Genomes of skipper butterflies reveal extensive convergence of wing patterns

Wenlin Li,a,b Qian Cong,a,b,c Jinhui Shen,a,b Jing Zhang,a,b Winnie Hallwachs,d Daniel H. Janzen,e and Nick V. Grishin,a,b,d,2

*Department of Biophysics, University of Texas Southwestern Medical Center, Dallas, TX 75390-9050; †Department of Biochemistry, University of Texas Southwestern Medical Center, Dallas, TX 75390-9050; ‡Department of Biology, University of Pennsylvania, Philadelphia, PA 19104; and §Howard Hughes Medical Institute, University of Texas Southwestern Medical Center, Dallas, TX 75390-9050

Contributed by Daniel H. Janzen, January 4, 2019 (sent for review December 17, 2018; reviewed by John Shuey and James Walters)

For centuries, biologists have used phenotypes to infer evolution. For decades, a handful of gene markers have given us a glimpse of the genotype to combine with phenotypic traits. Today, we can sequence entire genomes from hundreds of species and gain yet closer scrutiny. To illustrate the power of genomics, we have chosen skipper butterflies (Hesperiidae). The genomes of 250 representative species of skippers reveal rampant inconsistencies between their current classification and a genome-based phylogeny. We use a dated genomic tree to define tribes (six new) and subtribes (six new), to overhaul genera (nine new) and subgenera (three new), and to display convergence in wing patterns that fooled researchers for decades. We find that many skippers with similar appearance are distantly related, and several skippers with distinct morphology are close relatives. These conclusions are strongly supported by different genomic regions and are consistent with some morphological traits. Our work is a forerunner to genomic biology shaping biodiversity research.

Significance

With decreasing cost, biodiversity genomics can be done by small laboratories, not large centers. Here, we overview the genomic landscape of an entire family of animals. By sequencing 250 representatives of skipper butterflies, we infer their phylogeny and search for genotypic determinants of phenotypic traits. We find that wing patterns are frequently convergent. This likely mimetic convergence is diversified, resulting in five distinct parallel patterns. Each of the five patterns occurs within at least two genera as well as in more distant relatives diverged more than 20 Mya. This work offers an example of how each of the nearly 6,000 families of animals will be studied in the future.


The authors declare no conflict of interest.

Open Access

This article is distributed under the Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 (CC BY-NC-ND).

Published online March 15, 2019.

PNAS is a product of the National Academy of Sciences.
Fig. 1. Time-calibrated genomic tree of Hesperiidae constructed from nuclear gene coding regions of 250 species. The tree is cut in half: First and second segments are shown above and below. Colored horizontal lines delineate major taxonomic groups: red, subfamilies; orange (broken in several places), tribes; yellow shaded red (Eudaminae only), subtribes; and green shaded yellow (Eudaminae only), genera. New taxa are highlighted in yellow. New genera are in red font highlighted in yellow, and those used in different genus-species combinations than previously are in red font. Taxa placed in a different tribe or subtribe than previously are in blue font highlighted in green. Segments of the tree corresponding to different tribes are highlighted in different colors. Illustrated Hesperiidae species are associated with names in the tree by numbers in brown font. See Si Appendix for details.
subfamily of its own, sister to all other Hesperiidae except Coeliadinae. The next split in the genomic tree is different from that appearing in a number of DNA studies, but in agreement with the morphological view. The Spreading skippers (mostly dicot feeders) and the Grass skippers (mostly monocot feeders) are sisters. Both Spreading and Grass skippers have been previously divided into several subfamilies. genomic data very strongly support (bootstrap above 99%) monophyly of these subfamilies. The two latest diverging subfamilies (Hesperiinae and Trapezitinae) have split off by about 50 Mya (Fig. 1). Pyrginae, as defined by Warren et al. (3), diversified from their common ancestor before that time (>55 Mya). Moreover, while other subfamilies except Pyrginae (i.e., Eudaminae, Heteropterinae, and Hesperiinae) are well separated from each other, Pyrginae sensu Warren have split into three compact groups, within a short time and before divergence of Grass skippers into subfamilies. Thus, these three groups within Warren’s Pyrginae are no less distinct than Hesperiidae subfamilies, and we treat them as such. Accordingly, we unify the sister tribes Tagiadiini and Celaenorrhinini in a subfamily Tagiadiinae, and restate the subfamily Pyrrhopyginae. The latter are strikingly distinct in appearance from other skippers, and have been traditionally considered a subfamily. They diverged from their common ancestor with Pyrginae by about 55 Mya, before the divergence of Grass skippers. Thus, genomic data suggest that the family Hesperiidae consists of nine subfamilies (Fig. 1 and SI Appendix), and diversification into subfamilies occurred about 40 Mya.

Mitogenomes and Cytochrome Oxidase I Barcodes. In addition to a nuclear genome tree, we constructed a tree from mitogenomes. The resulting tree recapitulates major phylogenetic groups of the nuclear genome tree, but with weaker support. All of the subfamilies and tribes are composed of the same species in either mitogenome or nuclear genome phylogenies. Next, using mitogenomes of 250 species as a backbone, we increased taxonomic coverage by adding 290 species with cytochrome oxidase I (COI) barcodes only (SI Appendix, Fig. S1). These species grouped with mitogenomes of their expected relatives. Most of these barcode-only species were placed according to their current classification, with several exceptions discussed in SI Appendix. We used this mitogenome + barcode tree together with the nuclear genome tree (Fig. 1) as the basis for our proposed classification of Hesperiidae (SI Appendix, Taxonomic Appendix). Differences in barcodes and morphology suggested that 30 subspecies as defined by Evans (10–12) are more likely to be species, and such cases were analyzed in detail for wing pattern and genital characters (SI Appendix, Taxonomic Appendix).

Eudaminae Tribes and Subtribes. Previous studies refrained from defining tribes in the subfamily Eudaminae. We focused on this subfamily and attempted to delineate tribes consistently with the age of the tribes in Pyrginae (sensu stricto). The four Pyrginae tribes (Carcharodini, Achlyodini, Erynnini, and Pyrgini) outlined by Warren et al. (3) diverged around 42 Mya. Our genomic tree agrees with Warren’s, except that it places Grais, Tosta, Morvinina, Myriniina, Xypsina, Pseudodrophylas, Mimia, Eracon, and Spionoidea in Achlyodini; Corniphallus in Carcharodini; Cito in Erynnini; and Jera in Pyrrhopyginae. Unexpectedly, Emmelus belongs to Eudaminae. Conversely, Cabirus is not Eudaminae (4), but Pyrginae; Achlyodini. All Pyrginae tribes received 100% bootstrap support and are conspicuous groups in the subfamily.

To achieve consistent classification, we cut the dated genomic tree around the time of Pyrginae divergence into tribes and find four Eudaminae phylogenetic groups supported by 100% bootstrap that we define as tribes. Two of the tribes that form best-separated groups are described in Table 1 as Oleidiini and Entheini. The two others are closely related sisters Eudamini and Phociidini that diverged about 40 Mya, and are given a tribal rank due to their morphological distinction.

These four Eudaminae tribes correspond to groups with similar morphology (10, 11). For example, Entheini is largely the “B. Augiades group” of Evans. He defined it by the “divergent” third segment of palpi that stems from the outer edge of the second segment. Inconsistently, Evans included in this group three genera that have the central third segment: Phocides and Hypercypothrix, that do not belong to Entheini, and Cabirus, that is not even in Eudaminae (Fig. 1). Interestingly,

### Table 1. Description of new tribes and subtribes of Hesperiidae

<table>
<thead>
<tr>
<th>New tribe or subtribe</th>
<th>Type genus</th>
<th>Diagnostic characters*</th>
<th>ZooBank ID†</th>
</tr>
</thead>
<tbody>
<tr>
<td>Entheini Grishin, new tribe</td>
<td>Entheus Hübnerr, [1819]</td>
<td>276655.26:A192G; 85.28:C1767; 378.19:G1099C; 374.14:G1169T; keys to B.3a in ref. 10, but exclude B.8</td>
<td>303C1FDD-07C8-4919-900E-EA3D6347E5DD</td>
</tr>
<tr>
<td>Loboclina Grishin, new subtribe</td>
<td>Loboclus Moore, 1884</td>
<td>208.2:G1453; T1465; 18312.8:A619C; G620A; keys to B.4 in ref. 8 or C.5, C.10a, C.15.2 or C.18 in ref. 11</td>
<td>C606FC35-32D3-4655-FA5A-86BC36668BFA</td>
</tr>
<tr>
<td>Cephise Grishin, new subtribe</td>
<td>Cephise Evans, 1952</td>
<td>CO1.bc.a447, CR4G, 7479A; genitalia and palpi as described in ref. 17 (pp. 182–183) for Cephise are diagnostic</td>
<td></td>
</tr>
<tr>
<td>Telemiades Grishin, new subtribe</td>
<td>Telemiades Hübnerr, [1819]</td>
<td>536.149:G1488C; 997.8:B514T; 860.7:A748G; 3001.3:C1773T; keys to B.2, B.3, C.7a (exclude C.7.b), E.6a or E.9 in refs. 10 and 11</td>
<td>4A6E059C-8BF2-4CB4-8651-7E1A0C4C5C91</td>
</tr>
<tr>
<td>Oleidiini Grishin, new tribe</td>
<td>Oleides Hübnerr, [1825]</td>
<td>1139.19:T5652A; 851.8:C423A; G434A; 11945.11:G391A; 65.4:C330A; tuft of scales by anal fold from the base of hindwing, either above or below</td>
<td>CF9C3D29-523A-4D17-B140-9A69CFA98731</td>
</tr>
<tr>
<td>Typhedaria Grishin, new subtribe</td>
<td>Typhedanrus Butler, 1870</td>
<td>1341.12:T2558C; 499.8:G307T; 3466.8:T2304A; C2309G, A2505C; tuft of scales by anal fold from the base of hindwing, either above or below</td>
<td>D356F893-674F-869C-133D9BBFBD58</td>
</tr>
<tr>
<td>Netrocoryinei, new tribe</td>
<td>Netrocoryne C. &amp; R. Felder, [1867]</td>
<td>22B.20:399A; 804.14:T439G; 275215.7:C925G; 998.8:G308A; keys to B.2, C.3, B.7a (exclude C.7.b), E.6a or E.9 in refs. 10 and 11</td>
<td>DE6F1F08-03CF-4F8E-9592-D18A46188ABD</td>
</tr>
<tr>
<td>Jeri Grishin, new tribe</td>
<td>Jeri Lindsay, 1925</td>
<td>103.23:TS964; 420.27:G308A; 673.5:R355; 425.5:G1558T; keys to E.5 in ref. 11; forewing cell 3/5 of costa</td>
<td>AF3B5CEA-80B0-4CE2-AD40-ED87C93C400</td>
</tr>
<tr>
<td>Pythonidina Grishin, new tribe</td>
<td>Pythonides Hübnerr, [1819]</td>
<td>274.29:G397A; 3478.6:T116C; 7895.5:G916A; 925.10:G199C; in ref. 11, keys to C.4, E.4, G.9.1 or, if unicus undivided, then C.37a or 40d</td>
<td>CB690271-5483-4B5A-A73C-27BDC5E23DE5</td>
</tr>
<tr>
<td>Clítina Grishin, new subtribe</td>
<td>Clítina Evans, 1953</td>
<td>CO1.bc.B29; 81A, 169A, 266A, 302T, A353T, A321T; keys to E.52 or E.13.8 in ref. 11</td>
<td>971885ED-EF97-4EA3-8112-657279BA4778</td>
</tr>
<tr>
<td>Bitterlini Grishin, new tribe</td>
<td>Bitterlinia Kirby, 1871</td>
<td>2627.8A:14597; 141.4:C104A; 33738.38:G1337, G134C; keys to K.4.35 in ref. 12</td>
<td>D621FE81-AD45-8940-ED04159BD7A</td>
</tr>
<tr>
<td>Pichiranhini, new tribe</td>
<td>Pichirani Hübnerr, [1819]</td>
<td>596.8:C1601G; 144.41:G201C; 83.15:G8658A, T8657G; keys to K.27a in ref. 12</td>
<td>94668BD2-79F3-4EB5-80B1-7F5ACBC56511</td>
</tr>
</tbody>
</table>

*See SI Appendix for the lists of genera included in each taxon and sequences of protein-coding regions with diagnostic characters.
†ZooBank registration URL given for each taxon should be preceded by http://www.zoobank.org/.
Phareas that has a divergent third segment (and was included in the Augiades group) is not in Entheini. Its males possess, absent in Entheini, tufts of hair-like scales in the groove along the hindwing vein 1A+2A. Our tree places this unusual skipper in Phochidini, implying that its peculiar palp is convergent.

Oileidini is sister to other Eudaminae (Fig. 1). Genera in this tribe were grouped with some Pyrginae genera by Evans (11), and the tribe may be intermediate in morphology between Pyrginae and Eudaminae. This is the smallest tribe (six genera) and is characterized by tufts of hair-like scales in the groove along the hindwing vein 1A+2A in males, either below (Oileides) or above (the others). Similar structures in Phareas (Phochidini) are found on both sides of the hindwing.

The sister tribes Eudamini and Phochidini are separated from each other by a short branch and could be one tribe. However, each is strongly monophyletic, and Phochidini skippers stand out morphologically and ecologically: Their forewing veins R4 and R5 originate near each other, hindwing tornus is usually expanded (not lobed), skippers hold wings spread flat when resting, many are crepuscular, and many species are sexually dimorphic. Our Phochidini is the “D. Celaenorrhinus group” of Evans (10) after adding Phochides, Phareas, and Emmelus and removing Cephise, which has lobed or tailed hindwing tornus, and Celaenorrhinus, whose males have a tuft of long scales on the hind tibiae that fits into a thoracic pouch (not found in Eudamini). The genomic tree suggests that Oileides is polyphyletic: one species together with Aurina belong to Phochidini.

Eudamini is the largest and most diverse tribe. It encompasses more than half of the subfamily. The genomic tree reveals groupings within the tribes that are described here as subtribes (Table 1). One of them, subtribe Cephisina, is monotypic for the genus Phocides, which diverged from its sister tribe Teleminadia 35 Mya and is unique in its morphological features (17). Teleminadia includes three genera: Teleminadia with its close sister Polygenus and Ectomes, into which we sink Hypocryptorhina, Heronia, Polyhymnis, Chrysoplectrum, and Speculum (see Uncanny Divergence Within a Genus). Along with Loboclea, Lobocleina unifies genera with the arcuate antennal club from the “C. Urbanus” group of Evans plus Venada, Aguna, and Zeutus (Table 2). Others belong to the “crown” group of Eudamiae. It includes an array of skippers that have been largely misclassified due to widespread and possibly mimetic convergence as detailed in Widespread Convergence in Wing Patterns and Shapes.

**Eudamiae Genera and Subgenera.** We define a major phylogenetic cluster of species with a common ancestor existing within a certain timeframe as being a genus (18). We cut a dated phylogenetic tree at a time point to maximize agreement with the current classification. This neither splits nor merges most genera that are well defined by morphological features, and we treat as genera the groups of species supported by the cut branches. The time 15 Mya corresponds to such a point. It keeps well-known genera Aguna, Udranomia, and Urbanus proteinus group unsplit. However, it separates traditional and morphologically distinct pairs of genera such as Epargyreus and Chiodaes. We attempt to reduce the number of monotypic genera, unless the genus is strongly distinct, because we wish to indicate relationships to other species by the name of a genus. As a result (Fig. 1 and SI Appendix), we delineated 50 Eudamiae genera, 4 of which are described as new in Table 2. The number of monotypic genera decreased from 10 to 4: Nephelaeus, Spadilepeia, and Zeutus (19), plus Emmelus transferred from Pyrginae. These four genera diverged from their sister taxa at least 18 Mya and are morphologically distinct.

Within some genera, we see groups of species that may be defined as subgenera (SI Appendix), and three new subgenera are described (Table 2). Some of the subgenera, such as Thoebys, have been used as genera for decades, but their genomic and morphological distinctness is smaller compared with most genera.

**Widespread Convergence in Wing Patterns and Shapes.** The most unexpected result of this study is the astounding number of misplacements of species into genera they do not belong to. The
genera themselves, proposed over the years of classic entomological studies, mostly stood the test of genomic data: 55 genera were recognized before our work, and we revise them to be 50. We eliminated several monotypic genera for which visual morphological differences hindered close relationships with other species, and merged several phenotypically diverse but genotypically close genera (SI Appendix). Apparently, phenotypic distinction (e.g., wing shape, such as a tailed hindwing, or wing pattern, such as a pale stripe across the forewing) may be indicative of genetic differentiation. However, placement of a species into a genus by its dominant to human eye phenotypic feature is more problematic. For example, before our work, many tailed skippers were placed in the genus Urbanus based on the tail. Genomic data imply that half of them do not belong there, and we transferred them to three other genera. One genus is named here (Spicaua, Table 2), while the others have not previously included tailed skippers. Moreover, we transferred some skippers without long tails from Astraptes to Urbanus. They were formerly misclassified due to similar wing patterns consisting of shiny metallic-cyan wing bases and white forewing spots.

We found misclassifications to be widespread across Euadamiae and attribute them to convergence driven by the selection for large mimicry rings (20–22). This convergence is not confined to one or two basic patterns, but is more diverse. Some patterns are common in both the Old and New World, and, in addition to skippers, include butterflies from other families and even moths, flies, and beetles (20, 23). We find five different phenotypes (Fig. 2) that parallel each other in two genera (Telegonus and Cecropterus) and their more distant relatives (outgroups): (i) greenish bases of brown wings, white stripe on the forewing, and hindwing with a white tail and margins; (ii) metallic-blue wing bases, and forewing with white stripe (20); (iii) brown forewing with a yellow stripe across and apical white spots; (iv) cream-white, semitranslucent spots on the brown forewing; and (v) brown wings, and hindwing with yellow tornus.

At least four of these phenotypes are not ancestral and thus are convergent. Curiously, every species (10 out of 10) was placed by its appearance in a wrong genus before our study (crossed out genus in Fig. 2). In retrospect, assignment to a correct genus could have been possible through detailed comparison of male genitalia. In Telegonus, the dorsal side of the valva is concave in the middle and forms a mouth-like structure with two “kissing lips.” In Cecropterus, the valva is dorsally and terminally convex, without a “kiss.” These genitalic features agree with genomic phylogeny and reinforce our conclusions.

Uncanny Divergence Within a Genus. Before our work, each of the three genera Ectomis, Hypocryptothrix, and Heronia consisted of single species of unique appearance. No taxonomist had ever thought of them being a group. To our surprise, all phylogenetic trees we obtained (even from COI barcodes only) revealed only a slight divergence among these three and two other genera, Polythrix and Chrysoplectrum, suggesting that it is best to place all these skippers in a single genus, Ectomis. Moreover, their most divergent phylogenetic lineage was part of Polythrix, and is described here as a new subgenus Asina (Table 2). Barcode divergence among the subgenus Ectomis is within 10% and is less than within the genus of swallowtail butterflies Pierorus (24), which some researchers consider a subgenus of Papilio...

Despite limited genetic differentiation, the expanded Ectomis contains species of profound phenotypic divergence. All skippers in the former genus Polythrix (subsumed by Ectomis) are tailed. Others are not, although their hindwing is usually lobed at the tornus. While most Ectomis are brown skippers with a pale forewing band frequently divided into spots, some vary from solid dark brown to dark metallic green with a forewing central spot, or tawny with many white spots. Some are even part of the mimetic complex with brilliant blue thorax and wing bases above (20). Males of some species possess tufts of hairlike scales on the forewing below, while others have a double row of yellow spots on the hind tibiae. Male genitalia are as diverse as wing shapes and patterns, and the phylogenetic closeness is not apparent from genitalia. The valva varies from a simple curled plate without elaborations (in Ectomis cythna) to very complex with several processes (in “Heronia” labriatica). It would be of interest to investigate genetic mechanisms for such a rapid phenotypic divergence within Ectomis.

Connecting Genotype to Phenotype. Telegonus chalco (Fig. 2) and Telegonus brevicauda are sister species that do not look alike. T. brevicauda belongs to the “blue-wing-base” mimicry complex (Figs. 1–3) and is nearly indistinguishable from its more distant relative Urbanus tucuti (Fig. 1). We compared genomes of these three species to find rapidly diverging proteins between the two sisters (SI Appendix, Methods). Genes involved in the circadian clock system, transcription regulation, wing morphogenesis, fatty acid, and vitamin metabolism (SI Appendix, Tables S4 and S5) stood out. The differences in morphogenesis genes may be related to the hindwing tail development in T. chalco. Next, we looked for fixed mutations in the uniquely patterned T. chalco compared with the two similar species. Xanthine dehydrogenase

Fig. 2. Convergent wing patterns in Telegonus, Cecropterus, and others. Previously, they were placed in crossed out genera. Genitalic valvae are shown for the first two genera.

Fig. 3. Uncanny divergence within Ectomis. These skippers belonged to five genera listed above each image. Type species for these genera are marked with red asterisks.
had the largest number of such mutations (SI Appendix, Table S6). This enzyme is involved in pterin metabolism, and pterin derivatives were suggested to be responsible for wing coloration in pierid butterflies (25). The T. chalco-specific mutations may alter interactions between the enzyme and the transport proteins that deliver the enzyme to different pigment granules (26), and thus affect the wing color in T. chalco. Some of these mutations are present in other skipper species with extensive white scaling on hindwings.

Discussion: A Broad Picture. Today, and more so tomorrow, genome sequencing provides an efficient and cost-effective way to gain rapid insights about biodiversity. Given a taxonomic group, comparative analysis of genomic sequence along with morphology and ecology will be rich in discoveries. Even now, this can be done within a small laboratory. With the family Hesperiidae, we have shown how such a project can be accomplished and have reported some of the results that can be expected. Not only is this family interesting as a diverse group of butterflies, it also emphasizes several generalities. First, while the reference genomes required freshly collected specimens, the bulk of the project was done with old museum samples. Even specimens collected a century ago yielded usable genomic data. Second, we provided a genome-based phylogeny of the family and reclassified it taxonomically. We found that many species were classified differently than previously thought. Other families may display the same problem when examined genetically.

Third, we encountered many examples of phenotypic convergence and divergence, and mined genomic data for the links between genotype and phenotype. With the ever-decreasing cost of sequencing, we expect that, soon, any phylogenetic project will start from sequencing of genomes.

Materials and Methods

For freshly collected specimens, DNA was extracted from tissue of a specimen (minus wings and genitilia that were kept in an envelope) field-stored in a vial with RNA later. For pinned and dry specimens from museum collections, DNA was extracted from a whole abdomen or a leg. We have previously reported details of methods for DNA extraction, genomic library preparation, next-generation sequencing, and computational analysis of nuclear and mitochondrial genomes (13, 27). Phylogenetic trees were constructed with the programs RAxML (28) and ASTRAL (15), and were dated using the procedure described in SI Appendix, where further details of the methods are given. Genomic data obtained in the project have been deposited in the SRA database (accession SRP147939), and COI barcodes have GenBank accessions MH357724 through MH357835. ZooBank registration for this work, published on March 15, 2019, is BASB2CA-A8BA-488-BDF4-688FA8D4F3D.

ACKNOWLEDGMENTS. We thank the Area de Conservación Guanacaste, Costa Rica (ACG) parataxonomists who found and reared many specimens used in this study. We are grateful to Robert K. Robbins, John M. Burns, and Brian Harris (National Museum of Natural History, Smithsonian Institution); Blanca Huertas, David Lees, and Geoff Martin (National History Museum, London); David Grimail and Courtney Richenbacher (American Museum of Natural History); Andrew Warren and Andrei Sourakov (McGuire Center for Lepidoptera and Biodiversity); Edward G. Riley, Karen Wright, and John Oswald (Texas A&M University Insect Collection); John Rawlins (Carnegie Museum of Natural History); Weiping Xie (Los Angeles County Museum of Natural History); Tristian Maior and Thomas Baquaran (The Field Museum of Natural History); Vincent L. (California Academy of Sciences); James R. Reddell (University of Texas Biodiversity Center); Wolfram Mey (Berlin Museum für Naturkunde) for granting access to the collections under their care; and for stimulating discussions: Bernard Hermier, Ernst Brockmann, and Paul A. Opler for specimens and generously sharing COI barcode sequences through the Barcode of Life Data System (BOLD) database (www.boldsystems.org); Jim P. Brock, William R. Dempwolf, Bernard Hermier, John MacDonald, and James A. Scott for specimens; Gerard Lamas, Bernard Hermier, and Jonathan Pelham for fruitful discussions; Bernard Hermier, Lisa N. Kinch, and R. Dustin Schaeffer for critical reading of the manuscript; Texas Parks and Wildlife Department (Natural Resource Program Director David H. Riskind) for Research Permit 08-03Rev; and US National Park Service for the research permits: Big Bend (Raymond Skiles for BIBE-2004-SCI-0011 and Yellowstone (Erik Oberg and Annie Carlson for YELL-2017-SCI-7076). We acknowledge the Texas Advanced Computing Center, the University of Texas at Austin for providing high-performance computing resources. The study was supported, in part, by National Institutes of Health Grants GM094575 and GM127390 (to N.V.G.) and the Welch Foundation Grant I-1505 (to N.V.G.).