EVALUATING SPECIES INTERACTIONS AS A DRIVER OF PHYTOPHAGOUS INSECT DIVERGENCE

A PREPRINT

Bruno A. S. de Medeiros
Museum of Comparative Zoology
Harvard University
brunoasdemedeiros@gmail.com

Brian D. Farrell
Museum of Comparative Zoology
Harvard University
bfarrell@oeb.harvard.edu

November 13, 2019

ABSTRACT

Plants and their specialized flower visitors provide valuable insights into the evolutionary consequences of species interactions. In particular, antagonistic interactions between insects and plants have often been invoked as a major driver of diversification. Here we use a tropical community of palms and their specialized insect flower visitors to understand whether antagonisms lead to higher population divergence. Interactions between the palms *Syagrus coronata* and *Syagrus botryophora* and the weevils that visit their flowers range from brood pollination to florivory and commensalism. We use genomics to test the role of insect-host interactions in the early stages of diversification of nine species of beetles associated with these plams by using a model of isolation by environment. We find a surprising number of cryptic species, which in pollinating weevils coexist across a broad geographical range but are always associated with different hosts for non-pollinators. The degree to which insect populations are structured by the genetic divergence of plant populations varies. This variation is uncorrelated with the kind of interaction, showing that, at least in this system, antagonistic interactions are not associated with higher genetic differentiation. It is likely that more general aspects of host use, affecting plant-associated insects regardless of the outcomes of their interactions, are more important drivers of population divergence.

Keywords Insect-plant interactions · Speciation · Phylogeography · Curculionidae · Arecaceae

1 Introduction

Insects comprise about two thirds of the 1.5 million described species of animals [86], and current estimates predict that another 4 million insect species remain unknown [76]. This spectacular diversity is thought to be in a large degree a result of interactions with plants [27, 25, 79, 40, 55]. Antagonism between plants and insects could lead to accelerated rates of diversification, with the diversity of defenses among plants resulting from host specialization that in turn may spur radiations in insects circumventing those defenses [40, 70, 21, 55]. Alternatively, mutualism could also lead to highly specialized interactions and thereby promote diversification in insects and plants [83, 45, 46]. Early divergences between species or populations of insects that feed on plants are often associated with exploitation of new host plants, either in sympatry [18, 6] or allopatry [53, 4, 26]. While host races following a switch to a new plant seem to be common, insect co-divergence with a plant species is also an important mechanism generating insect diversity. In these cases, population divergence of specialized insects is correlated with population-level differentiation of plants in which they specialize, suggesting synchronous divergence that could be the result of coevolution. This pattern has been observed in brood pollinators (specialized pollinators that breed on their host plants) [54, 80, 5, 73, 72, 22], but not in all cases evaluated (see [36, 22]). Co-differentiation is also seen in some non-pollinator herbivores [81] and communities of herbivores and commensals specialized on pitcher plants [67, 68].

The relative contribution of ongoing coevolution and host switches to the generation of insect diversity is currently unknown. It is also unclear whether the outcomes of insect-plant interactions along the mutualism-antagonism spectrum

have an effect on the rates of population differentiation, especially considering that they may often be spatially and temporally variable [78]. Theory predicts that coevolution can lead to stronger genetic differentiation when compared to spatial isolation alone in the case of antagonism but not in mutualism [47, 85]. Under this hypothesis, genetic isolation between plant populations may be a better predictor of insect isolation in antagonists than in mutualists. Finally, it is also possible that the main driver of insect specialization and diversification is not divergent selection imposed by interactions, but sensory biases that act independently of the outcome of the interaction with host plants [41, 42]. A direct comparison between insects with different modes of interaction across scales of plant divergence can help resolve the role of interactions and the relative contribution of long-term coevolution and new host shifts in accelerating insect differentiation and the formation of new species. Here we perform this comparison by taking advantage of the variation in insect-plant interactions found in communities of palm-associated weevils distributed across the same geographic range and interacting with the same plants. If coevolution, and antagonism in particular, is a major driver of insect population divergence, we would expect that antagonist species would show higher levels of population divergence, and particularly of divergence associated with their hosts.

Palms in the genus *Syagrus*, one of the closest relatives of the coconut [59, 57], produce large inflorescences that are visited by dozens of insect species [71, 61, 60, 15]. The most abundant flower visitors of these Neoptropical palms are specialized weevils (family Curculionidae). We recently described the community of insects associated with the seasonally dry forest palm *Syagrus coronata*, showing that many weevil species are broadly distributed throughout the plant geographical range. Some of them are brood pollinators, while others are antagonists breeding on flowers or seeds and some are commensals breeding on decaying plant tissues [15]. *Syagrus coronata* shares many species of weevil with *Syagrus botryophora*, a parapatric palm specialized on rainforests and diverged from *S. coronata* early in the history of the genus *Syagrus*, about 20 million years ago [57, 56, 59]. The weevil species shared by the two plants are likely a result of relatively recent host shifts instead of longterm codiversification (de Medeiros et al, in preparation). We used double-digest RADseq (ddRAD) [64, 17] to obtain genome-wide genetic markers for several populations of both plant species as well as nine species of weevils broadly distributed across the range of one or both palms (Figure S1). These nine species are all attracted to flowers and locally specialized on their host plants. They mate and lay eggs on their hosts and are distributed through a similar geographical range, but differ in the kind of interaction with plants as a result of variation in their roles as pollinators as adults and whether their larvae breed on live or decaying tissues (Table 1). We first use the genomic data to delimit weevil species and better understand the diversity of these little-known insects. Then, we test models of isolation by environment to ask whether the kind of interaction with host plants is associated with differences in the degree of isolation by geographical distance or isolation by host plant. We specifically address whether antagonistic interactions lead to stronger levels of differentiation in relation to mutualism and commensalism, and whether this operates only at the level of plant species or also between plant populations.

Table 1: Weevil species included in the study. References indicate sources natural history information.

Species	Pollinator	Larval breeding
Anchylorhynchus trapezicollis [15]	Yes	Developing seeds
Remertus rectinasus [15]	No	Developing seeds
Microstrates bondari [8]	No	Live male flowers
Microstrates ypsilon [15]	No	Live male flowers
Andranthobius bondari [15]	No	Decaying male flowers
Celetes impar [15]	No	Decaying peduncular bract
Celetes decolor [15]	No	Decaying floral branches
Dialomia polyphaga [15]	No	Damaged inflorescences
Phytotribus cocoseae [9, 7]	No	Decaying bracts

Cryptic host-associated species

We initially assembled genomic datasets by filtering low-coverage loci and genotyping each individual separately. Visualization of patterns of missing data revealed that, for some of the weevil species, certain ddRAD loci are shared within groups of samples, with very few loci recovered across groups (Figure S2). This pattern could be a artifact resulting from batch effects during ddRAD library preparation, because samples in a batch are pooled before size selection and PCR amplification [64]. Alternatively, it could be a consequence of cryptic, deeply differentiated taxa contained within each species as traditionally recognized by morphology [66]. To test which is the case, we recorded the number of loci shared, average sequence divergence and batch identity for each pair of samples in each morphospecies. We found that samples processed in the same batch do share more loci, but extreme levels of missing data are only explained by deep sequence divergence, sometimes above 2.5% (Table 2, Figure S2). We note that, in all cases, splitting

Table 2: Effect of percent sequence distance and shared library batch on number of shared RAD loci (thousands). All samples of *R. rectinasus* were prepared in the same batch. *p-value<0.01

Morphospecies	Intercept	Distance	Batch	R^2
Anchylorhynchus trapezicollis	7.4*	-1.5*	0.3	0.46
Andranthobius bondari	3.2*	-0.8*	0.7*	0.15
Celetes decolor	3.7*	-0.6*	1.0	0.15
Celetes impar	7.2	-0.4	0.2	0.002
Dialomia polyphaga	4.1	-2.9	0.6	0.09
Microstrates bondari	1.6	0.4	1.4*	0.08
Microstrates ypsilon	3.8	-0.7	0.1	0.01
Phytotribus cocoseae	18.5*	-24.2*	1.4*	0.21
Remertus rectinasus	3.4	-0.6*	–	0.04

samples into operational taxonomic units (OTU) at this level of sequence divergence results in groups with very high genetic differentiation from each other as measured by G_{ST} (Figure S2). In all but one case, these clusters separate populations on each host plant (Figure 1). For all kinds of interactions, there is evidence for negligible to zero gene flow between these populations on the two different host plant species. In the case of the pollinator *Anchylorhynchus trapezicollis*, we find three genetic clusters, with one of them in both host species and broadly sympatric with the other two (Figure 1). By comparing the morphology of the two most abundant clusters in sympatry and allopatry, we found differences in the length of ventral plumose hairs and in male secondary sexual characters (Figure S3). The former might be involved in pollen adherence [15], suggesting the coexistence of an effective pollinator with a closely related exploiter of the weevil-plant mutualism, as found in other brood pollinators such as *Yucca* moths [14]. These diverged genetic clusters represent cryptic, previously unrecognized species. Hereafter, we treat each one as a separate species, highlighting that they need to be properly described in the future. In general, we also recommend caution in studies of little known organisms in which cryptic species might be common [77], noting that we were only able to distinguish OTUs because samples were individually barcoded and not pooled by location.

A principal component analysis of the genetic variation of each OTU reveals little spatial congruence among weevil species and variable congruence with the genetic variation of their host plants (Figure 1). We found evidence for genetic clusters in 6 of the 13 weevil species (Figure 1, Figure S4) and investigated whether there is gene flow between these clusters by using a model of isolation-with-migration based on the site frequency spectrum (Figure S5). We found that, in all cases, models including migration had higher support than those that did not (Table S1). Populations of *Anchylorhynchus trapezicollis* OTU 1 and *Remertus rectinasus* on different host plants have much deeper divergence and smaller migration rates than those interacting with *Syagrus coronata* alone (Table S1), indicating that there are well-delimited host races even in these cases that divergence is shallow enough to enable assembly of ddRAD datasets.

2 Interactions do not predict patterns of isolation

Following evidence for significant gene flow between all populations in each OTU, we assessed the role of geography, plant host or climate as genetic barriers for each species of weevil. We used matrices of geographical distance, host plant genetic distance and climatic distance between weevil populations as explanatory variables for the genetic covariance between weevils in a Bayesian model of isolation by distance and environment [11, 10]. With model choice by cross validation (Table S2), we found isolation by geographical distance in most cases, and isolation associated with plant host genetic divergence only in species interacting with both host palms or with *Syagrus coronata* only. The latter pattern was found for the three kinds of interaction, while more common for brood pollinators and antagonists. Climate is a significant but weak barrier to gene flow only for a few weevil species (Figure 2, Table S3). The divergence between populations of *S. botryophora* is very small when compared to *S. coronata*, and only in some of species interacting with the latter we find a significant role of plant host genetic distance as a barrier to weevil gene flow. The biomes on which each plant species specialize have undergone very different dynamics during climate cycles, with seasonally dry forests (*S. coronata*) being more unstable than wet forests by the coast (*S. botryophora*) [3, 13]. Previous studies in part of the range of *S. coronata* revealed high levels of population structure [74], and here we found much larger genetic distances between populations for *S. coronata* than for *S. botryophora*. While a detailed analysis of the population history of these palm species will be published elsewhere (de Medeiros et al, in preparation), this suggests that *S. coronata* had more opportunities for population isolation throughout climatic cycles, with consequences for its associated insects.

The degree of isolation by distance and by environment in these weevils co-distributed throughout the same range and interacting with the same plants varies widely, and this variation is largely unrelated to the kind of interaction

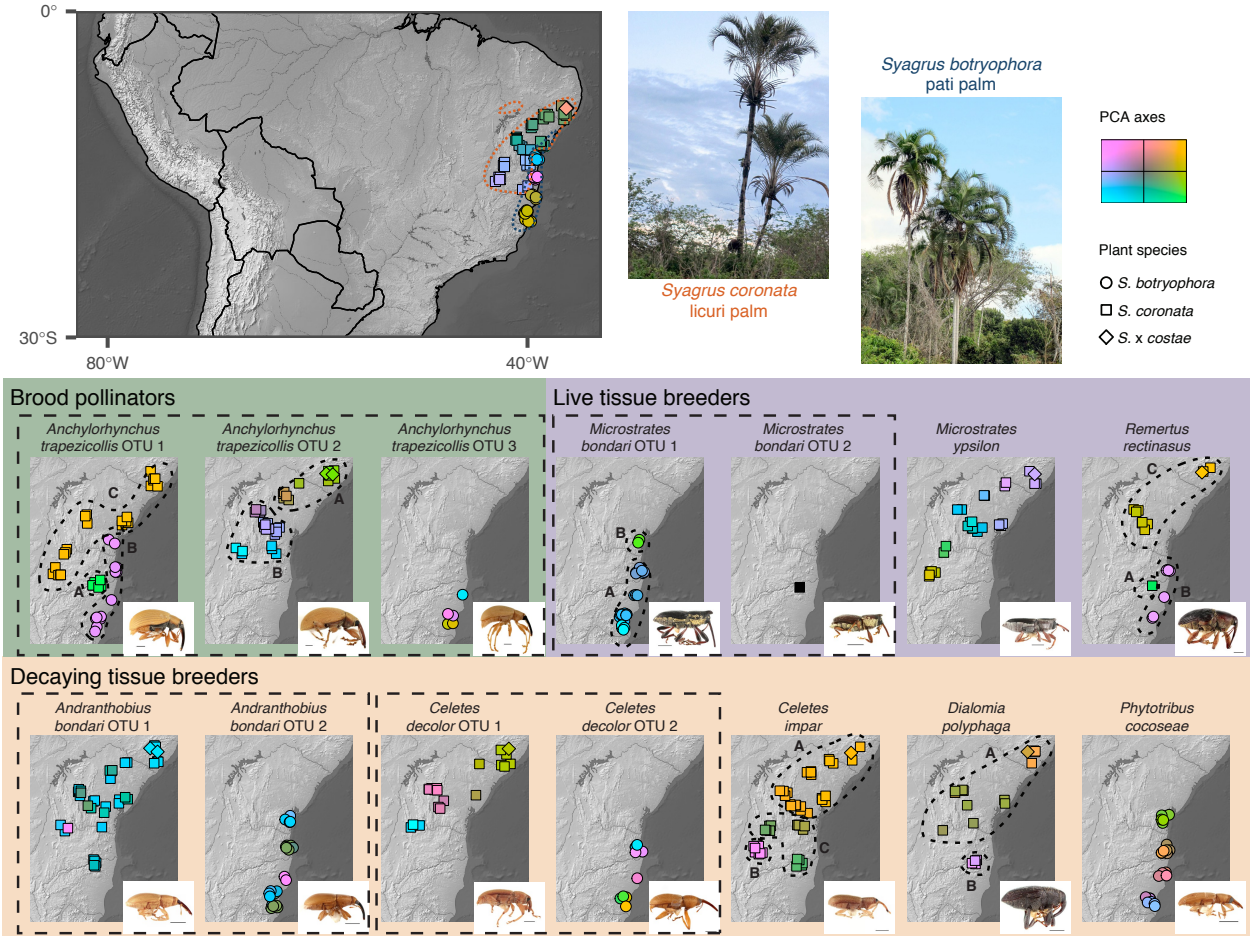

Figure 1: Principal component analyses of weevil and palm genetic diversity. In each case, the first two PC axes are color-coded following the legend provided. A small jitter was added to enable visualization. Dashed boxes enclose taxa previously considered to be the same species. Large map includes known palm distributions [59, 30] enclosed in dashed lines. PCA results are independent for *S. botryophora* and *S. coronata*. *Syagrus* × *costae* is a hybrid of *S. cearensis* and *S. coronata* [59] and is considered here as a population of the latter. Small maps show PCA results for each weevil species, with populations enclosed in black dashed lines following k-means clustering results. Letters indicate population labels in Table S1. Scales 1 mm in insect images.

with their hosts. All species previously thought to interact with both hosts are actually comprised of cryptic species or highly divergent populations, each specialized on a single plant. At a finer level, host population divergence is a barrier to weevil gene flow for some weevil species, encompassing all kinds of interaction. Considering that even species breeding on decaying and thus undefended tissues show this pattern, it is unlikely that coevolution and adaptation to plant defenses is a universal source of divergent selection and a necessary condition to explain insect diversification. A recent review found that most studies on candidate genes for host adaptation in phytophagous insects focus on resistance or detoxification of plant secondary metabolites [82], but the actual source of selection might be in other aspects of host use. Divergence following host shifts is pervasive in phytophagous insects and their parasitoids [26], despite the large variation in interaction outcomes. Even though coevolution might be an important driver of diversification in some cases [79, 1, 85, 35, 45, 55], evolution without reciprocal adaptation might be sufficient to explain many or most cases of insect specialization. One possibility is that sensory biases, unrelated to performance on host plants, are a more general driver of divergence across phytophagous insects [41, 42]. This hypothesis seems consistent with our findings, in which all species mate on flowers and are likely responding to the evolution of floral traits as sexual signals. Flowers, in turn, evolve in response to a network of species [33], including herbivores feeding on vegetative tissues [65].

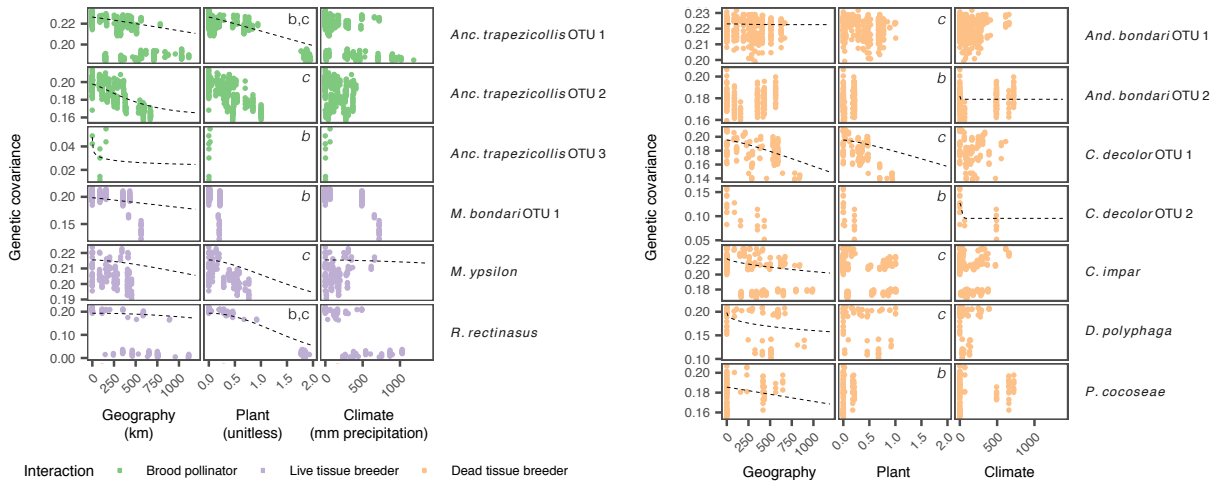

Figure 2: Effects of the three distance matrices on pairwise genetic covariance between samples for weevil species. When included in the best model, dashed lines show the marginal effects of each distance implied by average parameter estimates. Geographical distance measured in kilometers, plant distances as Euclidean distance in plant PCA and climate distances as differences in annual precipitation. Plant panels include names of plants that each weevil species interacts with: *Syagrus botryophora* (b) or *Syagrus coronata* (c).

3 Concluding remarks

We studied patterns of isolation by distance and by environment in nine morphospecies of weevils associated with flowers of two palm species, which turned out to be 14 weevil species after cryptic species were identified. Host plant species was a very strong barrier to gene flow in all cases, with a different species or a highly divergent population on each host. Both geography and host plants, but usually not climate, are important in determining genetic differentiation, with variation between insect species being weakly related to the kind of interaction with their host plants. Insect-plant antagonistic coevolution does not seem to be required for insect specialization and the generation of barriers to gene flow, and other aspects of insect-host interactions, such as sensory biases, should be investigated in studies of phytophagous insect diversification.

4 Acknowledgements

We thank Harri Lorenzi and Larry Noblick for sharing palm location data and assisting with identification. Curators Renato de Melo Silva (SPF) and Sonia Casari (MZSP) allowed collection access and deposit of specimens. Sergio A. Vanin assisted with weevil specimens and identification. Tauana Cunha, Elton Antunes, Filipe Gudin, Gabriel Pimenta, Zhengyang Wang, Luiz Fonseca, José I. L. Moura, Fr. Francisco José de Paula and the Queiroz family assisted with fieldwork. Collections in Vale Natural Reserve, Córrego do Veado Biological Reserve, Decobrimento National Park and Sooretama Biological Reserve were made with assistance from park authorities and permit from ICMBIO #39704-7. Work in the MCZ was assisted by Adrian Magaña, Scout Leonard, Dylan Ryals, Whit Farnum and Amie Jones. Alyssa Hernandez produced SEMs in Fig. S3. We thank Bauer core staff for help with sequencing and library preparation. Gideon Bradburd allowed access to the development version of BEDASSLE2. The computations in this paper were run on the FASRC Cannon cluster supported by the FAS Division of Science Research Computing Group at Harvard University. The first author received a Jorge Paulo Lemann Fellowship for Research in Brazil. This work was funded by NSF DEB #1355169 and #1601356 and the William F. Milton Fund.

5 Materials and Methods

5.1 Sampling

We sampled insects and plants from 13 populations of *Syagrus coronata* (including *S.* × *costae*, hybrids with *S. cearensis* [59]) and five populations of *Syagrus botryophora* throughout the distribution of both plant species (Figure 1). Whole inflorescences were bagged and excised, with insects were aspirated stored in 95% ethanol. Leaf tissues were collected from the sampled plant and other individuals in the vicinity. For this study, we chose nine specialized

weevil species that we previously identified [15] that engage into different kinds of interaction with their host plants and which have widespread geographical distributions. All populations in which at least one individual was sampled were included.

5.2 DNA extraction and library preparation

We extracted DNA from insects and prepared double-digest RAD-seq libraries [64] as described in de Medeiros & Farrell [17]. Some of the individuals were extracted destructively, but for others we digested full bodies separated at the base of the pronotum and preserved the remaining cuticle. We included a step of whole-genome amplification prior to library preparation for samples yielding less than 150 ng of DNA [17]. For plants, DNA was extracted from leaf tissues using the E.Z.N.A. HP Plant DNA Mini Kit (Omega Biotek) following the manufacturer protocol and libraries were prepared with the same enzymes and protocol as for insects, but from 300–1000 ng of genomic DNA without whole-genome amplification. Barcoded libraries were sequenced on Illumina systems, in several runs pooled with unrelated libraries. The minimum sequence length was single-end 100 bp, and all sequences were trimmed to this length prior to assembly.

5.3 Initial dataset assembly

Sequences were demultiplexed by inline barcodes and assembled using ipyrad v. 0.7.24 [19, 20]. For insects, sequences were entirely assembled de novo, but removing reads of potential endosymbionts by using the ipyrad option *denovo−reference* with reference sequences including genomes of known weevil symbionts [2] as well as *Rickettsia* and *Wolbachia*. We assembled datasets separately for each insect morphospecies. For plants, we generated a single dataset for both species, and sequences were assembled either by mapping to the draft genome assembly of the coconut [84] or denovo for unmapped reads, using the ipyrad option *denovo+reference*. Reads were clustered within and between samples at 85% identity, and only loci with coverage ≥ 12 in a sample were retained for statistical base calling using ipyrad. Initially, we retained all samples and all loci present in at least four samples and we used Matrix Condenser [17, 16] to visualize patterns of missing data. We then removed samples with excessive missing data from the datasets, since with whole-genome amplification these are more likely to include contaminants and amplification artifacts [17]. Instead of choosing an arbitrary threshold for filtering, we flagged for removal outliers as observed in the histogram view of Matrix Condenser.

5.4 Assesing missing data

For each insect morphospecies, we calculated the following pairwise quantities: (1) number of loci sequenced incommon for each pair of samples, (2) the average pairwise nucleotide distance using the R package phangorn v. 2.4.0 [69], and (3) whether the two samples were prepared in the same batch. Wetested whether sequence distance and batch effects are negatively associated with the number of common loci byfitting a regression on distance matrix [50, 52] implemented in the R packageecodist v. 2.0.1 [31].

5.5 Assembly of final datasets

After confirming that sequence distance is associated with fewer of loci, we split the datasets for each morphospecies into clusters separated by at least 2.5% nucleotide differences using the R package dendextend v. 1.8.0 [29]. To further confirm if clusters thus obtained consist of highly isolated populations, we used the R packages mmod v.1.3.3 [29] andadegenet v.2.1.1 [43, 44] to calculate G_{ST} between these clusters using all loci present in at least one individual per cluster. In the case of *Anchylorhynchus trapezicollis*, clusters were sympatric across a broad range, so we compared the morphology of individuals with preserved cuticle to confirm their divergence with an independent source of data. Sequencing statistics are available in table S4.

5.6 Population structure

We used bwa-mem v.0.7.15 [51] to map reads on the consensus sequence for each RAD loci in the final dataset. Alignment files in bam format were used as input to ANGSD v.0.920 [48] and PCAngsd v.0.973 [58] to calculate the genetic covariance matrices for each insect and plant species, as well as genotype probabilities. Principal component analyses of these variances were clustered by the k-means method with scripts modified from the R package adegenet v.2.1.1 [43, 44]. For each insect species, the optimal number of clusters was chosen by minimizing the Bayesian Information Criteria (BIC)[49].

5.7 Modeling Isolation with Migration

We used ANGSD and dadi v.1.7.0 [34] to generate the multidimensional site frequency spectrum for each morphospecies with more than one k-mean cluster. We then implemented models of isolation-with-migration [37] (Figure S5) in fastsimcoal v.2.6.0.3 [23] and inferred parameters from the site frequency spectra [24]. All simulations were done with a mutation rate of 3e-9, compatible with other insects [62], and inferred parameters were then scaled by the mutation rate (Figure S5). For each model, we ran 100 independent searches of the maximum likelihood parameters and selected the best model by the Akaike Information Criterion.

5.8 Isolation by distance and environment

We used BEDASSLE2 v.0.0.0.9000 [11, 10] to infer the effects of climate, geographical distance and host plant genetic distance on the genetic covariance of weevil populations. We generated valid [32], Euclidean distances for explanatory variables as follows. We first projected collection localities to UTM Zone 24S using the R package sf v.0.8-0 [63] and then calculated the Euclidean distance between them. For climatic distance, we downloaded records of *Syagrus coronata* and *Syagrus botryophora* from GBIF [30] using the R package rgbif v.1.3.0 [12], cleaned them with the R package CoordinateCleaner v.2.0-11[87] then used the R package raster v.3.0-7 [38] to extract bioclimatic variables [39] for these localities and used PCA to find that the first PC explained 90.9% of the variance in the dataset and Annual precipitation (bio12) had a very high loading on this component (Figure S6). Therefore, we used the Euclidean distance in Annual Precipitation as climatic distance. For plant host genetic distances, we did a PCA based on genetic covariances for each species. We then obtained the centroid of each population in the first 3 principal components and calculated the Euclidean distances between centroids. For each weevil species with 3 or more populations sampled, we called genotypes with probability >= 0.8 and filtered the dataset to one site per RAD locus to avoid linked sites. For cross-validation, we split datasets in 10 partitions with 50 replicates and chose the simplest model among those with highest and overlapping 95% confidence intervals for explanatory power. We ran the selected BEDASSLE2 model on the full dataset, with 4 chains of 20,000 generations each and used the R package shinystan v.2.5.0 [28] to evaluate convergence and rstan v.19.2 [75] to summarize estimates.

References

[1] David M. Althoff, Kari A. Segraves, and Marc T J Johnson. "Testing for coevolutionary diversification: linking pattern with process." In: *Trends in Ecology & Evolution* 29.2 (Mar. 2014), pp. 82–9. DOI: 10.1016/j.tree.2013.11.003.

[2] Hisashi Anbutsu et al. "Small genome symbiont underlies cuticle hardness in beetles." In: *Proceedings of the National Academy of Sciences* (2017), p. 201712857. DOI: 10.1073/pnas.1712857114.

[3] Daniel M. Arruda et al. "Vegetation cover of Brazil in the last 21 ka: New insights into the Amazonian refugia and Pleistocenic arc hypotheses." In: *Global Ecology and Biogeography* 27.1 (2018), pp. 47–56. DOI: 10.1111/geb.12646.

[4] Robin K. Bagley et al. "History, geography and host use shape genomewide patterns of genetic variation in the redheaded pine sawfly (Neodiprion lecontei)." In: *Molecular Ecology* 26.4 (Feb. 2017), pp. 1022–1044. DOI: 10.1111/mec.13972.

[5] A. Bain et al. "Geographic structuring into vicariant species-pairs in a wide-ranging, high-dispersal plant–insect mutualism: the case of Ficus racemosa and its pollinating wasps." In: *Evolutionary Ecology* 30.4 (2016), pp. 663–684. DOI: 10.1007/s10682-016-9836-5.

[6] Stewart H Berlocher and Jeffrey L. Feder. "Sympatric speciation in phytophagous insects: Moving beyond controversy?" In: *Annual Review of Entomology* 47 (Jan. 2002), pp. 773–815. DOI: 10.1146/annurev.ento.47.091201.145312.

[7] Gregório G. Bondar. "Notas Entomológicas da Baía. IX." In: *Revista de Entomologia* 13.1 (1942), pp. 1–38.

[8] Gregório G. Bondar. "Notas Entomológicas da Baía. VI." In: *Revista de Entomologia* 11.3 (1940), pp. 842–861.

[9] Gregório G. Bondar. "Notas Entomológicas da Baía. VIII." In: *Revista de Entomologia* 12.3 (1941), pp. 427–470.

[10] Gideon Bradburd. *BEDASSLE 2.0*. 2019. URL: https://github.com/gbradburd/bedassle.

[11] Gideon S. Bradburd, Peter L. Ralph, and Graham M. Coop. "Disentangling the effects of geographic and ecological isolation on genetic differentiation." In: *Evolution* 67.11 (Nov. 2013), pp. 3258–73. DOI: 10.1111/evo.12193.

[12] Scott Chamberlain et al. *rgbif: Interface to the Global Biodiversity Information Facility API*. 2019. URL: https://cran.r-project.org/package=rgbif.

[13] Gabriel C. Costa et al. "Biome stability in South America over the last 30 kyr: Inferences from long-term vegetation dynamics and habitat modelling." In: *Global Ecology and Biogeography* (2017), pp. 1–13. DOI: 10.1111/geb.12694.

[14] Clive T. Darwell, Kari A. Segraves, and David M. Althoff. "The role of abiotic and biotic factors in determining coexistence of multiple pollinators in the yucca–yucca moth mutualism." In: *Ecography* 40.4 (2017), pp. 511–520. DOI: 10.1111/ecog.02193.

[15] Bruno A S de Medeiros et al. "Flower visitors of the licuri palm (Syagrus coronata): brood pollinators coexist with a diverse community of antagonists and mutualists." In: *Biological Journal of the Linnean Society* 126.4 (Mar. 2019), pp. 666–687. DOI: 10.1093/biolinnean/blz008.

[16] Bruno A. S. de Medeiros. *Matrix Condenser*. 2018. URL: https://github.com/brunoasm/matrix_condenser.

[17] Bruno A. S. de Medeiros and Brian D. Farrell. "Whole-genome amplification in double-digest RADseq results in adequate libraries but fewer sequenced loci." In: *PeerJ* 6 (July 2018), e5089. DOI: 10.7717/peerj.5089.

[18] Michele Drès and James Mallet. "Host races in plant-feeding insects and their importance in sympatric speciation." In: *Philosophical Transactions of the Royal Society B: Biological Sciences* 357.1420 (Apr. 2002), pp. 471–492. DOI: 10.1098/rstb.2002.1059.

[19] Deren A. R. Eaton. "PyRAD: Assembly of de novo RADseq loci for phylogenetic analyses." In: *Bioinformatics* 30.13 (2014), pp. 1844–1849. DOI: 10.1093/bioinformatics/btu121.

[20] Deren A. R. Eaton and Isaac Overcast. *ipyrad v. 0.6.8*. 2017. URL: https://github.com/dereneaton/ipyrad.

[21] Paul R. Ehrlich and Peter H. Raven. "Butterflies and plants: a study in coevolution." In: *Evolution* 18.4 (1964), pp. 586–608.

[22] Anahí Espíndola, Bryan C. Carstens, and Nadir Alvarez. "Comparative phylogeography of mutualists and the effect of the host on the genetic structure of its partners." In: *Biological Journal of the Linnean Society* 113.4 (2014), pp. 1021–1035. DOI: 10.1111/bij.12393.

[23] Laurent Excoffier and Matthieu Foll. "fastsimcoal: A continuous-time coalescent simulator of genomic diversity under arbitrarily complex evolutionary scenarios." In: *Bioinformatics* 27.9 (May 2011), pp. 1332–1334. DOI: 10.1093/bioinformatics/btr124.

[24] Laurent Excoffier et al. "Robust Demographic Inference from Genomic and SNP Data." In: *PLoS Genetics* 9.10 (2013). DOI: 10.1371/journal.pgen.1003905.

[25] Brian D. Farrell. ""Inordinate Fondness" explained: Why are there so many beetles?" In: *Science* 281.5376 (July 1998), pp. 555–559. DOI: 10.1126/science.281.5376.555.

[26] Andrew A. Forbes et al. "Revisiting the particular role of host shifts in initiating insect speciation." In: *Evolution* 71.5 (2017), pp. 1126–1137. DOI: 10.1111/evo.13164.

[27] Douglas J. Futuyma and Anurag A. Agrawal. "Macroevolution and the biological diversity of plants and herbivores." In: *Proceedings of the National Academy of Sciences* 106.43 (2009), pp. 18054–18061.

[28] Jonah Gabry. *shinystan: Interactive Visual and Numerical Diagnostics and Posterior Analysis for Bayesian Models*. 2018. URL: https://cran.r-project.org/package=shinystan.

[29] Tal Galili. "dendextend: An R package for visualizing, adjusting and comparing trees of hierarchical clustering." In: *Bioinformatics* 31.22 (2015), pp. 3718–3720. DOI: 10.1093/bioinformatics/btv428.

[30] GBIF.org. *GBIF Ocurrence Download*. 2019. DOI: 10.15468/dl.lprfwo. URL: https://doi.org/10.15468/dl.lprfwo.

[31] Sarah C Goslee and Dean L Urban. "The ecodist package for dissimilarity-based analysis of ecological data." In: *Journal Of Statistical Software* 22.7 (2007), pp. 1–19.

[32] Gilles Guillot et al. "Validity of covariance models for the analysis of geographical variation." In: *Methods in Ecology and Evolution* 5.4 (2014), pp. 329–335. DOI: 10.1111/2041-210X.12167.

[33] Paulo R. Guimarães et al. "Indirect effects drive coevolution in mutualistic networks." In: *Nature* 550.7677 (2017), pp. 511–514. DOI: 10.1038/nature24273.

[34] Ryan N. Gutenkunst et al. "Inferring the joint demographic history of multiple populations from multidimensional SNP frequency data." In: *PLoS Genetics* 5.10 (2009). DOI: 10.1371/journal.pgen.1000695.

[35] David H. Hembry, Jeremy B. Yoder, and Kari Roesch Goodman. "Coevolution and the Diversification of Life." In: *The American Naturalist* 184.4 (Oct. 2014), pp. 425–438. DOI: 10.1086/677928.

[36] David H. Hembry et al. "Non-congruent colonizations and diversification in a coevolving pollination mutualism on oceanic islands." In: *Proceedings of the Royal Society B: Biological Sciences* 280.1761 (May 2013), p. 20130361. DOI: 10.1098/rspb.2013.0361.

[37] Jody Hey and Rasmus Nielsen. "Multilocus methods for estimating population sizes, migration rates and divergence time, with applications to the divergence of Drosophila pseudoobscura and D. persimilis." In: *Genetics* 167.2 (June 2004), pp. 747–60. DOI: 10.1534/genetics.103.024182.

[38] Robert J Hijmans. *raster: Geographic Data Analysis and Modeling*. 2019. URL: https://cran.r-project.org/package=raster.

[39] Robert J. Hijmans et al. "Very high resolution interpolated climate surfaces for global land areas." In: *International Journal of Climatology* 25.15 (2005), pp. 1965–1978. DOI: 10.1002/joc.1276.

[40] Niklas Janz. "Ehrlich and Raven revisited: Mechanisms underlying codiversification of plants and enemies." In: *Annual Review of Ecology, Evolution, and Systematics* 42.1 (Dec. 2011), pp. 71–89. DOI: 10.1146/annurev-ecolsys-102710-145024.

[41] T. Jermy. "Evolution of insect-plant relationships - a devil's advocate approach." In: *Entomologia Experimentalis et Applicata* 66.1 (1993), pp. 3–12. DOI: 10.1111/j.1570-7458.1993.tb00686.x.

[42] Tibor Jermy. "Evolution of insect/host plant relationships." In: *The American Naturalist* 124.5 (1984), pp. 609–630.

[43] Thibaut Jombart. "Adegenet: A R package for the multivariate analysis of genetic markers." In: *Bioinformatics* 24.11 (2008), pp. 1403–1405. DOI: 10.1093/bioinformatics/btn129.

[44] Thibaut Jombart and Ismaïl Ahmed. "adegenet 1.3-1: New tools for the analysis of genome-wide SNP data." In: *Bioinformatics* 27.21 (2011), pp. 3070–3071. DOI: 10.1093/bioinformatics/btr521.

[45] Katrina M Kaur et al. "Using text-mined trait data to test for cooperate-and-radiate co-evolution between ants and plants." In: *PLOS Computational Biology* 15.10 (Oct. 2019). Ed. by Jeremy M. Beaulieu, e1007323. DOI: 10.1371/journal.pcbi.1007323.

[46] Atsushi Kawakita et al. "Mutualism favours higher host specificity than does antagonism in plant-herbivore interaction." In: *Proceedings of the Royal Society B: Biological Sciences* 277.1695 (2010), pp. 2765–2774. DOI: 10.1098/rspb.2010.0355.

[47] Michael Kopp and Sergey Gavrilets. "Multilocus genetics and the coevolution of quantitative traits." In: *Evolution* 60.7 (2006), pp. 1321–1336.

[48] Thorfinn Sand Korneliussen, Anders Albrechtsen, and Rasmus Nielsen. "ANGSD: Analysis of Next Generation Sequencing Data." In: *BMC Bioinformatics* 15.1 (Dec. 2014), p. 356. DOI: 10.1186/s12859-014-0356-4.

[49] Chih Lee, Ali Abdool, and Chun Hsi Huang. "PCA-based population structure inference with generic clustering algorithms." In: *BMC Bioinformatics* 10.SUPPL. 1 (2009), pp. 1–13. DOI: 10.1186/1471-2105-10-S1-S73.

[50] Pierre Legendre, Francois-Joseph Lapointe, and Philippe Casgrain. "Modeling brain evolution from behavior: A permutational regression approach." In: *Evolution* 48.5 (1994), p. 1487. DOI: 10.2307/2410243.

[51] Heng Li and Richard Durbin. "Fast and accurate short read alignment with Burrows-Wheeler transform." In: *Bioinformatics* 25.14 (July 2009), pp. 1754–1760. DOI: 10.1093/bioinformatics/btp324.

[52] Jeremy W. Lichstein. "Multiple regression on distance matrices: A multivariate spatial analysis tool." In: *Plant Ecology* 188.2 (2007), pp. 117–131. DOI: 10.1007/s11258-006-9126-3.

[53] Catherine R. Linnen and Brian D. Farrell. "A test of the sympatric host race formation hypothesis in Neodiprion (Hymenoptera: Diprionidae)." In: *Proceedings of the Royal Society B: Biological Sciences* 277.May (May 2010), pp. 3131–3138. DOI: 10.1098/rspb.2010.0577.

[54] Min Liu et al. "Contrasting genetic responses to population fragmentation in a coevolving fig and fig wasp across a mainland-island archipelago." In: *Molecular Ecology* 22.17 (2013), pp. 4384–4396. DOI: 10.1111/mec.12406.

[55] John L. Maron, Anurag A. Agrawal, and Douglas W. Schemske. "Plant–herbivore coevolution and plant speciation." In: *Ecology* 100.7 (2019), pp. 1–11. DOI: 10.1002/ecy.2704.

[56] Alan W. Meerow et al. "Phylogenetic analysis of seven WRKY genes across the palm subtribe Attaleinae (Arecaceae) identifies Syagrus as sister group of the coconut." In: *PloS ONE* 4.10 (Jan. 2009), e7353. DOI: 10.1371/journal.pone.0007353.

[57] Alan W. Meerow et al. "Phylogeny and historical biogeography of the cocosoid palms (Areaceae, Arecoideae, Cocoseae) inferred from sequences of six WRKY gene family loci." In: *Cladistics* 31 (2015), pp. 509–534. DOI: 10.111/cla.12100.

[58] Jonas Meisner and Anders Albrechtsen. "Inferring population structure and admixture proportions in low-depth NGS data." In: *Genetics* 210.2 (2018), pp. 719–731. DOI: 10.1534/genetics.118.301336.

[59] Larry R. Noblick. "A revision of the genus Syagrus (Arecaceae)." In: *Phytotaxa* 294.1 (2017), pp. 001–262. DOI: 10.11646/phytotaxa.294.1.1.

[60] Luis A Núñez-Avellaneda, Carolina Isaza, and Gloria Galeano. "Ecología de la polinización de tres especies de Oenocarpus (Arecaceae) simpátricas en la Amazonia colombiana." In: *Revista de Biologia Tropical* 63.1 (2015), pp. 35–55.

[61] Luis A Núñez-Avellaneda and R. Rojas-Robles. "Biología reproductiva y ecología de la polinización de la palma milpesos O enocarpus bataua en los Andes colombianos." In: *Caldasia* 30.1 (2008), pp. 101–125.

[62] Ann-Marie Oppold and Markus Pfenninger. " Direct estimation of the spontaneous mutation rate by short-term mutation accumulation lines in Chironomus riparius." In: *Evolution Letters* 1.2 (2017), pp. 86–92. DOI: 10.1002/evl3.8.

[63] Edzer Pebesma. *sf: Simple Features for R*. 2018. URL: https://cran.r-project.org/package=sf.

[64] Brant K. Peterson et al. "Double Digest RADseq: An inexpensive method for de novo SNP discovery and genotyping in model and non-model species." In: *PLoS ONE* 7.5 (May 2012), e37135. DOI: 10.1371/journal.pone.0037135.

[65] Sergio E. Ramos and Florian P. Schiestl. "Rapid plant evolution driven by the interaction of pollination and herbivory." In: *Science* 364.6436 (2019), pp. 193–196. DOI: 10.1126/science.aav6962.

[66] Benjamin E R Rubin, Richard H Ree, and Corrie S. Moreau. "Inferring phylogenies from RAD sequence data." In: *PloS ONE* 7.4 (Jan. 2012), e33394. DOI: 10.1371/journal.pone.0033394.

[67] Jordan D. Satler and Bryan C. Carstens. "Do ecological communities disperse across biogeographic barriers as a unit?" In: *Molecular Ecology* 26.13 (2017), pp. 3533–3545. DOI: 10.1111/mec.14137.

[68] Jordan D. Satler and Bryan C. Carstens. "The Sarracenia alata pitcher plant system and obligate arthropod inquilines should be considered an evolutionary community." In: *Journal of Biogeography* 46.2 (Feb. 2019), pp. 485–496. DOI: 10.1111/jbi.13498.

[69] Klaus P. Schliep. "phangorn: phylogenetic analysis in R." In: *Bioinformatics* 27.4 (Feb. 2011), pp. 592–3. DOI: 10.1093/bioinformatics/btq706.

[70] Meredith C. Schuman et al. "How does plant chemical diversity contribute to biodiversity at higher trophic levels?" In: *Current Opinion in Insect Science* 14 (2016), pp. 46–55. DOI: 10.1016/j.cois.2016.01.003.

[71] Ilse Silberbauer-Gottsberger, Sergio Antonio Vanin, and Gerhard Gottsberger. "Interactions of the Cerrado palms Butia paraguayensis and Syagrus petraea with parasitic and pollinating insects." In: *Sociobiology* 60.3 (Nov. 2013), pp. 306–316. DOI: 10.13102/sociobiology.v60i3.306-316.

[72] Christopher Irwin Smith et al. "Comparative phylogeography of a coevolved community: Concerted population expansions in Joshua trees and four Yucca moths." In: *PLoS ONE* 6.10 (2011). DOI: 10.1371/journal.pone.0025628.

[73] Christopher Irwin Smith et al. "Distinguishing coevolution from covicariance in an obligate pollination mutualism: Asynchronous divergence in Joshua tree and its pollinators." In: *Evolution* 62.10 (2008), pp. 2676–2687. DOI: 10.1111/j.1558-5646.2008.00500.x.

[74] M. C.P. Souza et al. "Phylogeography of the palm Syagrus coronata (Martius) Beccari (Arecaceae): distribution in the "Caatinga" and Atlantic forest domains." In: *Revista Brasileira de Botanica* 41.4 (2018), pp. 849–857. DOI: 10.1007/s40415-018-0498-0.

[75] Stan Development Team. *{RStan}: the {R} interface to {Stan}*. 2019. URL: http://mc-stan.org/.

[76] Nigel E. Stork et al. "New approaches narrow global species estimates for beetles, insects, and terrestrial arthropods." In: *Proceedings of the National Academy of Sciences* 112.24 (2015), pp. 7519–7523. DOI: 10.1073/pnas.1502408112.

[77] Torsten H. Struck et al. "Finding evolutionary processes hidden in cryptic species." In: *Trends in Ecology and Evolution* 33.3 (2017), pp. 153–163. DOI: 10.1016/j.tree.2017.11.007.

[78] John N. Thompson. *The geographic mosaic of coevolution*. Chicago: The University of Chicago Press, 2005, xii+443p.

[79] John N. Thompson, Kari A Segraves, and David M. Althoff. "Coevolution and Macroevolution." In: *Evolutionary Developmental Biology*. Ed. by L. Nuño de la Rosa and G. B. Müller. 2017, pp. 1–13. DOI: 10.1007/978-3-319-33038-9. URL: http://link.springer.com/10.1007/978-3-319-33038-9.

[80] Enwei Tian et al. "Lack of genetic isolation by distance, similar genetic structuring but different demographic histories in a fig-pollinating wasp mutualism." In: *Molecular Ecology* 24.23 (2015), pp. 5976–5991. DOI: 10.1111/mec.13438.

[81] Hirokazu Toju and Teiji Sota. "Phylogeography and the geographic cline in the armament of a seed-predatory weevil: Effects of historical events vs. natural selection from the host plant." In: *Molecular Ecology* 15.13 (2006), pp. 4161–4173. DOI: 10.1111/j.1365-294X.2006.03088.x.

[82] Kim L. Vertacnik and Catherine R. Linnen. "Evolutionary genetics of host shifts in herbivorous insects: insights from the age of genomics." In: *Annals of the New York Academy of Sciences* 1389.1 (2017), pp. 186–212. DOI: 10.1111/nyas.13311.

[83] George D Weiblen and Erin L Treiber. "Evolutionary origins and diversification of mutualism." In: *Mutualism* (2015), pp. 37–56. DOI: 10.1093/acprof:oso/9780199675654.003.0003.

[84] Yong Xiao et al. "The genome draft of coconut (Cocos nucifera)." In: *GigaScience* 6.11 (2017), pp. 1–11. DOI: 10.1093/gigascience/gix095.

[85] Jeremy B. Yoder and Scott L Nuismer. "When does coevolution promote diversification?" In: *The American naturalist* 176.6 (Dec. 2010), pp. 802–17. DOI: 10.1086/657048.

[86] Zhi Qiang Zhang. "Animal biodiversity: An introduction to higher-level classification and taxonomic richness." In: *Zootaxa* 12.3148 (2011), pp. 7–12.

[87] Alexander Zizka et al. "CoordinateCleaner: Standardized cleaning of occurrence records from biological collection databases." In: *Methods in Ecology and Evolution* 10.5 (May 2019). Ed. by Tiago Quental, pp. 744–751. DOI: 10.1111/2041-210X.13152.