Zootaxa 4750 (3): 328–348
https://www.mapress.com/j/zt/

Article

ISSN 1175-5326 (print edition)
ZOOTAXA
ISSN 1175-5334 (online edition)

https://doi.org/10.11646/zootaxa.4750.3.2
http://zoobank.org/urn:lsid:zoobank.org:pub:B5B2E53D-8950-469C-92B4-AE6B8CBA96B5

Multilocus phylogeny of *Gryllus* field crickets (Orthoptera: Gryllidae: Gryllinae) utilizing anchored hybrid enrichment

DAVID A. GRAY[1,6], DAVID B. WEISSMAN[2], JEFFREY A. COLE[3], EMILY MORIARTY LEMMON[4] & ALAN R. LEMMON[5]

[1] *Department of Biology, California State University Northridge, 18111 Nordhoff Street, Northridge, CA 91330-8303, USA. E-mail: dave.gray@csun.edu*
[2] *Department of Entomology, California Academy of Sciences, San Francisco, CA 94118, USA. E-mail: gryllus@gmail.com*
[3] *Department of Biology, Pasadena City College, Pasadena, CA, 91106, and Natural History Museum of Los Angeles County, 900 Exposition Boulevard, Los Angeles, CA 90007, USA. E-mail: jacole@pasadena.edu*
[4] *Department of Biological Science, Florida State University, Tallahassee, FL, 32306, USA. E-mail: chorusfrog@bio.fsu.edu*
[5] *Department of Scientific Computing, Florida State University, Tallahassee, FL, 32306, USA. E-mail: alemmon@fsu.edu*
[6] *Correspondence: dave.gray@csun.edu*

Abstract

We present the first comprehensive molecular phylogeny of *Gryllus* field cricket species found in the United States and Canada, select additional named *Gryllus* species found in Mexico and the Bahamas, plus the European field cricket *G. campestris* Linnaeus and the Afro-Eurasian cricket *G. bimaculatus* De Geer. *Acheta*, *Teleogryllus*, and *Nigrogryllus* were used as outgroups. Anchored hybrid enrichment was used to generate 492,531 base pairs of DNA sequence from 563 loci. RAxML analysis of concatenated sequence data and Astral analysis of gene trees gave broadly congruent results, especially for older branches and overall tree structure. The North American *Gryllus* are monophyletic with respect to the two Old World taxa; certain sub-groups show rapid recent divergence. This is the first Anchored Hybrid Enrichment study of an insect group done for closely related species within a single genus, and the results illustrate the challenges of reconstructing the evolutionary history of young rapidly diverged taxa when both incomplete lineage sorting and probable hybridization are at play. Because *Gryllus* field crickets have been used extensively as a model system in evolutionary ecology, behavior, neuro-physiology, speciation, and life-history and life-cycle evolution, these results will help inform, interpret, and guide future research in these areas.

Key words: phylogenetics, evolution, phylogenomics

Introduction

Gryllus field crickets are typically large, loud, conspicuous species, which can be locally abundant in a wide variety of habitats (Weissman & Gray 2019). Because of this, they have been extensively studied in many contexts. For example, a Web of Science search (v5.32 on April 29, 2019) for "*Gryllus* " returned 2914 results, dating back to 1903, distributed among many sub-fields of biology; publication numbers in the top 10 sub-fields are shown in Fig. 1 (note: a single study can be represented in multiple sub-fields).

In some cases, interpretation of the results of such studies depends upon the evolutionary historical relationships among the various species (Alexander 1968; Alexander & Bigelow 1960; Blankers *et al.* 2015, 2017, 2018; Crone *et al*. 2007; Desutter-Grandcolas & Robillard 2003; Gray & Cade 2000; Gray *et al.* 2016a, b; Harrison 1983, 1985; Jang *et al.* 2008, 2009; Veen *et al.* 2013). Well-resolved phylogenetic hypotheses for *Gryllus* have been a challenge, however. The principal issue has been adequate taxon sampling for phylogenetic analysis in the absence of detailed alpha taxonomy. The secondary issue has been generation of sufficient molecular data, particularly nuclear gene data, from multiple loci to adequately resolve relationships especially among closely related taxa. We address both of these issues in this paper, but first briefly review the history of taxonomic and molecular work to better place our current effort into context.

 Accepted by D. Rentz: 7 Feb. 2020; published: 12 Mar. 2020

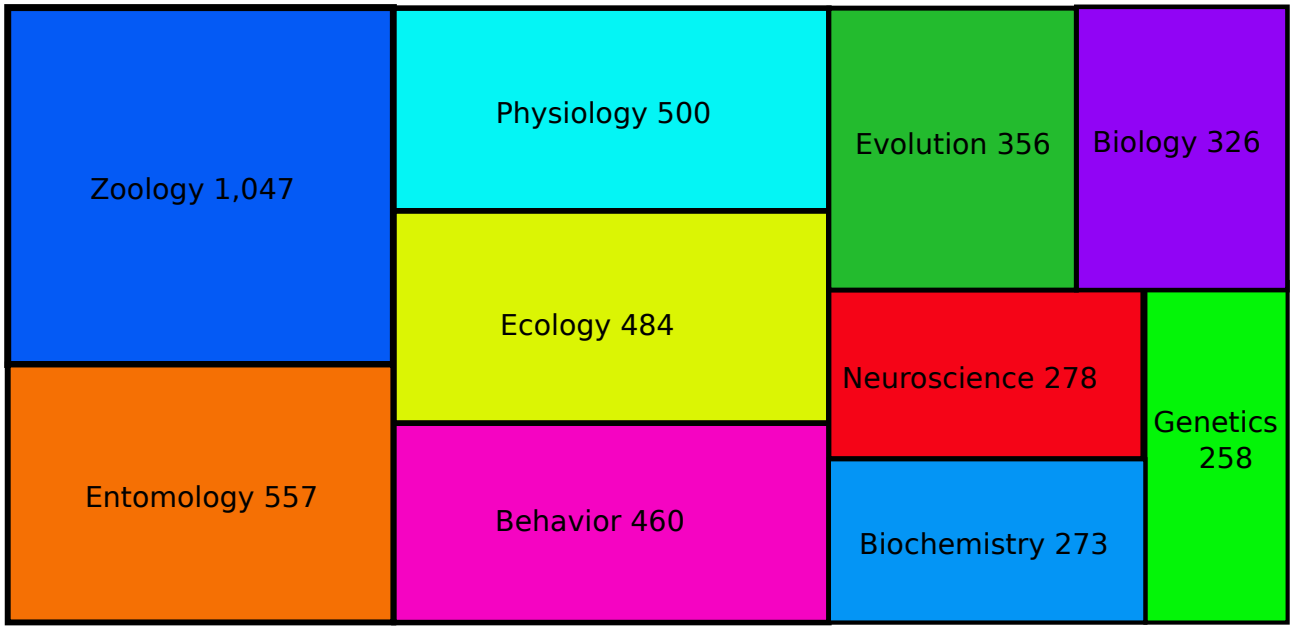

Fig. 1. Top 10 categories of publications involving *Gryllus* field crickets; data are from the Web of Science and represent 2914 records from 1903 – 2019.

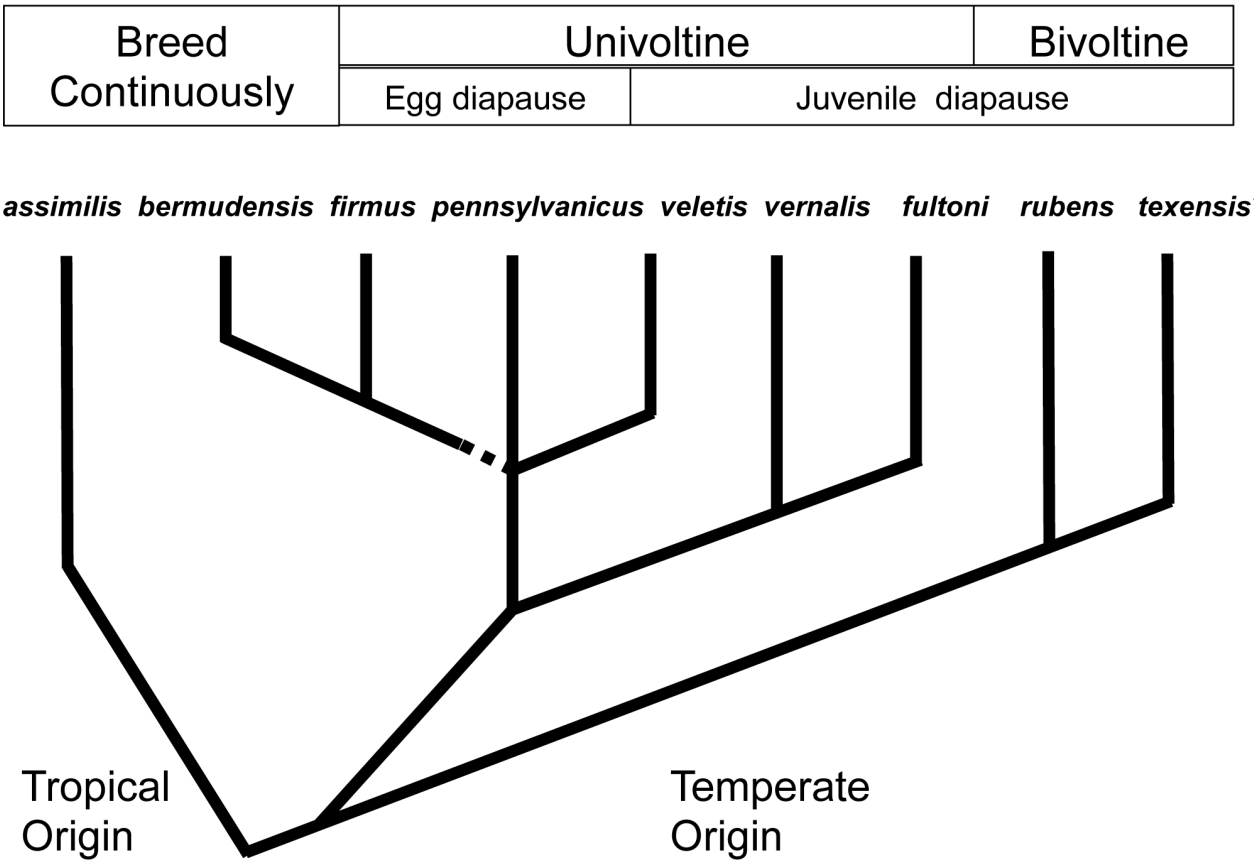

FIG. 2. Alexander's first phylogeny of US *Gryllus* principally based on life-cycle considerations. Redrawn from Alexander (1968). * Note: *G. texensis* was labeled as *G. integer* in Alexander's original figure, see Cade & Otte (2000).

Weissman & Gray (2019) present a detailed history of *Gryllus* taxonomy in North America, with discussion of each species individually. In brief, the first recognized US taxa, *G. assimilis* (Fabricius), was named in 1775, followed by *G. pennsylvanicus* Burmeister in 1838; *G. lineaticeps* Stål in 1861; and *G. personatus* Uhler in 1864.

Scudder subsequently described *G. vocalis, G. rubens, G. firmus, G. armatus,* and *G. integer* (Scudder, 1901, 1902). However, a review of Western Hemisphere specimens led Rehn and Hebard (1915) to conclude that all were variants of a single plastic species, *G. assimilis*. This early work was based entirely on morphology, without consideration of life cycle, ecology, or song, which later proved to be essential characters delimiting species. Once the importance of such characters was realized, 9 additional taxa were described, bringing the US total to 19 prior to Weissman & Gray's revision (2019). That revision considered all US species, but focused principally west of the Mississippi River, and resulted in the description of 17 new species, bringing the US total to 35, with the synonymizing of *G. alogus* Rehn under *G. vocalis*. Weissman & Gray (2019) also discussed, without formally naming, an additional set of problematic populations that may represent at least 7 additional lineages that further work should address. Publication of that revision means, for the first time, that phylogenetic consideration of US *Gryllus* could include all likely species, both named and potential.

The first phylogeny of US *Gryllus*, to our knowledge, was published by Alexander (1968, Fig. 2), and was based principally upon consideration of life cycle. Although not based on a formal analysis, Alexander nonetheless grouped some species pairs in ways generally supported by subsequent analyses of molecular data (e.g. *G. vernalis* Blatchley and *G. fultoni* (Alexander); *G. firmus* and *G. pennsylvanicus*; *G. rubens* and *G. texensis* Cade & Otte [as *G. integer* Scudder]). Notably, Alexander's placement of *G. veletis* (Alexander & Bigelow) with *G. pennsylvanicus*, which was foundational to the 'allochronic speciation' hypothesis (Alexander & Bigelow 1960), has not been supported by subsequent genetic data.

The first molecular phylogeny of US *Gryllus* came from the Harrison lab (Harrison & Bogdanowicz 1995), based upon mtDNA restriction sites (Fig. 3A). That work was followed soon thereafter by an analysis of mtDNA sequence variation in the Cytochrome b and 16S Ribosomal RNA genes (Fig. 3B) (Huang *et al.* 2000). The taxon sampling in the two studies overlapped only modestly, making comparisons difficult. Both place *G. firmus*, *G. pennsylvanicus*, and *G. ovisopis* Walker together, and both place Old World *G. bimaculatus* and *G. campestris* together. Huang *et al.* (2000) show the US *Gryllus* monophyletic with respect to the Old World *Gryllus* and outgroups.

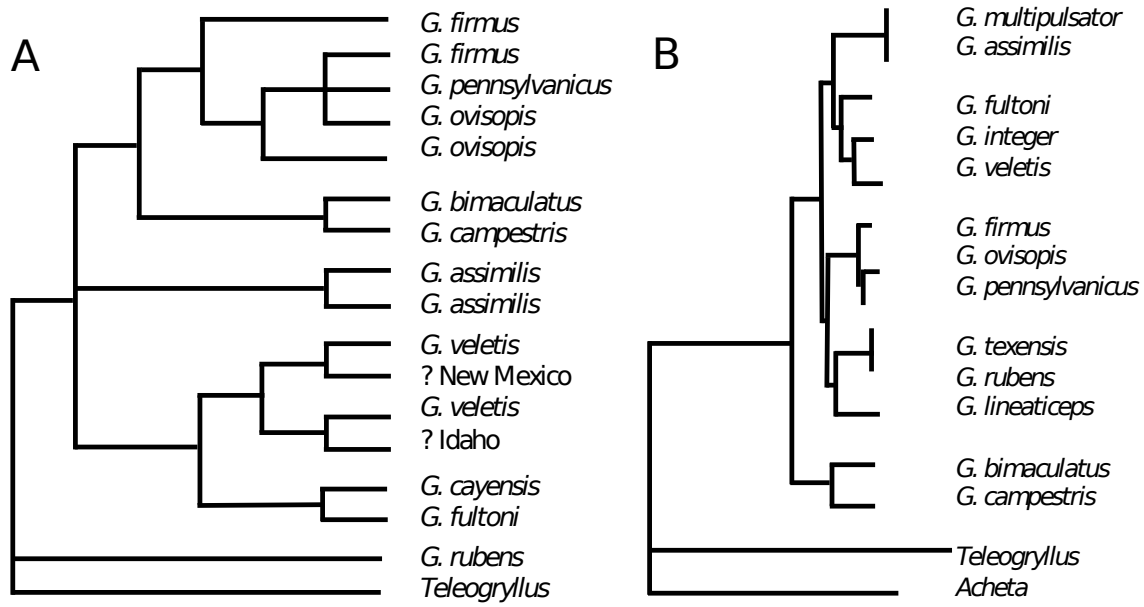

FIG. 3. Two mtDNA phylogenies of US *Gryllus*. (A) Strict consensus of two shortest trees from maximum parsimony analysis of mtDNA restriction sites (redrawn from Harrison & Bogdanowicz 1995). (B) Minimum evolution tree for ~1500 bp of combined Cytochrome b and 16S mtDNA sequence data (redrawn from Huang *et al.* 2000, with '*G. assimilis* CA' renamed *G. multipulsator* following Weissman *et al.* 2009).

In this work, we analyze the phylogenetic relationships of all named US taxa (35 species), several putative US taxa (7 additional lineages or "Clades" of unnamed US taxa), several named New World species from Mexico and the Bahamas, as well as feral or potentially feral species from the pet trade, and outgroups. The data were generated

using the 'anchored hybrid enrichment' technique (Lemmon *et al.* 2012), which enriches genomic DNA for loci likely to be conserved across taxa of interest and extending into less conserved more variable regions. This approach thus facilitates identification of multiple loci that can together provide insight at both shallow and deep phylogenetic scales (Lemmon *et al.* 2012, Lemmon & Lemmon 2012, 2013). In addition, we use BEAST v. 2.3.2 (Drummond & Bouckaert 2014; Bouckaert *et al.* 2014) in an attempt to identify the approximate timing of divergence in certain groups.

Methods

Taxon Sampling

Representative samples of all named US *Gryllus* species are included in this work, as are select additional named New World species, plus two Old World *Gryllus: G. bimaculatus* and *G. campestris*. We use *Acheta domesticus* (Linnaeus), *Teleogryllus emma* (Ohmachi & Matsuura), and *Nigrogryllus sibiricus* (Chopard) as outgroups. In total, our analysis consisted of 94 samples; species and collection details are provided in Table 1. Named species correspond to Weissman & Gray (2019). The putative potential additional lineages are: (1) *G. firmus* Texas, potentially distinct from *G. firmus* of Florida (type locality of North Carolina) and the southeastern seaboard; (2-4) *G. montis* Weissman & Gray Clades 1, 3, & 4, potentially distinct from the type *G. montis* Clade 2; (5-7) the 'tulare' 'mormoni' and 'mohave' forms of *G. saxatilis* Weissman & Gray (Note: 'tulare' 'mormoni' and 'mohave' names are disclaimed as unavailable per ICZN).

TABLE 1. Cricket species, ID determination, source, and sample code for individuals analyzed in this study. Taxa are listed roughly in the order that they appear in Figs. 7-10.

Species	ID Det.	Source	Sample Code
Teleogryllus emma	DAG, Y. Jang	South Korea, Daegu, Dalseong-gun, Yugasa, 25-viii-2012	2012-078
Acheta domesticus	DAG	Commercial pet store	2017-045
Nigrogryllus sibiricus	DAG, Y. Jang	South Korea, Daegu, Dalseong-gun, Yugasa, 25-viii-2012	2012-225
G. bimaculatus	W.H. Cade	Zimbabwe, Harare (culture)	1999-101
G. campestris	T. Tregenza	Spain, Asturias, Oviedo	2017-044
G. alexanderi	DBW	Mexico, Clarion Island, iii-2005	G451
G. brevicaudus	DBW	USA, CA, San Mateo Co., Jasper Ridge, 18-v-2016	G3393
G. insularis	DBW	Mexico, Guadalupe Island, 10-vi-2000	G274
G. bryanti	P.A. DeLuca, K.A. Judge, DBW	Bahamas, Eleuthera Island, 10-v-2012	GBM05
G. veintinueve	DBW, DAG	USA, OK, Love Co., 33.98385° -96.97518°, 15-xi-2015	2015-055
G. veintinueve	DBW	USA, TX, Howard Co., Big Springs State Park, 30-vi-2009 (S09-71)	G1330
G. locorojo	DBW, DAG	Commercial pet store	G2159
G. multipulsator	DBW	USA, AZ, Pima Co., Gila Bend, 1-viii-2009 (S09-103)	G1414
G. assimilis	DBW	USA, TX, Brewster Co., Big Bend National Park, Rio Grande Village, 28-v-2016 (S16-12)	G3373
G. thinos	DBW	USA, TX, Cameron Co., Boca Chica State Park, 10-vi-2007 (S07-25)	G1209
G. thinos	DBW	USA, TX, Kleberg Co., Padre Island National Seashore, 11-vi-2011 (S11-35)	G2018
G. ovisopis	K.A. Judge, DAG	USA, FL, Alachua Co., Gainesville, 23-viii-2003	2016-035
G. ovisopis	K.A. Judge, DAG	USA, FL, Alachua Co., Gainesville, 23-viii-2003	2018-001

.....*Continued next page*

TABLE 1. (continued)

Species	ID Det.	Source	Sample Code
G. firmus Florida	DBW	USA, FL, Alachua Co., Gainesville, 23-viii-2003 (S03-85)	G62
G. pennsylvanicus	DBW	USA, UT, Wayne Co., Hanksville, 12-ix-2004 (S04-128)	G368
G. pennsylvanicus	DBW	USA, VT, Addison Co., Middlebury College, 5-x-2008 (S08-74)	G710
G. pennsylvanicus	DAG, DBW	USA, TX, Lubbock Co., Lubbock 18-ix-2013 (S13-80)	G2708
G. firmus Texas	DBW	USA, TX, Kinney Co., Brackettville, 7-ix-2010 (S10-63)	G1920
G. firmus Texas	DBW	USA, TX, Bastrop Co., Bastrop State Park, 9-ix-2010 (S10-67)	G1915
G. firmus Texas	DBW	USA, TX, Fayette Co., 3.7 km S Schulenburg, 9-ix-2010 (S10-65)	G1917
G. firmus Texas	DBW	USA, TX, Jefferson Co., Sea Rim State Park, 10-vi-2011 (S11-29)	G2029
G. firmus Texas	DBW	USA, TX, Matagorda Co., Matagorda Island, 13-vii-2013 (S13-59)	G2715
G. rubens	DAG	USA, FL, Jackson Co., Marianna 28-ix-1999	d437
G. regularis	DAG, DBW	USA, AZ, Yavapai Co., Agua Fria National Monument, 9-viii-2016	2016-037
G. texensis	DBW	USA, TX, Brewster Co., Big Bend National Park, Rio Grande Village, 28-v-2016 (S16-12)	G3382
G. integer	DBW	USA, CA, Tulare Co., 6 m E Lemon Cove 29-vi-2016 (S16-21)	G3416
G. integer	DAG, DBW	USA, UT, Garfield Co., 37.63094° - 110.72110	2003-039
G. armatus	DBW	USA, CA, San Bernardino Co., near Goffs 23-vii-2016 (S16-32)	G3439
G. armatus	DBW	USA, TX, Presidio Co., Presidio 27-v-2016 (S16-5)	G3374
G. vernalis	DBW	USA, IN, Crawford Co., 38° 12.312 -86° 18.246, 4-vi-2003 (S03-62)	G31
G. vernalis	DBW	USA, IL, Johnson Co., Ferne Clyffe State Park, 8-vii-2014 (S14-35)	G2754
G. fultoni	DBW	USA, IN, Crawford Co., 38° 12.312 -86° 18.246, 4-vi-2003 (S03-62)	G34
G. cayensis	DAG, DBW	USA, FL, Miami-Dade Co., Long Pine Key, Everglades National Park, 14-v-2018	2018-002
G. planeta	DBW	USA, TX, Jeff Davis Co., Davis Mts., 1-vii-2015 (S15-61)	G3088
G. montis (clade 1)	DBW	USA, AZ, Cochise Co., Southwest Research Station, 20-viii-2012 (S12-103)	G2416
G. montis (clade 1)	DBW	USA, AZ, Cochise Co., Chiricahua National Monument, 2-vi-2013 (S12-21)	G2464
G. veletis	DBW	USA, UT, Juab Co., Nephi, 24-v-2015 (S15-23)	G2958
G. veletis	DBW	USA, IN, Spencer Co., Dale, 4-vi-2003 (S03-60)	G30
G. veletis	DBW	USA, OK, Texas Co., Guymon, 1-vii-2009 (S09-77)	G1345
G. veletis	DBW	USA, TX, Brewster Co., Alpine, 2-vii-2015 (S15-73)	G3075
G. cohni	DBW	Mexico, Sinaloa, 20 km S Mazatlán, 23-vii-2014 (S14-53)	G2776

.....Continued next page

TABLE 1. (continued)

Species	ID Det.	Source	Sample Code
G. cohni	DAG, DBW	USA, AZ, Yavapai Co., Agua Fria National Monument, 9-viii-2016	2016-041
G. cohni	DBW	Mexico, Baja California Sur, 8.5 km W La Paz, 25-viii-1995 (S95-81)	G101
G. vocalis	DBW	USA, NM, Bernalillo Co., Albuquerque, 11-vii-2012	G3335
G. vocalis	DBW	USA, AZ, Maricopa Co., Gila Bend, 30-vii-2015 (S15-111)	G3227
G. vocalis	DAG, DBW	USA, CA, Los Angeles Co., California State University Northridge, 11-v-2016	2016-036
G. personatus	DBW	USA, CO, Otero Co., La Junta, 2-vii-2009 (S09-82)	G1357
G. staccato	DAG, DBW	USA, AZ, Yavapai Co., Agua Fria National Monument, 9-viii-2016	2016-034
G. lineaticeps	DAG	USA, CA, San Joaquin Co., Tracy, 10-ix-2016	2016-033
G. lineaticeps	DBW	Mexico, Baja California Norte, Guadalupe Island, 21-ii-2008 (S08-9)	G647
G. chisosensis	DBW	USA, TX, Brewster Co, Big Bend National Park, 28-v-2016 (S16-13)	G3400
G. veletisoides	DBW	USA, CA, Santa Clara Co., Los Gatos, 10-v-2006 (S06-30)	G568
G. veletisoides	DBW	USA, CA, Santa Cruz Co., Santa Cruz, 13-ix-2015 (S15-120)	G3334
G. montis (clade 2)	DBW	USA, AZ, Pima Co., Kitt Peak, 8-vi-2013 (S13-36)	G2491
G. montis (clade 2)	DAG, DBW	USA, AZ, Santa Cruz Co., Sycamore Canyon, 8-iv-2004	2004-073
G. montis (clade 2)	DAG, DBW	USA, AZ, Santa Cruz Co., Santa Rita Mts., 19-viii-2005	2005-012
G. montis (clade 2)	DBW	USA, AZ, Cochise Co., Bisbee, 1-vi-2013 (S13-18)	G2471
G. montis (clade 2)	DBW	USA, AZ, Cochise Co., Ramsey Canyon Preserve, 1-vi-2013 (S13-17)	G2475
G. montis (clade 3)	DBW	USA, AZ, Coconino Co., 4.8 m N Sedona, 15-vi-2007 (S07-60)	G1097
G. montis (clade 3)	DBW	USA, AZ, Mohave Co., Hualapai Mt. Rec Area, 16-vi-2007 (S07-62)	G1151
G. montis (clade 3)	DBW	USA, AZ, Pima Co., Mt. Lemmon Rec Area, 27-vi-2009 (S09-50)	G1353
G. montis (clade 3)	DBW	USA, AZ, Graham Co., Mt. Graham, 10-vi-2012 (S12-18)	G2241
G. transpecos	DBW	USA, TX, Brewster Co., Big Bend National Park, Panther Junction, 2-vii-2015 (S15-68)	G3062
G. transpecos	DBW	USA, TX, Jeff Davis Co., Davis Mts., 1-vii-2015 (S15-61)	G3083
G. lightfooti	DAG, DBW	USA, NM, Dona Ana Co., Jornada 8 km E Las Cruces, culture	2016-038
G. lightfooti	DBW	Mexico, Sonora, 6 km W Guaymas, 7-vi-2013 (S13-35)	G2665
G. sotol	DBW	USA, NM, Dona Ana Co., Organ Mts., Aguirre Springs Campground, 19-v-2017 (S17-4)	G3090
G. sotol	DBW	USA, NM, Dona Ana Co., Organ Mts., Aguirre Springs Campground, 19-v-2017 (S17-4)	G3509
G. sotol	DBW	USA, NM, Dona Ana Co., Organ Mts., Aguirre Springs Campground, 19-v-2017 (S17-4)	G3493

.....Continued next page

TABLE 1. (continued)

Species	ID Det.	Source	Sample Code
G. vulcanus	DAG, DBW	USA, NM, Cibola Co., El Malpais National Monument, 21-vii-2016	2016-025
G. vulcanus	DBW	USA, NM, Lincoln Co., Valley of Fires State Park, 13-vi-2007 (S07-46)	G1030
G. longicercus	DAG, DBW	USA, AZ, Yuma Co., Kofa Mts., 6-viii-2016	2016-039
G. longicercus	DBW	USA, NM, Chaves Co., 11.7 km E Roswell, 28-vi-2009 (S09-59)	G1433
G. longicercus	DBW	USA, TX, Presidio Co., Shafter 27-v-2016 (S16-4)	G3386
G. montis (clade 4)	DBW	USA, NM, Catron Co., 33° 48.705 -109° 01.658, 14-vi-2007 (S07-54)	G1123
G. montis (clade 4)	DBW	USA, NM, Catron Co., 3.5 km S Reserve, 14-vi-2007 (S07-53)	G1048
G. navajo	DBW	USA, AZ, Coconino Co., 79 km E Flagstaff, 15-vi-2007 (S07-56)	G1067
G. navajo	DAG, DBW	USA, UT, Emery Co., Goblin Valley State Park, 18-v-2016	2016-040
G. saxatilis	DBW	USA, UT, Millard Co., Fillmore, 20-v-2017 (S17-7)	G3484
G. leei	DBW	USA, UT, Millard Co., 3.3 km NW Flowell, 20-v-2017 (S17-6)	G3475
G. leei	DBW	USA, UT, Millard Co., 13.7 km NW Fillmore, 18-v-2001 (S01-28)	G147
G. saxatilis (‘tulare’)	DBW	USA, CA, Tulare Co., 6 m E Lemon Cove 29-vi-2016 (S16-21)	G3422
G. saxatilis (‘tulare’)	DBW	USA, CA, Tulare Co., Hwy 190 16 km E Springville, 29-v-2009 (S09-34)	G1293
G. saxatilis (‘mormoni’)	DBW	USA, CA, Sacramento Co., Folsom, 19-vii-2015 (S15-87)	G3180
G. saxatilis (‘mormoni’)	DBW	USA, CA, Placer Co., Folsom Lake SRA, Beals Pt, 6-v-2013 (S13-10A)	G2725
G. saxatilis	DBW	USA, CA, Santa Clara Co., Mt. Hamilton, 27-viii-2015 (S15-114)	G3310
G. saxatilis (‘mohave’)	DBW	USA, AZ, Mohave Co., 2.9 km NW Hualapai Mt. Park, 16-vi-2007 (S07-64)	G1131
G. saxatilis (‘mohave’)	DBW	USA, CA, Inyo Co., Big Pine 27-vi-2016 (S16-18)	G3431
G. makhosica	DBW	USA, SD, Jackson Co., Badlands National Park, Cedar Pass, 3-vii-2009 (S09-89)	G1340

AHE Marker Development

Following the general approach of Lemmon *et al.* (2012), Hamilton *et al.* (2016), and Haddad *et al.* (2018), we developed an Anchored Hybrid Enrichment kit for Orthoptera, using the published transcriptomes of three species and low coverage whole genome reads from four species. The species and their sources are given in Table 2. Indexed Illumina libraries with insert size 200-500bp were prepared for the WGS samples following Prum *et al.* (2015). Those libraries were pooled and sequenced in the Translational Laboratory at the Florida State University College of Medicine. Two Illumina HiSeq2500 sequencing lanes (paired-end 200bp) produced ~120Gb of raw reads. The overlapping read pairs were merged following Rokyta *et al.* (2012). Consensus sequences were formed from assembly clusters containing at least 100 mapped reads. The set of consensus sequences for all individuals at each locus (the homologs) were then compared by computing the percentage of shared 20mers, and a pairwise distance matrix between all homologs was generated for each locus. Using this distance matrix, homologs were clustered

into orthologous sets (using a neighbor-joining approach), with the homologs being joined in an order reflecting their similarity (most similar joined first). Homologs originating from the same sample were never joined into the same orthology cluster. Each orthology cluster representing at least 50% of the samples were treated as a separate locus downstream.

Transcriptomes were scanned for the 941 AHE probe region sequences of the red flour beetle, *Tribolium castaneum* from Haddad *et al.* (2018). The transcript best matching to each of the *Tribolium* reference were identified and utilized downstream, so long as the pairwise sequence similarity exceeded 55%. Likewise, the WGS reads were mapped to the *Tribolium* reference sequences and the best matching read was retained (again, at least 55% sequence similarity was required). In order to obtain longer sequences for probe design, the best matching read sequences were extended up to 2000 bp in each direction using the WGS reads (see Hamilton *et al.* 2016 for details).

For each AHE locus, the *Tribolium* reference sequence and the best matching transcript and extended read for each species were aligned using MAFFT (Katoh & Standley 2013). Alignments were visually inspected in Geneious R9 (Biomatters Ltd., Kearse *et al.* 2012), and the largest region containing the *Tribolium* reference sequence but not containing an intron longer than 150bp was selected. Poorly aligned sequences within each selected region were masked. These regions were then profiled for repeat regions which were masked (see Hamilton *et al.* 2016 for details). Sequences from different loci that shared one or more 20-mers were considered to be overlapping. Loci were removed such that no loci overlapped. Loci containing fewer than 50% of the seven species were also removed. These filters resulted in the retention of 496 loci averaging 370bp. Probes of length 120bp were tiled uniformly across all sequences in the alignments (except *Tribolium*) at 10x coverage. A total of 39203 probes covered 183,233 target bases.

AHE Data/Sequencing

AHE data were collected in collaboration with the FSU Center for Anchored Phylogenomics (www.anchoredphylogeny.com), following Prum *et al.* (2016). Extracted DNA was fragmented to a size distribution ranging from 150-400bp using a Covaris ultrasonicator. Libraries with 8bp indexes were then prepared using a Beckman Coulter FXp liquid handling robot. After QC using Qubit, libraries were pooled in groups of 16 for sequencing on 2 lanes of a HiSeq2500 (PE150 protocol), which produced 93.3 Gb of raw data. Sequencing was performed at the Translational Laboratory at the Florida State University College of Medicine.

Bioinformatics

Reads were demultiplexed with no mismatches tolerated and quality filtered using the CASAVA high-chastity setting. Overlapping reads were merged following Rokyta *et al.* (2012). During this process, sequencing errors in overlapping regions were corrected and adapters were removed. Following Hamilton *et al.* (2016), reads were then assembled using a quasi-de-novo approach with the four WGS-derived sequences from the probe design serving as references. Assembly clusters containing fewer than 100 reads were removed from downstream analyses. Orthology was assessed using a neighbor-joining approach based on pairwise distances among consensus sequences derived from the assembly clusters. Alignments of orthologous sets of sequences were aligned using MAFFT (Katoh & Standley 2013), then masked to remove misaligned regions and misplaced sequences (see Hamilton *et al.* 2016 for methodological details).

The final alignments were comprised of 563 loci, 94 samples and 492,531 sites, 92,284 of which were variable and 45,262 were parsimony informative. The alignment for each locus contained an average of 879 sites. Only 10.22% of the bases contained gaps or ambiguous characters.

A maximum likelihood phylogeny was estimated using RAxML (2.2.3; Stamatakis 2006) from an alignment generated by concatenating the individual locus alignments. In the RAxML analyses, the GTR+Rates model was assumed, with a different set of model parameters being estimated for each locus (i.e. partitioned by locus). Support values were estimated using 100 bootstrap replicates. A species tree was also estimated using the pseudo-coalescent approach implemented in ASTRAL-II (Mirarab & Warnow 2015). As input for this analysis, we used gene trees estimated using RAxML (as described above but with one tree being estimated for each locus).

To estimate approximate divergence times among taxa, we used BEAST v. 2.3.2 (Drummond & Bouckaert 2014; Bouckaert *et al.* 2014). A dataset for divergence time estimation that consisted of 89 new world *Gryllus* exemplars was pared down from the 94 exemplar anchored hybrid enrichment dataset. Outgroups from related genera and the two European species *G. bimaculatus* and *G. campestris* were removed from the chronogram dataset as use of outgroups is not recommended with divergence time estimation, especially in analyses that employ a coalescent model as we used here (Drummond & Bouckaert 2014); furthermore trial analyses that included these taxa failed to root properly. Site substitution models and partitioning for the 563 loci in this dataset were found with PartitionFinder v. 2.1.1 (Lanfear *et al.* 2012, 2016) under the Bayesian Information Criterion, using the supercomputer resources at the CIPRES Science Gateway (Miller *et al.* 2010).

A Beast XML input file was created using BEAUTi v. 2 (Bouckaert *et al.* 2014). Unlinked subsets with independent site substitution models were set up according to the PartitionFinder results, with gamma category counts of 4 and proportions of invariant sites of 0.1 estimated as required in partitions. In order to reduce the number of parameters in this large analysis, site substitution rates were not estimated. This analysis assumed a single relaxed log normal clock model with the automatic set clock rate option turned off. The coalescent constant population size tree model was chosen as appropriate for this dataset as many exemplars represent intraspecific variation. The uncorrelated lognormal relaxed clock prior was set to a mean of 0.03±3.0 in real space. Four priors calibrated the analysis to geologic time. Two normally distributed priors were based upon molecular divergence time estimates: the divergence of Florida *G. firmus* from *G. pennsylvanicus* had a prior of 1.25 ± 0.25 MYA, based on a estimate of 0.2 MYA from mtDNA (Maroja *et al*. 2009) but potentially up to 2 MYA based on ~100 loci nDNA (L. Maroja, pers. comm. to DAG); second, a prior of 0.525 ± 0.116 MYA was used for the split of *G. rubens* and *G. texensis* (estimated at 0.35 to 0.70 MYA, Blankers *et al.* 2018). We also used a normally distributed prior for the split between *G. brevicaudus* Weissman, Rentz, & Alexander and *G. insularis* Scudder at 7±2 MYA based on the age of Guadalupe Island (Gonzalez-Garcia *et al*. 2003). A fourth calibration point was a uniform prior describing the split of *G. vulcanus* Weissman & Gray and *G. longicercus* Weissman & Gray using a minimum age estimate of the Valley of Fires lava fields in New Mexico, USA of 0.0052 MYA (Dunbar 1999) with the upper bound of the prior set in the present. All other priors remained at default settings.

Four replicate analyses ran at CIPRES for 3 x 10^8 generations each. Convergence was assessed using Tracer v. 1.6 (Rambaut & Drummond 2013). The resulting tree blocks were summarized into MCC trees with mean node heights using TreeAnnotator v. 2.3.2 (Rambaut & Drummond 2014) after discarding 25% burnin. The MCC trees were visualized and annotated in FigTree v. 1.4.4 (Rambaut 2009).

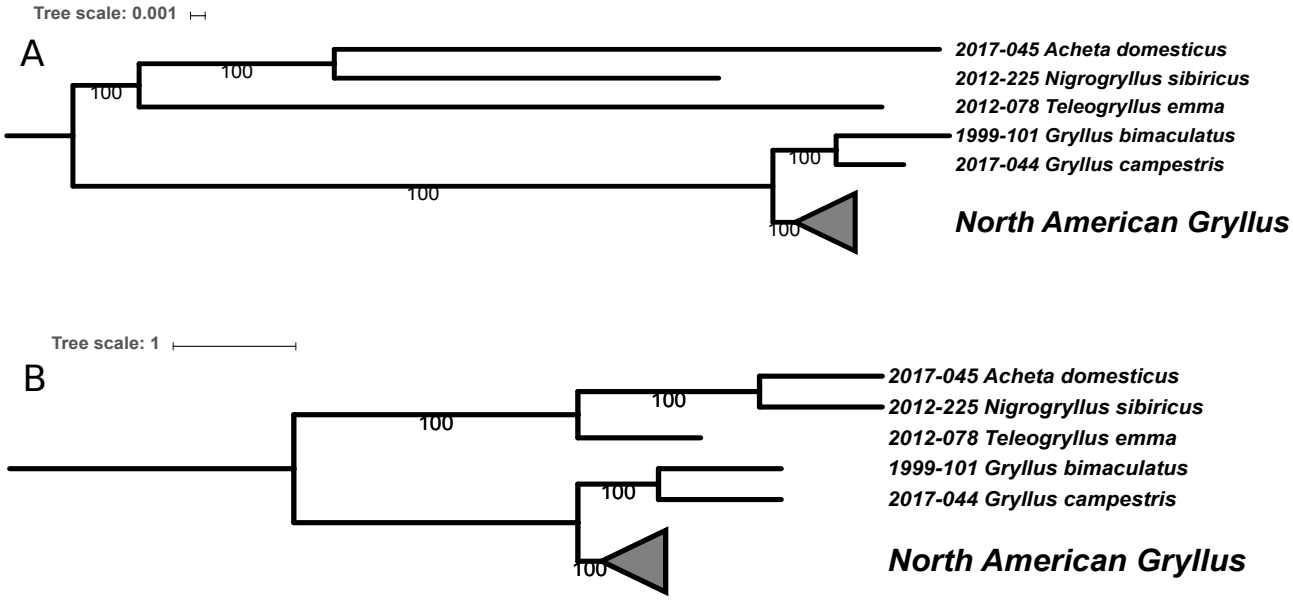

FIG. 4. Monophyly of North American *Gryllus* is strongly supported by (A) RAxML analysis of concatenated data (492,531 bp) with bootstrap values, and by (B) Astral analysis of gene trees (563 loci) with quadripartition branch support values.

Results

Both RAxML analysis of concatenated data and Astral analysis of gene trees strongly support monophyly of the North American *Gryllus* (Fig. 4). The analyses also support the placement of *Nigrogryllus sibiricus* outside of *Gryllus* despite similar male genitalia (i.e., supporting the synonymy of *Gryllus nigrohirsutus* Alexander under *Nigrogryllus sibiricus*).

Overall tree structure of the in-group taxa is likewise well supported, and generally concordant under both RAxML and Astral analytical approaches (Figs. 5, 6). Both approaches mostly resulted in the same Groups of taxa, with similar arrangements, with the major exception of within the *G. montis* Group (see below). The Astral analysis in general resulted in lower support values than the RAxML concatenated analysis.

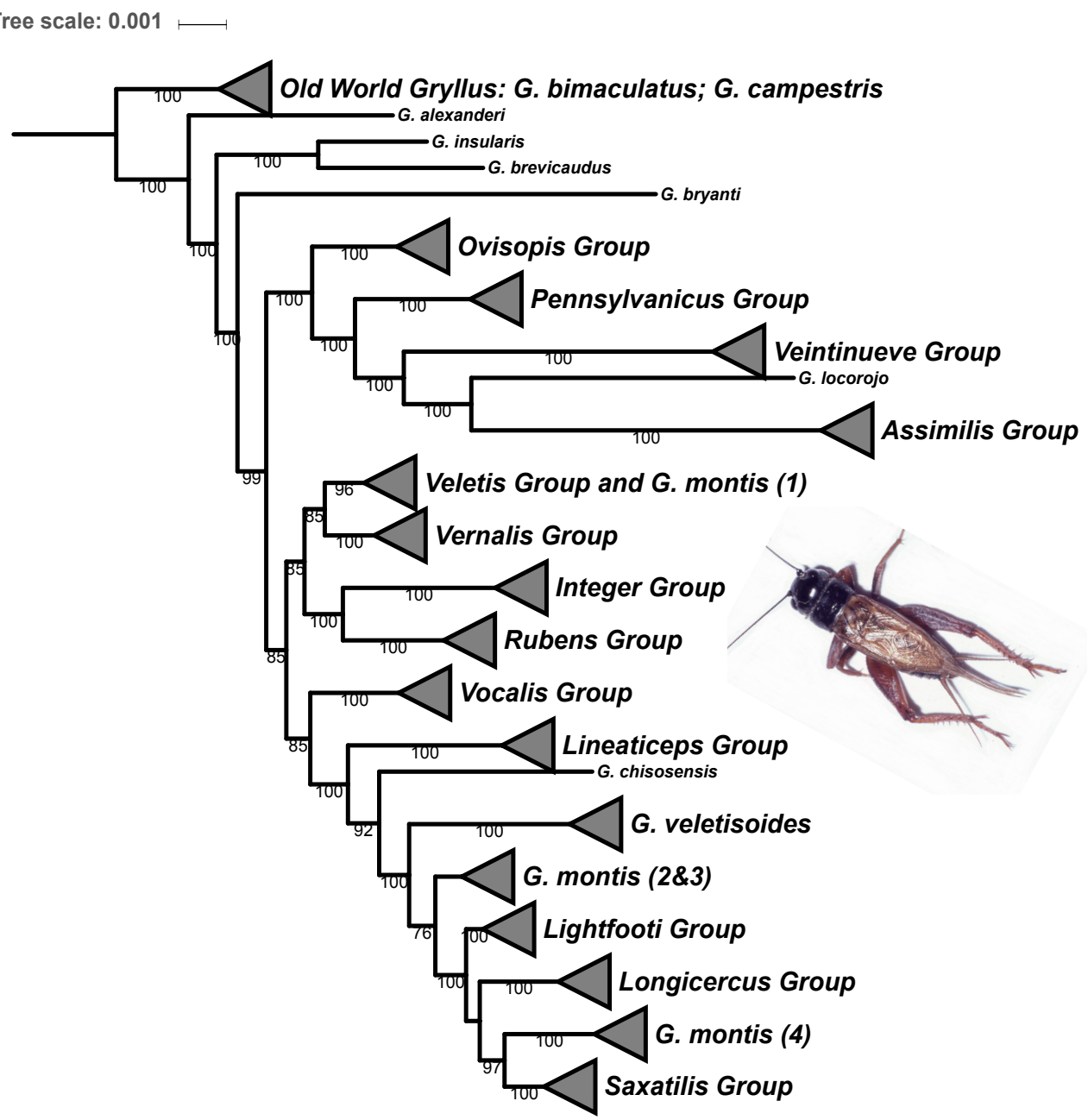

FIG. 5. Overview of in-group taxa relationships based on RAxML concatenated analysis, with bootstrap support values. The named Groups correspond to Weissman & Gray (2019).

The details of the in-group analyses are presented in Figs. 7-10, which connect to each other as indicated by arrows. Terminals and interior branch support values less than 50% are color coded in black; interior branch support values from 50-100% are coded low to high from red to bright green using iTOL (Letunic & Bork 2016). Support

value color codes: ≥ 90% = bright green; ≥ 80% and <90% = green; ≥ 70% and <80% = olive; ≥ 60% and <70% = orange/brown; ≥ 50% and <60% = red; <50% = black. Numerical support values are available in the Nexus format treefiles provided on Dryad (Gray *et al.* 2019).

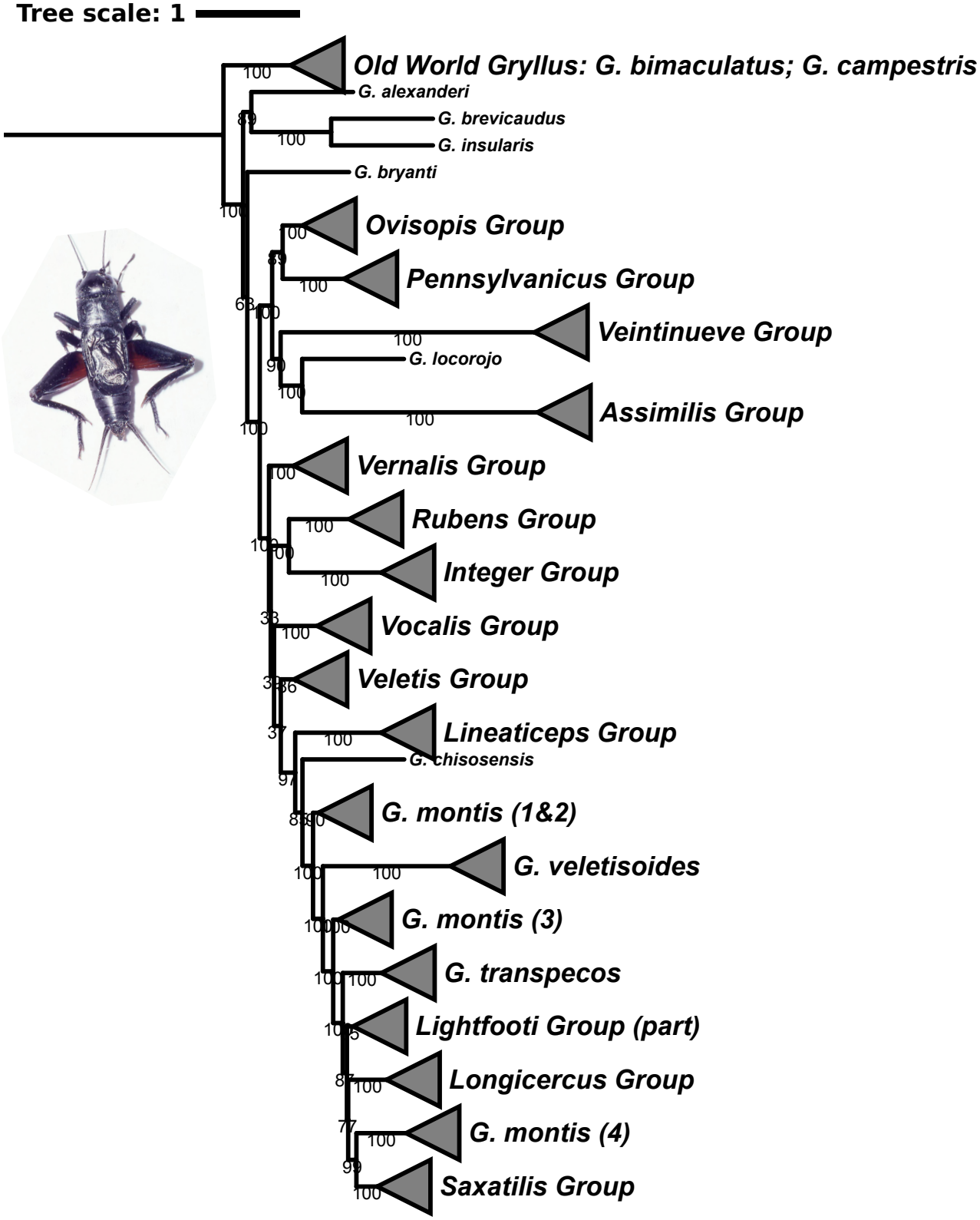

FIG. 6. Overview of in-group taxa relationships based on Astral gene-tree analysis, with quadripartition branch support values. The named Groups correspond to Weissman & Gray (2019).

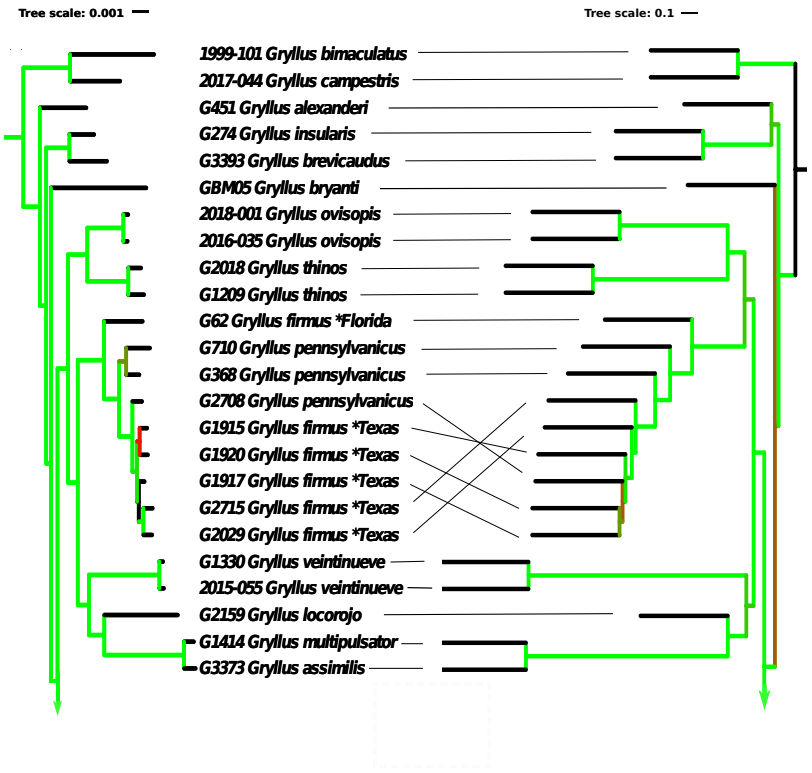

FIG. 7. Phylogeny of *Gryllus* panel 1 of 4. RAxML tree (left) with taxon names; Astral tree (right) with thin lines connecting tips to taxon names. Support value color codes: ≥ 90% = bright green; ≥ 80% and <90% = green; ≥ 70% and <80% = olive; ≥ 60% and <70% = orange/brown; ≥ 50% and <60% = red; <50% = black.

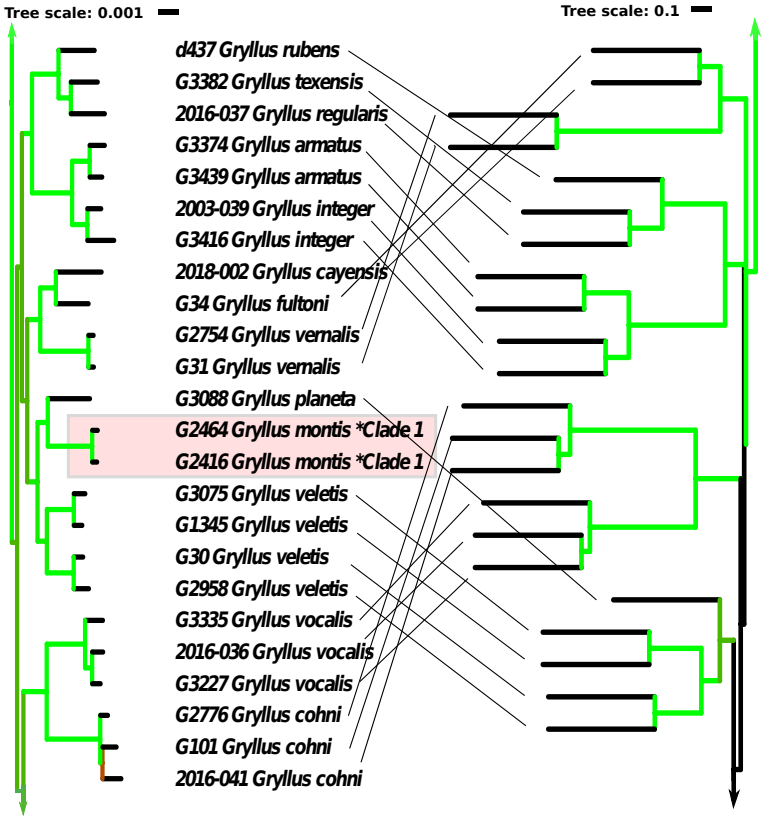

FIG. 8. Phylogeny of *Gryllus* panel 2 of 4. RAxML tree (left) with taxon names; Astral tree (right) with thin lines connecting tips to taxon names. Support value color codes: ≥ 90% = bright green; ≥ 80% and <90% = green; ≥ 70% and <80% = olive; ≥ 60% and <70% = orange/brown; ≥ 50% and <60% = red; <50% = black.

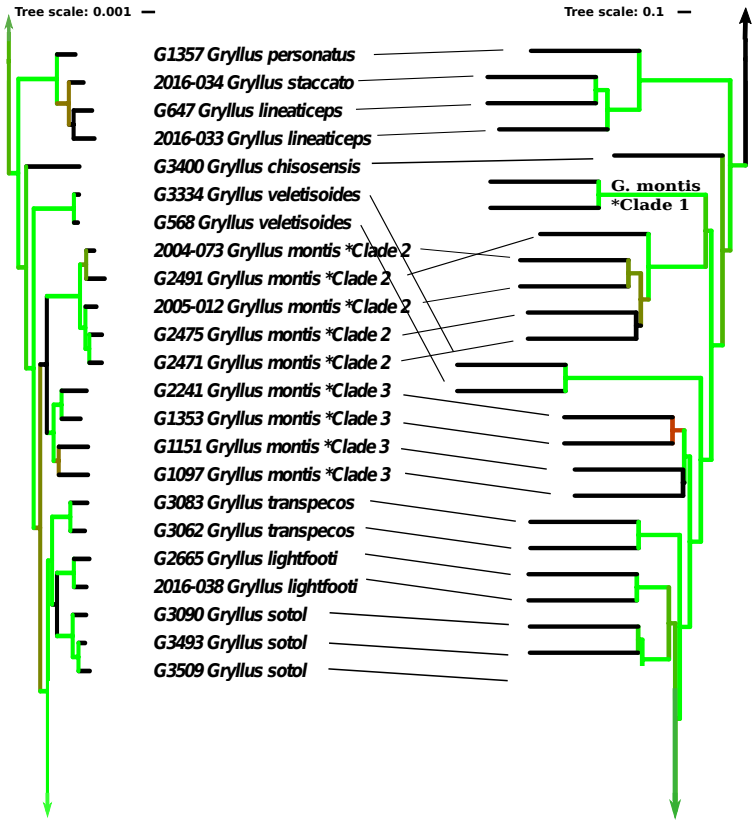

FIG. 9. Phylogeny of *Gryllus* panel 3 of 4. RAxML tree (left) with taxon names; Astral tree (right) with thin lines connecting tips to taxon names. Support value color codes: ≥ 90% = bright green; ≥ 80% and <90% = green; ≥ 70% and <80% = olive; ≥ 60% and <70% = orange/brown; ≥ 50% and <60% = red; <50% = black.

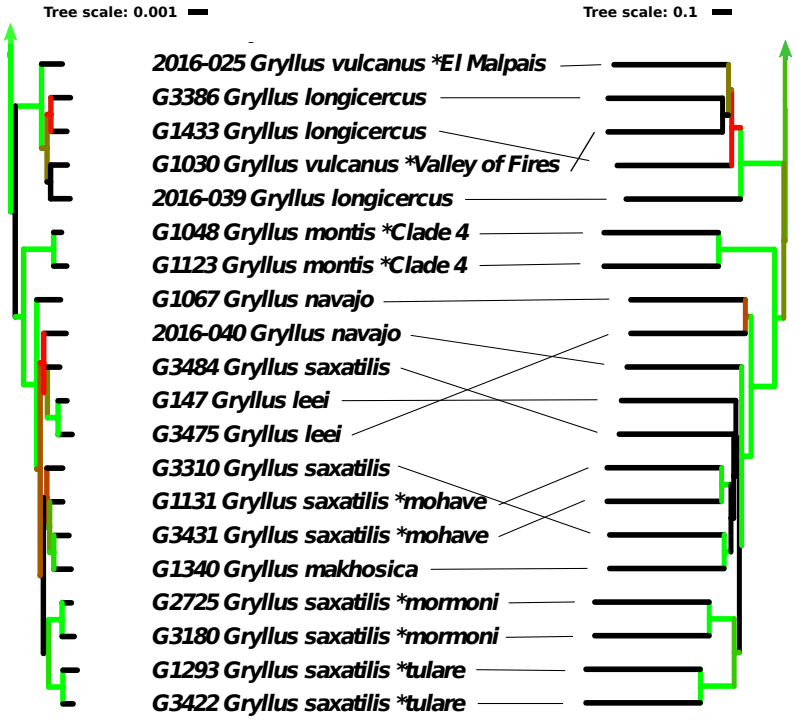

FIG. 10. Phylogeny of *Gryllus* panel 4 of 4. RAxML tree (left) with taxon names; Astral tree (right) with thin lines connecting tips to taxon names. Support value color codes: ≥ 90% = bright green; ≥ 80% and <90% = green; ≥ 70% and <80% = olive; ≥ 60% and <70% = orange/brown; ≥ 50% and <60% = red; <50% = black.

The BEAST coalescent time calibrated tree is presented in Fig. 11.

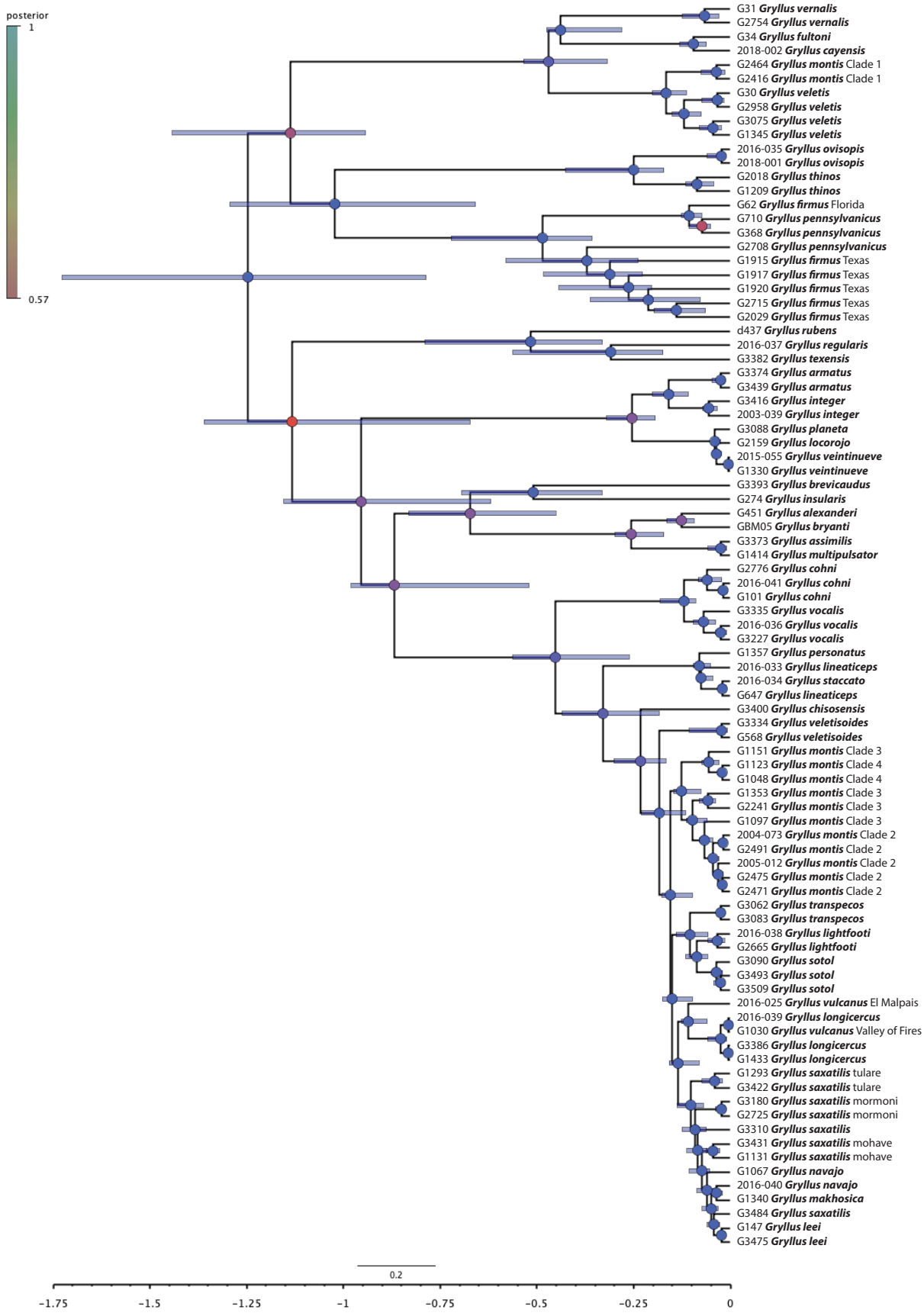

FIG. 11. Time calibrated tree from BEAST.

Finally, we present in Fig. 12 what we consider to be our 'best' working hypothesis of relationships among the described species and promising candidate lineages. It is based on the RAxML analysis of concatenated data, but with only a single terminal per species.

Discussion

Our results provide a valuable framework for future comparative studies of *Gryllus*, however certain caveats are essential. (i) Despite very large amounts of data, certain areas of the overall tree framework, i.e. the relationships among some species Groups, are poorly supported. For example, relationships among the *Veletis, Vocalis*, and (*Rubens* + *Integer*) Groups only have 30-40% support in the Astral analysis. (ii) Placement of Clades 1-3 of *G. montis* differs substantially between RAxML and Astral analyses, particularly so for Clade 1. (iii) The *Longicercus* Group is well supported, but the relationship of the two *G. vulcanus* samples from different lava flows suggests separate transitions from a *G. longicercus*-like ancestor. (iv) Relationships among the named species within the *Saxatilis* Group are poorly recovered. In fact, the unnamed 'tulare' and 'mormoni' lineages are better supported than are the named lineages *G. leei* Weissman & Gray, *G. navajo* Weissman & Gray, and *G. saxatilis*; the unnamed 'mohave' lineage is not well supported and *G. makhosica* Weissman & Gray is impossible to judge given our inclusion of only one sample. Given that the different lineages within the *Saxatilis* Group appear to be both (i) recently diverged (Fig. 11) and (ii) separated principally by geography and substrate type, with minimal song divergence (Weissman & Gray 2019; Talavera *et al.* in prep.), prezygotic reproductive isolation may be too weak to prevent intermittent gene flow, especially given that *G. saxatilis* is geographically widespread and some individuals are flight capable.

The coalescent time tree (Fig. 11) recovers mostly the same Groups of closely related sister-taxa as both the RAxML and Astral analyses. The exceptions are that *G. planeta* Weissman & Gray would be separated from the Veletis Group and *G. montis* Clade 4 would be placed with *G. montis* Clades 2 and 3, rather than with the Saxatilis Group. Tree topography differs substantially for the deeper nodes, with generally poor support in the BEAST tree. The time tree does suggest rapid recent diversification at the species level, with most species level diversity generated within the last 1 mya. Given estimates of rapid speciation in *Laupala* crickets (Mendelson & Shaw 2005) and Gorochov's (2019) conclusion that *Gryllus* in North America are recent and morphologically rather uniform, this might appear reasonable, however we caution that (i) such estimates are hard to interpret, especially as the calibration point priors used here are widely variable in the nature of data (e.g. molecular studies v. geological minimum ages of lava) and (ii) that our data do not represent a random sampling of genomic diversity (as do the AFLPs of Mendelson & Shaw) but rather are by their nature mostly highly conserved sequence data from exons with most of the recent phylogenetic signal coming from introns and intergenic regions. Thus, although we do include the coalescent time tree here, we urge caution and expect that future comparative genomic research with closely related sister species will greatly help refine and clarify the timing of this radiation.

Consistent with previous analyses, the egg-diapausing species *G. ovisopis*, *G. pennsylvanicus*, and *G. firmus* group together. *G. thinos* Weissman & Gray, however, lacks an egg diapause but is nonetheless placed within this group, suggesting trait reversal in the immediate ancestor of *G. thinos*. *G. firmus*, as found in Texas, appears likely to be a separate lineage from the type *G. firmus* as found in Florida and along the Atlantic seaboard. We also note that the *G. pennsylvanicus* sample from Texas [G2708, from Lubbock Stop#13-80] appears closer to *G. firmus* Texas samples than to the *G. pennsylvanicus* samples from Vermont (G710) and Utah (G368), dramatically so in both the Astral and Beast analyses. The eastern *G. firmus*-*G. pennsylvanicus* species pair have been the subject of extensive evolutionary genetic analyses (Larson *et al.* 2013, 2014; Maroja *et al.* 2009, 2015); clearly the situation described here calls for future work along the boundary between *G. pennsylvanicus* and what is currently called *G. firmus* in Texas (see Weissman & Gray 2019 p. 61 for further discussion).

The situation with *G. montis* also clearly needs future work. In discussion of these lineages, Weissman & Gray (2019) suggested that a history of hybridization may have complicated our understanding of these sky-island taxa. In that paper they note that *G. montis* Clade 1 has mtDNA like *G. veletis*, suggesting introgression. In all analyses except the BEAST time-tree, *G. montis* Clade 4 is associated with the *Saxatilis* Group, all of which have nymphs with distinct transverse stripes; based on collections in August 2019, we now know that *G. montis* Clade 4 nymphs are likewise striped, supporting this association. Because Astral analyses represent concordance among gene trees, the differences between RAxML and Astral trees might simply reflect a history of introgression as suggested by the mitochondrial-nuclear discordance discussed by Weissman & Gray (2019).

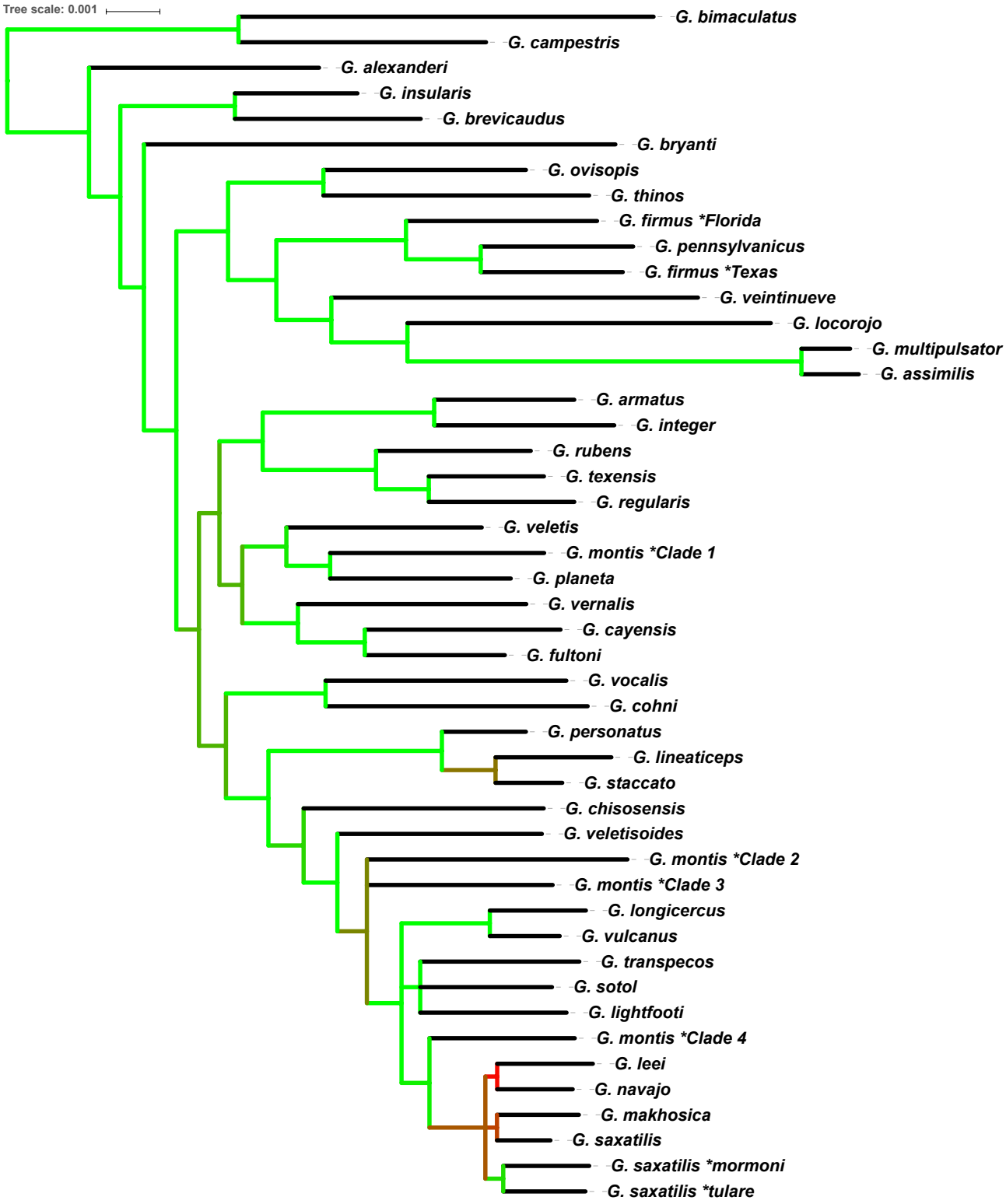

FIG. 12. Our current hypothesis of species relationships among the named *Gryllus* species as well as the candidate lineages (denoted with an *) of putative species based on the RAxML analysis of concatenated data. Color coding indicates levels of bootstrap support from 50% (red) – 100% (bright green); branches with <50% bootstrap support collapsed.

Given that this is the first within-genus species level phylogeny of insects generated by the Anchored Hybrid Enrichment (AHE) approach it is relevant to discuss strengths and weaknesses, as well as resulting levels of tree support and discordance between the RAxML and Astral analyses. The premise of the AHE approach is to relatively easily generate large amounts of data for phylogeny reconstruction at both deep and shallow scales. This is because

the highly conserved target regions ought to recover deep nodes while the more variable introns and intergenic regions provide resolution at shallow nodes. Indeed this approach has been successful in a variety of taxa including invertebrates, e.g. spiders (Hamilton *et al.* 2016) and Cerambycid beetles (Haddad *et al*. 2018). One of the limitations of the approach as applied here was the lack of genomic resources for probe design in crickets. The cricket genome is ca. 2 GB and no published reference genome is available. This caused us to rely on a combination of published transcriptomic resources as well as *de novo* whole genome sequencing of four widely divergent orthopterans (Table 2). Had we used four Gryllinae instead, we may have achieved better resolution within *Gryllus*.

TABLE 2. Summary table of genomic resources.

Type	SampleID	Source File	Citation	Taxon
Transcriptome	|19660	tr2.fasta	Berdan *et al.* 2016	*Gryllus rubens* (Gryllinae)
Transcriptome	|19661	GAWZ01.1.fasta	Misof *et al.* 2014	*Gryllotalpa* sp. (Gryllotalpidae)
Transcriptome	|19662	GAUX01.1.fsa_nt	Misof *et al.* 2014	*Ceuthophilus* sp. (Rhaphidophoridae)
WGS: 180 million reads; 36.0 Gb	|10936	N/A	This study	*Gryllus longicercus* (Gryllinae)
WGS: 124 million reads; 24.8 Gb	|16303	N/A	This study	*Hoplosphyrum boreale* (Mogoplistinae)
WGS: 112 million reads; 22.4 Gb	|16307	N/A	This study	*Stenopelmatus piceiventris* (Stenopelmatidae)
WGS: 96 million reads; 19.2 Gb	|16308	N/A	This study	*Neduba* sp. (Tettigoniidae)

It is worth questioning, however, what level of tree support to expect. With phylogenomic data, bootstrap support values of 100% are commonplace, reflecting the massive amounts of data used and the fact that the taxon sets analyzed are typically more divergent (i.e. genera within families, etc.) than was the case here. Such an effect can be seen in our data: support for deeper nodes is high, typically 100%; this is true for outgroup/ingroup structure by both RAxML and Astral analyses (Fig. 4) and for the deeper nodes and our species Groups particularly with the RAxML analysis (Fig. 5). Where support values fall off somewhat is for the species relationships within our named Groups, although support is still quite strong, i.e. >90%, for most branches (Figs. 7-10, in bright green). What is revealing is that the support levels for very closely related species, e.g. the *Saxatilis* Group, are poor and for certain groups especially there is substantial discordance between the RAxML and the Astral analyses.

It is perhaps tempting to view these lower support levels and gene-tree conflicts as indicative of a lack of strong phylogenetic signal due to limitations of the probe design discussed above; researchers familiar with phylogenomics in general, but unfamiliar with *Gryllus*, may well favor this interpretation. However, to us, having worked on the systematics and behavior of *Gryllus* for several decades, such discordance makes biologically meaningful sense. (1) Most *Gryllus* species are either fully or occasionally flight capable (Weissman & Gray 2019), which could facilitate gene flow among distant groups. (2) Many if not most *Gryllus* species have large population sizes, which reduces the rate of lineage sorting due to drift. (3) In laboratory settings at least, several sets of sister taxa will mate and can produce viable offspring (Izzo & Gray 2011; Gray 2005; Larson *et al.* 2012; Veen *et al.* 2013; Gray *et al.* 2016b), which suggests the potential for introgression in nature. (4) Both nuclear-mitochondrial discordance (see Weissman & Gray 2019) and genomic tests for introgression (i.e. ABBA-BABA, Durand *et al.* 2011) among related *Gryllus* taxa often suggest past introgression (DA Gray, unpublished). In other words, our understanding of the reproductive biology of many *Gryllus* suggests that the lower support values, especially from the Astral gene tree quartets, is likely to be a true reflection of a mosaic of reticulation across the genomes of a number of taxa.

Despite these caveats and calls for future work, we hope that our hypothesis of species/lineage relationships among these difficult taxa will nonetheless be useful to future evolutionary study of the group.

Acknowledgments

The following colleagues provided samples, advice and/or assistance: T.J. Walker, K.A. Judge, Y. Jang, D.C. Lightfoot, T. Tregenza, W.H. Cade, P.A. De Luca. We thank Michelle Kortyna, Chris Zdyrski, Sean Holland and Jesse Cherry in the Center for Anchored Phylogenomics for assistance with data collection and analysis. We are also grateful to the many Federal and State agencies that provided permission to collect specimens over a number of years.

Data Accessibility

Nexus format treefiles as well as fasta alignment files will be made available in Dryad (Gray *et al.* 2019).

REFERENCES

Alexander, R.D. (1968) Life cycle origins, speciation, and related phenomena in crickets. *Quarterly Review of Biology*, 43, 1–41.
https://doi.org/10.1086/405628

Alexander, R.D. & Bigelow, R.S. (1960) Allochronic speciation in field crickets, and a new species, *Acheta veletis*. *Evolution*, 14, 334–346.
https://doi.org/10.1111/j.1558-5646.1960.tb03095.x

Berdan, E.L., Blankers, T., Waurick, I., Mazzoni, C.J. & Mayer, F. (2016) A genes eye view of ontogeny: de novo assembly and profiling of the *Gryllus rubens* transcriptome. *Molecular Ecology Resources*, 16 (6), 1478–1490.
https://doi.org/10.1111/1755-0998.12530

Blankers, T., Hennig, R.M. & Gray, D.A. (2015) Conservation of multivariate female preference functions and preference mechanisms in three species of trilling field crickets. *Journal of Evolutionary Biology*, 28, 630–641.
https://doi.org/10.1111/jeb.12599

Blankers, T., Gray, D.A. & Hennig, R.M. (2017) Multivariate phenotypic evolution: Divergent acoustic signals and sexual selection in *Gryllus* field crickets. *Evolutionary Biology*, 44, 43–55.
https://doi.org/10.1007/s11692-016-9388-1

Blankers, T., Vilaça, S.T., Waurick, I., Gray, D.A., Hennig, R.M., Mazzoni, C.J., Mayer, F. & Berdan, E.L. (2018) Demography and selection shape transcriptomic divergence in field crickets. *Evolution*, 72, 553–567.
https://doi.org/10.1101/193839

Bouckaert, R., Heled, J., Kühnert, D., Vaughan, T., Wu, C.-H., Xie, D., Suchard, M.A., Rambaut, A. & Drummond, A.J. (2014) BEAST 2: A software platform for Bayesian evolutionary analysis A. Prlic (Ed). *PLoS Computational Biology*, 10, e1003537.
https://doi.org/10.1371/journal.pcbi.1003537

Cade, W.H. & Otte, D. (2000) *Gryllus texensis* n. sp.: a widely studied field cricket (Orthoptera; Gryllidae) from the southern United States. *Transactions of the American Entomological Society (1890-)*, 126 (1), 117–123.

Crone, E.J., Zera, A.J., Anand, A.N., Oakeshott, J.G., Sutherland, T.D., Russell, R.J., Harshman, L.G., Hoffmann F.G. & Claudianos, C. (2007) Jhe in *Gryllus assimilis*: cloning, sequence-activity associations and phylogeny. *Insect Biochemistry and Molecular Biology*, 37, 1359–1365.
https://doi.org/10.1016/j.ibmb.2007.08.005

Desutter-Grandcolas, L. & Robillard, T. (2003) Phylogeny and the evolution of calling songs in *Gryllus* (Insecta, Orthoptera, Gryllidae). *Zoologica Scripta*, 32, 173–183.
https://doi.org/10.1046/j.1463-6409.2003.00107.x

Drummond, A.J. & Bouckaert, R.R. (2014) *Bayesian evolutionary analysis with BEAST 2*. Cambridge University Press, Cambridge, 260 pp.
https://doi.org/10.1017/CBO9781139095112

Dunbar, N.W. (1999) Cosmogenic ^{36}Cl-determined age of the Carrizozo lava flows, south-central New Mexico. *New Mexico Geology*, 21, 25–29.

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.
https://doi.org/10.1093/molbev/msr048

Gray, D.A. (2005) Does courtship behavior contribute to species-level reproductive isolation in field crickets? *Behavioral Ecology*, 16 (1), 201–206.
https://doi.org/10.1093/beheco/arh144

Gray, D.A. & Cade, W.H. (2000) Sexual selection and speciation in field crickets. *Proceedings National Academy of Sciences*

U S A, 97, 14449–14454.
https://doi.org/10.1073/pnas.97.26.14449
Gray, D.A., Gabel, E., Blankers, T. & Hennig, R.M. (2016a) Multivariate female preference tests reveal latent perceptual biases. *Proceedings Royal Society, B*, 283 (1842). [published online]
https://doi.org/10.1098/rspb.2016.1972
Gray, D.A., Gutierrez, N.J., Chen, T.L., Gonzalez, C., Weissman, D.B. & Cole, J.A. (2016b) Species divergence in field crickets: genetics, song, ecomorphology, and pre- and postzygotic isolation. *Biological Journal of the Linnaean Society*, 117, 192–205.
https://doi.org/10.1111/bij.12668
Gray, D.A., Weissman, D.B., Cole, J.A., Lemmon, E.M. & Lemmon, A.R. (2019) Multilocus phylogeny of *Gryllus* field crickets (Orthoptera: Gryllidae: Gryllinae) utilizing anchored hybrid enrichment. *Dryad*, Dataset. [in press]
Gonzalez-Garcia, J.J., Prawirodirdjo, L., Bock, Y. & Agnew, D. (2003) Guadalupe Island, Mexico as a new constraint for Pacific plate motion. *Geophysical Research Letters*, 30 (16), 1872.
https://doi.org/10.1029/2003GL017732
Gorochov, A.V. (2019) Preliminary hypothesis on the history of Gryllinae (Orthoptera: Gryllidae) in America and new taxa of both the subtribe Anurogryllina and the genus *Megalogryllus*. *Amurian Zoological Journal*, 11 (4), 279–308.
https://doi.org/10.33910/2686-9519-2019-11-4-279-308
Haddad S., Shin, S., Lemmon, A.R., Lemmon, E.M., Svacha, P., Farrell, B., Sìlipinìski, A., Windsor, D. & McKenna, D.D. (2018) Anchored hybrid enrichment provides new insights into the phylogeny and evolution of longhorned beetles (Cerambycidae). *Systematic Entomology*, 43, 68–89.
https://doi.org/10.1111/syen.12257
Hamilton, C.A., Lemmon, A.R., Moriarty Lemmon, E. & Bond, J.E. (2016) Expanding Anchored Hybrid Enrichment to resolve both deep and shallow relationships within the spider Tree of Life. *BMC Evolutionary Biology*, 16 (1), 212
https://doi.org/10.1186/s12862-016-0769-y
Harrison, R.G. (1983) Barriers to gene exchange between closely related cricket species. I. Laboratory hybridization studies. *Evolution*, 37, 245–251.
https://doi.org/10.1111/j.1558-5646.1983.tb05534.x
Harrison, R.G. (1985) Barriers to gene exchange between closely related cricket species. II. Life cycle variation and temporal isolation. *Evolution*, 39, 244–259.
https://doi.org/10.2307/2408360
Harrison, R.G. & Bogdanowicz, S.M. (1995) Mitochondrial DNA phylogeny of North American field crickets: perspectives on the evolution of life cycles, songs, and habitat associations. *Journal of Evolutionary Biology*, 8, 209–232.
https://doi.org/10.1046/j.1420-9101.1995.8020209.x
Huang, Y., Ortí, G., Sutherlin, M., Duhachek, A. & Zera, A. (2000) Phylogenetic relationships of North American field crickets inferred from mitochondrial DNA data. *Molecular Phylogenetics and Evolution*, 17, 48–57.
https://doi.org/10.1006/mpev.2000.0815
Izzo, A.S. & Gray, D.A. (2011) Heterospecific courtship and sequential mate choice in sister species of field crickets. *Animal Behaviour*, 81 (1), 259–264.
https://doi.org/10.1016/j.anbehav.2010.10.015
Jang, Y., Gerhardt, H.C. & Choe, J.C. (2008) A comparative study of aggressiveness in eastern North American field cricket species (genus *Gryllus*). *Behavioral Ecology and Sociobiology*, 62, 1397–1407.
https://doi.org/10.1007/s00265-008-0568-6
Jang, Y., Won, Y.-J. & Choe, J.C. (2009) Convergent and divergent patterns of morphological differentiation provide more evidence for reproductive character displacement in a wood cricket *Gryllus fultoni* (Orthoptera: Gryllidae). *BMC Evolutionary Biology*, 9, 27.
https://doi.org/10.1186/1471-2148-9-27
Katoh, K. & Standley, D.M. (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution*, 30 (4), 772–780.
https://doi.org/10.1093/molbev/mst010
Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., Thierer, T., Ashton, B., Meintjes, P. & Drummond, A. (2012) Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*, 28 (12), 1647–1649.
https://doi.org/10.1093/bioinformatics/bts199
Lanfear, R., Calcott, B., Ho, S.Y.W. & Guindon, S. (2012) PartitionFinder: combined selection of partitioning schemes and substitution models for phylogenetic analyses. *Molecular Biology and Evolution*, 29, 1695–1701.
https://doi.org/10.1093/molbev/mss020
Lanfear, R., Frandsen, P.B., Wright, A.M., Senfeld, T. & Calcott, B. (2016) PartitionFinder2: new methods for selecting partitioned models of evolution for molecular and morphological phylogenetic analysis. *Molecular Biology and Evolution*, 34, 772–773.
https://doi.org/10.1093/molbev/msw260
Larson, E.L., Hume, G.L., Andrés, J.A. & Harrison, R.G. (2012) Post-mating prezygotic barriers to gene exchange between

hybridizing field crickets. *Journal of Evolutionary Biology*, 25 (1), 174–186.
https://doi.org/10.1111/j.1420-9101.2011.02415.x
Larson, E.L., Becker, C.G., Bondra, E.R. & Harrison, R.G. (2013) Structure of a mosaic hybrid zone between the field crickets *Gryllus firmus* and *G. pennsylvanicus*. *Ecology and Evolution*, 3 (4), 985–1002.
https://doi.org/10.1002/ece3.514
Larson, E.L., White, T.A., Ross, C.L. & Harrison, R.G. (2014) Gene flow and the maintenance of species boundaries. *Molecular Ecology*, 23 (7), 1668–1678.
https://doi.org/10.1111/mec.12601
Lemmon, A.R. & Lemmon, E.M. (2012) High-throughput identification of informative nuclear loci for shallow-scale phylogenetics and phylogeography. *Systematic Biology*, 61, 745–761.
https://doi.org/10.1093/sysbio/sys051
Lemmon, E.M. & Lemmon, A.R. (2013) High-throughput genomic data in systematics and phylogenetics. *Annual Review of Ecology, Evolution, and Systematics*, 44, 99–121.
https://doi.org/10.1146/annurev-ecolsys-110512-135822
Lemmon, A.R., Emme, S.A. & Lemmon, E.M. (2012) Anchored hybrid enrichment for massively high-throughput phylogenomics. *Systematic Biology*, 61, 727–744.
https://doi.org/10.1093/sysbio/sys049
Letunic, I. & Bork, P. (2016) Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Research*, 44, W242–W245.
https://doi.org/10.1093/nar/gkw290
Maroja, L.S., Andrés, J.A. & Harrison, R.G. (2009) Genealogical discordance and patterns of introgression and selection across a cricket hybrid zone. *Evolution*, 63, 2999–3015.
https://doi.org/10.1111/j.1558-5646.2009.00767.x
Maroja, L.S., Larson, E.L., Bogdanowicz, S.M. & Harrison, R.G. (2015) Genes with restricted introgression in a field cricket (*Gryllus firmus*/*Gryllus pennsylvanicus*) hybrid zone are concentrated on the X chromosome and a single autosome. *G3: Genes, Genomes, Genetics*, 5 (11), 2219–2227.
https://doi.org/10.1534/g3.115.021246
Mendelson, T.C. & Shaw, K.L. (2005). Rapid speciation in an arthropod. *Nature*, 433 (7024), 375–376.
https://doi.org/10.1038/433375a
Miller, M., Pfeiffer, W. & Schwartz, T. (2010) Creating the CIPRES Science Gateway for inference of large phylogenetic trees. *In*: *Gateway Computing Environments Workshop (GCE), 2010*. IEEE, New Orleans, Louisiana, pp. 1–8.
Mirarab, S. & Warnow, T. (2015) ASTRAL-II: coalescent-based species tree estimation with many hundreds of taxa and thousands of genes. *Bioinformatics*, 31 (12), i44–i52.
https://doi.org/10.1093/bioinformatics/btv234
Misof, B., Liu, S., Meusemann, K., Peters, R.S., Donath, A., Mayer, C., Frandsen, P.B., Ware, J., Flouri, T., Beutel, R.G., Niehuis, O., Petersen, M., Izquierdo-Carrasco, F., Wappler, T., Rust, J., Aberer, A.J., Aspöck, U., Aspöck, H., Bartel, D., Blanke, A., Berger, S., Böhm, A., Buckley, T.R., Calcott, B., Chen, J., Friedrich, F., Fukui, M., Fujita, M., Greve, C., Grobe, P., Gu, S., Huang, Y., Jermiin, L.S., Kawahara, A.Y., Krogmann, L., Kubiak, M., Lanfear, R., Letsch, H., Li, Y., Li, Z., Li, J., Lu, H., Machida, R., Mashimo, Y., Kapli, P., McKenna, D.D., Meng, G., Nakagaki, Y., Navarrete-Heredia, J.L., Ott, M., Ou, Y., Pass, G., Podsiadlowski, L., Pohl, H., Von Reumont, B.M., Schütte, K., Sekiya, K., Shimizu, S., Slipinski, A., Stamatakis, A., Song, W., Su, X., Szucsich, N.U., Tan, M., Tan, X., Tang, M., Tang, J., Timelthaler, G., Tomizuka, S., Trautwein, M., Tong, X., Uchifune, T., Walzl, M.G., Wiegmann, B.M., Wilbrandt, J., Wipfler, B., Wong, T.K.F., Wu, Q., Wu, G., Xie, Y., Yang, S., Yang, Q., Yeates, D.K., Yoshizawa, K., Zhang, Q., Zhang, R., Zhang, W., Zhang, Y., Zhao, J., Zhou, C., Zhou, L., Ziesmann, T., Zou, S., Li, Y., Xu, X., Zhang, Y., Yang, H., Wang, J., Wang, J., Kjer, K.M. & Zhou, X. (2014) Phylogenomics resolves the timing and pattern of insect evolution. *Science*, 346 (6210), 763–767.
https://doi.org/10.1126/science.1257570
Prum, R.O., Berv, J.S., Dornburg, A., Field, D.J., Townsend, J.P., Lemmon, E.M. & Lemmon, A.R. (2015) A fully resolved, comprehensive phylogeny of birds (Aves) using targeted next generation DNA sequencing. *Nature*, 526, 569–573.
https://doi.org/10.1038/nature15697
Rambaut, A. (2009) FigTree. Version 1.3.1. Available from: http://tree.bio.ed.ac.uk/software/figtree/ (accessed 12 January 2020)
Rambaut, A. & Drummond, A.J. (2013) Tracer. Version 1.6. Available from: http://tree.bio.ed.ac.uk.software/tracer/ (accessed 12 January 2020)
Rambaut, A. & Drummond, A.J. (2014) TreeAnnotator v. 2.1.2. Available from: https://code.google.com/p/beast-mcmc/ (accessed 12 January 2020)
Rehn, J.A.G. & Hebard, M. (1915) The genus *Gryllus* (Orthoptera) as found in America. *Proceedings of the Academy of Natural Sciences Philadelphia*, 67, 293–322.
Rokyta, D.R., Lemmon, A.R., Margres, M.J. & Arnow, K. (2012) The venom-gland transcriptome of the eastern diamondback rattlesnake (*Crotalus adamanteus*). *BMC Genomics*, 13, 312.
https://doi.org/10.1186/1471-2164-13-312
Scudder, S.H. (1901) The species of *Gryllus* on the Pacific coast. *Psyche*, 9, 267–270.

https://doi.org/10.1155/1901/239671
Scudder, S.H. (1902) The species of *Gryllus* found in the United States east of the Sierra Nevadas. *Psyche*, 9, 291–296. https://doi.org/10.1155/1902/35962
Stamatakis, A. (2006) RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics*, 22 (21), 2688–2690. https://doi.org/10.1093/bioinformatics/btl446
Veen, T., Faulks, J., Tyler, F., Lloyd, J. & Tregenza, T. (2013) Diverse reproductive barriers in hybridising crickets suggests extensive variation in the evolution and maintenance of isolation. *Evolutionary Ecology*, 27, 993–1015. https://doi.org/10.1007/s10682-012-9610-2
Weissman, D.B. & Gray, D.A. (2019) Crickets of the genus *Gryllus* in the United States (Orthoptera: Gryllidae: Gryllinae). *Zootaxa*, 4705 (1), 001–277. https://doi.org/10.11646/zootaxa.4705.1.1
Weissman, D.B., Walker, T.J. & Gray, D.A. (2009) The Jamaican field cricket *Gryllus assimilis* and two new sister species (Orthoptera: Gryllidae). *Annals of the Entomological Society of America*, 102, 367–380. https://doi.org/10.1603/008.102.0304