent selection between host plants.

Cross-rearing study to test the genetic basis of trait differences

Rearing diet (*K. paniculata* or *U. speciosa*) did not significantly affect adult beak length in individuals associated with either host. The beak lengths of adult macropterous females did not differ between insects reared on their natal vs. non-natal host seeds ($F_{1,152} = 0.32$, P = 0.57; Figure 4, Table 2). In contrast, the insect's natal host (i.e., the host from which their parent was collected) did have an effect on beak length ($F_{1,152} = 56.89$, P<0.001). This pattern of beak length differentiation between *K. paniculata* and *U. speciosa*-origin bugs for laboratory-reared insects was similar to the pattern observed in insects from our natural collections, with *Ungnadia*-associated insects having 11% longer beaks than *Koelreuteria*-associated insects. Interestingly, the beak length of laboratory-reared insects was shorter than that of individuals collected from natural populations, with beak length reductions on average of 14.1% in *Ungnadia*-associated adults and 9% in *Koelreuteria*-associated adults.
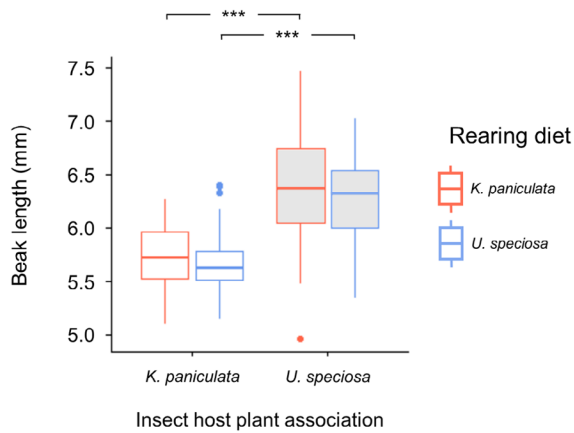
Figure 4 Box plots of effect of rearing diet host plants (*Koelreuteria paniculata* in red vs. *Ungnadia speciosa* in blue) on *Jadera haematoloma* beak length (mm). The fill (white vs. grey) represents host association. The upper and lower boxes indicate the first and third quartile, the thicker line in between indicates the median value, and the whiskers show 1.5× the interquartile range. The dots are outliers. The asterisks indicate significant differences (Tukey's test: P<0.05).

Table 2 Evolved trait differentiation between *Jadera haematoloma* host associations: GLMM results of phenotypic plasticity in beak length of macropterous female insects using host plant origin and rearing diet as predictor variables

Predictor	SS	d.f.	MS	F	P
(Intercept)	1310.83	1	1310.83	8989.06	<0.001
Origin (O)	8.30	1	8.30	56.89	<0.001
Diet (D)	0.05	1	0.05	0.32	0.57
O*D	0.04	1	0.04	0.28	0.60
Error	22.02	151	0.15		

Host-associated differences in *Jadera haematoloma* seed preference

Individuals collected from *U. speciosa* were significantly more likely to visit *U. speciosa* seeds than any other host-associated population ($\chi^2 = 7.20$, d.f. = 2, P = 0.03). Of the 211 preference trials conducted on insects collected from *U. speciosa*, 68.4% resulted in a *U. speciosa* visitation. In contrast, insects collected from *S. saponaria* var. *drummondii* (57.4%, n = 129) or *K. paniculata* (54.3%, n = 162) exhibited lower visitation rates to the novel *U. speciosa* seeds.

The three host-associated populations differed in the durations spent on hosts (Figure 5). The main effects of host association ($F_{2,1260} = 3.98$, P = 0.02) and seed type ($F_{3,1260} = 70.66$, P<0.001) were significant predictors of insect feeding time. Most important to the current study, the interaction between host-association and seed type was also highly significant ($F_{6,1260} = 6.65$, P<0.001).

Multiple pairwise comparison across host associations indicated that individuals spent significantly different amounts of time between each of the four host seeds. *Ungnadia*-associated insects spent most of their time feeding on *U. speciosa* seeds (least square mean ± SE= 4.047 ± 0.151 observations), followed by *S. saponaria* var. *drummondii* seeds (2.708 ± 0.172), and then *K. paniculata* seeds (1.844 ± 0.204), with *C. halicacabum* being their least preferred host seed (0.166 ± 0.375; Figure 5). This preference for their resident seed was so strong that *Ungnadia*-origin bugs spent on average nearly 50% more time feeding on those seeds than on the seeds of their second most preferred host *S. saponaria* var. *drummondii* (Tukey's test: P<0.001).

A different pattern was observed in *Sapindus*-associated insects, which spent their time feeding equally on *Sapindus* and *Ungnadia* (*S. saponaria* var. *drummondii*: 3.642± 0.234, *U. speciosa*: 3.611 ± 0.227, *K. paniculata*: 1.32 ± 0.351, *C. halicacabum*: 0.98 ± 0.0408 observations; Figure 5). The observed pattern of increased time spent on both regionally native sapinds decreased the time *Sapindus*-associated insects spent on the non-native *K. paniculata* compared to *Ungnadia*-associated insects (Tukey's test: P = 0.02). *Sapindus*-associated insects spent similar amounts of time on both regionally non-native sapinds, *K. paniculata* and *C. halicacabum*, which was significantly less than the time they spent on both regionally native species (Figure 5).

For *Koelreuteria*-associated insects, individuals did not demonstrate a preference towards their resident host seeds and spent a similar amount of time on all three regionally available hosts (*K. paniculata*: 2.905 ± 0.204, *S. saponaria* var. *drummondii*: 3.357 ± 0.172, *U. speciosa*: 3.504 ±0.151, *C. halicacabum*: 0.794 ± 0.372 observations; Figure 5). Specifically, they spent equal time on the two regionally native sapinds, *U. speciosa* and *S. saponaria* var. *drummondii*, and a similar amount on the non-native *K. paniculata* (Figure 5). Although they did not exhibit a strong preference for their natal host seed, individuals collected from *K. paniculata* did spend more time on *K. paniculata* than individuals from other host associations (Tukey's test: *Koelreuteria* – *Ungnadia*, P = 0.0033; *Koelreuteria* – *Sapindus*, P = 0.0026). Similar to other host-associated populations, *Koelreuteria*-associated insect preference for *C. halicacabum* was greatly reduced compared to the other seed options (all pairwise comparisons that include *C. halicacabum*, Tukey's test: P<0.001).
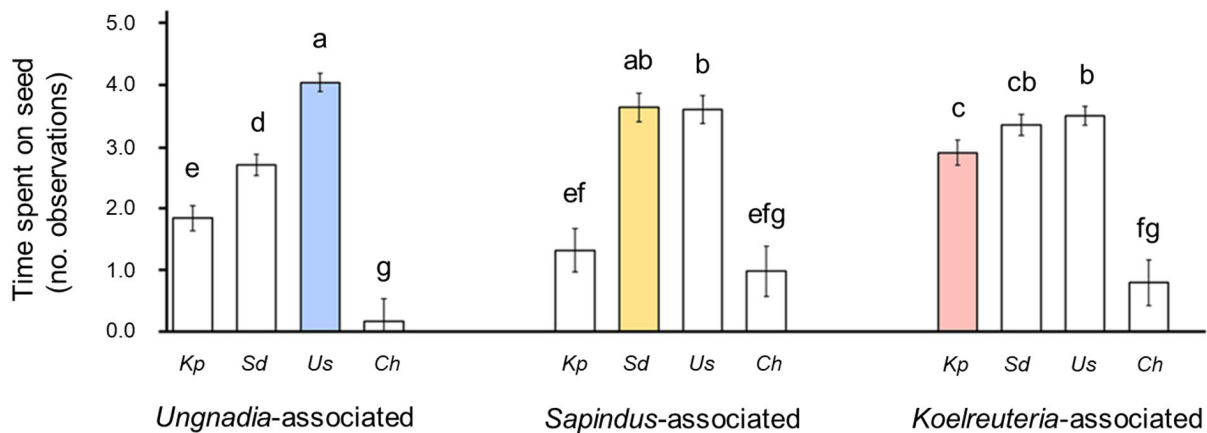
Figure 5 Relative host seed preference (mean ± SE time spent on a seed) for each *Jadera haematoloma* host association, based on a four-way choice experiment. Seeds from each of four plants – *Koelreuteria paniculata* (Kp), *Sapindus saponaria* var. *drummondii* (Sd), *Ungnadia speciosa* (Us), and *Cardiospermum halicacabum* (Ch) – were offered simultaneously to adult insects originating from one of three host plants. Colored columns represent the matching seed to *J. haematoloma* host association. Means capped with different letters are significantly different (Tukey's test: P<0.05).

Discussion

In this study, we describe the recent discovery of Mexican buckeye (*U. speciosa*), as another host plant in *J. haematoloma*'s diet in a region of the Chihuahuan desert including west Texas and southeastern New Mexico, USA. Although this is a new discovery, this host association is likely not recent, but represents an older association that has gone undetected in the rugged Chihuahuan desert region until now. *Ungnadia* differs from *J. haematoloma*'s previously recognized hosts in its plant and seed chemistry, seed capsule morphology, ecology, and evolutionary relationship to other members of the Sapindaceae. Most important for a granivorous insect may be the ability to detect and overcome the defensive secondary metabolites produced by their given host. For *Jadera* feeding on sapindaceous plants, this defensive chemistry comes in the form of toxic cyanolipids. *Ungnadia speciosa* is distinctive from *J. haematoloma*'s other hosts as it produces a unique variant of the toxin (Mikolajczak et al., 1970a; Nahrstedt, 1985). *Ungnadia speciosa* is also a unique member of the plant family Sapindaceae in comparison to previously reported hosts, based on a multilocus phylogeny, including seven chloroplast DNA regions and one nuclear ribosomal DNA region (Buerki et al., 2009). The native host in west Texas discussed here (*Sapindus*), as well as the other recorded native hosts from south Texas (*Serjania*) and east Texas (*Cardiospermum*), are all members of the infrafamily group Paullinia within the Sapindoideae, whereas the non-native *Koelreuteria* is a member of the infrafamily group Koelreuteria within the Sapindoideae. Interestingly,

the newly discovered host *U. speciosa* is part of a third infrafamily group, Delavaya, within the same subfamily Sapindoideae (Buerki et al., 2009, 2010).

The propensity of *J. haematoloma* populations to experience rapid host plant-associated adaptation (e.g., Carroll & Boyd, 1992; Carroll et al., 2001; Cenzer, 2016) is interesting considering *U. speciosa*'s relatively distant relationship to the other known host plants (Buerki et al., 2009, 2010). This new host discovery creates a unique opportunity to study adaptive divergence, early host race formation, and – potentially – host-associated speciation, as *Ungnadia*-associated bugs exhibit an appreciable degree of morphological and behavioral differences from the other local, sympatric host-associated populations. Given that *Ungnadia*-associated insects overlap geographically with two previously recorded host plants – the native western soapberry tree (*S. saponaria* var. *drummondii*) and the non-native goldenrain tree (*K. paniculata*) – this study system offers an excellent opportunity to explore the speciation continuum with populations exhibiting different degrees of divergence.

The differences we observed in mouthpart length among hosts ecologically relevant to feeding on species with larger or smaller seed capsules occurred without concomitant change in overall body size. Furthermore, this phenotypic divergence is maintained in areas of geographic overlap with other host-associated populations, potentially through the development of reproductive isolation in the form of habitat isolation, which we detected in our seed preference trials. These differences in *J. haematoloma* populations are predicted by the observed

differences in seedpod size among hosts, with *Ungnadia* seeds exhibiting the largest distance between seed pod and seed, and *Ungnadia*-associated *J. haematoloma* exhibiting the largest beak length. Together, our results suggest an additional set of populations of a model insect in evolutionary biology has divergence in adaptation to the newly discovered host plant *U. speciosa*.

Historical biogeography of *Ungnadia speciosa*

The regional component of this interaction is intriguing as *U. speciosa*'s range is completely encapsulated within the larger *J. haematoloma* distribution that spans most of the USA, Mexico, and Central America; however, *Ungnadia* does not appear to be attacked equally by *J. haematoloma* across its range. Inferences based on historical herbarium records suggest that *U. speciosa* may have undergone a contemporary range expansion from west to east associated with its adoption by the nursery and horticulture trade. Our search of herbaria records indicated a clustering of early collections of *U. speciosa* occurring in the Trans-Pecos region of Texas and the southern counties of New Mexico in the late 19th century.

Interestingly, although multiple paleolithic human caches of *U. speciosa* seeds in the Trans-Pecos region of Texas have been ^{14}C-dated to as early as 8440 BC (Adovasio & Fry, 1976), we have found no Texas records of *U. speciosa* occurring east of the Pecos River prior to 1948. The stark absence of early eastern *U. speciosa* records is unlikely an artifact of sampling bias as plant collections should have favored those locales based on much higher human populations. Furthermore, the comparisons of *S. saponaria* var. *drummondii* herbarium records lend support to this contemporary range expansion hypothesis. When contrasting these records, unlike *U. speciosa*, we observe a complete lack of geo-spatial patterning, with the earliest county occurrences appearing simultaneously across the region. Furthermore, the earliest occurrences of *S. saponaria* var. *drummondii* in central and east Texas predate the first occurrences of *U. speciosa* by more than a century. Taken together, our findings suggest that the lack of early eastern *U. speciosa* records is not the consequence of sampling biases by herbarium collectors, but rather evidence consistent with a recent range expansion. This finding is surprising, as the literature indicates a native range for *U. speciosa* that includes central Texas (Little, 1976). A potential alternative hypothesis is that pockets of *U. speciosa* have been present in the eastern part of its range the entire time, but the species has become more common with human modification to the landscape and human-aided dispersal. These alternative hypotheses of range expansion vs. changes in abundance require further testing, as we cannot

definitively differentiate between them with the current data.

Regardless, extensive insect and plant sampling across the eastern distribution of *U. speciosa* has failed to yield any evidence of a host association with *J. haematoloma* since the early 1990s, and this was verified again in the current study. What could account for this pattern? We speculate that *J. haematoloma*'s original association with *U. speciosa* predates European settlement and this old association evolved in what was once a narrower ancestral home range of *U. speciosa* in west Texas and southern New Mexico (and likely, north-central Mexico, although that area was not surveyed in the current study). Furthermore, the eastern region where *J. haematoloma* has yet to utilize *U. speciosa* as a host plant may represent a circumstance in which *J. haematoloma* has failed to keep pace behaviorally, ecologically, or evolutionarily with the range expansion of *U. speciosa*.

Ungnadia speciosa as a source of divergent selection

A major and defining component of sapindaceous seed composition is the presence of toxic cyanolipids, which can account for as much as 50% of their overall seed oil weight (Mikolajczak et al., 1970a; Seigler & Kawahara, 1976). Unlike *J. haematoloma*'s previously known host plants in this region, *S. saponaria* var. *drummondii* and the contemporary novel host plant *K. paniculata*, which share a similar cyanolipid profile (cyanolipid schemes II and III), *U. speciosa* seed oil has a profile unique to this plant (scheme IV) (Mikolajczak et al., 1970b; Table 3, Figure S5). This cyanolipid profile difference in *U. speciosa* is the result of a shift of the chemical structure of the molecule where a double bond in the parent propyl chain changes

Table 3 Cyanolipid content of *Jadera haematoloma*'s host plant seeds. Each cyanolipid scheme represents different esterification of a fatty acid chain unique to the Sapindaceae

Host plant	Cyanolipids			
	I	II	III	IV
Cardiospermum halicacabum[1]	X	Trace		
Serjania brachycarpa[2]		X		
Sapindus saponaria var. *drummondii*[1,3]		X	X	
Koelreuteria paniculata[1,2]		X	X	
Ungnadia speciosa[2]		Trace	Trace	X

See Figure S5 for structural comparison diagrams of each cyanolipid scheme. Trace levels represent <1.5% oil weight.
[1] Seigler & Kawahara (1976).
[2] Mikolajczak et al. (1970a).
[3] Mikolajczak et al. (1970b).

from the second carbon to the first, along with a loss of a hydroxyl group located on the substituent methyl group, which would normally provide an additional site of future esterification for the cyanolipid scheme II (Figure S5). Given this difference in seed chemistry between hosts, adaptation to seed chemistry may be critical for *J. haematoloma* success on *U. speciosa*. Future studies of insect feeding on host seeds with different plant chemistry is an area requiring further investigation that could reveal a major axis of ecological differentiation distinguishing *J. haematoloma* populations.

Soapberry seed chemistry is also important to the life history of *J. haematoloma*. Soapberry bugs bioaccumulate cyanolipids, exploiting them for their own purposes in predator defense and as a conspecific attractant when forming insect aggregations (Ribeiro, 1989; Aldrich et al., 1990). The integral role that these distinctive host plant compounds play in soapberry bug life history further supports cyanolipid differences in host plants as a likely source of divergent selection and a candidate feeding cue in host preference, as secondary metabolites in such situations are often phagostimulants directing host choice (Chapman, 2003).

Host-associated divergence in *Jadera haematoloma*

In the west Texas and southern New Mexico region, we found *J. haematoloma* associated with three host plants: native Mexican buckeye, native western soapberry tree, and non-native goldenrain tree. Outcomes from our host preference experiments demonstrated that *J. haematoloma* associated with *Sapindus* and *Ungnadia*, the two native host plants, exhibited a strong preference for their natal host, such that the time spent on *U. speciosa* by its associated bugs was 3× higher than what would be predicted by a null hypothesis of no preference and equal time spent on each host seed. Interestingly, insects collected from the recently introduced Asian host *K. paniculata* did not prefer their natal host, instead using all three similarly, and hence showing the highest overall usage of *K. paniculata*. Furthermore, little time was spent by any insect population on the host seed *C. halicacabum*, which is a common host in east Texas but not available in the west Texas region. Studies of the influence of maternal diet on offspring host preferences in this insect species have shown no strong effects (Carroll et al., 1998). Preference for their natal hosts represents not just a form of cryptic phenotypic differentiation, but also a potential limit to gene flow among host-associated populations by creating habitat isolation. Such preference has been shown to reduce movement and gene flow between ecologically divergent populations in many insect herbivores, including *Belonocnema* gall wasps (Egan et al., 2012b, 2013), *Timema* stick insects (Nosil,

2007), *Neochlamisus* leaf beetles (Funk, 1998; Egan et al., 2011), and *Rhagoletis* fruit flies (Feder et al., 1994).

The reproductive isolation between ecotypes generated by host-associated adaptation provides a foundation for other forms of phenotypic variation to accumulate and persist (Nosil, 2007). One phenotype under direct selection from their host plant is *J. haematoloma* beak length (Dingle et al., 2009). Our study provides a similar comparison with the inclusion of a previously undescribed host association. Our findings of host-associated phenotypic divergence are consistent with previous studies that have found differences in mouthpart size between insects that associate between the native *S. saponaria* var. *drummondii* and the non-native *K. paniculata* (Carroll & Boyd, 1992). Beak length is an adaptive response to host plant morphology, as it directly affects the insect's ability to access their food source, which is protected by large air-filled seedpods (Carroll & Loye, 1987). The depth in which a feeding *J. haematoloma* beak must penetrate the seedpod to access the interior of the seed is greater in *U. speciosa* seed pods than in the fleshy drupe and thick seed coat of the insects' previously identified native host *S. saponaria* var. *drummondii*. This difference provides an ecologically divergent landscape likely responsible for a localized *Ungnadia*-associated shift in insect beak length. Intriguingly, despite a consistent pattern of divergence in the observed beak lengths expressed on each host, it was only significantly different between all three hosts in the comparison of macropterous individuals with fully developed wings. This difference occurred independent of insect body size, as demonstrated in our analysis of thorax width and tibia length. This decoupling of beak length and body size is most consistent with genetic differentiation rather than allometric variations based on diet, which could be present as a consequence of differences in food quality between host plants. An additional line of evidence consistent with divergence in beak length having a genetic basis is that differences between hosts were maintained without a plastic response when insects were reared on alternative diets to their host origin. Along these lines, previously observed plasticity to non-native seeds masks some of the underlying genetic differentiation (Cenzer, 2017), so our results may be an underestimate of the total genetic differences in beak length.

Some of the inconsistency in beak length divergence between sex and wing morphology can perhaps be best explained by the insect's natural history. Adult feeding in male insects is often limited to predation of fallen seeds, thus removing the source of directional selection on their beak length as their food source is no longer protected by fruit coverings (Carroll & Boyd, 1992). Moreover, brachypterous females, who reach reproductive maturity sooner than macropterous females, lack the ability to diapause, resulting in an elevated emphasis on reproduction

(Carroll et al., 2003). With oviposition and increased mate encounters occurring on the ground, these flightless females may also spend more time feeding on unprotected seeds than their long-winged counterparts. Macropterous females, on the other hand, experience diapause initiated by photoperiod and food availability (Fowles et al., 2015). This extended longevity in volant females during low food availability allows them to put off reproduction while searching for more productive sites. Earlier access to seeds while they are still developing on the host plant is a clear advantage for individuals who are equipped with host-matching beak lengths. Furthermore, the flight muscles of macropterous individuals undergo histolysis in a proportion of the population upon engagement in feeding and reproduction, leading them to become marooned on their host plant (Dingle & Winchell, 1997). With both wing morphology and beak length phenotypes having previously been linked to strong genetic underpinnings (Dingle et al., 2009), a pattern of consistent differentiation between insect host associations may represent an adaptive multivariate suite of insect traits that are responding to localized host environments.

Future directions

Until now, the study of host race formation in *J. haematoloma* has been limited to the divergent populations shifting onto non-native introduced host plants. Here we find similar results on a newly discovered, but likely older, host association between *J. haematoloma* and *U. speciosa* occurring in sympatry with both its known native host plant and its introduced host plant. *Jadera haematoloma*'s range of host species is much broader in the south-central and southwestern USA than the well-known populations in Florida associated with native *Cardiospermum* and non-native *Koelreuteria* (Carroll et al., 1997, 2001; Dingle et al., 2009; Cenzer, 2016, 2017). The addition of *U. speciosa* creates the opportunity for additional pairwise comparisons along the speciation continuum in both allopatric and sympatric areas. Past and contemporary host shifts have facilitated expansions that have resulted in secondary contact between previously disjunct populations while also driving range expansion northward into areas outside the range of their ancestral hosts (Hoffman & Steiner, 2005). The discovery of this regional host association with *U. speciosa* mirrors the evolutionary history of past host shifts within the species. Exploration of the underlying genetic architecture could provide insight into both the repeatability and spatial components of host race formation. Our assessment of adaptive divergence in beak length and host preference suggests barriers to gene flow in this system that, although likely permeable, provide groundwork for such future eco-evolutionary investigations.

Author Contributions

Mattheau Steven Comerford: Conceptualization (lead); Data curation (lead); Formal analysis (lead); Funding acquisition (lead); Investigation (lead); Methodology (equal); Validation (lead); Visualization (lead); Writing-original draft (lead); Writing-review & editing (equal). Scott P Carroll: Conceptualization (equal); Funding acquisition (equal); Methodology (equal); Writing-review & editing (equal). Scott P Egan:Conceptualization (equal); Funding acquisition (equal); Methodology (equal); Project administration (lead); Supervision (lead); Writing-original draft (equal); Writing-review & editing (equal).

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

Adovasio JM & Fry GF (1976) Prehistoric psychotropic drug use in Northeastern Mexico and Trans-Pecos Texas. Economic Botany 30: 94–96.

Aldrich JR, Carroll SP, Lusby WR, Thompson MJ, Kochansky JP & Waters RM (1990) Sapindaceae, cyanolipids, and bugs. Journal of Chemical Ecology 16: 199–210.

Bates D, Maechler M & Walker S (2015) Fitting linear mixed-effects models using lme4. Journal of Statistical Software 67: 1–48.

Berlocher SH & Feder JL (2002) Sympatric speciation in phytophagous insects: moving beyond controversy? Annual Review of Entomology 47: 773–815.

Buerki S, Forest F, Acevedo-Rodríguez P, Callmander MW, Nylander JAA et al. (2009) Plastid and nuclear DNA markers reveal intricate relationships at subfamilial and tribal levels in the soapberry family (Sapindaceae). Molecular Phylogenetics and Evolution 51: 238–258.

Buerki S, Lowry PP II, Alvarez N, Razafimandimbison SG, Küpfer P & Callmander MW (2010) Phylogeny and circumscription of Sapindaceae revisited: molecular sequence data, morphology

and biogeography support recognition of a new family, Xanthoceraceae. Plant Ecology and Evolution 143: 148–159.

Carroll SP (1988) Contrasts in reproductive ecology between temperate and tropical populations of *Jadera haematoloma*, a mate-guarding Hemipteran (Rhopalidae). Annals of the Entomological Society of America 81: 54–63.

Carroll SP (2008) Facing change: forms and foundations of contemporary adaptation to biotic invasions. Molecular Ecology 17: 361–372.

Carroll SP & Boyd C (1992) Host race radiation in the soapberry bug: natural history with the history. Evolution 46: 1052–1069.

Carroll SP, Dingle H & Klassen SP (1997) Genetic differentiation of fitness-associated traits among rapidly evolving populations of the soapberry bug. Evolution 51: 1182–1188.

Carroll SP, Dingle H, Famula TR & Fox CW (2001) Genetic architecture of adaptive differentiation in evolving host races of the soapberry bug, *Jadera haematoloma*. Genetica 112: 257–272.

Carroll SP & Fox CW (2007) Dissecting the evolution impacts of plant invasions: bugs and beetles as native guides. Global Change Biology 13: 1644–1657.

Carroll SP, Klassen SP & Dingle H (1998) Rapidly evolving adaptations to host ecology and nutrition in the soapberry bug. Evolutionary Ecology 12: 955–968.

Carroll SP & Loye JE (1987) Specialization of *Jadera* species (Hemiptera: Rhopalidae) on the seeds of Sapindaceae (Sapindales), and coevolutionary responses of defense and attack. Annals of the Entomological Society of America 80: 373–378.

Carroll SP & Loye JE (2012) Soapberry bug (Hemiptera: Rhopalidae: Serinethinae) native and introduced host plants: biogeographic background of anthropogenic evolution. Annals of the Entomological Society of America 105: 671–684.

Carroll SP, Marler M, Winchell R & Dingle H (2003) Evolution of cryptic flight morph and life history differences during host race radiation in the soapberry bug, *Jadera haematoloma* Herrich-Schaeffer (Hemiptera: Rhopalidae). Annals of the Entomological Society of America 96: 135–143.

Cenzer ML (2016) Adaption to an invasive host is driving the loss of a native ecotype. Evolution 70: 2296–2307.

Cenzer ML (2017) Maladaptive plasticity masks the effects of natural selection in the red-shouldered soapberry bug. American Naturalist 190: 521–533.

Chapman RF (2003) Contact chemoreception in feeding by phytophagous insects. Annual Review of Entomology 48: 455–484.

Christensen RHB (2019) Ordinal – regression models for ordinal data. R package v.2019.12-10. https://CRAN.R-project.org/package=ordinal (accessed 10 February 2020).

Christensen RHB & Brockhoff PB (2013) Analysis of sensory ratings data with cumulative link models. Journal de la Société Française De Statistique 154: 58–79.

Coyne JA & Orr HA (2004) Speciation. Sinauer Associates, Sunderland, MA, USA.

Craig TP, Itami JK, Abrahamson WJ & Horner JD (1993) Behavioral evidence for host-race formation in *Eurosta solidaginis*. Evolution 47: 1696–1710.

Dingle H, Carroll SP & Famula TR (2009) Influence of genetic architecture on contemporary local evolution in the soapberry

bug, *Jadera haematoloma*: artificial selection on beak length. Journal of Evolutionary Biology 22: 2031–2040.

Dingle H & Winchell R (1997) Juvenile hormone as a mediator of plasticity in insect life histories. Archives of Insect Biochemistry and Physiology 35: 359–373.

Drès M & Mallet J (2002) Host races in plant-feeding insects and their importance in sympatric speciation. Philosophical Transactions of the Royal Society of London B 357: 471–492.

Edelaar P & Bolnick DI (2012) Non-random gene flow: an underappreciated force in evolution and ecology. Trends in Ecology and Evolution 12: 659–665.

Egan SP, Hood GR, Feder JL & Ott JR (2012a) Divergent host-plant use promotes reproductive isolation among cynipid gall wasp populations. Biology Letters 8: 605–608.

Egan SP, Hood GR & Ott JR (2012b) Testing the role of habitat isolation among ecologically divergent gall wasp populations. International Journal of Ecology 2012: 809897.

Egan SP, Hood GR, DeVela G & Ott JR (2013) Parallel patterns of morphological and behavioral variation among host-associated populations of two gall wasp species. PLoS One 8: e54690.

Egan SP, Janson EM, Brown CG & Funk DJ (2011) Postmating isolation and genetically variable host use in ecologically divergent host forms of *Neochlamisus bebbianae* leaf beetles. Journal of Evolutionary Biology 24: 2217–2229.

Feder JL, Chilcote CA & Bush GL (1988) Genetic differentiation between sympatric host races of *Rhagoletis pomonella*. Nature 336: 61–64.

Feder JL, Opp SB, Wlazlo B, Reynolds K, Go W & Spisak S (1994) Host fidelity is an effective premating barrier between sympatric races of the apple maggot fly. Proceedings of the National Academy of Sciences of the USA 91: 7990–7994.

Fowles TM, Coscarón MC, Panizzi AR & Carroll SP (2015) Scentless plant bugs (Rhopalidae). True Bugs (Heteroptera) of the Neotropics (ed. by AR Panizzi & J Grazia), pp. 607–637. Springer, Dordrecht, The Netherlands.

Funk DJ (1998) Isolating a role for natural selection in speciation: host adaptation and sexual isolation in *Neochlamisus bebbianae* leaf beetles. Evolution 52: 1744–1759.

Funk DJ (2012) Of 'host forms' and host races: terminological issues in ecological speciation. International Journal of Ecology 2012: 506957.

Funk DJ, Filchak KE & Feder JL (2002) Herbivorous insects: model systems for the comparative study of speciation ecology. Genetica 116: 251–267.

Geiger AP & Saltz JB (2019) Strong and weak cross-sex correlations govern the quantitative-genetic architecture of social group choice in *Drosophila melanogaster*. Evolution 74: 145–155.

Gildenhuys E, Ellis AG, Carroll SP & Le Roux JJ (2013) The ecology, biogeography, history and future of two globally important weeds: *Cardiospermum halicacabum* Linn. and *C. grandiflorum* Sw. NeoBiota 19: 45–65.

Gilman EF & Watson DG (2011) *Ungnadia speciosa*: Mexican buckeye. ENH-814. Environmental Horticulture Department, UF/IFAS Extension, Institute of Food and Agricultural Sciences, University of Florida, Gainesville, FL, USA.

Hill RT (1887) The topography and geology of the Cross Timbers and surrounding regions in Northern Texas. American Journal of Science s3–33: 291–303.

Hoffman RL & Steiner WE Jr (2005) *Jadera haematoloma*, another insect on its way north (Heteroptera: Rhopalidae). Banisteria 26: 7–10.

Hood GR, Zhang L, Hu EG, Ott JR & Egan SP (2019) Cascading reproductive isolation: plant phenology drives temporal isolation among populations of a host-specific herbivore. Evolution 73: 554–568.

Kuznetsova A, Brockhoff PB & Christensen RHB (2017) lmerTest package: tests in linear mixed effects models. Journal of Statistical Software 82: 1–26.

Lara-Villalon M, Vanoye-Eligio V, Solis MA, Sanchez-Ramos G & Chacon-Hernandez JC (2017) The navel orangeworm, *Amyelosis transitella* (Walker) (Lepidoptera: Pyralidae), discovered in Northeastern Mexico feeding on Sapindaceae. Proceedings of the Entomological Society of Washington 119: 601–605.

Length RV (2021) Emmeans: estimated marginal means, aka least-squares means. R package v.1.5.5-1. https://cran.r-project.org/web/packages/emmeans/index.html (accessed 12 February 2020).

Little EL Jr (1976) Atlas of United States Trees, Vol. 3: Minor Western Hardwoods. US Department of Agriculture, Washington, DC, USA.

Medina RF (2017) Host races. Reference Module in Life Sciences. Elsevier 2021: https://doi.org/10.1016/B978-0-12-809633-8.12248-4 (accessed 28 April).

Mikolajczak KL, Smith CR Jr & Tjarks LW (1970a) Cyanolipids of *Koelreuteria paniculata* Laxm. seed oil. Lipids 5: 672–677.

Mikolajczak KL, Smith CR Jr & Tjarks LW (1970b) Cyanolipids of *Cardiospermum halicacabum* L. and other sapindaceous seed oils. Lipids 5: 812–817.

Nahrstedt A (1985) Cyanogenic compounds as protecting agents for organisms. Plant Systematics and Evolution 150: 35–47.

Nosil P (2007) Divergent host plant adaptation and reproductive isolation between ecotypes of *Timema cristinae* walking sticks. American Naturalist 169: 151–162.

Nosil P, Crespi BJ & Sandoval CP (2002) Host-plant adaptation drives the parallel evolution of reproductive isolation. Nature 417: 440–443.

Nosil P, Vines TH & Funk DJ (2005) Reproductive isolation caused by natural selection against immigrants from divergent habitats. Evolution 59: 705–719.

R Core Team (2019) R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.

Ribeiro ST (1989) Group effects and aposematism in *Jadera haematoloma* (Hemiptera: Rhopalidae). Annals of the Entomological Society of America 82: 466–475.

Seigler DS & Kawahara W (1976) New reports of cyanolipids from sapindaceous plants. Biochemical Systematics and Ecology 4: 263–265.

Via S (1999) Reproductive isolation between sympatric races of pea aphids. I. Gene flow restriction and habitat choice. Evolution 53: 1446–1457.

Vines RA (1984) Trees of Central Texas. University of Texas Press, Austin, TX, USA.

Wisz MS, Pottier J, Kissling WD, Pellissier L, Lenoir J et al. (2013) The role of biotic interactions in shaping distributions and realised assemblages of species: implications for species distribution modelling. Biological Reviews 88: 15–30.

Wood TK (1993) Speciation of the *Enchenopa binotata* complex (Insect: Homoptera: Membracidae). Evolutionary Patterns and Processes (ed. by DR Lees & D Edwards), pp. 299–317. Academic Press, New York, NY, USA.

Wood TK, Tilmon KJ, Shantz AB, Harris CK & Pesek J (1999) The role of host-plant fidelity in initiating insect race formation. Evolutionary Ecology Research 1: 317–332.

Xie X, Rull J, Michel AP, Velez S, Forbes AA et al. (2007) Hawthorn-infesting populations of *Rhagoletis pomonella* in Mexico and speciation mode plurality. Evolution 61: 1091–1105.

Zhang L, Hood GR, Roush AM, Shzu SA, Comerford MS et al. (2021) Asymmetric, but opposing reductions in immigrant viability and fecundity promote reproductive isolation among host-associated populations of an insect herbivore. Evolution 75: 476–489.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Figure S1 Earliest *Jadera haematoloma* native host plant record per county.

Figure S2 Density map of each *Jadera haematoloma* host plant occurrence.

Figure S3 Host preference experimental set-up with 16 Petri dish preference arenas.

Figure S4 Box plots of native host-associated seed traits: the combined measurement of seedcoat thickness and the distance from the surface of the seedpod to the seed within.

Figure S5 Structural diagrams of the four cyanolipid schemes in the seed oil of the Sapindaceae.

Table S1 List of host plants and insect sampling sites included in the study.

Table S2 List of herbaria used in data collection and their corresponding no. of records of each host plant.

Table S3 Test of allometric thorax divergence: fixed-effects GLMM results of phenotypic differentiation in insect (*Jadera haematoloma*) thorax width between host association, sex, and wing polymorphism as predictors and collection site as a random effect.

Table S4 Test of allometric tibia divergence: fixed-effects GLMM results of phenotypic differentiation in insect (*Jadera haematoloma*) tibia length between host association, sex, and wing polymorphism as predictors and collection site as a random effect.