Systematic Entomology

Systematic Entomology (2020), **45**, 594–605 DOI: 10.1111/syen.12415

Paraphyletic species no more – genomic data resolve a Pleistocene radiation and validate morphological species of the *Melanoplus scudderi* complex (Insecta: Orthoptera)

JEN-PAN HUANG[1], JOVONN G. HILL[2], JOAQUÍN ORTEGO[3] and L. LACEY KNOWLES[4]

[1]Biodiversity Research Center, Academia Sinica, Taipei, Taiwan, [2]Mississippi Entomological Museum, Department of Entomology and Plant Pathology, Mississippi State University, Mississippi State, MS, U.S.A., [3]Department of Integrative Ecology, Estación Biológica de Doñana (EBD-CSIC), Seville, Spain and [4]Department of Ecology and Evolutionary Biology, University of Michigan, Ann Arbor, MI, U.S.A.

Abstract. Rapid speciation events, with taxa generated over a short time period, are among the most investigated biological phenomena. However, molecular systematics often reveals contradictory results compared with morphological/phenotypical diagnoses of species under scenarios of recent and rapid diversification. In this study, we used molecular data from an average of over 29 000 loci per sample from RADseq to reconstruct the diversification history and delimit the species boundary in a short-winged grasshopper species complex (*Melanoplus scudderi* group), where Pleistocene diversification has been hypothesized to generate more than 20 putative species with distinct male genitalic shapes. We found that, based on a maximum likelihood molecular phylogeny, each morphological species indeed forms a monophyletic group, contrary to the result from a previous mitochondrial DNA sequence study. By dating the diversification events, the species complex is estimated to have diversified during the Late Pleistocene, supporting the recent radiation hypothesis. Furthermore, coalescent-based species delimitation analyses provide quantitative support for independent genetic lineages, which corresponds to the morphologically defined species. Our results also showed that male genitalic shape may not be predicted by evolutionary distance among species, not only indicating that this trait is labile, but also implying that selection may play a role in character divergence. Additionally, our findings suggest that the rapid speciation events in this flightless grasshopper complex might be primarily associated with the fragmentation of their grassland habitats during the Late Pleistocene. Collectively, our study highlights the importance of integrating multiple sources of information to delineate species, especially for a species complex that diversified rapidly, and whose divergence may be linked to ecological processes that create geographic isolation (i.e. fragmented habitats), as well as selection acting on characters with direct consequences for reproductive isolation (i.e. genitalic divergence).

Introduction

Rapid diversification events, or radiations, are among the most investigated biological phenomena because they can be useful to infer the proximate causes and constraints that generate and maintain biodiversity (Losos & De Queiroz, 1997; Rundell & Price, 2009). Yet, the study of ecological and phenotypic divergence among closely related taxa that are of primary interest under such scenarios is often confounded by the extreme difficulty of not only confidently identifying species boundaries using molecular data (i.e. many putative species are either poly-

Correspondence: L. Lacey Knowles, Department of Ecology and Evolutionary Biology, University of Michigan, Ann Arbor, MI 48109, U.S.A. E-mail: knowlesl@umich.edu

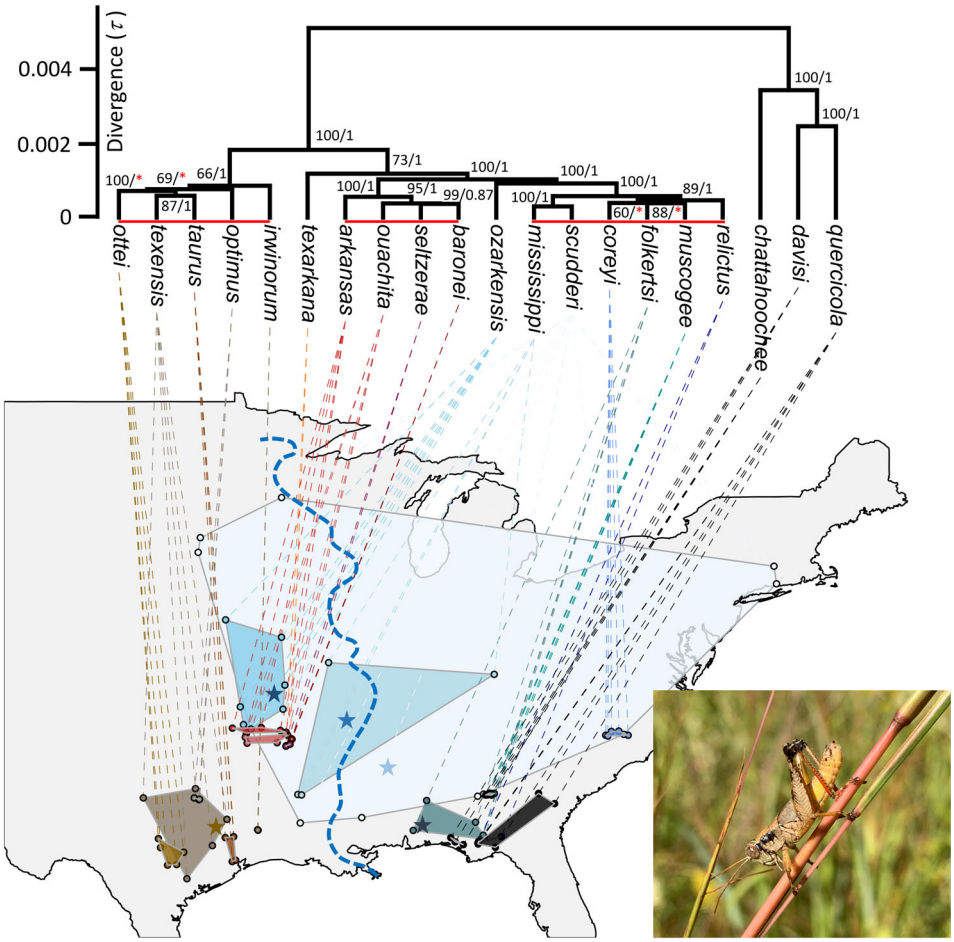

Fig. 1. An estimated species tree of 20 species from the *Melanoplus scudderi* species group and their approximate geographic distribution, as marked by the differently coloured areas based on geographical coordinates from Hill (2015). The sampling sites for geographically widespread species are indicated by asterisks. Numbers on nodes of the tree are bootstrapping support values from the SVDQUARTETS analysis/posterior support values from the BPP analysis. Red asterisks indicate conflicting nodes between species tree topologies reconstructed using SVDQUARTETS (shown in this figure) and BPP (available in Fig. S6). Divergence time is expressed in substitutions per site (τ). The groups of putative species lineages tested in separate species delimitation analyses are marked by red bars at the tree tips. The Mississippi river is identified by the thick blue dashed line. Bottom right: adult *Melanoplus relictus* on *Schizachyrium scoparium* in Thomas Co., Georgia, U.S.A. (image taken by J. Hill). [Colour figure can be viewed at wileyonlinelibrary.com].

phyletic or paraphyletic), but also obtaining a robust reconstruction of the underlying history (i.e. conflicting gene tree topologies can complicate phylogenetic estimation) (e.g. the Lake Victoria cichlids: Wagner *et al.*, 2013; Darwin's finches: Farrington *et al.*, 2014; Cadena *et al.*, 2018). Moreover, without the statistic confidence from molecular or phenotypic data to support inferred putative species, the proposition of a radiation and high species diversity itself may be questionable (e.g. artifacts of taxonomic inflation; Cadena *et al.*, 2018). That is, the underlying mechanisms leading to speciation can be confounded with those that are responsible for population subdivision (Smith *et al.*, 2013; Sukumaran & Knowles, 2017). When different phylogenetic histories are reconstructed because of choices of different datasets and the criteria for identifying taxa, this contributes to further conflicting inferences of diversification processes and history.

Such challenges, including conflicting results across data types, are exemplified in the short-wing grasshoppers belonging to the *Melanoplus scudderi* (Uhler, 1964) species complex (Fig. 1). There are more than 20 recognized putative species in the *M. scudderi* species complex according to a recent revision (Hill, 2015). The complex is geographically widespread. However, most species have restricted geographic distributions (Fig. 1). They are recognized by a unique genitalic shape (Fig. 2), a key diagnostic character in the genus *Melanoplus* (e.g. Knowles & Otte, 2000), and in Orthoptera more generally (e.g. Cohn *et al.*, 2013), whose divergence by sexual selection may play a role in speciation (e.g. Eberhard, 1996), including in grasshoppers (e.g. Márquez & Knowles, 2007; Knowles *et al.*, 2016). However, the putative species are not monophyletic, based on analyses of mitochondrial loci (Hill, 2015). It has been hypothesized that there has been insufficient time for lineage

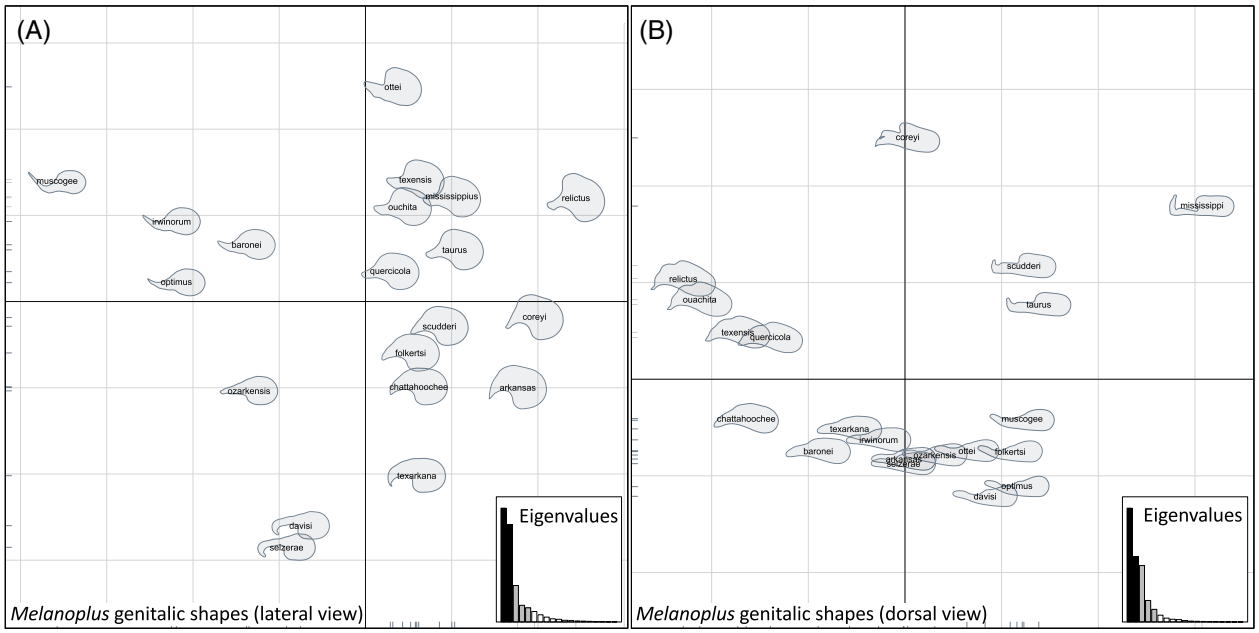

Fig. 2. The position of putative species in the morphospace based on the principal component analyses results of the shape analyses of the lateral (A) and dorsal (B) views of the male genitalia (see Figs S10, S11 for the original shape outlines and Figs S13, S14 for a characterization the mean and variance of shape differences associated with each axis). [Colour figure can be viewed at wileyonlinelibrary.com].

sorting to result in reciprocal monophyly of described taxa. For example, if the reductions of grasslands and expansion of forests in the Late Pleistocene facilitated the rapid diversification in the species group (see Hill, 2015), the lack of monophyly would be expected. On the other hand, it is possible that the divergences among these morphologically distinct taxa represent the incipient stages of speciation (see Huang & Knowles, 2016), without the reproductive isolation necessary for sustained evolutionary independence over time.

The study of the systematics of the *M. scudderi* species complex is also more generally representative of the challenges that arise when the biology of the group seems incongruent with the taxon's traits. For example, before the recent thorough revision (Hill, 2015), only one morphological species was described. Such a systematic proposal is suspect on several fronts. First, biogeographically it would certainly be a puzzling observation that a flightless species would be so widely distributed (see Knowles & Otte, 2000), not to mention the questions that such a range raises about dispersal mechanisms given it seems unlikely that a flightless species would traverse major physical barriers (e.g. the Mississippi river; Fig. 1), thereby preventing the initiation of species differentiation. Moreover, the lack of morphological divergence (or potential 'cryptic' diversity) largely seems to reflect an exclusive focus on easily visible external characters (e.g. wing patterns and body coloration). In fact, when internal characters, and specifically the male genitalia, were examined (Hill, 2015), a diversity of potential taxa were revealed. As such, the 'cryptic' species diversity in the *M. scudderi* species complex, as with other studies, has a direct impact on biodiversity estimates, with downstream implications for ecological and evolutionary studies (Bickford *et al.*, 2007). On the other hand, whether distinct genitalic shapes correspond to species boundaries is not always clear or rarely empirically evaluated [e.g. tests of correlated evolution of the genitalic complex (Márquez & Knowles, 2007); tests of lock-and-key hypothesis (Masly, 2012)]. Likewise, not only in *Melanoplus*, but in other insects as well, whether genital divergence reflects taxonomic oversplitting has been a persistent concern (Masly, 2012). Note that when two species are morphologically and/or ecologically indistinguishable and genitalic differentiation is the only diagnostic character to separate different species, it is conventionally assumed that divergence in male genitalia may function as a reproductive barrier.

In this study, we harnessed the power of next-generation sequencing (McCormack *et al.*, 2013; Andrews *et al.*, 2016) to generate thousands of nuclear loci for estimating phylogenetic relationships and testing hypothesized species boundaries in the *M. scudderi* species complex. Specifically, we: (i) estimated the phylogenetic relationships and times of diversification in the species complex; (ii) tested for a correspondence between morphologically identified species boundaries and genetic lineages detected under a coalescent-model; and (iii) used this framework evaluate the evolution of male genitalia, and, specifically, whether divergence in these characters is correlated with phylogenetic distance, where deviations could suggest the action of selection during the divergence of the putative taxa. We discuss our results with respect to what they highlight in particular about species boundaries and the evolutionary diversification of the *M. scudderi* species complex, as well as more general issues surrounding the reconciliation and interpretation of discordance between data types when studying speciation and species delimitation.

Materials and methods

Species sampling and genomic data

Genomic data were collected from 96 individuals, with three to five representatives per putative species sequenced for each of 20 putative species from the *M. scudderi* species complex (collection and voucher information are in Table S1); only males were sequenced so that individuals could be assigned to putative taxa based on the diagnostic genitalic characters (see Hill, 2015). Specifically, we applied a reduced-representation library to contend with large genomes of *Melanoplus* grasshoppers (*c.* 4 GB). Likewise, we applied analyses that are robust to missing data (i.e. SVDQUARTETS) or carefully selected variable and complete molecular datasets among individual samples for analyses that may (or may not) be sensitive to missing data (i.e. BPP analyses). Detailed methods are provided in the following.

Genomic DNA was extracted from tissues stored in absolute alcohol using Qiagen DNeasy following the manufacturer's Animal Tissue Protocol (Qiagen, Hilden, Germany). Two reduced representation libraries were constructed following a double digest restriction-site associate DNA sequencing protocol (ddRADseq; for details see Peterson *et al.*, 2012); each library contained a total of 48 individuals from all 20 putative species. Briefly, for each library, individuals (with a total genomic DNA of *c.* 160 ng) were double-digested using restriction enzymes *Eco*RI and *Mse*I and uniquely tagged with a 10 bp barcode. The digested and tagged products were then pooled and size-selected for 350–450 bp fragments using a Pippin Prep (Sage Science, Beverly, MA, U.S.A.). Size-selected fragments were PCR-amplified with iProof™ High-Fidelity DNA Polymerase (Bio-Rad, Hercules, CA, U.S.A.). DNA quantification and cleaning with Agencourt AMPure XP (Beckman Coulter, CA, Brea, U.S.A.) occurred after every step in the library construction procedure. Each genomic library was sequenced on an Illumina HiSeq2000 (Illumina, CA, San Diego, U.S.A.) at the Centre for Applied Genomics (Toronto, ON, Canada) to generate 150 bp single-end reads.

The sequence data were processed using PYRAD v.3.0.6 pipeline (Eaton, 2014) to identify single nucleotide polymorphisms (SNPs). Specifically, after demultiplexing, low-quality sites (quality score > 20) were converted to Ns, and reads with a minimum coverage depth ($m = 5$) were assembled into putative alleles. Loci with more than two alleles per individual sample were discarded, given they violate diploid expectations. A de novo assembly and alignment of different loci across all the sampled individuals were then performed for sequences with a minimum similarity of 88% and a maximum of two low-quality sites per locus (assemblies using minimum sequence similarities of 85% and 90% were also generated and did not differ significantly). Downstream sites after site 115 were excluded for all aligned clusters to remove low-quality bases near the 3′ ends (see Figs S1, S2). Additionally, we excluded loci containing one or more heterozygous sites shared across more than 10 individuals (or approximately two putative species) to account for incomplete lineage sorting and/or potential interspecific gene flow between closely related taxa while avoiding indiscriminate clustering of paralogues (i.e. shared heterozygous sites across too many samples most probably represent clustering of paralogues with a fixed difference rather than true heterozygous sites; Eaton & Ree, 2013). A concatenated unlinked SNP file (i.e. with one randomly selected SNP from each locus) and a file with individually aligned loci in separate blocks were generated for loci that were present in at least ten individuals.

A total of 167 078 647 raw reads passed the initial quality control, with an average of 1 740 403 good reads per sample (SE = 81 147). Clustering analyses within samples identified an average of 59 615 clusters (SE = 1710) with a minimum of 5× and mean 18× per cluster sequencing depth (Table S2). The final dataset contained 93 720 variable loci with an average of 29 644 (SE = 830) loci per individual; a summary of pre- and post-processing read numbers is given in Table S2. The number of loci shared between individuals is not correlated with their phylogenetic relatedness, indicating that the amount of missing data between individuals was not caused by mutations in the restriction sites (Huang & Knowles, 2016b; see Fig. S7 and the Results section), and instead reflects the finite number of reads available for two lanes of sequencing on a Illumina HiSeq2000.

Phylogenetic and divergence time estimation

Phylogenetic relationships were estimated from the genomic data using two approaches: one in which the data across loci were concatenated, and one in which potential genealogical discord across loci is accommodated (i.e. using a coalescent-based analysis; Knowles & Kubatko, 2010). An unlinked SNP file from the pyRAD pipeline (Eaton, 2014) was used for phylogenetic reconstruction (a total of 94 846 loci).

For the analyses based on concatenation of the data, one SNP, which was shared across most of the analysed individuals, was randomly chosen for each locus to build a concatenated unlinked SNP dataset. A maximum likelihood (ML) tree was estimated using the GTRGAMMA model implemented in RAXML v.8.1.17 (Stamatakis, 2014). We also performed a fast bootstrapping analysis with 100 replicates in RAXML with the model GTRCAT to assess nodal support and the reconstructed ML tree was visualized in FIGTREE v.1.4.2 (http://tree.bio.ed.ac.uk/software/figtree/) using the midpoint rooting option.

An estimate of phylogenetic relationships based on the unlinked SNP dataset was generated using the coalescent-based analysis of SVDQUARTETS (Chifman & Kubatko, 2014) implemented in PAUP* v.4.0a (Swofford, 2002). Specifically, an exhaustive search of all possible quartets options was specified for estimating a bifurcating tree from the sampled quartets using Quartet FM (Fiduccia & Mattheyses, 1982). Each SNP was treated as an independent locus and 1000 bootstrapping replicates were analysed based on the same search options as described earlier to estimate branch support. The estimated tree was visualized using FIGTREE; note that the same node as in the ML phylogenetic analysis based on concatenated data was used to root the tree.

Times of divergence among the 20 putative species of the *M. scudderi* species complex were estimated using the program BPP v.4 (Flouri *et al.*, 2018). The species tree estimated with

SVDQUARTETS with species (not individuals) representing operational taxonomic units was used as a fixed species tree topology to estimate divergence times (i.e. applying the A00 model in BPP). A series of R scripts were applied to edit and convert the data file from pyRAD (.loci file) analysis into a BPP input file (see Huang 2016, Huang, 2018). For the data used in this analysis: (i) only sequence clusters, or loci, present in a at least one representative per taxon were retained; and (ii) the maximum pairwise sequence divergence between samples was set to 8% following Hill (2015) based on analyses of mitochondrial sequence data; the common threshold was chosen to facilitate comparison across studies. The resulting dataset contained 4132 variable loci (Fig. S2). Branch lengths were estimated using the full set of 4132 loci, and using four sets of 1000 randomly chosen loci, where the loci were mutually exclusive among the four 1000 loci datasets. The BPP analyses were run for 10^5 generations with a sampling frequency of 2 and a burn-in period of 10^4, and we specified the diploid variable to indicate that all the sequences from each individual were unphased. We also estimated the species tree topology (i.e. using the A01 model in BPP) based on the full 4132 loci dataset using BPP. The distributions of θ and τ priors were informed from point estimates based on the number of segregation site per site (Figs S2, S3).

Species delimitation

Three recently diversified clades (Fig. 1), each with four to six morphologically well-differentiated putative species, were selected for species delimitation analyses. It is theoretically possible to perform species delimitation with 20 species using the program BPP (v.4; Flouri *et al.*, 2018), given the compatibility with short reads of unlinked neutral loci from ddRADseq libraries (see Flouri *et al.*, 2018). However, such an analysis would be challenging because parallel computation was not available when we conducted the analyses. Furthermore, given that the main idea of our study is to test whether recently diverged lineages can be supported by genomic data as distinct species, the exclusion of distantly related lineages with respect to each focal clade of interest is reasonable (Fig. 1).

For BPP analyses, we randomly chose 100 loci from the 4132 loci dataset for each of the separate analyses of the different subsets of taxa (see Fig. 1 for the identities of putative taxa included for each subset); the analyses were repeated five times for each subset using randomly chosen, and mutually exclusive, sets of 100 loci each. Priors for θ and τ values were informed from point estimates based on the completed 4132 loci dataset (Fig. S4). To account for phylogenetic uncertainty among putative species (see Results section; Fig. 1), we used the unguided option in BPP to estimate the supports of evolutionary independence for each analysis (A11 model; Yang & Rannala, 2014). The BPP analyses were run for 5×10^5 generations with a sampling frequency of ten and a burn-in period of 5×10^4.

Morphological analyses of male genitalia

Outlines of the lateral and dorsal view of the male genitalia, specifically the shape of the apical valves of the aedeagus, were collected from digital images of specimens (data available as supplementary data in Dryad; see also Hill, 2015); for details regarding outline data collection see the protocol detailed in Márquez & Knowles (2007). We analysed one individual per species, a standard approach applied in similar comparative phenotypic and phylogenetic studies on grasshoppers and crickets (e.g. Márquez & Knowles, 2007; Oneal & Knowles, 2013; Knowles & Cohn, 2016), as well as in other animal groups [e.g. catfish (Silva *et al.*, 2016) and turtles (Dickson & Pierce, 2019)]. Shape analyses were performed on the outline coordinates using the R package MOMOCS (Bonhemme *et al.*, 2014). Specifically, elliptical Fourier analyses were performed (Kuhl & Giardina, 1982) and the best number of harmonics was chosen based on the harmonic power that described the analysed shapes. Principal component analyses (PCA) were applied to quantitatively investigate the differences in the outlines of male genitalic shapes among putative species.

To examine the evolution of the male genitalia, we tested whether there is a significant correlation between phylogenetic distance and morphological difference in male genitalic shapes. Specifically, phylogenetic distance between pairs of putative species was calculated using the 'cophenetic' function implemented in APE (Paradis & Schliep, 2019) based on the branch lengths in the reconstructed species tree, where the branch lengths were estimated using BPP (Fig. 1). The difference in male genitalic shape between pairs of species was represented by Euclidean distance calculated based on the principal component (PC1 and PC2) coordinates from the PCAs. The 'phylomorphospace' function implemented in PHYTOOLS (Revell, 2012) was used to visualize the correlation between phylogenetic distance and shape difference in the male genitalia. Similarly, the PC1 and PC2 coordinates from PCAs were used as continuous phenotypic datasets for generating a phylomorphospace plot, and visualized using the 'Phenogram' function implemented in PHYTOOLS for each of the four retained PCs.

We also tested whether the divergence in male genitalic shape among the putative species was consistent with a Brownian motion model. For these analyses, we extracted the first two PCs from the PCAs of the lateral and dorsal shape variation of the male genitalia. Multidimensional phylogenetic signal tests using the extracted PCs with 1000 permutations were used to test a Brownian motion model with the 'fast.SSC' function in the R package RPHYLOPARS (Goolsby *et al.*, 2017). The function uses a fast ancestral state reconstruction algorithm to calculate the sum of square changes between ancestral and descendant nodes/tips based on a model described in Klingenberg & Gidaszewski (2010). Specifically, we tested: (i) all four extracted PCs; (ii) the two PCs from lateral shape; and (iii) the two PCs from dorsal shape; and (iv) each of the individual PCs. A scaled sum of squared changes = 1 indicates the evolution of the continuous trait data can be explained by Brownian motion. A value > 1 indicates that the evolution of the trait data is more conserved than expected given their phylogenetic distance, while a value < 1 implies that the trait evolution is more labile between taxa than expected by their phylogenetic distance.

Data availability

All raw sequence data are archived in the NCBI sequence read archive database under bioproject PRJNA574680. R scripts and input data for SVDQUARTETS, BPP, and shape analyses are available in the Dryad data repository (doi: 10.5061/dryad.dncjsxkvg).

Results

Monophyly of putative species and divergence time estimates

Estimates of the phylogenetic relationships among individuals based on the analysis of the concatenated data (Fig. S5) show that morphological identified putative species were monophyletic with high bootstrap support (> 90%). The species tree estimated using the coalescent-based approach of SVDQUARTETS were generally consistent with the relationships inferred from the concatenated data (Fig. 1). Moreover, the estimated relationships correspond to the geography of putative species distributions. For example, taxa east of the Mississippi river [i.e. *Melanoplus quercicola* (Hebard 1918), *Melanoplus davisi* (Hebard 1918) and *Melanopolus chattahoochee* (Hill, 2015)] form a well-supported (bootstrap values = 100) monophyletic group. Among the 17 remaining putative species, which are west of the Mississippi river, clear geographic structure [especially among *Melanoplus irwinorum* (Hill, 2015), *Melanoplus optimus* (Hill, 2015), *Melanoplus ottei* (Hill, 2015), *Melanoplus taurus* (Hill, 2015), and *Melanoplus texensis* (Hart 1906)] is evident based on the phylogenetic relationships, with relatively recent divergences (see Fig. 1).

A species tree topology search using BPP resulted in similar species tree topology with high posterior probability support (Figure S6), with few minor differences (compare with Fig. 1). The most probable species tree topology from BPP analysis (Fig. S6) was observed in 86.7% of the post-burn-in species trees. The subject dataset did not show biased sampling effect of shared loci between closely related species (Fig. S7), because we specifically selected loci that do not have any missing taxon (see Materials and methods for details).

The divergence time estimates were consistent across four independent subsets of 1000 loci (Figs S8, S9). Assuming a generation time of 1 year, the common ancestor of the 17 western species may have originated 3333000 years ago (333.3 kya) (based on the genomic mutation rate of 2.9×10^{-9} estimated for *Heliconius melpomene*; Keightley *et al.*, 2015) or 62.5 kya (based on a genomic mutation rate of 1.6×10^{-8} estimated for *Aedes aegypti*; Sherpa *et al.*, 2018) ($\tau = 2\mu t$, where τ is the estimated population divergence in substitutions per site, μ is the per site mutation rate per generation, and t is the absolute divergence time in years). The time to the most recent common ancestor for the species from the *M. scudderi* species group (i.e. the common ancestor of *M. scudderi*, *M. mississippi*, *M. coreyi*, *M. muscogee*, *M. relictus*, and *M. folkertsi*) that dispersed across the Mississippi river is estimated to be *c.* 0.0006 τ units, or about 100–18.8 kya, depending on the different mutation rate calibrations applied. The most recent speciation event occurred between *M. mississippi* and *M. scudderi* ($\tau = 0.0003$), or around 50 or 9.4 kya, whereas the initial split between the eastern (three species) and the western (17 species) lineages is *c.* 0.0048 τ units (Fig. 1), or 150 or 700 kya depending on which calibrated mutation rate is applied, both of which are consistent with divergence during the Mid-Pleistocene (126–781 kya).

Morphologically diagnosed species boundaries correspond to those delimited genetically

All morphologically diagnosed species boundaries are supported by genetic-based delimitation analyses in BPP analyses with high posterior probability support (posterior probabilities > 0.99). The results were consistent across five replicate datasets with 100 mutually exclusive and randomly chosen loci (see Materials and methods) for each of the analysed subsets of taxa (see Fig. 1 for lineages analysed in each subset). Specifically, from all the post burn-in samples of the Markov chain Monte Carlo searches, < 1% of the species delimitation results lumped species. The only exception involved analyses of the subset *M. mississippi*, *M. scudderi*, *M. coreyi*, *M. folkertsi*, *M. muscogee* and *M. relictus*, in which 16% of the species delimitation results lumped *M. scudderi* and *M. mississippi* into one species. Note that these are the only two species that dispersed across the Mississippi river, and represent the most recent speciation event (see Fig. 1). Although support for the delimited taxa in this group was lower, it was nevertheless still relatively high; the probability that all six of them are independently evolving lineages is *c.* 84%.

Evolution of male genitalia

Morphological dissimilarity was not correlated with phylogenetic distance among species (Fig. 3). When visualizing the phenotypic data and the phylogenetic relationship among species using dorsal and lateral genitalic shapes (Fig. 3), there is no detectable phylogenetic structure on phenotypic trait values. For example, sister species *M. scudderi* and *M. mississippi* have similar genitalic shapes, while another sister species pair, *M. folkertsi* and *M. muscogee*, exhibit very different genitalic shapes. On the other hand, distantly related species such as *M. davisi* and *M. seltzerae* have very similar shaped genitalia (Figs 2,3, S10, S11).

Tests of the four PC values (PC1 and PC2 from both the lateral and dorsal genitalic shape analyses) suggest that evolution of genitalic shape is more labile than expected by their phylogenetic relationships among species (total sum of squared changes = 0.380 17; scaled sum of squared changes = 0.543 03); however, this deviation from a Brownian motion model is not statistically significant ($P = 0.158$). Analyses of the lateral and dorsal genitalic shape variables showed similar results. Specifically, total sum of squared changes = 0.623 96 and 0.380 17, and scaled sum of squared changes = 0.557 29 and 0.543 03 were estimated for the lateral

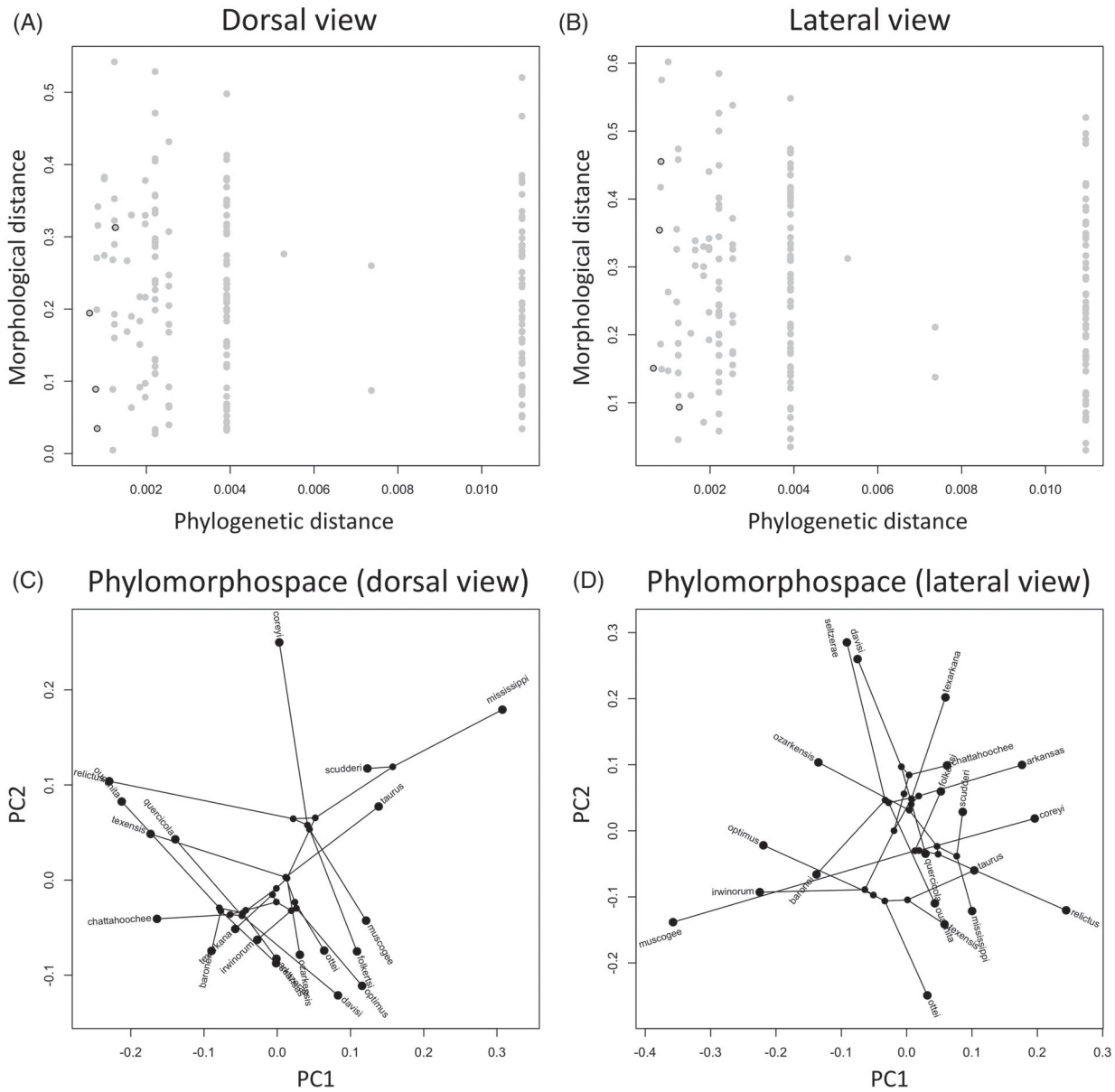

Fig. 3. Plots for association between phylogenetic relationships and dissimilarities between male genitalic shapes. (A, B) Plots of phylogenetic distance between species against the shape dissimilarity between species calculated based on the Euclidean distance derived from principal component 1 (PC1) and PC2 coordinates; (C, D) phylomorphospace plots derived using the PC1 and PC2 values from the shape analyses and the reconstructed species tree (Fig. 1) using the *phylomorphospace* function in PHYTOOLS.

and dorsal shape variables, respectively, neither of which is significant ($P = 0.57$ and 0.17, respectively). Analysing each individual PC also resulted in the same inference; all the tests resulted in values of scaled sum of squared changes *c*. 0.5 and *P*-values > 0.1.

The shape analyses (Fig. S12) showed that > 20 harmonics are sufficient to describe lateral and dorsal shape variation in the genitalia (Fig. S13). The first two axes of the PCA explained > 50% of the shape variation across putative species (Figs 2, S14); all shapes were correctly centred with a common point of origin (Fig. S13).

Discussion

The study of recent biological radiation events has been extremely insightful with regard to the origin and structuring of global biodiversity patterns (e.g. the adaptive radiations in the Darwin's finches and the Lake Victoria cichlids), but rampant conflicts among data types can obscure the systematic frameworks upon which such studies rely. Our study of the recent radiation of a grasshopper species complex demonstrates how, with the resolution provided by genomic data and powerful analytical approaches, we can move beyond the confusion

surrounding phylogenetic relationships and species boundaries when diversification has occurred rapidly in the recent past. Specifically, our analysis not only confirms the validity of utilizing variation in genitalic shape to delimit recently diverged evolutionary entities, by showing that distinct male genitalic shapes correspond to independently evolving lineages identified by the genomic data, but also shows that divergence in these characters occurred rapidly and recently. In the following, we discuss potential factors that may have contributed to the diversification of taxa in the *M. scudderi* species complex, as well as the importance of integrating phenotypic and genomic data into insect systematics, and of delimiting species of *Melanoplus* grasshoppers in particular.

Multiple evolutionarily independent lineages radiated in the Late Pleistocene

The Pleistocene glacial cycles have promoted speciation in many North American faunas, including grasshoppers (Knowles, 2001; Knowles & Richards, 2005; Carstens & Knowles, 2007). However, many of such studies have focused on speciation in montane taxa, where the effect of glacial cycles are manifested in shifts in the altitudinal distribution of suitable habitats. For members of the *M. scudderi* species complex from the eastern and central regions of southern North America, the geography of species divergence occurred on a fundamentally different landscape. Yet, diversification in many ways mirrors that of montane *Melanoplus* grasshoppers, i.e. rapid and recent speciation (Fig. 1).

With regard to geographic isolation, our study may fit with others that have invoked the fragmentation, or the horizontal restructuring of habitat types from the great plains during the Late Pleistocene as a potentially important factor in the divergence of species from the southeastern United States (e.g. Satler & Carstens, 2016). Fossil pollen records indicate that during the change from a dry to a wet climate during the Late Pleistocene to the Early Holocene, including during the last glacial maximum (LGM), the southeastern United States was dominated by prairie and savannah habitat types (Russel et al., 2009). However, this once continuously distributed grassland was subsequently fragmented into multiple patches after the LGM by expanding forest habitats. Furthermore, with the wetter climate, potentially increased water flow in the rivers and their drainage systems (Bentley Sr. *et al.*, 2016) may have created additional physical barriers that further contributed to the isolation of grassland inhabitants (Hill, 2015). As such, and as grassland-dependent species, populations of taxa from the *M. scudderi* species complex may have become recently fragmented into local grassland patches, contributing to the divergence into multiple species as past widespread ancestral populations became locally isolated (Hill, 2015). Such a scenario is consistent with several aspects of the history of diversification inferred from the genomic data. For example, the reconstructed species tree reveals an initial divergence between an eastern lineage, which is composed of *M. davisi*, *M. quercicola*, and *M. chattahoochee*, and a western lineage, which contains the remaining 17 species from the complex (Fig. 1). Subsequent to the isolation of the eastern and western lineages, a western lineage, but not an eastern one, dispersed across the Mississippi river in Late Pleistocene (ranging from 100 to 18.8 kya; see the Results section), consistent with the hypothesis that previously expansive areas of grassland on the North American Coastal Plain formed in the Mississippi river delta because of the lowered sea level in the Pleistocene (Bentley Sr. *et al.*, 2016), possibly facilitating dispersal of the flightless *Melanoplus* grasshoppers across a main geographic barrier, the Mississippi river. Many terrestrial organisms from the Appalachian mountains and the southeastern coastal region of the United States have retreated into hypothesized climatic refugia in the glacial periods during the Pleistocene (e.g. Walker *et al.*, 2009). This climatic constraint may help to explain why the eastern lineage did not expand westwards, while the western lineage expanded eastwards (Fig. 1). Note that the two geographically widespread lineages, *M. scudderi* and *M. mississippi*, which are distributed on both sides of the Mississippi river (Fig. 1; Hill, 2015), originate from western ancestors.

Evolution of genitalia variation and applications for delimiting species

Genitalic morphology has been an important, although sometimes debated, character in taxonomic studies in Orthoptera (e.g. Cohn *et al.*, 2013). With differentiation in size, shape and genitalic structures, evolutionary changes in genitalic morphology between species have been hypothesized to play an important role in intraspecific recognition during copulation, and may be under strong sexual selection (e.g. Eberhard, 1996; Masly, 2012; Oneal & Knowles, 2015; Fujisawa *et al.*, 2019). The genetic distinctiveness of the species revealed from the coalescent-based delimitation analyses indicates little (if any) gene flow between species with different male genitalic shapes, where genetic distinctiveness is maintained despite recent species divergence. In particular, species with different male genitalic shapes form reciprocal monophyletic groups in the *M. scudderi* species complex (Fig. S5) and their evolutionary independence is statistically supported by delimitation analysis based on the genomic data.

The correspondence between phenotypic and genetic divergence not only corroborates hypotheses about species boundaries, but also implies that divergence in genitalic shapes reflects selectively driven divergence. For example, when considering past analyses based on mitochondrial genes, it is clear that the evolution of male genitalic shape in the *M. scudderi* species complex outpaces the sorting of shared ancestral mitochondrial nucleotide variation after species divergence (cf. Hill, 2015: figs 49, 50). Specifically, different genetic groups (Figs 1, S5) can be unambiguously distinguished by examining the male genitalic shape, while mitochondrial sequence data failed to identify the different evolutionary entities (Hill, 2015). Moreover, some recent speciation events (< 10 kya), such as the case of the sister taxa *M. mississippi* and *M. scudderi*,

can be inferred based on male genitalic shape (Hill, 2015). Furthermore, these sister species have extensive geographic overlap (Fig. 1), while their genetic uniqueness is still well maintained (as supported by BPP results). That is, the lack of genetic distinctiveness in mitochondrial genetic variation reflects insufficient time for sorting, and as such is not in conflict with a recent divergence history because mitochondrial DNA simply cannot effectively discriminate species of rapid diversification origin (Hickerson *et al.*, 2006; Wagner *et al.*, 2013). Such species diversity can, however, be successfully identified by carefully examining male genitalia variation or using genomic data from hundreds of thousands of loci, as our results show.

Conflicting results in systematic studies happen frequently when different data types are used to delimit species or to reconstruct their evolutionary history (Hillis, 1987; Sites & Marshall, 2004; Cohen, 2018). However, when the basis of this conflict is examined, the conflict itself, while inconvenient for species delimitation, can be informative about the speciation process and help make robust, biologically meaningful inferences about taxonomy (Yang & Rannala, 2014; Huang & Knowles, 2016; Sukumaran & Knowles, 2017). Specifically, because morphological traits subject to strong natural or sexual selection, such as the genitalic morphology (Eberhard, 1996; Masly, 2012), can evolve rapidly, species-specific phenotypes may become evident before neutral genetic markers show differentiation (e.g. Sites & Marshall, 2004; Wagner *et al.*, 2013; Solis-Lemus *et al.*, 2014). Likewise, the conflicts among data types and between molecular datasets revealed in the systematics of the *M. scudderi* species complex have collectively provided insights into its diversification history and, thus, such differences should not hamper taxonomic decisions.

Determining whether changes in genitalic shape are the cause or a consequence of speciation is difficult (e.g. Oneal & Knowles, 2013). The geographic distribution pattern of species in the *M. scudderi* species complex implies a dominant role for allopatric speciation, potentially in association with the rapid and simultaneous fragmentation of grassland habitats during the Late Pleistocene (see earlier). Currently, there is not much geographic overlap between species (Fig. 1), which indicates that secondary contacts between diverged lineages have been rare. It is possible that species split may have simply resulted from geographic isolation and that observed changes in genitalic shape between species have accumulated after speciation. However, the differences in genitalic shape are not a simple function of the phylogenetic distance among species (Figs 3, S15). Furthermore, although not statistically significant, comparison with expectations under a Brownian motion model indicate a pattern of evolutionary lability in genitalic shapes across species of the *M. scudderi* complex (note that the lack of significance does not mean genitalia are evolving neutrally; see Harmon, 2019). Together, the analyses suggest that the genitalia may have played a significant role in promoting speciation in the study system. Irrespective of whether differences in male genitalic shape are a proximate cause of speciation in the *M. scudderi* species complex, the genitalic shape is still a good diagnostic character.

Are we oversplitting geographic/morphological taxa into 'species'?

The boundary between geographic/ecological forms and species has been blurred in taxonomic and evolutionary studies, with recent advances in obtaining and analysing genetic data stirring more controversy (Freudenstein *et al.*, 2017; Sukumaran & Knowles, 2017). Genomic data have become a commonly used data type in systematics and evolutionary biology, where new models and associated computer programs have been rapidly developed. However, recent simulation-based as well as empirical studies show the limitations of relying upon genetic data alone for delimiting species (Sukumaran & Knowles, 2017; Huang, 2018). For example, the commonly used multispecies coalescent model in species delimitation confounds genetic structure within species with genetic divergence associated with species boundaries. By showing that phenotypic and genetic boundaries delimit the species lineages, our study avoids the misleading inferences that can occur when relying only on genetic data.

Lastly, with regard to concerns of potential oversplitting of taxa based on differences in the genitalia, our work validates the identity of the species diagnosed by genitalic differences (see also Knowles, 2001b for another example in *Melanoplus*). The only case in which there was some suggestion that species may be oversplit, and what might be considered as sibling species (Mayr, 1942), involves the divergence between *M. scudderi* and *M. mississippi* (see Fig. 1). *Melanoplus scudderi* and *M. mississippi* are morphologically similar, but their geographic distributions largely overlap (Fig. 1). Although we have only genotyped one population for each taxon and we do not know if these two putative sibling species are cohesive across their respective geographic distributions, our results suggest a possible example of cryptic divergence, especially given that our molecular data show statistically significant genetic divergence between the two species, which form two reciprocally monophyletic genetic groups of individuals (Fig. S5). Consequently, we conclude that species status should be conferred to all the genetically differentiated lineages, which also correspond to unique male genitalic shapes in the *M. scudderi* species complex.

Author contributions

LLK conceived the study. JH collected the samples and took the image data. JO prepared the ddRADseq libraries. J-PH processed and analysed the data. All authors wrote and read the manuscript.

Supporting Information

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

Table S1. Collection and voucher information of samples from putative species in the *Melanoplus scudderi* complex.

Table S2. Summary of pyRAD filtering results.

Figure S1. The distributions of site variation (A) and the position of the variable site of the retained loci after processing (B).

Figure S2. Summaries of the genetic data of retained sequences after filtering the ddRADseq data.

Figure S3. The distribution of estimated theta values from the 4132 retained variable loci used in phylogenetic and species delimitation analyses.

Figure S4. The distributions of estimated theta values and maximum pairwise genetic differences from each of the three groups of lineages (see Fig. 1) for each of the separate species delimitation analyses.

Figure S5. A maximum likelihood phylogenetic estimate of the 96 sampled individuals based on a concatenated dataset of SNPs.

Figure S6. The best species tree topology from BPP v.4 analyses.

Figure S7. A pairwise estimation of shared loci between samples from a 4132 loci dataset.

Figure S8. The estimated species tree and divergence times using a total of 4132 variable loci.

Figure S9. The estimated species tree and divergence times using four different sets of 1000 variable loci.

Figure S10. Outlines of the lateral view of the male genitalia shapes.

Figure S11. Outlines of the ventral view of the male genitalia shapes.

Figure S12. Results from elliptical Fourier analyses using different number of harmonics and the shape data for the lateral (A) and ventral (B) views of the male genitalia.

Figure S13. An overlay of all the outlines of the lateral (A) and ventral (B) male genitalia shape variation across individuals.

Figure S14. The first five PCs from shape analyses based on the lateral (A) and ventral (B) views of the male genitalia shape.

Figure S15. Phenograms based on individual PCs and the reconstructed species tree.

Acknowledgements

We thank Qixin He, Rob Massatti, and Marcelo J. Sturaro for their advice during the preparation of the ddRADseq libraries. Two reviewers provided comments and suggestions that improved the clarity of this manuscript. This research was funded by the National Science Foundation (NSF DEB 16-55607, to LLK), a grant from the Orthopterists' Society (to JGH), the National Institute of Food and Agriculture, and the Hubbell Orthopteran Research Endowment, Museum of Zoology, University of Michigan. JO's research visit to the University of Michigan was funded by a fellowship from the University of Castilla-La Mancha. The authors declare there are no conflicts of interest.

References

Andrews, K.R., Good, J.M., Miller, M.R., Luikart, G. & Hohenlohe, P.A. (2016) Harnessing the power of RADseq for ecological and evolutionary genomics. *Nature Reviews Genetics*, **17**, 81–92.

Bentley, S.J. Sr, Blum, M.D., Maloney, J., Pond, L. & Paulsell, R. (2016) The Mississippi River source-to-sink system: perspectives on tectonic, climatic, and anthropogenic influences, Miocene to Anthropocene. *Earth-Science Reviews*, **153**, 139–174.

Bickford, D., Lohman, D.J., Sodhi, N.S. *et al.* (2007) Cryptic species as a window on diversity and conservation. *Trends in Ecology and Evolution*, **22**, 148–155.

Bonhemme, V., Picq, S., Gaucherel, C. & Claude, J. (2014) Momocs: outline analysis using R. *Journal of Statistical Software*, **56**, 1–24.

Cadena, C.D., Zapata, F. & Jiménez, I. (2018) Issues and perspectives in species delimitation using phenotypic data: Atlantean evolution in Darwin's finches. *Systematic Biology*, **67**, 181–194.

Carstens, B.C. & Knowles, L.L. (2007) Shifting distributions and speciation: species divergence during rapid climate change. *Molecular Ecology*, **16**, 619–627.

Chifman, J. & Kubatko, L. (2014) Quartet inference from SNP data under the coalescent model. *Bioinformatics*, **30**, 3317–3324.

Cohen, B.L. (2018) Match and mismatch of morphological and molecular phylogenies: causes, implications, and new light on cladistics. *Zoological Journal of the Linnean Society*, **184**, 516–527.

Cohn, T.J., Swanson, D.R. & Fontana, P. (2013) *Dichopetala and New Related North American Genera: A Study in Genitalic Similarity in Sympatry and Genitalic Differences in Allopatry (Tettigoniidae: Phaneropterinae: Odonturini). Museum of Zoology*. University of Michigan, Miscellaneous Publications. NO. 203, Ann Arbor, Michigan.

Dickson, B.V. & Pierce, S.E. (2019) Functional performance of turtle humerus shape across an ecological adaptive landscape. *Evolution*, **73**, 1265–1277.

Eaton, D.A. (2014) PyRAD: assembly of *de novo* RADseq loci for phylogenetic analyses. *Bioinformatics*, **30**, 1844–1849.

Eaton, D.A. & Ree, R.H. (2013) Inferring phylogeny and introgression using RADseq data: an example from flowering plants (Pedicularis: Orobanchaceae). *Systematic Biology*, **62**, 689–706.

Eberhard, W.G. (1996) *Female Control: Sexual Selection by Cryptic Female Choice*. Princeton University Press, Princeton, NJ, USA.

Farrington, H.L., Lawson, L.P., Clark, C.M. & Petren, K. (2014) The evolutionary history of Darwin's finches: speciation, gene flow, and introgression in a fragmented landscape. *Evolution*, **68**, 2932–2944.

Fiduccia, C.M. & Mattheyses, R.M. (1982) A linear time heuristic for improving network partitions. *Proceedings of ACM/IEEE Design Conference*, pp. 175–181. https://doi.org/10.1109/DAC.1982.1585498

Flouri, T., Jiao, X., Rannala, B. & Yang, Z. (2018) Species tree inference with bpp using genomic sequences and the multispecies coalescent. *Molecular Biology and Evolution*, **35**, 2585–2593.

Freudenstein, J.V., Broe, M.B., Folk, R.A. & Sinn, B.T. (2017) Biodiversity and species concept – lineages are not rnough. *Systematic Biology*, **66**, 644–656.

Fujisawa, T., Sasabe, M., Nagata, N., Takami, Y. & Sota, T. (2019) Genetic basis of species-specific genitalia reveals role in species diversification. *Science Advance*, **5** eaav9939.

Goolsby, E.W., Bruggeman, J. & Ané, C. (2017) Rphylopars: fast multivariate phylogenetic comparative methods for missing data and within-species variation. *Methods in Ecology and Evolution*, **8**, 22–27.

Harmon L. (2019) Phylogenetic Comparative Methods: Learning From Tree. https://doi.org/10.32942/osf.io/e3xnr

Hickerson, M.J., Meyer, C.P. & Moritz, C. (2006) DNA barcoding will often fail to discover new animal species over broad parameter space. *Systematic Biology*, **55**, 729–739.

Hill, J.G. (2015) Revision and biogeography of the *Melanoplus Scudderi* species group with the description of 21 new species and establishment of the Carnegiei and Davisi species groups. *Transactions of the American Entomological Society*, **141**, 252–350.

Hillis, D.M. (1987) Molecular versus morphological approaches to systematics. *Annual Review in Ecology and Systematics*, **18**, 23–42.

Huang, J.-P. (2018) What have been and what can be delimited as species using molecular data under the multi-species coalescent model? A case study using Hercules beetles. *Insect Systematics and Diversity*, **2**, 3.

Huang, J.-P. (2016) Parapatric genetic introgression and phenotypic assimilation: testing conditions for introgression between Hercules beetles (Dynastes, Dynastidae). *Molecular Ecology*, **25**, 5513–5526.

Huang, J.-P. & Knowles, L.L. (2016) The species versus subspecies conundrum: quantitative delimitation from integrating multiple data types within a single Bayesian approachin Hercules beetles. *Systematic Biology*, **65**, 685–699.

Huang, H. & Knowles, L.L. (2016b) Unforeseen consequences of excluding missing data from next-generation sequences: simulation study of RAD sequences. *Systematic Biology*, **65**, 357–365.

Keightley, P.D., Pinharanda, A., Ness, R.W. *et al.* (2015) Estimation of the spontaneous mutation rate in *Heliconius melpomene*. *Molecular Biology and Evolution*, **32**, 239–243.

Klingenberg, C.P. & Gidaszewski, N.A. (2010) Testing and quantifying phylogenetic signals and homoplasy in morphometric data. *Systematic Biology*, **59**, 245–261.

Knowles, L.L. (2001) Did the Pleistocene glaciations promote divergence? Tests of explicit refugial models in montane grasshoppers. *Molecular Ecology*, **10**, 691–701.

Knowles, L.L. (2001b) Genealogical portraits of speciation in montane grasshoppers (genus *Melanoplus*) from the sky islands of the Rocky Mountains. *Proceedings of the Royal Society London Series B.*, **268**, 319–324.

Knowles, L.L. & Cohn, T.J. (2016) Tests of the role of sexual selection in genitalic divergence with multiple hybrid clines. *Journal of Orthoptera Research*, **25**, 75–82.

Knowles, L.L. & Kubatko, L.S., co-editors (2010) *Estimating Species Trees: Practical and Theoretical Aspects*. Wiley-Blackwell.

Knowles, L.L. & Otte, D. (2000) Phylogenetic analysis of montane grasshoppers from western North America (genus *Melanoplus*, Acrididae: Melanoplinae). *Annals of the Entomological Society of America*, **93**, 421–431.

Knowles, L.L. & Richards, C. (2005) Importance of genetic drift during Pleistocene divergence as revealed by analysis of genomic variation. *Molecular Ecology*, **14**, 4023–4032.

Knowles, L.L., Chappell, T.M., Marquez, E.J. & Cohn, T.J. (2016) Tests of the role of sexual selection in genitalic divergence with multiple hybrid clines. *Journal of Orthoptera Research*, **25**, 75–82.

Kuhl, F.P. & Giardina, C.R. (1982) Elliptic Fourier features of a closed contour. *Computer Graphics and Image Processing*, **18**, 236–258.

Losos, J.B. & De Queiroz, K. (1997) Evolutionary consequences of ecological release in Caribbean *Anolis* lizards. *Biological Journal of the Linnean Society*, **61**, 459–483.

Márquez, E.J. & Knowles, L.L. (2007) Correlated evolution of multivariate traits: detecting co-divergence across multiple dimensions. *Journal of Evolutionary Biology*, **20**, 2334–2348.

Sites, J.W. Jr & Marshall, J.C. (2004) Operational criteria for delimiting species. *Annual Review of Ecology, Evolution, and Systematics*, **35**, 199–227.

Masly, J.P. (2012) 170 years of "lock-and-key": genital morphology and reproductive isolation. *International Journal of Evolutionary Biology*, **2012**, 24736.

Mayr, E. (1942) *Systematics and the Origin of Species from the Viewpoint of a Zoologist*. Columbia University Press, New York.

McCormack, J.E., Hird, S.M., Zellmer, A.J., Carstens, B.C. & Brumfield, R.T. (2013) Applications of next-generation sequencing to phylogeography and phylogenetics. *Molecular Phylogenetics and Evolution*, **66**, 526–538.

Oneal, E. & Knowles, L.L. (2013) Ecological selection as the cause and sexual differentiation as the consequence of species divergence? *Proceedings of the Royal Society B: Biological Sciences*, **280**, 20122236.

Oneal, E. & Knowles, L.L. (2015) Paternity analyses in wild-caught and laboratory-reared Caribbean cricket females reveal the influence of mating environment on post-copulatory sexual selection. *Journal of Evolutionary Biology*, **28**, 2300–2307.

Paradis, E. & Schliep, K. (2019) ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics*, **35**, 526–528.

Peterson, B.K., Weber, J.N., Kay, E.H., Fisher, H.S. & Hoekstra, H.E. (2012) Double digest RADseq: an inexpensive method for *de novo* SNP discovery and genotyping in model and non-model species. *PLoS One*, **7**, e37135.

Revell, L.J. (2012) phytools: an R package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution*, **3**, 217–223.

Russel, D.A., Rich, F.J., Schneider, V. & Lynch-Stieglitz, J. (2009) A warm thermal enclave in thelate Pleistocene of the south-eastern United States. *Biological Reviews*, **84**, 173–202.

Rundell, R. & Price, T.D. (2009) Adaptive radiation, nonadaptive radiation, ecological speciation and nonecological speciation. *Trends in Ecology and Evolution*, **24**, 394–399.

Satler, J.D. & Carstens, B.C. (2016) Phylogeographic concordance factors quantify phylogeographic congruence among co-distributed species in the *Sarracenia alata* pitcher plant system. *Evolution*, **70**, 1105–1119.

Sherpa, S., Rioux, D., Goindin, D., Fouque, F., François, O. & Després, L. (2018) At the origin of a worldwide invasion: unraveling the genetic makeup of the Caribbean bridgehead populations of the dengue vector *Aedes aegypti*. *Genome Biology and Evolution*, **10**, 56–71.

Silva, G.S.C., Roxo, F.F., Lujan, N.K., Tagliacollo, V.A., Zawadzki, C.H. & Oliveira, C. (2016) Transcontinental dispersal, ecological opportunity and origins of an adaptive radiation in the Neotropical catfish genus *Hypostomus* (Siluriformes: Loricariidae). *Molecular Ecology*, **25**, 1511–1529.

Smith, B.T., Ribas, C.C., Whitney, B.M., Hernández-Baños, B.E. & Klicka, J. (2013) Identifying biases at different spatial and temporal scales of diversification: a case study in the Neotropical parrotlet genus *Forpus*. *Molecular Ecology*, **22**, 483–494.

Solis-Lemus, C., Knowles, L.L. & Ané, C. (2014) Bayesian species delimitation combining multiple genes and traits in a unified framework. *Evolution*, **69**, 492–507.

Stamatakis, A. (2014) RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, **30**, 1312–1313.

Sukumaran, J. & Knowles, L.L. (2017) Multispecies coalescent delimits structure, not species. *Proceedings of the National Academy of Sciences, USA*, **114**, 1607–1612.

Swofford D. L. (2002) PAUP: Phylogenetic Analysis Using Parsimony (and Other Methods), Version 4.0 Beta 10. Sinauer Association, Sunderland, MA.

Wagner, C.E., Keller, I., Wittwer, S. *et al.* (2013) Genome-wide RAD sequence data provide unprecedented resolution of species boundaries and relationships in the Lake Victoria cichlid adaptive radiation. *Molecular Ecology*, **22**, 787–798.

Walker, M.J., Stockman, A.K., Marek, P.E. & Bond, J.E. (2009) Pleistocene glacial refugia across the Appalachian Mountains and coastal plain in the millipede genus *Narceus*: evidence from population genetic, phylogeography, and paleoclimatic data. *BMC Evolutionary Biology*, **9**, 25.

Yang, Z. & Rannala, B. (2014) Unguided species delimitation using DNA sequence data from multiple loci. *Molecular Biology and Evolution*, **31**, 3125–3135.

Accepted 4 December 2019
First published online 6 January 2020