



Nuclear genomes distinguish cryptic species suggested by their DNA barcodes and ecology

Daniel H. Janzen^{a,1}, John M. Burns^b, Qian Cong^{c,d}, Winnie Hallwachs^a, Tanya Dapkey^a, Ramya Manjunath^e, Mehrdad Hajjibabaei^e, Paul D. N. Hebert^e, and Nick V. Grishin^{c,d,f}

^aDepartment of Biology, University of Pennsylvania, Philadelphia, PA 19104-6018; ^bDepartment of Entomology, National Museum of Natural History, Smithsonian Institution, Washington, DC 20013-7012; ^cDepartment of Biophysics, University of Texas Southwestern Medical Center, Dallas, TX 75390-9050; ^dDepartment of Biochemistry, University of Texas Southwestern Medical Center, Dallas, TX 75390-9050; ^eDepartment of Integrative Biology, University of Guelph, Guelph, ON, Canada N1G 2W1; and ^fHoward Hughes Medical Institute, University of Texas Southwestern Medical Center, Dallas, TX 75390-9050

Contributed by Daniel H. Janzen, January 16, 2017 (sent for review December 12, 2016; reviewed by May R. Berenbaum and John Shuey)

DNA sequencing brings another dimension to exploration of biodiversity, and large-scale mitochondrial DNA cytochrome oxidase I barcoding has exposed many potential new cryptic species. Here, we add complete nuclear genome sequencing to DNA barcoding, ecological distribution, natural history, and subtleties of adult color pattern and size to show that a widespread neotropical skipper butterfly known as *Udranomia kikkawai* (Weeks) comprises three different species in Costa Rica. Full-length barcodes obtained from all three century-old Venezuelan syntypes of *U. kikkawai* show that it is a rainforest species occurring from Costa Rica to Brazil. The two new species are *Udranomia sallydaleyae* Burns, a dry forest denizen occurring from Costa Rica to Mexico, and *Udranomia tomdaleyi* Burns, which occupies the junction between the rainforest and dry forest and currently is known only from Costa Rica. Whereas the three species are cryptic, differing but slightly in appearance, their complete nuclear genomes totaling 15 million aligned positions reveal significant differences consistent with their 0.00065-Mbp (million base pair) mitochondrial barcodes and their ecological diversification. DNA barcoding of tropical insects reared by a massive inventory suggests that the presence of cryptic species is a widespread phenomenon and that further studies will substantially increase current estimates of insect species richness.

cryptic species | ACG | butterflies | DNA barcoding | genomics

The 35-y ongoing biodiversity inventory of the estimated 15,000 species of moths and butterflies of Area de Conservación Guanacaste (ACG)—a large complex tropical conserved wildland in northwestern Costa Rica (1–4)—poses many questions of core interest to evolution, ecology, and conservation. The intense collecting (3) followed by mitochondrial DNA barcoding (4, 5) often suggests that what was known as a single species may comprise several species. But is there really more than one species? Subsequent study often finds characters from morphology, ecology, or natural history that covary with the barcodes and thereby confirm a complex of several species that may not differ from each other in their external appearance (3–10). *Udranomia kikkawai*, a small and long-known Costa Rican skipper butterfly (Hesperiidae), is one exemplar among many. Here, we unravel its biology and taxonomy, in part by augmenting barcoding with complete genomic analysis. Every species, even cryptic, carries a unique set of biological traits and should be recognized as a unit of biodiversity.

Act One

Udranomia kikkawai was first encountered in the ACG dry forest caterpillar inventory when it began in 1978. An adult was reared from a small caterpillar feeding only on the youngest leaves of *Ouratea lucens* (Ochnaceae), a common dry forest understory evergreen shrub. Today, 1,303 reared wild-caught ACG caterpillars later (among 650,000+ reared wild-caught caterpillars of 7,000+ species), *U. kikkawai* still has the same limited diet. In 1999, the ACG inventory discovered its caterpillar in immedi-

ately adjacent ACG rainforest, feeding only on the newest leaves of *Cespedesia spathulata* (Ochnaceae), a common tree of forest edges and middle-aged succession. Today, almost 888 wild-caught rainforest caterpillars later, they are still just as restricted in their diet. *U. kikkawai* occurs from Mexico to Brazil (11). It is one of the 500+ species of HesperIIDae now known to live in ACG (10), and it is visibly distinct from the two other species in ACG that belong to the same genus: *Udranomia orcinus* and *Udranomia eurus*.

A large, common, and showy hesperiid named *Astrartes fulgurator* in 1775, spans the neotropics and frequents ACG. In 2004, DNA barcoding a large sample of this skipper and correlating the barcodes with food plants, caterpillar color patterns, and subtle differences in adult patterns and size revealed 10 species under that one name, just in ACG (6). In 2004, stimulated by this stunning discovery, the inventory began barcoding many thousands of specimens amassed over 26 y (4). The barcodes of dry forest *U. kikkawai* and rain forest *U. kikkawai* formed two distinct clusters, shallowly separated in a neighbor-joining (NJ) phenogram (SI Appendix, Fig. S1 and Table S1) by just 1.5% (10 bp). They were dubbed *U. kikkawai*DHJ01 and *U. kikkawai*DHJ02, respectively, as with hundreds of other “barcode splits” encountered in the inventory (3–11). Examination of adults (Fig. 1) revealed no reliable interspecific differences in their color patterns. Genitalia

Significance

Thirteen years of mitochondrial DNA barcoding of 15,000+ species of Lepidoptera and their parasitoids living in Area de Conservación Guanacaste, northwestern Costa Rica, indicate several thousand cases where barcodes combined with ecology suggest unrecognized cryptic species, substantially increasing species counts. Here, we show that the slightly different barcodes of three extremely similar parapatric-sympatric species of butterflies covary not only with ecology and subtle morphological traits but also with nuclear genomes—a finding that we predict will be commonplace and a method that we predict will be widely used. The barcodes of the century-old type specimens of *Udranomia kikkawai* from Venezuela reveal that this name applies to one of the three Costa Rican cryptic species; the others we describe as new.

Author contributions: D.H.J., J.M.B., W.H., and N.V.G. designed research; D.H.J., J.M.B., Q.C., W.H., T.D., R.M., M.H., P.D.N.H., and N.V.G. performed research; D.H.J., J.M.B., Q.C., W.H., R.M., and N.V.G. analyzed data; D.H.J., J.M.B., Q.C., W.H., and N.V.G. wrote the paper.

Reviewers: M.R.B., University of Illinois at Urbana-Champaign; and J.S., The Nature Conservancy.

The authors declare no conflict of interest.

Data deposition: The barcode sequences reported in this paper have been deposited in the GenBank database (accession nos. KY421070, KY421071, and KY421072).

¹To whom correspondence should be addressed. Email: djanzen@sas.upenn.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1621504114/-DCSupplemental.

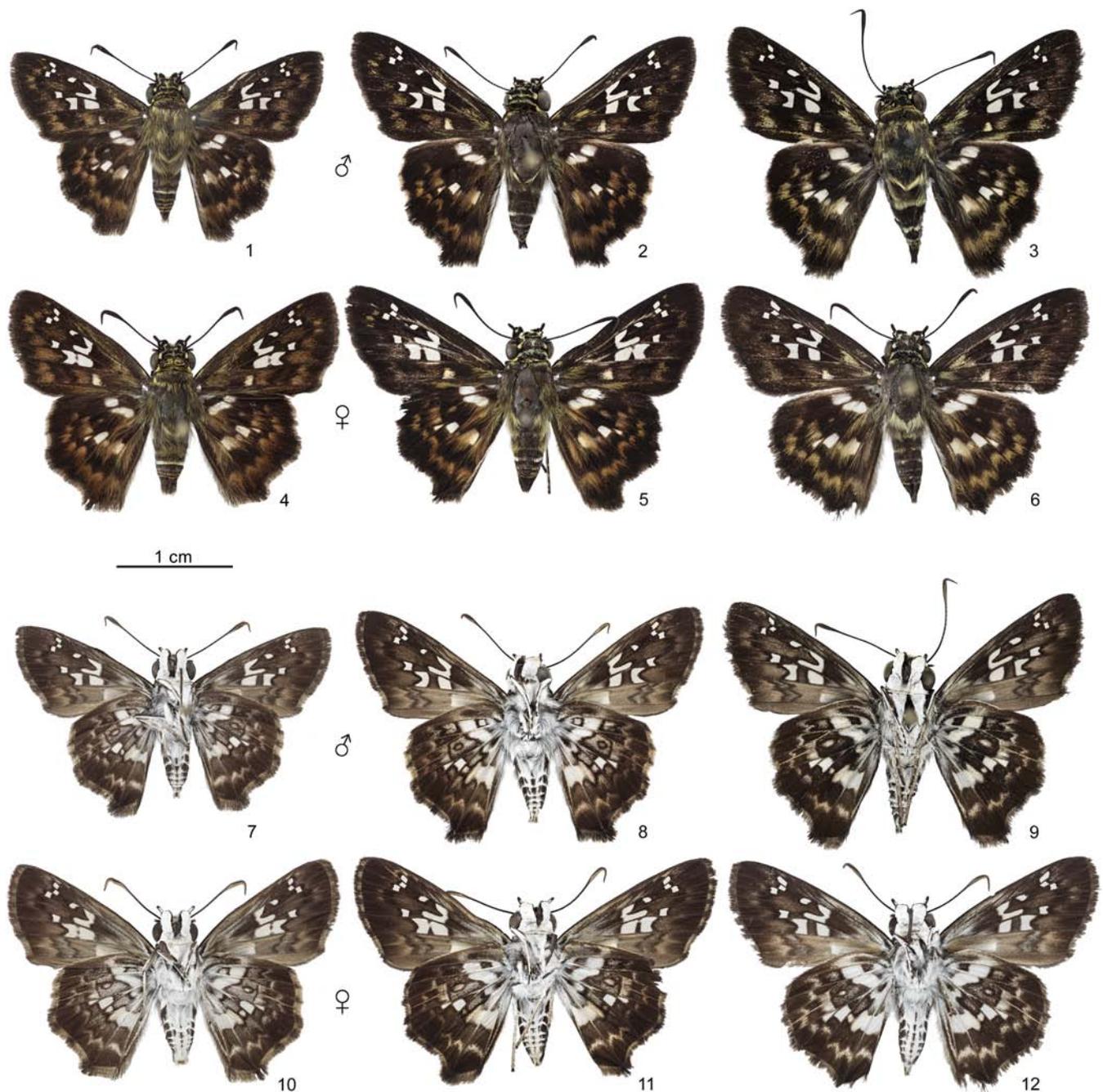


Fig. 1. Adults of the three species of the *U. kikkawai* complex in ACG, Costa Rica. Column 1, *U. sallydaleyae* (i.e., DHJ01); column 2, *U. tomdaleyi* (i.e., DHJ03); column 3, *U. kikkawai* (i.e., DHJ02). Rows 1 and 3, males; rows 2 and 4, females. Rows 1 and 2, dorsal views; rows 3 and 4, ventral views of the same specimens. Voucher codes (in parentheses) of these six specimens: 1, 7 (14-SRNP-12550); 2, 8 (14-SRNP-20533); 3, 9 (15-SRNP-71056); 4, 10 (14-SRNP-12541); 5, 11 (14-SRNP-20517); and 6, 12 (15-SRNP-30791). To avoid confusion with the elaborate spot patterns of the wings, pinholes have been artificially removed. Specimens 1, 7 and 2, 8 are holotypes. Photos are by Karie Darrow.

(SI Appendix, Figs. S2 and S3) showed none of the various differences that often characterize hesperiid species.

Whereas some taxonomists are reluctant to recognize species distinguished only by “invisible characters” such as DNA barcodes and ecology, even when those characters covary, we felt that two species were involved. Moreover, each one feeds on just one of two species of *Ochnaceae*, limited to dry forest and rainforest. There was one anomalous specimen among 43 (voucher code 01-SRNP-17596), whose barcode differed by about 2.5% from the barcodes of both *U. kikkawai*DHJ01 and -DHJ02. It

was reared from a caterpillar eating *O. lucens* found side by side with numerous *U. kikkawai*DHJ01, on the boundary between dry forest and rainforest. This odd specimen was baptized *U. kikkawai*DHJ03.

Today, 342 DHJ01, 164 DHJ02, and 171 DHJ03 have been reared and DNA barcoded. DHJ03 (Fig. 1) occupies only the intergrade between dry forest and rainforest (Fig. 2). DHJ03 is a low-density dry-forest differentiate mixed in with DHJ01 along its eastern edge adjoining the rainforest. Their caterpillars eat the same species of food plant and even cooccur on the same

Table 1. *Udranomia kikkawai* species complex reared in Area de Conservación Guanacaste, northwestern Costa Rica

Characters	<i>U. sallydaleyae</i> Burns, new species	<i>U. tomdaleyi</i> Burns, new species	<i>U. kikkawai</i> (Weeks)
Barcodes NJ tree (<i>SI Appendix</i> , Fig. S1):		55C, 91G, 178C, 235C, 284T, 340C, 361C, 367C, 382G, 514G, 526C	19C, 304C, 508C, 571C
diagnostic base pairs (bp)	52G, 403C, 562A		
% bp that differ between species	2.6% from <i>U. tomdaleyi</i> , 1.5% from <i>U. kikkawai</i>	3.3% from <i>U. kikkawai</i> , 2.6% from <i>U. sallydaleyae</i>	1.5% from <i>U. sallydaleyae</i> , 3.3% from <i>U. tomdaleyi</i>
Genome ML tree (Fig. 3):			
% genes that differ between species	14% from <i>U. tomdaleyi</i> , 22% from <i>U. kikkawai</i>	17% from <i>U. kikkawai</i> , 14% from <i>U. sallydaleyae</i>	22% from <i>U. sallydaleyae</i> , 17% from <i>U. tomdaleyi</i>
Known range	Mexico to Costa Rica	Costa Rica	Costa Rica to Brazil
Ecosystem (Fig. 2)	Dry forest	Dry/rain forest intergrade zone	Rain forest
Larval foodplant	<i>Ouratea lucens</i>	<i>Ouratea lucens</i>	<i>Cespedesia spathulata</i>
Adult (Fig. 1)			
Forewing length (mm)	♂, 13.55, 0.584, 12.7 – 14.7, 21	♂, 14.08, 0.565, 13.0 – 14.9, 11	♂, 14.27, 0.960, 11.7 – 15.6, 21
Sex, Mean, SD, Range, N	♀, 13.73, 0.432, 13.0 – 14.4, 21	♀, 14.63, 0.494, 13.8 – 15.2, 13	♀, 14.39, 0.691, 12.8 – 15.6, 25
Ground color of both wings	Lighter brown	Dark brown	Dark brown
Dorsal, outer row of spots on both wings	Coppery	Coppery	Yellowish/yellow
Two showy mid-hindwing spots on median pale band	Little to no overlap	Little to no overlap	Appreciable overlap
Ventral pattern elements of hindwing	Duller; less contrast	Good contrast reduced distally, leaving dark outer area	Bold; high contrast
Holotype ♂ voucher code	14-SRNP-12550	14-SRNP-20533	
Type locality	Costa Rica, ACG, Guanacaste Prov., Sector Santa Rosa, Cafetal, 280m, 10.8583, -85.6109	Costa Rica, ACG, Guanacaste Prov., Sector Orosi, Vado Rio Espavelar, 275m, 10.9468, -85.5670	Venezuela, Suapure
Paratypes (N, Sex)	140 ♂, 140 ♀	83 ♂, 78 ♀	
Type specimen deposition	U. S. Nat. Mus. Natural Hist.	U. S. Nat. Mus. Natural Hist.	Mus. Comp. Zool., Harvard Univ.

U. S. Nat. Mus. Natural Hist., United States National Museum of Natural History; Mus. Comp. Zool., Harvard Univ., Museum of Comparative Zoology, Harvard University.

from beige to cream to white. The basal band consists primarily of two spots that are especially prominent when they are semihyaline. The irregular medial band consists primarily of seven spots, of which a small triangular one in cell M_3-Cu_1 and a larger rectangular one in cell Cu_1-Cu_2 are almost always semihyaline so that they also show dorsally in the middle of the hindwing (Fig. 1). These two contiguous spots generally overlap little if at all in *Udranomia sallydaleyae* and *Udranomia tomdaleyi* but noticeably in *U. kikkawai* (Fig. 1).

The pale bands of the ventral hindwing (Fig. 1) connect via a weak to strong stripe of more or less white scales (some hair-like) near the inner margin of the wing. Between the pale bands is a band of brown ground color that may exhibit grayish (to the naked eye) overscaling (more extensive in females than in males) and may enclose a small pale spot in the discal cell. Distal to the medial pale band is a band of brown ground color distally bordered by a faint, irregular, submarginal row of small, beige, opaque spots, of which those in cells Cu_2-2A and $2A-3A$ always appear and sometimes suggest proximally pointing chevrons. Dorsally (Fig. 1), in all three species, this submarginal row of spots is better developed but blurry, and there is a corresponding row on the dorsal forewing; dorsally, on both wings, these spots (and some more proximal scaling) are generally coppery in *U. sallydaleyae* and *U. tomdaleyi* but yellowish to yellow in *U. kikkawai*.

On the ventral hindwing (Fig. 1), in *U. sallydaleyae*, especially in females, grayish-looking scales over the brown ground color between (and sometimes also distal to) the two pale bands are usually so extensive that the proximal two-thirds of the wing look pallid. This heightens the lighter aspect of *U. sallydaleyae*, in which the brown ground color of both wings usually looks a little paler than that of the other two species. Compared with those species, the pale bands are dull. The pale bands of *U. kikkawai*

are bright, well-developed, and prominent, especially against the brown ground color, which looks dark, even when overscaled with gray. Like the two pale spotbands, the faint submarginal row of spots develops most fully in this species. Altogether, the different elements of the complex ventral hindwing pattern stand out, in distinctively sharp contrast with one another. Despite a dark brown ground color, the ventral hindwing pattern of *U. tomdaleyi*, although strong, is often not as bold as that of *U. kikkawai* because the lightness of the pale bands is less intense; the bands themselves are not as well-developed, nor is the faint submarginal row of spots. As a result, the outer (and especially the lower) part of the hindwing may look more uniform and dark.

By mutual transplants of eggs and larvae, we determined that *U. kikkawai*, the rainforest species, survives to adulthood if fed only leaves of *O. lucens* in laboratory confinement, but the field collections show no sign of population-level invasion of dry forest and its large stands of *O. lucens*. Equally, both *U. sallydaleyae* and *U. tomdaleyi* survive to adulthood when fed only new leaves of *C. spathulata* in the laboratory, but show no sign of having invaded the adjacent rainforest ecosystem in which this plant is common. The caterpillars and pupae of all three are similar and colored as the light pinkish yellows of the new expanding leaves of both species of food plants (2, 11).

Whereas DNA barcoding is cheap and fast for large numbers of specimens, a deep dive into the nuclear genome, although slow and expensive, was worthwhile in this exemplary case to exclude the possibility of irregularities with barcodes (16). Therefore, we obtained genomic sequence reads for three specimens of each species from the *U. kikkawai* complex and single specimens of *U. orcinus* and *U. eurus* as outgroups (Fig. 3), assembled their protein-coding regions (15,152,393 aligned positions, 14,476 genes), and constructed a maximum-likelihood tree from

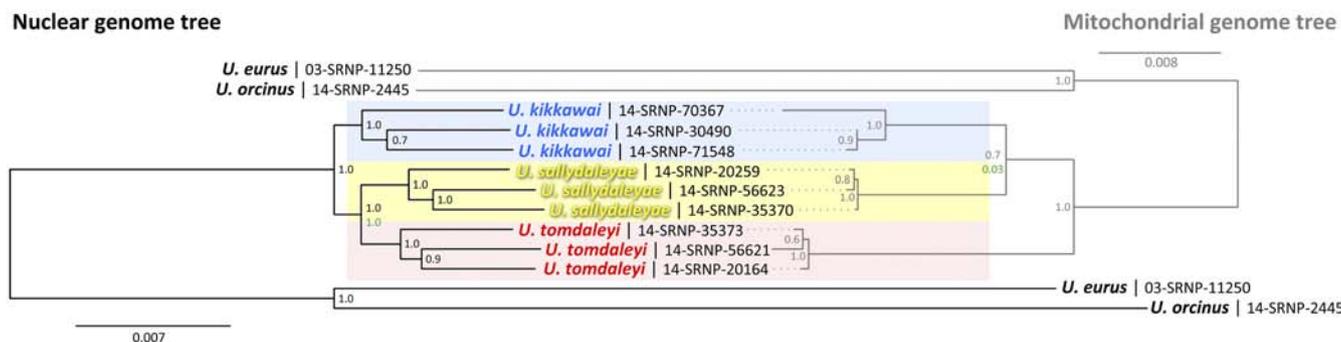


Fig. 3. Maximum-likelihood trees of *Udranomia*. Shown are nuclear genome (left, black, 15 million bp) and mitogenome (right, gray, 7,000 bp) trees constructed by RAxML (12) from protein-coding regions. RAxML bootstrap fractions are shown by the nodes and PhyloBayes (13) bootstrap (green) is shown for the two sister species in the *U. kikkawai* species complex (grouping in mitogenomes is not statistically supported).

both nuclear and mitochondrial genes. The mitogenome tree based on 6,891 positions recapitulated the barcode phenogram (Fig. 3, Right). The nuclear genome tree (Fig. 3, Left) recovered all three species as monophyletic with 100% bootstrap using both RAxML (12) and PhyloBayes (13), supporting the conclusions based on barcodes and ecology. We found protein-coding genes that are conserved within, but divergent among, the three species as measured by a fixation index (17) larger than 0.8. In all three species, GO term enrichment analysis (18) detected proteins that participate in circadian temperature homeostasis, transcription regulation, circadian behavior, copulation, and eclosion rhythm as the most statistically significant (*SI Appendix, Table S2*). Previously, we found circadian clock proteins to be speciation hotspots in several pairs of closely related species (18, 19), and the divergence in proteins involved in copulation is expected to correlate with speciation. Thus, the nuclear genomic regions that diverged the most between the three *Udranomia* species may be responsible for the mechanics of their speciation in terms of both reproductive isolation and adaptation to different ecological conditions. This further supports the distinctiveness of the three *Udranomia* species.

The tree built from concatenated nuclear genes differs in topology from the mitogenome tree. Using the F4 statistic (20, 21), we detected statistically significant ($P < 0.001$) introgression between *U. sallydaleyae* and *U. tomdaleyi* that may bias the nuclear tree and explain why they are grouped together. Thus, we selected genes that give trees in which all three species are monophyletic (i.e., have phylogenetic signal and show no introgression), resulting in an alignment of 397 nuclear genes. The tree built from them matches the topology of the mitogenome tree, placing *U. tomdaleyi* at the root.

Act Three

Even though *U. kikkawai* comprises three species, without the barcoding data that suggested splits it would still be considered just one, occurring in both dry forest and rainforest through much of the Neotropics. When only a few specimens were bar-coded, the apparent barcode split was so shallow (less than 2%) that it could have been ignored. Indeed, the rainforest species and the dry forest species are placed in the same barcode index number (BIN) (BOLD:AAA5981) by one of the standard ways to group DNA barcodes at the species level (22, 23). When the single specimen of *U. tomdaleyi* turned up among the mass of *U. kikkawai* (rainforest) and *U. sallydaleyae* (dry forest) specimens, it was viewed as a peculiar outlier, pseudogene, sequencing error, etc., and would have remained that way if haphazardly increasing the sample size had detected no more of them. However, as it is wont to do, the ACG inventory heavily sampled other sites that appeared to have the same ecological circumstances as the site where the first *U. tomdaleyi* was encountered. Sampling

went backward through specimens previously accumulated by the inventory and forward through new ones explicitly collected and reared for barcoding. The result thoroughly demonstrated that *U. tomdaleyi* is a real species that is ecologically specialized to live on the fragile join of dry forest with rainforest (24). Such species are rarely sought among tropical species because it is difficult to imagine how they would come to be, because ecosystem interfaces are commonly destroyed by the expanding agroscape (24), and because it can be difficult to see what ecosystem traits really matter to a given species when most museum collection data come from one-time hit-and-run collecting expeditions.

Act Four

The *U. kikkawai* complex ranges from Mexico to Brazil (11, 25). Determining the distribution of its three species outside ACG (and whether the complex includes additional species) requires extensive DNA barcoding of specimens throughout its range. Current data are sketchy. Specimens with *U. sallydaleyae* barcodes are known from dry forest in the states of Campeche and Quintana Roo in Mexico (26). *Udranomia kikkawai*, in addition to the type series from rainforest in Suapure, Venezuela, is known from the states of Para, Mato Grosso do Sul, Minas Gerais, and Sao Paulo in Brazil (11). In the Brazilian Cerrado, *U. kikkawai* caterpillars feed on young leaves of *Ouratea spectabilis* and *Ouratea hexasperma* (Ochnaceae) and take about 1 mo to develop from egg to adult (11), just as they do in ACG.

Discussion

How abundant are barcode splits among some 15,000 species of Lepidoptera, and fly and wasp parasitoids of their caterpillars, now DNA barcoded by the ACG inventory? As many as 4,500 species have barcode splits. At least 1,000 of these splits are as shallow as the 1–2% difference between the skipper butterflies *U. kikkawai* and *U. sallydaleyae* (NJ tree in *SI Appendix, Fig. S1* and refs. 4, 6, 10). Barring further analysis, we cannot confidently say how many of the barcode splits reflect biological species. But from long experience, we hazard that 10–20% of the traditional, morphologically based “single” species will turn out to be two or more (3, 4, 8–10). Sound evidence for species status of problematic barcode clusters separated by shallow splits comes from covariance between those clusters and independent characters. We demonstrate here that molecular data from nuclear DNA support those from mitochondrial DNA.

Altogether, covariance between traits as disparate as DNA barcodes, nuclear genomes, ecology, and wing facies shows that there are indeed three distinct biological species comprising a long unsuspected *U. kikkawai* complex in ACG and beyond (i.e., Mexico to Brazil). Due to new tools and their widespread use, it is increasingly clear that diverse geographic regions and

taxonomic groups contain many cryptic species. Their prior obscurity stems largely from our limited, and therefore biased, visual perception of the living world.

Forty-five protein-coding genes associated with specific functions differ among the species of the *U. kikkawai* complex (SI Appendix, Table S2). None of these genes directly affects an externally visible morphological trait. Some of them relate to various aspects of behavior, including copulation, that can support differentiation that leads to speciation. Under such circumstances, which are doubtless common to many organisms, offspring species may, by chance, not noticeably differ outwardly from one another or from their parent species; and these cryptic species may endure in sympatry. A similar result is even more likely when, for any of various reasons (including mimicry and camouflage), selection favors retention of the parental appearance. Although the evolution and persistence of cryptic species frustrates some taxonomists, it stimulates others.

Materials and Methods

Procedures for collecting and barcoding specimens were described previously (6). Barcodes of *U. kikkawai* type specimens were amplified in nine overlapping segments. Primers for them are given in GenBank entries KY421070, KY421071, and KY421072. Methods for DNA extraction, library preparation, next-generation sequencing, and computational analysis of

complete nuclear and mitochondrial genomes have been reported previously (19). See SI Appendix for details.

ACKNOWLEDGMENTS. We gratefully acknowledge the unflagging support of the team of ACG parataxonomists (3, 5) who found and reared the specimens used in this study and the team of biodiversity managers who protect and manage the ACG forests that host these insects and their food plants. We thank Donald J. Harvey for dissecting genitalia and Karie Darrow for photographing some of them, as well as adult specimens. We are indebted to Phil Perkins, Rachel Hawkins, and Naomi Pierce (Museum of Comparative Zoology, Harvard University) for the loan of *U. kikkawai* type series specimens for DNA analysis essential for this study. The two new species are named in honor of Mr. and Mrs. Tom and Sally Daley of Moraga, CA, in recognition of their many years of diligent management of the Guanacaste Dry Forest Conservation Fund financial bureaucracy. This study has been supported by US National Science Foundation Grants BSR 9024770 and DEB 9306296, 9400829, 9705072, 0072730, and 0515699; and grants from the Wege Foundation, International Conservation Fund of Canada, Jessie B. Cox Charitable Trust, Blue Moon Fund, Guanacaste Dry Forest Conservation Fund, Area de Conservación Guanacaste, Permian Global, Smithsonian Institution (National Museum of Natural History), US Department of Agriculture Systematic Entomology Laboratory, individual private donors, the University of Pennsylvania (D.H.J. and W.H.), National Institutes of Health Grant GM094575, and the Welch Foundation Grant I-1505 (to N.V.G.). This study has been supported by the Government of Canada through its ongoing support of the Canadian National Collection, Genome Canada, the Biodiversity Institute of Ontario, Centre for Biodiversity Genomics, the Ontario Genomics Institute, and the Natural Sciences and Engineering Research Council of Canada.

- Area de Conservacion Guanacaste (2016) Area de Conservacion Guanacaste, northwestern Costa Rica. Available at www.acguanacaste.ac.cr. Accessed December 31, 2016.
- Janzen DH, Hallwachs W (2016) Dynamic database for an inventory of the macrocaterpillar fauna, and its food plants and parasitoids, of Area de Conservacion Guanacaste (ACG), northwestern Costa Rica. Available at janzen.bio.upenn.edu/caterpillars/database.lasso. Accessed December 31, 2016.
- Janzen DH, et al. (2009) Integration of DNA barcoding into an ongoing inventory of complex tropical biodiversity. *Mol Ecol Resour* 9(Suppl 1):1–26.
- Janzen DH, Hallwachs W (2016) DNA barcoding the Lepidoptera inventory of a large complex tropical conserved wildland, Area de Conservacion Guanacaste, northwestern Costa Rica. *Genome* 59:641–660.
- Janzen DH, Hallwachs W (2011) Joining inventory by parataxonomists with DNA barcoding of a large complex tropical conserved wildland in northwestern Costa Rica. *PLoS One* 6:e18123.
- Hebert PDN, Penton EH, Burns JM, Janzen DH, Hallwachs W (2004) Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator*. *Proc Natl Acad Sci USA* 101:14812–14817.
- Burns JM, Janzen DH, Hajibabaei M, Hallwachs W, Hebert PDN (2008) DNA barcodes and cryptic species of skipper butterflies in the genus *Perichares* in Area de Conservación Guanacaste, Costa Rica. *Proc Natl Acad Sci USA* 105:6350–6355.
- Smith MA, Wood DM, Janzen DH, Hallwachs W, Hebert PDN (2007) DNA barcodes affirm that 16 species of apparently generalist parasitoid flies (Diptera, Tachinidae) are not all generalists. *Proc Natl Acad Sci USA* 104:4967–4972.
- Smith MA, et al. (2008) Extreme diversity of tropical parasitoid wasps exposed by iterative integration of natural history, DNA barcoding, morphology, and collections. *Proc Natl Acad Sci USA* 105:12359–12364.
- Janzen DH, et al. (2011) Reading the complex skipper fauna of one tropical place. *PLoS One* 6:e19874.
- Bachtold A, et al. (2017) Integrative data helps the assessment of a butterfly within the *Udranomia kikkawai* species complex (Lepidoptera: Hesperidae): Immature stages, natural history, and molecular evidence. *Zool Anz* 266:169–176.
- Stamatakis A (2014) RAxML Version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30:1312–1313.
- Lartillot N, Lepage T, Blanquart S (2009) PhyloBayes 3: A Bayesian software package for phylogenetic reconstruction and molecular dating. *Bioinformatics* 25:2286–2288.
- Weeks AG (1906) New Lepidoptera. No. 2. *Canad Entomol* 38:174–178.
- Weeks AG (1911) *Illustrations of Diurnal Lepidoptera, with Descriptions* (University Press, Boston), Vol 2.
- Zakharov EV, Lobo NF, Nowak C, Hellmann JJ (2009) Introgression as a likely cause of mtDNA paraphyly in two allopatric skippers (Lepidoptera: Hesperidae). *Heredity* 102:590–599.
- Hudson RR, Slatkin M, Maddison WP (1992) Estimation of levels of gene flow from DNA sequence data. *Genetics* 132:583–589.
- Cong Q, Borek D, Otwinowski Z, Grishin NV (2015) Tiger Swallowtail genome reveals mechanisms for speciation and caterpillar chemical defense. *Cell Rep* 10:910–919.
- Cong Q, et al. (2016) Complete genomes of hairstreak butterflies, their speciation, and nucleo-mitochondrial incongruence. *Sci Rep* 6:24863.
- Patterson N, et al. (2012) Ancient admixture in human history. *Genetics* 192:1065–1093.
- Pickrell JK, Pritchard JK (2012) Inference of population splits and mixtures from genome-wide allele frequency data. *PLoS Genet* 8:e1002967.
- Ratnasingham S, Hebert PDN (2007) BOLD: The barcode of life data system. *Mol Ecol Notes* 7:355–364.
- Ratnasingham S, Hebert PDN (2013) A DNA-based registry for all animal species: The barcode index number (BIN) system. *PLoS One* 8:e66213.
- Janzen DH (1987) Habitat sharpening. *Oikos* 48:3–4.
- Evans WH (1952) *A Catalogue of the American Hesperidae Indicating the Classification and Nomenclature Adopted in the British Museum (Natural History), Part II, Pyrginae, Section 1* (British Museum, London).
- Centre for Biodiversity Genomics (2016) BOLD systems. Available at www.boldsystems.org. Accessed December 31, 2016.