Title: Historical isolation facilitates species radiation by sexual selection: insights from *Chorthippus* grasshoppers

Running title: species radiation by sexual selection

Target journal: Molecular Ecology: 5.855

Authors: Zachary J. Nolen[1], Burcin Yildirim[1], Iker Irisarri[2,3], Shanlin Liu[4,5], Clara Groot Crego[1], D. Amby[4], Frieder Mayer[6], M. Thomas P. Gilbert[4], Ricardo J. Pereira[1,4]*

1 Division of Evolutionary Biology, Faculty of Biology II, Ludwig- Maximilians-Universität München, Grosshaderner Strasse 2, 82152 Planegg-Martinsried, Germany.

2 Department of Biodiversity and Evolutionary Biology, Museo Nacional de Ciencias Naturales (MNCN-CSIC), c/ José Gutiérrez Abascal 2, 28006 Madrid, Spain

3 Department of Organismal Biology (Systematic Biology), Uppsala University, Norbyv. 18D, 75236 Uppsala, Sweden

4 Natural History Museum of Denmark, University of Copenhagen, Copenhagen, Denmark

5 College of Plant Protection, China Agricultural University, Beijing 100193, China

6 Museum für Naturkunde – Leibniz Institute for Evolution and Biodiversity Science, Berlin, Germany

* corresponding author: ricardojn.pereira@gmail.com

Abstract:

Theoretical and empirical studies have shown that species radiations are facilitated when a trait under divergent natural selection is also involved in sexual selection. It is yet unclear how quick and effective radiations are where sexual selection is unrelated to the ecological environment. We address this question using grasshopper species of the genus *Chorthippus*, which have evolved strong assortative mating while lacking noticeable eco-morphological

28 divergence. Mitochondrial genomes suggest that the radiation is relatively recent, dating to
29 the mid-Pleistocene, which leads to extensive incomplete lineage sorting throughout the
30 mitochondrial and the nuclear genomes. Nuclear data show extremely low genomic
31 differentiation among species, yet hybrids are absent in sympatric localities. Demographic
32 analyses shed some light into these seemingly contradictory patterns. The estimated
33 demographic model shows a long period of geographic isolation, followed by secondary
34 contact and extensive introgression. This suggests that an initial period of geographic
35 isolation might favor the coupling of male signaling and female preference, which currently
36 maintains species boundaries in the face of long-term gene flow. More generally, these results
37 suggest that sexual selection can lead to radiations without a primary role of divergent natural
38 selection, resulting in cryptic species that are genetically, morphologically and ecologically
39 similar, but otherwise behave mostly as good biological species.

40

41 **Keywords:** speciation, hybridization, gene flow, sexual selection, diversification

42

Introduction:

Speciation begins when populations accumulate genetic differences that result in reproductive barriers (Mayr, 1963), such as incompatible mating signals that prevent interbreeding, incompatible ecological interactions that render hybrids unfit in either extrinsic parental habitat, or incompatible gene interactions that cause intrinsic hybrid dysfunction (Presgraves, 2010). Although multiple barriers are expected to become coupled during the speciation process (Roger K. Butlin & Smadja, 2018), behavioral barriers usually evolve early during divergence, leading taxonomists to rank their importance "far ahead of all others" (Mayr, 1963). Such barriers are often maintained in the face of homogenizing gene flow, so that genomic patterns can be attributed more confidently to evolutionary processes causing lineage splitting, rather than to processes that happen after speciation (Harrison, 1993). Although sexual selection is generally accepted as an important driving force of such early behavioral incompatibilities, it remains debated whether results from sexual selection acting in concert with divergent natural selection (Servedio & Boughman, 2017), or if they can result from sexual selection alone (Kondrashov & Shpak, 1998; Takimoto, Higashi, & Yamamura, 2000; van Doorn, Dieckmann, & Weissing, 2004).

Emblematic examples of species radiation driven by divergent natural selection, termed "ecological radiations" (Nosil, 2012; Schluter, 2000), have been a major focus of evolutionary studies because they shed light onto the selective forces and the demographic scenarios that result in stable behavioral incompatibilities in the face of gene flow. For example, in radiations of sympatric species of Darwin finches (Grant & Grant, 2011), cichlid fishes (Kocher, 2004), and butterflies (Jiggins, 2017), consistent phenotype-environment correlations suggest the repeated action of natural selection driven by different ecological niches (seed sizes, lake depth, or anti-predator strategies) shaping trait values (beak shape, opsins, and color pattern). Because such ecologically relevant traits also function as mating cues, behavioral isolation emerges as an automatic byproduct of natural selection (Servedio, Doorn, Kopp, Frame, & Nosil, 2011). Population genomic studies of ecological radiations (e.g. Brawand et al., 2014; Heliconius Genome Consortium, 2012; Lamichhaney et al., 2015) showed that genes underlying such "magic traits" show high genetic differentiation between incipient species relative to the genomic background. This is consistent with the coupled action of natural and sexual selection in reproductive isolation. In contrast, patterns of genomic variation across most of the genome reflect the demographic history of these radiations. Phylogenetic analyses show that ecological radiations are usually recent (between

15 ka to 12 Ma in cichlids, Brawand et al., 2014; Irisarri et al., 2018) and thus ancestral variants segregate between species. In addition, significant D-statistics (Durand, Patterson, Reich, & Slatkin, 2011) are consistent with ancestral introgression (Heliconius Genome Consortium, 2012; Marques, Lucek, Sousa, Excoffier, & Seehausen, 2019; Meier et al., 2017). In some species, extensive interspecific gene flow in sympatry has eroded the phylogenetic signal of common ancestry of each nominal species to generate genetic clusters that reflect geographic proximity, rather than phenotypic similarity (Magalhaes, Lundsgaard-Hansen, Mwaiko, & Seehausen, 2012; Nosil, Crespi, & Sandoval, 2002; Ravinet et al., 2016; Renaut, Owens, & Rieseberg, 2014). Demographic modeling shows that gene flow usually proceeds a period of genetic isolation, suggesting that the fixation of adaptive alleles was facilitated by historical geographic isolation (Meier, Sousa, et al., 2017). Together, all these observations support the theoretical expectation that the concerted action of natural with sexual selection, maintains stable behavioral incompatibilities in localized regions of the genome, while neutral homogenizing gene flow occurs genome wide.

Nevertheless, species radiations can also occur without the clear evidence of divergent natural selection. In these so-called "non-ecological" radiations (Mayer, Berger, Gottsberger, & Schulze, 2010; Rundell & Price, 2009), sympatric species do not occupy different ecological niches (Gillespie, Howarth, & Roderick, 2001), and did not diverge in ecologically relevant traits (Schluter, 2000). The operational difficulty of classifying a radiation as non-ecological comes from the impossibility of measuring all ecological attributes of species, which means that ecological differences can escape detection when they are difficult to study, when their relative contribution for reproductive isolation is small, or restricted to certain life stages (Rundell & Price, 2009). Nevertheless, such radiations can still respond to other selective forces, for example when sexual selection drives the coevolution between mating cues and preference, and thus the emergence of behavioral incompatibilities (Ritchie, 2007; Snook, Chapman, Moore, Wedell, & Crudgington, 2009). This is particularly clear in radiations of birds, amphibians or insects, where males differ markedly in mating coloration, and/or mating song and females differ in mating preference, but these traits have no obvious ecological role in their current environment. Although these systems have deserved less attention from evolutionary biologists relative to well-studied ecological radiations, non-ecological radiations can lead to similar burst of species that maintain behavioral barriers, despite lacking geographic and ecological isolation. For example, the highest rate of speciation recorded in arthropods occurs in *Laupala* Hawaiian crickets (Mendelson & Shaw, 2005), where species remain reproductively isolated due to pulse rate of male courtship songs

associated to female preference for specific pulse rates (Otte, 1994), suggesting that sexual selection on this trait may be the driver of this radiation. Studies of selective introgression and quantitative trait locus mapping have shown that male signaling and female preference map to the same genomic region (Xu & Shaw, 2019), potentially facilitating speciation in the face of gene flow. Although these systems are common in nature (e.g. crows (Vijay et al., 2016), frogs (Yang, Servedio, & Richards-Zawacki, 2019), and field crickets (Blankers et al., 2018)), the demographic conditions associated to their radiation remain poorly understood, in particular how temporal changes on gene flow modulates speciation.

Chorthippus grasshoppers are ideally suited to answer this question and to provide general insights into whether sexual selection can maintain species boundaries without a primary role of environmental factors (Mayer et al., 2010). This genus comprises around 230 nominal species (Cigliaano, Braun, Eades, & Otte, 2020), including several species groups of morphologically similar species such the biguttulus group (Harz, 1975). These species strongly differ in male songs (Fig. 1) and in co-evolving female acoustic preferences, which are both genetically inherited. This acts as a strong pre-mating barrier between parental species (von Helversen & von Helversen, 1994), so that hybrids are rarely found in nature (Baur, 2006; Jacobs, 1963; Perdeck, 1958). By experimentally muting males and stimulating females with conspecific male songs, laboratory crosses have resulted in 2% of interspecific mating between *C. biguttulus* and *C. brunneus* (Gottsberger & Mayer, 2019). Although F1 males do not appear to be impacted by hybrid inviability or sterility (but see Finck and Ronacher 2017), they produce an intermediate song that is not appealing to either parental female nor to their F1 sisters, leaving them "behaviorally sterile" (Gottsberger & Mayer, 2019). F1 females are behaviorally and physiologically fertile (Gottsberger and Mayer 2019), and thus can potentially backcross to parental species. In addition to acoustic cues, chemical cues are also species-specific, and genetically inherited (Finck, Kuntze, & Ronacher, 2016), suggesting that behavioral incompatibilities might translate into multiple barriers to gene flow.

Although *Chorthippus* spp. offer an unusual opportunity to test how gene flow modulates behavioral isolation arising from sexual selection, evolutionary genetic studies have been rare. This is due to its large genome size (10 Gb (Gosalvez, López-Fernandez, & Esponda, 1980)) that is still challenging to assemble using modern sequencing methods. The draft genome of the grasshopper *Locusta migratoria* (6.5 Gb; (Wang et al., 2014)) has shown that expansion of genome size in grasshoppers is largely caused by an accumulation of

repetitive elements (e.g. transposons, retrotransposons), and by an expansion of intergenic and intronic regions, but coding regions having similar length to that observed in *Drosophila melanogaster*. Because 93.8% of the 17,307 predicted proteins in *Locusta* were detected using transcriptome data, and as they do not appear to be involved in major duplications (Wang et al., 2014), transcriptomes offer a reliable reference-free approach for population genomics, as already established for organisms with simpler genomes (Gayral et al., 2013).

In this study we used four species of the biguttulus group (*C. mollis*, *C. biguttulus*, *C. brunneus*, and *C. rubratibialis*) and its closely related genus *Pseudochorthippus parallelus* (previously named *Chorthippus parallelus*) to understand the demographic history of species radiation via sexual selection. By assembling the full mitochondria and of 12,735 independent nuclear genes from transcriptome data we test whether this radiation is recent, and whether behavioral phenotypes form monophyletic groups (i.e. are phylogenetically independent). Using SNPs, we test whether allopatric and sympatric species experienced hybridization and genetic introgression. Finally, using demographic modeling, we test whether speciation has occurred in the face of continuous gene flow or with periods of historical isolation.

Materials and Methods:

1. Specimen and Genomic Sampling

We focused on four species from the biguttulus group (D. Ragge, Reynolds, & Willemse, 1990): *C. mollis* (herein, *Cmol*), *C. biguttulus* (*Cbig*), *C. rubratibialis* (*Crub*), and *C. brunneus* (*Cbru*) (see Table S1 for sampling details and accession numbers). RNAseq data from *Cmol* (n = 20 individuals), *Cbig* (n = 18), and *Cbru* (n = 20) were obtained from a previous study (Berdan, Mazzoni, Waurick, Roehr, & Mayer, 2015). These populations were sampled in localities in which each species pair is sympatric or allopatric, allowing to test whether gene flow follows current geographic overlap between species. Additionally, we collected *Crub* (n = 16), and two parapatric subspecies of *Pseudochorthippus parallelus* (*Ppar*, n = 10; until recently named *Chorthippus parallelus*) to serve as a phylogenetic outgroup of the radiation and to avoid potential mapping biases arising from using any of the ingroup species.

Tissue samples were preserved in RNAlater (Qiagen). To sample the largest variety of transcripts possible, we extracted total RNA from whole body, after removing the head and digestive track of each individual to avoid over-representation of eye pigments and gut contaminants. We homogenized the samples using ceramic beads 1.4/2.8 mm (Precellys) and the standard Tri-Reagent (Sigma) protocol. Resuspended RNA pellets were purified with RNAeasy Mini columns (Qiagen), and final sample integrity and quantity was assessed with an Agilent 2100 BioAnalyzer. Enrichment for mRNA, library construction and paired end Illumina HiSeq2500 sequencing was performed by the BGI at a high coverage for the reference *Ppar* samples and at a lower coverage for *Crub* and the remaining *Chorthippus* spp..

2. Transcriptome Assembly and Mapping

We constructed a complete reference transcriptome for each of the 10 *Ppar* samples separately. We used Trinity *de novo* assembler v2.2.0 (Grabherr et al., 2011), which used Trimmomatic with default parameters for trimming and filtering the raw reads, and using a kmer of 31. We retained scaffolds with kmer coverage > 2, supported by > 3 reads, and with > 400 bp in length, retaining the longest isoform per gene. We evaluated individual assemblies by computing the contig N50 based on transcript contigs representing the top 90% of expressed transcripts. We assessed the completeness of each individual assembly, first by

191 estimating the length of our assembled genes relative to the complete transcriptome of
192 *Locusta migratoria* (Wang et al., 2014) using BLAST, and second by estimating the
193 percentage of complete, fragmented and missing genes relative to the single copy genes
194 expected for metazoans using BUSCO (Waterhouse et al., 2018).

195 We then constructed a single transcriptome reference for *Ppar* representing the
196 maximum variability observed in this species. First, we assigned genes to the mitochondrial
197 genome through BLASTN searches of the assembled scaffolds against the mitogenome of
198 *Chorthippus chinensis* (accession number NC_011095.1; Liu & Huang, 2008) and keeping
199 the longest scaffolds. Second, we compared the 10 individual assemblies using the Markov
200 Cluster algorithm (MCL), using the pipeline implemented in OrthoMCL (L. Li, Stoeckert, &
201 Roos, 2003) to distinguish "orthologous genes" identified across all individuals from
202 "singleton genes" identified only in some individuals, keeping the longest sequence for each
203 gene. Third, for the "orthologous genes", we identified genes potentially involved in
204 duplications, based on the distribution of coverage of whole genome reads from *Ppar*
205 (accession number PRJNA280771) mapped onto our reference transcriptome (median
206 coverage = 1.5x). We labelled as "single-copy genes" all genes with coverage up to quantile
207 80% (< 5x), and the remaining as "multi-copy genes" (> 5x). Fourth, for the "single-copy
208 genes" we identified untranslated regions (UTRs) and open reading frames (ORFs) using
209 TransDecoder v3.0.1 (Haas & Papanicolaou, 2012). For simplifying downstream analyses, all
210 reference genes were separated by 500 Ns and organized in four artificial chromosomes: 1)
211 single-copy genes (with and without ORFs), 2) multi-copy genes, 3) singleton genes, and 4)
212 mitochondrial genes. We also produced auxiliary files containing the coordinates of each
213 protein coding gene, and the coordinates of synonymous positions (i.e. UTRs and third codon
214 positions).

215 We processed raw sequencing reads using the Paleomix BAM pipeline v1.2.9 (Schubert
216 et al., 2014) with default settings. This process first removed adapters and trimmed low
217 quality bases with AdapterRemoval (Lindgreen, 2012). Overlapping paired reads were
218 collapsed into a single consensus read. Trimmed and collapsed reads were then mapped to the
219 reference transcriptome using BWA-MEM, discarding unmapped reads and reads below a
220 mapping quality threshold of 30 (H. Li & Durbin, 2009). In the final step, local realignment
221 was performed around indels using GATK IndelRealigner (McKenna et al., 2010). We
222 recorded the number filtered nucleotides, read length and coverage to assess potential
223 mapping biases between *Chorthippus* spp.

Because the mitochondrial genome evolves faster that the nuclear genome, we also assembled one reference mitochondrial genome for one individual of each of the four *Chorthippus* species. We used 10 million pairs of filtered RNAseq reads in MIRA v4.0.2 (Chevreux et al., 2004) and MITObim v1.8 (Hahn, Bachmann, & Chevreux, 2013) with the default parameters and using the mitochondrial genome from the transcriptome of *Pseudochorthippus parallelus* as a reference. The initial assemblies were annotated with MITOS revision 656 (Bernt et al., 2013) to correct gene orientations. Indels in protein-coding genes were identified by aligning their nucleotide sequences against orthologous sequences from other Acrididae (available at NCBI; Table S2). Corrected mitochondrial genomes were then used as references to map the reads from all the individuals from the corresponding species with BWA-MEM as explained above. The corrections introduced to the references were confirmed by a higher proportion of aligned reads compared to the initial assemblies. The final mitochondrial genomes were annotated with MITOS (revision 942).

3. Phylogenetic Relationships

As phylogenetic inference requires probabilistic estimates of gene sequences, we called sequences for each of the 84 individuals separately for reads mapping to the mitochondrial and nuclear references using ANGSD v0.921 (Korneliussen, Albrechtsen, & Nielsen, 2014). We called the most frequent base at each position with a minimum coverage threshold of 10x.

3.1. Mitochondrial DNA

In order to infer the timing of diversification of *Chorthippus* species, we estimated a mitochondrial time tree for our 84 individuals plus all the available mitogenomes of Acrididae (Table S2). Single gene alignments were constructed using MUSCLE (Edgar, 2004) as implemented in SEAVIEW v.4.7 (Gouy, Guindon, & Gascuel, 2010). Protein-coding genes were translated with the invertebrate mitochondrial genetic code, aligned at the amino acid level, and back-translated to nucleotides. The 13 protein-coding genes and 2 ribosomal RNA genes were concatenated into a single dataset, and alignment columns with >80% missing data were removed with BMGE v.1.12 (Criscuolo & Gribaldo, 2010). We inferred a maximum likelihood tree using IQTREE v.1.6.5 (Nguyen, Schmidt, von Haeseler, & Minh, 2015), with BIC-selected best-fit substitution models and partitions (Table S3) as identified using ModelFinder (Kalyaanamoorthy, Minh, Wong, Haeseler, & Jermiin, 2017). Branch support was assessed with 1,000 bootstrap replicates. The inferred tree topology was then

used to estimate the timing of diversification of nodes representing the most recent common ancestors of 1) *Chorthippus* and *Pseudochorthippus*, 2) the subspecies *P. p. parallelus* and *erythropus*, and 3) all focal *Chorthippus* species. We used the uncorrelated relaxed clock model implemented in BEAST v.1.10.4 (Suchard et al., 2018), with a normally distributed prior on the substitution rate with mean of 0.0115, and standard deviation of 0.33 substitutions per million year (Brower, 1994), which we believe is appropriate given the remarkably conserved mitochondrial rates of insects (e.g., 0.0115 in butterflies and 0.0133 in beetles; Papadopoulou, Anastasiou, & Vogler, 2010). The tree was fixed to the previously inferred maximum likelihood topology, and we used a Yule speciation tree prior. To facilitate convergence, tree topology, clock models and substitution parameters (GTR+Γ) were linked across genes. Two independent MCMC chains were run for 100 million generations, sampling every 10,000. Convergence was checked *a posteriori* with Tracer v.1.7.1 (Rambaut, Drummond, Xie, Baele, & Suchard, 2018), and all values had ESS values >300. The final divergence times were obtained from one of the chains using TreeAnnotator, after excluding the initial 10 million as burn-in.

3.2. Nuclear DNA

We inferred nuclear species trees under the multi-species coalescent model to accommodate incomplete lineage sorting expected for fast radiations. Maximum likelihood trees were inferred for all single copy genes (UTRs were excluded due to the prevalence of assembly errors in these regions) that included at least three individuals per sampling locality, and called genotypes for > 50% of the sequences. We used BIC-selected models in IQTREE (Nguyen et al., 2015), 1,000 replicates of ultrafast bootstrapping (UFBoot), and SH-like approximate likelihood ratio tests (SH-aLRT) (Guindon et al., 2010; Hoang, Chernomor, von Haeseler, Minh, & Vinh, 2018). To account for the uncertainty in gene tree estimation, we used the consensus of bootstrapped tree for downstream analyses. In addition, we performed likelihood mapping in IQTREE to assess the information content of each gene tree (Strimmer & Haeseler, 1997). The consensus gene trees were fed into ASTRAL II (Mirarab & Warnow, 2015). This method estimates the species tree that maximizes the agreement among independent gene trees. We calculated the quartet support for each internal node, which reflects the fraction of quartets supporting the main, the first and second alternative topologies, and thus the amount of gene tree discordance at that node that is caused by incomplete lineage sorting or by gene flow. From this analysis, we output the main species tree topology, posterior probabilities for this main topology, branch lengths in coalescent units ($T/4N_e$), and quartet supports. We built a species tree first considering each individual as a

terminal branch, and second one by treating individuals of the same sampling locality as populations (as in Table S1).

We also used the nuclear DNA data to estimate relative divergence times between *Chorthippus* and *Pseudochorthippus*, between all focal *Chorthippus* species and between the two subspecies *of Pseudochorthippus*. For this, we calculated Nei's measure of absolute divergence (Dxy) that, contrary to the branch length of the species tree, is independent of effective population size. We calculated Dxy per gene from the individual gene alignments used to estimate nuclear gene trees using the R package PopGenome (Pfeifer, Wittelsbürger, Ramos-Onsins, & Lercher, 2014; R Core Team, 2017). We plotted the distribution of average per gene Dxy for each pairwise comparison and calculated the median value for each population comparison.

4. Population Genetic Analyses

Recent studies have shown that genotype likelihoods are more reliable than standard genotype calls to estimate demographic parameters (da Fonseca et al., 2016), especially when coverage is low or variable, such as expected for RNAseq data. We thus estimated genotype likelihoods using ANGSD (Korneliussen et al., 2014). We used both trimmed and collapsed reads, excluding reads with the following criteria: mapping quality lower than 15 after adjustment, a base read quality lower than 20, and that have multiple hits. Additionally, we limited the analyses to sites that were present in at least 80% of the individuals, with a depth greater than twice the total number of individuals, in single copy genes, and excluding the 1st and 2nd codons positions that are generally involved in non-synonymous substitutions. All the specific commands were deposited in a public repository (https://github.com/zjnolen/chorthippus_radiation). Together, this filtering approach restricts transcriptome data to bi-allelic sites that in principle are not directly affected by selection, conforming to the assumptions of population genetic methods (Gayral et al., 2013).

4.1. Population Structure

To examine the distribution of genetic variance among individuals, we performed a Principal Component Analysis (PCA) with the program ngsPopgen (Fumagalli, Vieira, Linderoth, & Nielsen, 2014). We excluded non-variable sites by limiting the analysis to sites with a SNP p-value < 1.0 x 10^{-2} (-SNP_pval). We have determined this value to be sufficient to remove the majority of non-variable sites, while not distorting the site frequency spectrum based on all sites (Fig. S1).

We determined the most likely number of population clusters (K) and estimated admixture proportion between these clusters, based on Hardy-Weinberg and Linkage equilibrium. We used NGSadmix (Skotte, Korneliussen, & Albrechtsen, 2013), also restricting to sites with a SNP p-value < 1.0×10^{-2}. We performed this analysis considering K from two to nine, with 50 replicates, selecting the replicate with the highest likelihood for each K. We expect each species or sampling locality to form a cluster, with genetic admixture only in sympatric localities if hybridization is ongoing or common.

4.2. Genetic Variability

To estimate the relative effective population size of each sampling locality and whether it changes over time, we have calculated Watterson's Theta (Watterson, 1975) and Tajima's D (Tajima, 1989), respectively. For each of these summary statistics, we first calculated the site frequency spectrum (SFS) for each population, and then calculated Theta and Tajima's D per site in ANGSD (Korneliussen et al., 2014), combined values across sites within each gene, and plotted their distributions for the genes sampled across all populations.

We assessed population differentiation using Fst (Reynolds, Weir, & Cockerham, 1983). We calculated a two-dimensional SFS for each population pair, calculated Fst per site in ANGSD, per gene Fst, and plotted their distributions for genes showing at least 10 SNPs across all population pairs.

4.3. Genetic Introgression

To test for the presence of introgression, and whether it is more common in sympatric relative to allopatric *Chorthippus* populations, we used Patterson's D-statistics (Durand et al., 2011). This test weights the number of bi-allelic sites that have a different topology from a species tree, with an outgroup and three ingroup taxa (i.e. ABBA and BABA sites). We used the Abbababa2 method implemented in ANGSD, which extends this analysis to comparisons among populations (Soraggi, Wiuf, & Albrechtsen, 2018). We used all individuals of *Ppar* as the outgroup, and all possible combinations of *Chorthippus* populations that conformed to our estimated species tree, restricting to sites with a SNP p-value of less than $1e^{-6}$. We inferred significance using a p-value calculated from 10 windows of ~150,000 relevant sites. A similar proportion of ABBA and BABA sites are expected under the null hypothesis of incomplete lineage sorting driving discordance, while significantly different proportions must be explained by gene flow. When different populations of the same species are considered, finding gene flow in sympatric but not in allopatric comparisons suggests that introgression

has occurred only recently, while finding gene flow across both comparisons suggests also ancestral introgression.

We additionally tested if gene flow occurred between each pair of populations, and whether it occurred early during speciation or only after an initial period of geographic isolation. We selected four demographic models reflecting alternative speciation scenarios: 1) divergence without gene flow (parameters: time since divergence, Ne-pop1, Ne-pop2), 2) divergence with continuous gene flow (extra parameters: symmetric gene flow), 3) divergence with ancestral gene flow (extra parameters: time until cessation of gene flow), and 4) divergence with gene flow in secondary contact (same as model 3, but with gene flow only starting after secondary contact). We constructed the two-dimensional SFS using ANGSD for each of the 16,969 single-copy genes, and summed them into a single complete SFS per population pair. We folded each complete SFS and fit it to all four demographic models using the diffusion approximation methods implemented in δaδi v2.0.5 (Gutenkunst, Hernandez, Williamson, & Bustamante, 2009). To ensure our model fittings reached the global optima, we used the dadi-pipeline developed by (Portik et al., 2017), which performs several rounds of replicated optimizations using the highest likelihood parameters of the previous rounds as the initial parameters in subsequent rounds. We performed three replicates of optimization to ensure model fittings converged on a single set of parameter values. This procedure was repeated for all pairs of populations.

In order to account for the physical linkage between SNPs within the same gene, we built 100 bootstraps by resampling the gene SFS with replacement and summing them. We used these bootstraps to 1) select the best demographic model using the adjusted likelihood ratio test, and 2) estimate parameter uncertainties using the Godambe Information Matrix (GIM; Coffman, Hsieh, Gravel, & Gutenkunst, 2016). If speciation in *Chorthippus* occurred in sympatry, we expect higher support for models with continuous gene flow or with reduction of gene flow over time. Conversely, if speciation was facilitated by an initial period of geographic isolation, we expect higher support for the models without gene flow or with gene flow after secondary contact. If gene flow depends on the current geographic overlap between species, we expect higher migration rates between sympatric relative to allopatric populations.

Results:

1. Transcriptome Assembly and Mapping

The individual assemblies had similar levels of quality across the 10 high-coverage *Ppar* individuals. Scaffold sizes were relatively large across individuals (Fig. S2), with scaffolds showing a N50 around 2,200 bp for genes explaining 90% of the expression. Around 4,300 genes contain the complete gene length found in *Locusta migratoria*, and ~11,000 are at least 10% complete (Fig. S3). Based on expectations for metazoans, between 78-82% of the BUSCO genes were complete, 6.6-9.7% were fragmented, and 10-12% were missing (Table S4). After merging and filtering individual assemblies, our reference transcriptome contained: 12,735 single-copy genes with identified ORF; 4,235 single-copy genes without ORFs; 4,263 multi-copy genes; 18,623 singleton genes; and the 13 mitochondrial genes plus 2 rRNAs. Within the single-copy genes with identified ORFs, there are 20,604,833 synonymous positions.

The number of filtered nucleotides and read length was similar across *Chorthippus* spp. and higher in *Pseudochorthippus* (Fig. S5), reflecting a shorter library insert size and larger sequencing effort for the reference transcriptome. Accordingly, coverage for the single copy nuclear genes averaged 70x for the outgroup, and 25x for the ingroup, species. The reference mitogenomes lacked any stop codons and showed an individual coverage above 8,000x due to their high expression.

2. Phylogenetic Relationships

2.1. Mitochondrial DNA

The mitogenomes produced a well-supported topology, due to a relatively large number of substitutions (Fig. 2). *Pseudochorthippus parallelus* is highly divergent from *Chorthippus* spp. (bootstrap proportion (BP) = 100%), with the genera *Euchorthippus* and *Gomphocerus* being more closely related to *Chorthippus* than to *Pseudochorthippus* (BP = 100%). Interestingly, *Chorthippus chinensis* nested between these two genera, suggesting that they do not belong to the same genus as the European *Chorthippus* (BP = 100%). The focal *Chorthippus* spp. formed a monophyletic group (BP = 100%), yet none of the four species nor the seven sampling localities formed monophyletic clades (Fig. S4). Interestingly, the mitogenome of *Gomphocerippus rufus* nested within our focal species of *Chorthippus*.

The time-calibrated trees show that the diversification of Acrididae is relatively old, dating to some 13.6 Ma (95% HPD 16.2-11.1), and that *Pseudochorthippus* and *Chorthippus* diverged around 5.9 Ma (7.1-4.7), before the Pleistocene. We estimate that *Pseudochorthippus* subspecies and *Chorthippus* species diversified around the same time, between 476 to 506 ka respectively (see Table S5 for HPD).

2.2. Nuclear DNA

As expected for transcriptome data, the completeness of nuclear genes was generally consistent across species, reflecting general expression values in grasshoppers (Fig. S6). After filtering, we recovered 3,963 genes that were used to estimate nuclear gene trees.

The individual consensus trees often showed polytomies, reflecting the uncertainty on gene tree estimation (Fig. S7). The likelihood mapping analysis showed that most gene trees have enough phylogenetic signal to 60 -100% of the quartets (Fig. S8). No gene tree contained more than 20% partially resolved quartets.

The individual coalescent tree (Fig. S9) showed that individuals sampled in each locality generally form well supported monophyletic clade (posterior probability, PP > 0.9). The exceptions are *Cbig* from Berlin, which includes the monophyletic clade from Erlangen, and *Cbru* from Berlin, which includes a monophyletic clade from Spain. All species and genera are reciprocally monophyletic (PP = 1).

In agreement, the population coalescent tree shows that populations of every species form monophyletic clades (PP = 1; Fig. 3). *Cmol* has the oldest split within the *Chorthippus* radiation (PP = 1). The relationship between *Crub* and *Cbru* and *Cbig* is less clear, but most likely *Crub* is a sister species of *Cbru* (PP = 0.67). Coalescent branch lengths are shorter between *Chorthippus* species compared to that between the *Pseudochorthippus* populations.

The quartet support for the common ancestor of *Chorthippus* and *Pseudochorthippus* shows a complete prevalence of the main topology, reflecting no gene tree discordance at those nodes. In contrast, for all nodes within the radiation of *Chorthippus* the fraction of the main topology is only slightly larger (34 - 41%) than the fractions of alternative topologies (29 - 33%). The ancestor of *Crub* and *Cbru* has the highest amount of gene tree discordance, consistent with the lower support for this relationship.

As expected, absolute divergence (Dxy, Fig. 4A) was highest in all comparisons between *Chorthippus* and *Pseudochorthippus* populations (medians = 0.00471 - 0.00476).

Yet, divergence between *Chorthippus* species (medians = 0.00155 - 0.00170) was similar to that between *Pseudochorthippus* subspecies (median = 0.00142).

3. Population Genetic Analyses

3.1. Population Structure

PC1 explains a large fraction of the genetic variance (18.6 %) and separates *Pseudochorthippus* from *Chorthippus* individuals. PC2, 3 and 4 explain relatively similar fractions of variance (2.91, 2.55 and 2.11 %, respectively), with PC2 separating the two *P. parallelus* subspecies. Only PC3 and PC4 separate *Chorthippus* species, reflecting a genetic continuum between most species, with the exception of *Cmol* (Fig. S10).

The population structure analysis showed a marked difference between *Pseudochorthippus* and *Chorthippus* at K = 2, without genetic admixture (Fig. 5A, Figure S11). At K = 5, all species were assigned to their own distinct clusters, with admixture only observed between the sympatric *Cbig* and *Cbru* in Berlin. At K = 9, each locality formed its own cluster, without noticeable genetic admixture between species; the only exception are two populations of *Cbig* sampled in close geographic proximity.

3.2. Genetic Variability

The number of genes used to calculate Theta and Tajima's D ranged between 9,675 and 12,521, with 9,143 genes sampled across all populations for both parameters. The per-gene Theta values (Fig. S12) were generally higher for *Chorthippus* populations (medians between 0.011 and 0.013), intermediate for *Pseudochorthippus parallelus parallelus* (median = 0.007), and lowest for *P. p. erythropus* (median = 0.005). Mean Tajima's D is generally negative in all populations (medians between -1.093 and -0.039), reflecting an excess of singletons, with the exception of *P. p. erythropus* (median = 0.026), which conforms to the neutral expectation (Fig. 5B).

The number of genes for the Fst analysis varied from 9,228 to 12,390, and 8,935 genes were common across population comparisons. Per gene Fst showed a high variance across all comparisons, but distributions were skewed to three modes (Fig. 4B). All comparisons between *Pseudochorthippus* and *Chorthippus* populations showed the highest differentiation (medians = 0.678 - 0.727), the subspecies of *Pseudochorthippus* showed intermediate

differentiation (median = 0.371), and all comparisons between *Chorthippus* populations showed virtually no genomic differentiation (medians = 0.057 - 0.105).

3.3. Genetic Introgression

The D-statistics showed significant deviations from the null expectation of no gene flow between most populations of *Chorthippus* (27 out of the 35 possible combinations; Fig. S13, Table S6). Gene flow involved all species pairs within *Chorthippus*, except between *Cbig* and *Crub*. Notably, we found evidence of gene flow between the same species pairs, irrespective of whether we used allopatric or sympatric populations (Fig. 6A), consistent with ancestral gene flow prior to these species establishing their current geographic range.

The highest likelihood model for all population pair comparisons was the secondary contact model (Fig. 6B, Table S7). Inferred divergence times for all population pairs were similar, ranging from 0.00235/μ – 0.00302/μ (where μ is the number of mutations per site per generation for the transcriptome, Fig. S13B, Table S8). Times of secondary contact were more variable, ranging from 0.00035/μ – 0.00107/μ (or 13 – 35% of the total divergence time) in *Chorthippus* population pairs and 0.00137/μ (or 49% of the total divergence time) in the *Pseudochorthippus* subspecies (Fig. S13B, Table S8). Secondary contact times were most recent in each species with a sympatric comparison in the northernmost sympatric pair (Berlin 0.00041/μ – 0.00057/μ or 13 – 21% of the total divergence time). Migration rates were high between *Chorthippus* species (203.48μ – 633.33μ) relative to that found between parapatric subspecies of *Pseudochorthippus* (97.57μ) (Figure S13A, Table S8). Northernmost sympatric pairs exhibited the highest migration rates for their species comparisons (Berlin 350.12 – 633.33μ).

Discussion

The rapid radiation of Chorthippus species

The studied species of *Chorthippus* have been assumed to be highly divergent relative to the two hybridizing subspecies of *Pseudochorthippus parallelus,* largely because most of these *Chorthippus* species persist in broad sympatry, and hybridization is rare (Baur, 2006; Jacobs, 1963; Perdeck, 1958). Yet, recent studies considering the barcoding COI mitochondrial gene (Hawlitschek et al., 2017) have shown that these species are undistinguishable, either showing a phylogenetic limitation of this single marker, and/or a recent radiation of this highly specious genus. By reconstructing full mitochondrial time trees and comparing them to thousands of independent nuclear trees, we show that this radiation is relatively recent, resulting in a large amount of incomplete lineage sorting across the genome.

As expected for highly transcribed mitochondrial genes, using transcriptomic reads we were able to assemble full mitochondrial genomes with high depth that showed enough mutations to infer a well-resolved topology (Fig. 2, major clade BP = 100%) and to estimate time of divergence between focal clades (Fig. 2, Table S5). Our analyses of the Acridae mitogenomes, show that the genus *Chorthippus* is not monophyletic, and that a large taxonomic revision needs to be continued by systematists. In agreement with previous studies based on the COI mitochondrial gene and a much smaller representation of the species (Defaut, 2012; Vedenina & Mugue, 2011), we show that *Pseudochorthippus parallelus*, previously classified as *Chorthippus parallelus*, is in fact a highly divergent lineage with respect to other *Chorthippus*. We estimate it to have diverged from *Chorthippus* some 5.9 Ma, during the transition between the Miocene and Pliocene. This was followed by the split of the genus *Euchorthippus*, the Asian *Chorthippus chinensis*, the genus *Gomphocerus*, and finally the diversification of the focal species of European *Chorthippus*, some 506 ka. Interestingly, *Gomphocerippus rufus* is nested inside the European *Chorthippus* clade, suggesting that this lineage belongs to the same radiation and should instead be classified as *Chorthippus*. However, we highlight that these species are morphologically cryptic, and thus our finding needs to be followed up by taxonomic studies integrating recorded songs, museum specimens, and possibly transcriptomic data. None of the focal *Chorthippus* species is monophyletic in respect to the mitochondria. Likewise, subspecies of *Pseudochorthippus parallelus* do not form reciprocally monophyletic clades, and have also diverged some 476 ka. Such coincident divergent times within *Chorthippus* and within *Pseudochorthippus* fall within the mid-

Pleistocene. It is important to note that these times are coalescent times for the mitochondria, and thus the timing of speciation is necessarily more recent because the most recent common ancestor occurs within the ancestral population of all *Chorthippus* species. These estimates imply that *Chorthippus* species have experienced multiple episodes of contraction to southern Pleistocene glacial refugia followed by post-glacial expansion to northern latitudes where they occur now.

Our analyses of the nuclear transcripts show that the pattern we observe in the mitochondrial genome is also observed across thousands of independent nuclear genes. Nuclear gene trees show *Pseudochorthippus* subspecies usually form two reciprocally monophyletic clades (BP 100%), and *Chorthippus* species are paraphyletic (Fig. S7). Moreover, absolute divergence across nuclear genes (Dxy) is similar between *Pseudochorthippus* subspecies (median = 0.00142) and *Chorthippus* species (medians = 0.00155 - 0.00170), again supporting the observation that those lineages diversified around a similar time. *Pseudochorthippus* and *Chorthippus* are about 3.3 times more divergent than taxa within these genera (medians = 0.00471 - 0.00476), confirming that these lineages should be considered different genera. Considering these recent times since divergence for the *Chorthippus* species, and their typically large invertebrate effective population sizes (Leffler et al., 2012), it is not surprising that we find a large amount of incomplete lineage sorting across independent nuclear gene trees (Fig. S7), as we also observe in the mitochondrial tree (Fig. 2).

By accounting for incomplete lineage sorting through the multispecies coalescent, we estimate a well-supported species tree where every species forms a monophyletic clade, irrespective of if we consider the 84 individuals separately (Fig S9; PP > 0.9) or grouped by the seven sampling localities (Fig. 3, PP = 1.0). Our species tree shows that *C. mollis* was the first species to split, also in agreement with the differentiation reflected on the PCA analyses (Fig. S10). Most likely, *C. biguttulus* split next, followed by the divergence between *C. rubratibialis* and *C. brunneus*. Yet, the phylogenetic relationship between *C. rubratibialis* and the other two species is not strongly supported (PP = 0.67), and it is unlikely to be resolved with more data than the 12,735 genes used here. Based on the short branch lengths estimated in our species tree (in coalescent units of T/Ne), the radiation of *Chorthippus* spp. has likely occurred very rapidly and with very large effective population sizes. This had led to a large fraction of gene discordance reflected by the quartet supports (pie charts in Fig. 3), where only 34 to 41% of the quartets share the main topology.

Together, our results suggest that the radiation of *Chorthippus* species fits several key criteria for adaptive radiations (Schluter, 2000). First, species form a monophyletic clade (Fig. 2), together with a second genus (*Gomphocerippus*) that likely belonged to the same radiation. Second, the radiation is relatively recent (less than 506 ka; Table S5), similar to other well-studied adaptive radiations where ecological divergence is the main driver of reproductive isolation, which range from 200 ka to 10 Ma in cichlids (Irisarri et al., 2018) and from 2 to 18 Ma in stick insects (Riesch et al., 2017). Third, radiation led to a burst of monophyletic species (Fig. 3), and thus behavioral isolation evolved once in each nominal species, and this phylogenetic signal has not been completely eroded by homogenizing gene flow. Given that, in contrast with ecological radiations, species of *Chorthippus* are maintained by behavioral isolation in song traits that are not clearly linked to divergent natural selection, our results suggest that this system can provide key insights on how sexual isolation can drive species formation in the absence, or reduced contribution, of other selective forces.

Leaky genomic barriers between Chorthippus species, despite behavioral incompatibilities

Genomic studies of rapid radiations driven by ecological divergence have shown that, even when species rarely hybridize in nature today, ancestral gene flow has been common during speciation. It remains unclear whether gene flow occurs during radiations driven by sexual selection, such as in *Chorthippus*, especially given that in this system hybridization is rarely seen in nature, and that behavioral experiments with lab-bred interspecific hybrids revealed strong pre- and post-zygotic barriers to gene flow (Gottsberger & Mayer, 2007, 2019). Our results show that although ongoing hybridization seems to be rare in nature, it has led to ancestral introgression between species.

Although our PCA analyses seems to reflect a genetic gradient between most *Chorthippus* species (PC3 and 4; Fig, S10), our STRUCTURE analyses revealed well defined genetic clusters at the genus, species, and population levels (Fig. 5A). When we allow the maximum number of clusters (K=9), we do not see any admixture between clusters, with one exception of individuals of the same species (*C. biguttulus*) sampled in close geographic proximity (Berlin and Erlangen). Nevertheless, considering 4 to 8 clusters, we observe that some *C. biguttulus* individuals from Berlin contain up to 18% of the genome of the sympatric *C. brunneus* (Fig. S11), suggesting that some asymmetric backcrossing probably exists in this locality. In general, our results indicate that ongoing hybridization between *Chorthippus* species is too rare to be detected by our sampling (up to 28 individuals of 3 species in

598 sympatry). Yet, it is important to note that these species closely resemble each other in
599 morphology (D. R. Ragge & Reynolds, 1988), and thus sampling biases against intermediate
600 hybrids may lead to an underestimation of real hybridization rates in nature.

601 Our findings of similar times of divergence (Table S5), but much stronger reproductive
602 isolation in *Chorthippus* spp. (Fig. 5A) relative to that observed between subspecies of
603 *Pseudochorthippus parallelus,* would normally suggest a stronger level of genomic
604 differentiation. However, our results instead show that *Chorthippus* species are genomically
605 similar (all medians of pairwise Fst are between 0.057 and 0.105), being at least 3.5 times less
606 differentiated than the *Pseudochorthippus parallelus* subspecies (Fst = 0.371) that form a
607 clinal hybrid zone when meeting upon secondary contact (G. M. Hewitt, 1993). These low
608 levels of genetic differentiation can either be explained by unusually large effective
609 population sizes in *Chorthippus* spp. that resulted in incomplete lineage sorting in descendant
610 species, and/or by interspecific gene flow. While our estimates of diversity within populations
611 (theta; Fig. S12) are consistent with lower effective population sizes in *Pseudochorthippus*
612 relative to *Chorthippus*, these estimates are largely overlapping, and probably insufficient to
613 explain the observed differences in differentiation. The D-statistics offer a direct test of
614 interspecific gene flow, because an overrepresentation of ABBA or BABA topologies cannot
615 be solely explained by larger effective population sizes in *Chorthippus* species leading to
616 incomplete lineage sorting. Accordingly, out of the 35 comparisons, 27 show significant
617 deviations from the neutral expectations, indicating that there has been ancestral gene flow
618 among almost all species pairs (Fig. S13). Because our sampling often includes the same
619 species pair in allopatry and in sympatry, we used this to explicitly test if gene flow is only
620 observed in sympatry, consistent with recent gene flow. Our results show significant
621 deviations from the neutral expectation regardless of the geographic overlap (Fig. 6A),
622 suggesting that interspecific gene flow has started before the species established their current
623 distributions.

624 In summary, our indirect estimates of hybridization in nature show that interspecific
625 hybridization is rare between *Chorthippus* species, despite the absence of noticeable
626 geographic and ecological isolation. Nevertheless, during the radiation of these species, rare
627 hybridization seems to have resulted in introgression, reducing genomic differentiation
628 genome wide. Such lack of genomic integrity (Fst ~ 0.08), despite strong behavioral
629 incompatibilities, has also been observed in other systems where sexual selection plays a
630 more prominent role than natural selection, such as in crows (Vijay et al., 2016; Fst ~ 0.13)
631 and in warblers (Toews et al., 2016; Fst ~ 0.005). In such systems, low background genomic

divergence has proven useful to localize peaks of genomic differentiation that resist to gene flow, and that underlie behavioral isolation. Although the same approach cannot yet be applied to *Chorthippus* spp. due to the lack of a reference genome, our results suggest that that the loci underlying the strong behavioral isolation in these grasshopper species must also be resistant to the pervasive gene flow experienced by most of the genome. Follow up studies using admixture mapping in hybrid zones or QTL experiments are promising ways to identify genes underlying behavioral isolation.

Historical geographic isolation occurred during the evolution of behavioral isolation

In well-studied adaptive radiations behavioral isolation is facilitated by divergent natural (Servedio et al., 2011) and thus speciation can occur in the face of continuous, but reduced, gene flow. This model is harder to explain in systems where sexual selection is not functionally associated to ecological divergence because recombination can prevent the evolution of association between male signals and female preference (Felsenstein, 1981; Fry, 2003; Servedio et al., 2011), questioning how sexual selection alone can lead to speciation (Arnegard & Kondrashov, 2004; Kondrashov & Shpak, 1998; Takimoto et al., 2000; van Doorn et al., 2004). Using demographic modeling, we clarify the demographic conditions that have led to the strong behavioral isolation that maintains species boundaries in *Chorthippus*.

For all species pairs, our analyses rejected simpler models of allopatric speciation without gene flow, and a sympatric speciation model with continuous gene flow, even after accounting for the extra parameters using a likelihood ratio test (Table S7). When fitting two more complex models with a second time (T2) at which gene flow either ends or starts, we always find stronger support for the latter model, suggesting that all population pairs fit a model of initial geographic isolation followed by secondary contact (Fig. 6B). More biologically realistic demographic models would include multiple cycles of geographic isolation and secondary contact. Yet, adding additional parameters, such as asymmetric gene flow, did not substantially improve model fit after the likelihood ratio test. Given our estimated divergence times from the mitochondrial time tree (less than 506 ka; Table S5), *Chorthippus* species must have experienced multiple episodes of contraction to different Pleistocene glacial refugia that may have caused geographic isolation, as is well described in *Pseudochorthippus parallelus* (G. Hewitt, 2000). Future studies sampling these species at or near the described refugia in Eurasia may help testing this hypothesis.

Because there is no mutation rate (μ) available for transcriptomes, the demographic parameters were estimated in reference to μ, as the number of mutations per site per generation. Nevertheless, these estimates reflect comparable, relative measures of isolation and gene flow, as μ is likely the same for such closely related taxa. Our estimates of time show that *Chorthippus* species have spent most of their time in geographic isolation, with gene flow occurring only from 13 – 35% of the total time since divergence. Such recent estimates of secondary contact are also consistent with our estimates of Tajima's D (Fig. 5B), which are negative for all populations sampled in previously glaciated areas. The notable exception is *Pseudochorthippus parallelus erythropus*, which was sampled close to its Pleistocenic refugium (G. M. Hewitt, 1993). Gene flow following secondary contact was likely very high between all *Chorthippus* species (from 203.48μ – 633.33μ), leading to the observed low levels of genome wide differentiation (Fig. 4B). Migration estimates are higher in population pairs that are sympatric in Berlin, particularly between *C. biguttulus* and *C. brunneus*, corroborating our findings of backcrossing in the locality (Fig. 5, S11).

Together, our demographic analyses bring a consensus to our previous phylogenetic and population analyses. These suggest that the strong behavioral barriers observed today among *Chorthippus* species were likely facilitated by the long periods of geographic isolation, probably associated with the glacial periods that affected Europe since the mid-Pleistocene. Such periods provide opportunity for selective processes such as Fisherian sexual selection to lead to the coupling of male song and female preference, as observed in other species that arose via sexual selection (Xu & Shaw, 2019). Once that association is established within a species, behavioral barriers can resist the homogenizing effects of gene flow, because recombination between signal and preference loci is suppressed either by physical linkage (e.g. as observed in Hawaiian crickets, Xu & Shaw, 2019) or by assortative mating (e.g. as observed in field crickets, Blankers et al., 2018). Such periods of geographic isolation are unlikely to be unique to these species complexes that are predominantly maintained by behavioral isolation. Recent demographic modeling studies of emblematic adaptive radiations also maintained by divergent natural selection, such as the radiation of cichlid fishes (Meier, Sousa, et al., 2017) or of Caribbean anoles (Thorpe, Surget-Groba, & Johansson, 2010), confirm that geographic isolation favors adaptive processes that can lead to speciation. Yet, in contrast with those well studied cases of ecological radiations, our study in *Chorthippus* suggests that sexual selection can lead to radiations in the absence of strong divergent natural selection, resulting in cryptic species that are genetically, morphologically and ecologically similar but that otherwise generally behave as good biological species.

Acknowledgements:

We thank the Spanish and French authorities from the Gobierno de Aragón and the French Parc National des Pyrénées for the permission to sample the individuals used for this study. RJP was funded by the European Union's Horizon 2020 research and innovation programme, under the Marie Sklodowska-Curie grant agreement no. 658706. II was supported by a Juan de la Cierva-Incorporación postdoctoral fellowship (IJCI-2016- 29566) from the Spanish Ministry for Science and Competitiveness (MINECO). Computations were performed on resources provided by the Swedish National Infrastructure for Computing (SNIC) at Uppsala Multidisciplinary Center for Advanced Computational Science (UPPMAX) under Project SNIC 2019/8-67. Photographs in Fig. 1 were courtesy of C. Roestri and P. Fontana.

References:

Arnegard, M. E., & Kondrashov, A. S. (2004). Sympatric Speciation by Sexual Selection Alone Is Unlikely. *Evolution*, *58*(2), 222–237. doi: 10.1111/j.0014-3820.2004.tb01640.x

Baur, B. (2006). *Die Heuschrecken der Schweiz*. Bern Naturhistorisches Museum.

Berdan, E. L., Mazzoni, C. J., Waurick, I., Roehr, J. T., & Mayer, F. (2015). A population genomic scan in *Chorthippus* grasshoppers unveils previously unknown phenotypic divergence. *Molecular Ecology*, *24*(15), 3918–3930. doi: 10.1111/mec.13276

Bernt, M., Donath, A., Jühling, F., Externbrink, F., Florentz, C., Fritzsch, G., … Stadler, P. F. (2013). MITOS: Improved de novo metazoan mitochondrial genome annotation. *Molecular Phylogenetics and Evolution*, *69*(2), 313–319. doi: 10.1016/j.ympev.2012.08.023

Blankers, T., Vilaça, S. T., Waurick, I., Gray, D. A., Hennig, R. M., Mazzoni, C. J., … Berdan, E. L. (2018). Demography and selection shape transcriptomic divergence in field crickets. *Evolution*, *72*(3), 553–567. doi: 10.1111/evo.13435

Boughman, J. W. (2001). Divergent sexual selection enhances reproductive isolation in sticklebacks. *Nature*, *411*(6840), 944–948. doi: 10.1038/35082064

Brawand, D., Wagner, C. E., Li, Y. I., Malinsky, M., Keller, I., Fan, S., … Palma, F. D. (2014). The genomic substrate for adaptive radiation in African cichlid fish. *Nature*, *513*(7518), 375–381. doi: 10.1038/nature13726

Brower, A. V. (1994). Rapid morphological radiation and convergence among races of the butterfly *Heliconius erato* inferred from patterns of mitochondrial DNA evolution. *Proceedings of the National Academy of Sciences*, *91*(14), 6491–6495. doi: 10.1073/pnas.91.14.6491

Butlin, Roger K., & Smadja, C. M. (2018). Coupling, Reinforcement, and Speciation. *The American Naturalist*, *191*(2), 155–172. doi: 10.1086/695136

Chevreux, B., Pfisterer, T., Drescher, B., Driesel, A. J., Müller, W. E. G., Wetter, T., & Suhai, S. (2004). Using the miraEST Assembler for Reliable and Automated mRNA Transcript Assembly and SNP Detection in Sequenced ESTs. *Genome Research*, *14*(6), 1147–1159. doi: 10.1101/gr.1917404

Cigliaano, M. M., Braun, H., Eades, D. C., & Otte, D. (2020). Orthoptera Species File. Version 5.0/5.0. Retrieved May 11, 2020, from http://Orthoptera.SpeciesFile.org

Coffman, A. J., Hsieh, P. H., Gravel, S., & Gutenkunst, R. N. (2016). Computationally Efficient Composite Likelihood Statistics for Demographic Inference. *Molecular Biology and Evolution*, *33*(2), 591–593. doi: 10.1093/molbev/msv255

Criscuolo, A., & Gribaldo, S. (2010). BMGE (Block Mapping and Gathering with Entropy): a new software for selection of phylogenetic informative regions from multiple sequence alignments. *BMC Evolutionary Biology*, *10*(1), 210. doi: 10.1186/1471-2148-10-210

da Fonseca, R. R., Albrechtsen, A., Themudo, G. E., Ramos-Madrigal, J., Sibbesen, J. A., Maretty, L., … Pereira, R. J. (2016). Next-generation biology: Sequencing and data analysis approaches for non-model organisms. *Marine Genomics*, *30*, 3–13. doi: 10.1016/j.margen.2016.04.012

Defaut, B. (2012). Implications taxonomiques et nomenclaturales de publications récentes en phylogénie moléculaire: 1. Les Gomphocerinae de France (Orthoptera, Acrididae). *Matériaux Orthoptériques et Entomocénotiques*, *17*, 15–20.

Durand, E. Y., Patterson, N., Reich, D., & Slatkin, M. (2011). Testing for Ancient Admixture between Closely Related Populations. *Molecular Biology and Evolution*, *28*(8), 2239–2252. doi: 10.1093/molbev/msr048

Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, *32*(5), 1792–1797. doi: 10.1093/nar/gkh340

Felsenstein, J. (1981). Skepticism Towards Santa Rosalia, or Why Are There so Few Kinds of Animals? *Evolution*, *35*(1), 124–138. doi: 10.1111/j.1558-5646.1981.tb04864.x

Finck, J., Kuntze, J., & Ronacher, B. (2016). Chemical cues from females trigger male courtship behaviour in grasshoppers. *Journal of Comparative Physiology A*, *202*(5), 337–345. doi: 10.1007/s00359-016-1081-4

Fry, J. D. (2003). Multilocus Models of Sympatric Speciation: Bush Versus Rice Versus Felsenstein. *Evolution*, *57*(8), 1735–1746. doi: 10.1111/j.0014-3820.2003.tb00582.x

Fumagalli, M., Vieira, F. G., Linderoth, T., & Nielsen, R. (2014). ngsTools: methods for population genetics analyses from next-generation sequencing data. *Bioinformatics*, *30*(10), 1486–1487. doi: 10.1093/bioinformatics/btu041

Gayral, P., Melo-Ferreira, J., Glémin, S., Bierne, N., Carneiro, M., Nabholz, B., … Galtier, N. (2013). Reference-Free Population Genomics from Next-Generation Transcriptome Data and the Vertebrate–Invertebrate Gap. *PLoS Genetics*, *9*(4). doi: 10.1371/journal.pgen.1003457

Gillespie, R., Howarth, F., & Roderick, G. (2001). *Adaptive Radiation*. *Encyclopedia of Biodiversity*. Academic Press, New York.

Gosalvez, J., López-Fernandez, C., & Esponda, P. (1980). Variability of the DNA Content in Five Orthopteran Species. *Caryologia*, *33*(2), 275–281. doi: 10.1080/00087114.1980.10796840

Gottsberger, B., & Mayer, F. (2007). Behavioral sterility of hybrid males in acoustically communicating grasshoppers (Acrididae, Gomphocerinae). *Journal of Comparative Physiology A*, *193*(7), 703–714.

Gottsberger, B., & Mayer, F. (2019). Dominance effects strengthen premating hybridization barriers between sympatric species of grasshoppers (Acrididae, Orthoptera). *Journal of Evolutionary Biology*, *32*(9), 921–930. doi: 10.1111/jeb.13490

Gouy, M., Guindon, S., & Gascuel, O. (2010). SeaView Version 4: A Multiplatform Graphical User Interface for Sequence Alignment and Phylogenetic Tree Building. *Molecular Biology and Evolution*, *27*(2), 221–224. doi: 10.1093/molbev/msp259

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., … Regev, A. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology*, *29*(7), 644–652. doi: 10.1038/nbt.1883

Grant, P. R., & Grant, B. R. (2011). *How and Why Species Multiply: The Radiation of Darwin's Finches*. Princeton University Press.

Guindon, S., Dufayard, J.-F., Lefort, V., Anisimova, M., Hordijk, W., & Gascuel, O. (2010). New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0. *Systematic Biology*, *59*(3), 307–321. doi: 10.1093/sysbio/syq010

Gutenkunst, R. N., Hernandez, R. D., Williamson, S. H., & Bustamante, C. D. (2009). Inferring the Joint Demographic History of Multiple Populations from Multidimensional SNP Frequency Data. *PLOS Genetics*, *5*(10), e1000695. doi: 10.1371/journal.pgen.1000695

Haas, B., & Papanicolaou, A. (2012). TransDecoder. Retrieved from Github website: https://github. com/TransDecoder/TransDecoder

Hahn, C., Bachmann, L., & Chevreux, B. (2013). Reconstructing mitochondrial genomes directly from genomic next-generation sequencing reads—a baiting and iterative mapping approach. *Nucleic Acids Research*, *41*(13), e129–e129. doi: 10.1093/nar/gkt371

Harrison, R. G. (1993). Hybrids and hybrid zones: Historical perspective. *Hybrid Zones and the Evolutionary Process*, 3–12.

Harz, K. (1975). The Orthoptera of Europe. II. Series entomologica, Vol. 11. *Dr. W. Junk Publ. The Hague*.

Hawlitschek, O., Morinière, J., Lehmann, G. U. C., Lehmann, A. W., Kropf, M., Dunz, A., … Haszprunar, G. (2017). DNA barcoding of crickets, katydids and grasshoppers (Orthoptera) from Central Europe with focus on Austria, Germany and Switzerland. *Molecular Ecology Resources*, *17*(5), 1037–1053. doi: 10.1111/1755-0998.12638

Heliconius Genome Consortium. (2012). Butterfly genome reveals promiscuous exchange of mimicry adaptations among species. *Nature*, *487*(7405), 94–98. doi: 10.1038/nature11041

Hewitt, G. (2000). The genetic legacy of the Quaternary ice ages. *Nature*, *405*(6789), 907–913. doi: 10.1038/35016000

Hewitt, G. M. (1993). After the ice: parallelus meets erythropus in the Pyrenees. *Hybrid Zones and the Evolutionary Process*, 140–164.

Hoang, D. T., Chernomor, O., von Haeseler, A., Minh, B. Q., & Vinh, L. S. (2018). UFBoot2: Improving the Ultrafast Bootstrap Approximation. *Molecular Biology and Evolution*, *35*(2), 518–522. doi: 10.1093/molbev/msx281

Irisarri, I., Singh, P., Koblmüller, S., Torres-Dowdall, J., Henning, F., Franchini, P., … Meyer, A. (2018). Phylogenomics uncovers early hybridization and adaptive loci shaping the radiation of Lake Tanganyika cichlid fishes. *Nature Communications*, *9*(1), 1–12. doi: 10.1038/s41467-018-05479-9

Jacobs, W. (1963). Über das Singen der Feldheuschrecke Chorthippus biguttulus (L.) in verschiedenen Bereichen Mittel-und Westeuropas. *Zeitschrift Für Tierpsychologie*, *20*(4), 446–460.

Jiggins, C. D. (2017). *The Ecology and Evolution of Heliconius Butterflies*. Oxford University Press.

Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., Haeseler, A. von, & Jermiin, L. S. (2017). ModelFinder: Fast model selection for accurate phylogenetic estimates. *Nature Methods*, *14*(6), 587–589. doi: 10.1038/nmeth.4285

Kocher, T. D. (2004). Adaptive evolution and explosive speciation: The cichlid fish model. *Nature Reviews Genetics*, *5*(4), 288–298. doi: 10.1038/nrg1316

Kondrashov, A. S., & Shpak, M. (1998). On the origin of species by means of assortative mating. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *265*(1412), 2273–2278. doi: 10.1098/rspb.1998.0570

Korneliussen, T. S., Albrechtsen, A., & Nielsen, R. (2014). ANGSD: Analysis of Next Generation Sequencing Data. *BMC Bioinformatics*, *15*(1), 356. doi: 10.1186/s12859-014-0356-4

Lamichhaney, S., Berglund, J., Almén, M. S., Maqbool, K., Grabherr, M., Martinez-Barrio, A., … Andersson, L. (2015). Evolution of Darwin's finches and their beaks revealed by genome sequencing. *Nature*, *518*(7539), 371–375. doi: 10.1038/nature14181

Leffler, E. M., Bullaughey, K., Matute, D. R., Meyer, W. K., Ségurel, L., Venkat, A., … Przeworski, M. (2012). Revisiting an Old Riddle: What Determines Genetic Diversity Levels within Species? *PLoS Biology*, *10*(9). doi: 10.1371/journal.pbio.1001388

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics*, *25*(14), 1754–1760. doi: 10.1093/bioinformatics/btp324

Li, L., Stoeckert, C. J., & Roos, D. S. (2003). OrthoMCL: Identification of Ortholog Groups for Eukaryotic Genomes. *Genome Research*, *13*(9), 2178–2189. doi: 10.1101/gr.1224503

Lindgreen, S. (2012). AdapterRemoval: Easy cleaning of next-generation sequencing reads. *BMC Research Notes*, *5*(1), 337. doi: 10.1186/1756-0500-5-337

Liu, Y., & Huang, Y. (2008). Sequencing and analysis of complete mitochondrial genome of Chorthippus chinensis Tarb. *Chinese Journal of Biochemistry and Molecular Biology*, *4*.

Magalhaes, I., Lundsgaard-Hansen, B., Mwaiko, S., & Seehausen, O. (2012). Evolutionary divergence in replicate pairs of ecotypes of Lake Victoria cichlid fish. *Evolutionary Ecology Research*, *14*(4), 381–401.

Marques, D. A., Lucek, K., Sousa, V. C., Excoffier, L., & Seehausen, O. (2019). Admixture between old lineages facilitated contemporary ecological speciation in Lake Constance stickleback. *Nature Communications*, *10*(1), 1–14. doi: 10.1038/s41467-019-12182-w

Mayer, F., Berger, D., Gottsberger, B., & Schulze, W. (2010). Non-Ecological Radiations in Acoustically Communicating Grasshoppers? In M. Glaubrecht (Ed.), *Evolution in Action: Case studies in Adaptive Radiation, Speciation and the Origin of Biodiversity* (pp. 451–464). doi: 10.1007/978-3-642-12425-9_21

Mayr, E. (1963). Animal species and evolution. *Animal Species and Evolution*. Retrieved from https://www.cabdirect.org/cabdirect/abstract/19640100703

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., … DePristo, M. A. (2010). The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research*, *20*(9), 1297–1303. doi: 10.1101/gr.107524.110

Meier, J. I., Marques, D. A., Mwaiko, S., Wagner, C. E., Excoffier, L., & Seehausen, O. (2017). Ancient hybridization fuels rapid cichlid fish adaptive radiations. *Nature Communications*, *8*(1), 1–11. doi: 10.1038/ncomms14363

Meier, J. I., Sousa, V. C., Marques, D. A., Selz, O. M., Wagner, C. E., Excoffier, L., & Seehausen, O. (2017). Demographic modelling with whole-genome data reveals parallel origin of similar *Pundamilia* cichlid species after hybridization. *Molecular Ecology*, *26*(1), 123–141. doi: 10.1111/mec.13838

Mendelson, T. C., & Shaw, K. L. (2005). Rapid speciation in an arthropod. *Nature*, *433*(7024), 375–376. doi: 10.1038/433375a

Mirarab, S., & Warnow, T. (2015). ASTRAL-II: Coalecent-based species tree estimation with many hundreds of taxa and thousands of genes. *Bioinformatics*, *31*(12), i44–i52. doi: 10.1093/bioinformatics/btv234

Nguyen, L.-T., Schmidt, H. A., von Haeseler, A., & Minh, B. Q. (2015). IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies. *Molecular Biology and Evolution*, *32*(1), 268–274. doi: 10.1093/molbev/msu300

Nosil, P. (2012). *Ecological Speciation*. Oxford University Press.

Nosil, P., Crespi, B. J., & Sandoval, C. P. (2002). Host-plant adaptation drives the parallel evolution of reproductive isolation. *Nature*, *417*(6887), 440–443. doi: 10.1038/417440a

Otte, D. (1994). The Crickets of Hawaii: Origin. *Systematics and Evolution: Philadelphia, The Orthopterists Society*.

Papadopoulou, A., Anastasiou, I., & Vogler, A. P. (2010). Revisiting the Insect Mitochondrial Molecular Clock: The Mid-Aegean Trench Calibration. *Molecular Biology and Evolution*, *27*(7), 1659–1672. doi: 10.1093/molbev/msq051

Perdeck, A. C. (1958). The Isolating Value of Specific Song Patterns in Two Sibling Species of Grasshoppers (*Chorthippus brunneus* Thunb. and *C. biguttulus* L.). *Behaviour*, *12*(1/2), 1–75. Retrieved from JSTOR.

Pfeifer, B., Wittelsbürger, U., Ramos-Onsins, S. E., & Lercher, M. J. (2014). PopGenome: An Efficient Swiss Army Knife for Population Genomic Analyses in R. *Molecular Biology and Evolution*, *31*(7), 1929–1936. doi: 10.1093/molbev/msu136

Portik, D. M., Leaché, A. D., Rivera, D., Barej, M. F., Burger, M., Hirschfeld, M., … Fujita, M. K. (2017). Evaluating mechanisms of diversification in a Guineo-Congolian tropical forest frog using demographic model selection. *Molecular Ecology*, *26*(19), 5245–5263. doi: 10.1111/mec.14266

Presgraves, D. C. (2010). The molecular evolutionary basis of species formation. *Nature Reviews Genetics*, *11*(3), 175–180. doi: 10.1038/nrg2718

R Core Team. (2017). *R: A Language and Environment for Statistical Computing*. Retrieved from https://www.R-project.org/

Ragge, D. R., & Reynolds, W. J. (1988). The songs and taxonomy of the grasshoppers of the *Chorthippus biguttulus* group in the Iberian Peninsula (Orthoptera: Acrididae). *Journal of Natural History*, *22*(4), 897–929. doi: 10.1080/00222938800770611

Ragge, D., Reynolds, W., & Willemse, F. (1990). The songs of the European grasshoppers of the *Chorthippus biguttulus* group in relation to their taxonomy, speciation and biogeography (Orthoptera, Acrididae). *Boletin de Sanidad Vegetal. Plagas (Spain).(Spec. No. 20).*

Rambaut, A., Drummond, A. J., Xie, D., Baele, G., & Suchard, M. A. (2018). Posterior Summarization in Bayesian Phylogenetics Using Tracer 1.7. *Systematic Biology*, *67*(5), 901–904. doi: 10.1093/sysbio/syy032

Ravinet, M., Westram, A., Johannesson, K., Butlin, R., André, C., & Panova, M. (2016). Shared and nonshared genomic divergence in parallel ecotypes of *Littorina saxatilis* at a local scale. *Molecular Ecology*, *25*(1), 287–305. doi: 10.1111/mec.13332

Renaut, S., Owens, G. L., & Rieseberg, L. H. (2014). Shared selective pressure and local genomic landscape lead to repeatable patterns of genomic divergence in sunflowers. *Molecular Ecology*, *23*(2), 311–324. doi: 10.1111/mec.12600

Reynolds, J., Weir, B. S., & Cockerham, C. C. (1983). Estimation of the Coancestry Coefficient: Basis for a Short-Term Genetic Distance. *Genetics*, *105*(3), 767–779.

Riesch, R., Muschick, M., Lindtke, D., Villoutreix, R., Comeault, A. A., Farkas, T. E., … Nosil, P. (2017). Transitions between phases of genomic differentiation during stick-insect speciation. *Nature Ecology & Evolution*, *1*(4), 1–13. doi: 10.1038/s41559-017-0082

Ritchie, M. G. (2007). Sexual Selection and Speciation. *Annual Review of Ecology, Evolution, and Systematics*, *38*(1), 79–102. doi: 10.1146/annurev.ecolsys.38.091206.095733

Rundell, R. J., & Price, T. D. (2009). Adaptive radiation, nonadaptive radiation, ecological speciation and nonecological speciation. *Trends in Ecology & Evolution*, *24*(7), 394–399. doi: 10.1016/j.tree.2009.02.007

Schluter, D. (2000). *The Ecology of Adaptive Radiation*. OUP Oxford.

Schubert, M., Ermini, L., Sarkissian, C. D., Jónsson, H., Ginolhac, A., Schaefer, R., … Orlando, L. (2014). Characterization of ancient and modern genomes by SNP detection and phylogenomic and metagenomic analysis using PALEOMIX. *Nature Protocols*, *9*(5), 1056–1082. doi: 10.1038/nprot.2014.063

Servedio, M. R., & Boughman, J. W. (2017). The Role of Sexual Selection in Local Adaptation and Speciation. *Annual Review of Ecology, Evolution, and Systematics*, *48*(1), 85–109. doi: 10.1146/annurev-ecolsys-110316-022905

Servedio, M. R., Doorn, G. S. V., Kopp, M., Frame, A. M., & Nosil, P. (2011). Magic traits in speciation: 'magic' but not rare? *Trends in Ecology & Evolution*, *26*(8), 389–397. doi: 10.1016/j.tree.2011.04.005

Skotte, L., Korneliussen, T. S., & Albrechtsen, A. (2013). Estimating Individual Admixture Proportions from Next Generation Sequencing Data. *Genetics*, *195*(3), 693–702. doi: 10.1534/genetics.113.154138

Snook, R. R., Chapman, T., Moore, P. J., Wedell, N., & Crudgington, H. S. (2009). Interactions between the sexes: New perspectives on sexual selection and reproductive isolation. *Evolutionary Ecology*, *23*(1), 71–91. doi: 10.1007/s10682-007-9215-3

Soraggi, S., Wiuf, C., & Albrechtsen, A. (2018). Powerful Inference with the D-Statistic on Low-Coverage Whole-Genome Data. *G3: Genes, Genomes, Genetics*, *8*(2), 551–566. doi: 10.1534/g3.117.300192

Strimmer, K., & Haeseler, A. von. (1997). Likelihood-mapping: A simple method to visualize phylogenetic content of a sequence alignment. *Proceedings of the National Academy of Sciences*, *94*(13), 6815–6819. doi: 10.1073/pnas.94.13.6815

Suchard, M. A., Lemey, P., Baele, G., Ayres, D. L., Drummond, A. J., & Rambaut, A. (2018). Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. *Virus Evolution*, *4*(1). doi: 10.1093/ve/vey016

Tajima, F. (1989). Statistical Method for Testing the Neutral Mutation Hypothesis by DNA Polymorphism. *Genetics*, *123*(3), 585–595.

Takimoto, G., Higashi, M., & Yamamura, N. (2000). A Deterministic Genetic Model for Sympatric Speciation by Sexual Selection. *Evolution*, *54*(6), 1870–1881. doi: 10.1111/j.0014-3820.2000.tb01234.x

Thorpe, R. S., Surget-Groba, Y., & Johansson, H. (2010). Genetic Tests for Ecological and Allopatric Speciation in Anoles on an Island Archipelago. *PLOS Genetics*, *6*(4), e1000929. doi: 10.1371/journal.pgen.1000929

Toews, D. P. L., Taylor, S. A., Vallender, R., Brelsford, A., Butcher, B. G., Messer, P. W., & Lovette, I. J. (2016). Plumage Genes and Little Else Distinguish the Genomes of Hybridizing Warblers. *Current Biology*, *26*(17), 2313–2318. doi: 10.1016/j.cub.2016.06.034

van Doorn, G. S., Dieckmann, U., & Weissing, F. J. (2004). Sympatric Speciation by Sexual Selection: A Critical Reevaluation. *The American Naturalist*, *163*(5), 709–725. doi: 10.1086/383619

Vedenina, V., & Mugue, N. (2011). Speciation in gomphocerine grasshoppers: Molecular phylogeny versus bioacoustics and courtship behavior. *Journal of Orthoptera Research*, *20*(1), 109–125. Retrieved from JSTOR.

Vijay, N., Bossu, C. M., Poelstra, J. W., Weissensteiner, M. H., Suh, A., Kryukov, A. P., & Wolf, J. B. W. (2016). Evolution of heterogeneous genome differentiation across multiple contact zones in a crow species complex. *Nature Communications*, *7*(1), 1–10. doi: 10.1038/ncomms13195

von Helversen, O., & von Helversen, D. (1994). Forces driving coevolution of song and song recognition in grasshoppers. *Fortschritte Der Zoologie*, 253–253.

Wang, X., Fang, X., Yang, P., Jiang, X., Jiang, F., Zhao, D., … Kang, L. (2014). The locust genome provides insight into swarm formation and long-distance flight. *Nature Communications*, *5*(1), 1–9. doi: 10.1038/ncomms3957

Waterhouse, R. M., Seppey, M., Simão, F. A., Manni, M., Ioannidis, P., Klioutchnikov, G., … Zdobnov, E. M. (2018). BUSCO Applications from Quality Assessments to Gene Prediction and Phylogenomics. *Molecular Biology and Evolution*, *35*(3), 543–548. doi: 10.1093/molbev/msx319

Watterson, G. A. (1975). On the number of segregating sites in genetical models without recombination. *Theoretical Population Biology*, *7*(2), 256–276. doi: 10.1016/0040-5809(75)90020-9

Xu, M., & Shaw, K. L. (2019). Genetic coupling of signal and preference facilitates sexual isolation during rapid speciation. *Proceedings of the Royal Society B: Biological Sciences*, *286*(1913), 20191607. doi: 10.1098/rspb.2019.1607

Yang, Y., Servedio, M. R., & Richards-Zawacki, C. L. (2019). Imprinting sets the stage for speciation. *Nature*, *574*(7776), 99–102. doi: 10.1038/s41586-019-1599-z

Data Accessibility Statement:

Raw reads were deposited in the NCBI Sequence Read Archive (BioProject XXXX; BioSample Acc. XXXX). The dryad archive (doi: XXXX) contains the de novo assembly of the transcriptome, coordinates of protein coding genes, coordinates of neutral sites, list of Fst, Dxy, theta and Tajima's D per gene. Bioinformatic scripts are available in a public GitHub repository (https://github.com/zjnolen/chorthippus_radiation).

Author's Contributions:

The project was conceived by RJP. The data was produced by DBA, FM and RJP. The analyses were performed and interpreted by ZJN, BY, II, SL, CG and RJP. Reagents and analytical tools were contributed by MTPG and RJP. The first version of the manuscript was drafted by ZJN and RJP, and all remaining coauthors contributed to the final version.

Figures

Figure 1. Grasshopper species of the genus *Chorthippus* have similar distribution, morphology and habitat, but are highly divergent in male mating song (shown by oscillograms) and in associated female preference. Map depicts approximate species distributions (Ragge et al. 1990) and sampling sites.

Figure 2. Time calibrated tree based on full mitochondrial data shows that the radiation of *Chorthippus* species has occurred in the mid-Pleistocene and shows incomplete lineage sorting among species (colored squares). Dotted lines and grey boxes mark the estimated time of the initial divergence between *Pseudochorthippus* and *Chorthippus*, and of the recent radiation of *Chorthippus* species. See topology support and labels for taxa in Fig. S4.

Figure 3. Multispecies coalescent using 3,963 independent nuclear gene trees shows that species of *Chorthippus* rapidly radiated from a recent common ancestor and that behaviorally isolated species form monophyletic clades. Stars sign posterior probabilities of 1.0. Pie charts show the quartet support at each node and reflect the amount of discordant gene trees caused by incomplete lineage sorting and/or gene flow. Branch lengths reflect coalescent times of T/Ne.

Figure 4. Species of *Chorthippus* have diverged at a similar time as subspecies of *Pseudochorthippus* according to Dxy values of 3,963 genes (A), but 4.5 times lower levels of genomic differentiation according to Fst of 8,935 genes (B).

Figure 5. Sympatric species of *Chorthippus* sampled in central Europe show rare hybridization (left) and experienced a recent range expansion (right). Color of the bar plots show the individual assignment of each individual based on 554,849 SNPs, at three K levels of genetic structure (see all K's in Fig. S11). Violin plots show distribution of Tajima's D for 9,143 genes.

Figure 6. Extensive introgression between *Chorthippus* spp. follow a period of genetic isolation. A. D statistics is significant irrespective if those taxa were sampled in allopatry (orange) or sympatry (blue), consistent with ancestral gene flow; see Fig. S13 and Table S6 for all pairwise comparisons. B. Demographic modeling of polymorphisms considering four speciation scenarios (top) show that for all population pairs the observed SFS (bottom left) are

most similar to a SFS produced under a model of genetic isolation followed by secondary contact (bottom right); see Fig. S14 and Tables S7 and S8 for results on model and parameter estimation.