Genomics overrules mitochondrial DNA, siding with morphology on a controversial case of species delimitation

Carmen del R. Pedraza-Marron, Raimundo Silva, Jonathan Deeds, Steven M. Van Belleghem, Alicia Mastretta-Yanes, Omar Domınguez-Domı´nguez, Rafael A. Rivero-Vega, Loretta Lutackas, Debra Murie, Daryl Parkyn, Lewis H. Bullock, Kristin Foss, Humberto Ortiz-Zuazaga, Juan Narvaez-Barandica, Arturo Acero, Grazielle Gomes, and Ricardo Betancur-R

SEDAR74-RD70

March 2021

PROCEEDINGS B

royalsocietypublishing.org/journal/rspb

Research

Cite this article: Pedraza-Marrón C del R *et al*. 2019 Genomics overrules mitochondrial DNA, siding with morphology on a controversial case of species delimitation. *Proc. R. Soc. B* **286**: 20182924.
http://dx.doi.org/10.1098/rspb.2018.2924

Received: 2 January 2019
Accepted: 12 March 2019

Subject Category:
Genetics and genomics

Subject Areas:
genomics, evolution, taxonomy and systematics

Keywords:
species delimitation, SNP data, mito-nuclear discordance, Bayes factor delimitation, introgression

Author for correspondence:
Ricardo Betancur-R
e-mail: betanri@fishphylogeny.org

†These authors contributed equally to this work.

Electronic supplementary material is available online at http://dx.doi.org/10.6084/m9.figshare.c.4444466.

THE ROYAL SOCIETY PUBLISHING

Genomics overrules mitochondrial DNA, siding with morphology on a controversial case of species delimitation

Carmen del R. Pedraza-Marrón[1,†], Raimundo Silva[3], Jonathan Deeds[5], Steven M. Van Belleghem[1], Alicia Mastretta-Yanes[6], Omar Domínguez-Domínguez[7], Rafael A. Rivero-Vega[1], Loretta Lutackas[1,8], Debra Murie[9], Daryl Parkyn[9], Lewis H. Bullock[10], Kristin Foss[11], Humberto Ortiz-Zuazaga[2], Juan Narváez-Barandica[12], Arturo Acero[13], Grazielle Gomes[3,4] and Ricardo Betancur-R[14,†]

[1]Department of Biology, and [2]Computer Science Department, University of Puerto Rico, Río Piedras, PO Box 23360, San Juan, PR 00931, USA
[3]Instituto de Estudos Costeiros, Laboratório de Genética Aplicada, and [4]Instituto de Estudos Costeiros, Laboratório de Genética e Biologia Molecular, Universidade Federal do Pará, Campus Bragança, Alameda Leandro Ribeiro, S/N, Aldeia, 68600-000 Bragança, Pará, Brazil
[5]USFDA Center for Food Safety and Applied Nutrition, 5001 Campus Drive, College Park, MD 20740, USA
[6]CONACYT Research Fellow – Comisión Nacional para el Conocimiento y Uso de la Biodiversidad, México, Liga Periférico – Insurgentes Sur, No. 4903, 14010, México, DF, México
[7]Laboratorio de Biología Acuática, Facultad de Biología, Universidad Michoacana de San Nicolás de Hidalgo, Ciudad Universitaria, 58000, Morelia, Michoacán, México
[8]North Carolina Wildlife Resources Commission, Raleigh, NC 27699-1700, USA
[9]University of Florida, School of Forest Resources and Conservation, Program of Fisheries and Aquatic Sciences, Gainesville, FL 32653, USA
[10]Florida Fish and Wildlife Research Institute, 100 8th Ave SE, St. Petersburg, FL 33701, USA
[11]Florida Fish and Wildlife Conservation Commission, Farris Bryant Building, 620 S. Meridian St., Tallahassee, FL 32399-1600, USA
[12]Universidad del Magdalena, carrera 32 No. 22-08 Santa Marta, Magdalena, Colombia
[13]Universidad Nacional de Colombia Sede Caribe Cecimar, El Rodadero, Santa Marta, Colombia
[14]Department of Biology, The University of Oklahoma, 730 Van Vleet Oval, Room 314, Norman, OK 73019, USA

RB-R, 0000-0002-9512-5011

Species delimitation is a major quest in biology and is essential for adequate management of the organismal diversity. A challenging example comprises the fish species of red snappers in the Western Atlantic. Red snappers have been traditionally recognized as two separate species based on morphology: *Lutjanus campechanus* (northern red snapper) and *L. purpureus* (southern red snapper). Recent genetic studies using mitochondrial markers, however, failed to delineate these nominal species, leading to the current lumping of the northern and southern populations into a single species (*L. campechanus*). This decision carries broad implications for conservation and management as red snappers have been commercially over-exploited across the Western Atlantic and are currently listed as vulnerable. To address this conflict, we examine genome-wide data collected throughout the range of the two species. Population genomics, phylogenetic and coalescent analyses favour the existence of two independent evolutionary lineages, a result that confirms the morphology-based delimitation scenario in agreement with conventional taxonomy. Despite finding evidence of introgression in geographically neighbouring populations in northern South America, our genomic analyses strongly support isolation and differentiation of these species, suggesting that the northern and southern red snappers should be treated as distinct taxonomic entities.

1. Introduction

Delimitation of species—the basic unit of biological diversity—is of great interest across many fields in biology. The adoption of molecular information for species delimitation analyses has unveiled cryptic diversity across several taxa [1,2]. Initial approximations that integrated genetic markers, such as mitochondrial DNA (mtDNA) or scant nuclear DNA (nDNA) fragments, into traditional taxonomy provided greater resolution for a broad array of groups [1], from marine corals and fishes [3,4] to terrestrial fungi and mammals [5,6]. Mitochondrial and single nuclear markers, however, are not always efficient tools [7–9], and can at times fail to discriminate species correctly [10]. This is exemplified by the often incongruent genealogies inferred from different genetic loci that identify conflicting histories [11], which can ultimately arise from incomplete lineage sorting (ILS) or introgression [7,9,12] and reveal the history of the genes examined rather than that of the species [13]. Although mtDNA markers, widely used in molecular barcoding, have proven powerful at detecting cryptic species (e.g. fishes [3,14]), there are few examples in natural populations where mitochondrial-based approaches conflict with both conventional taxonomy and genomic inferences (e.g. sharks [15], lampreys [16], caddisflies [17]; see [18] for a review). Recently, the advent of high-throughput sequencing technologies has facilitated the generation of large-scale datasets with thousands of markers for high resolution of shallow evolutionary inferences [19], further allowing the elucidation of complex speciation scenarios (e.g. [17,20,21]). Uncovering signals of population and species differentiation with genome-wide molecular information is now becoming mainstream [2,22] and permits the rigorous validation of relationships that were previously inferred from single or few genetic loci.

Here, we address a controversial case of species delimitation of red snappers (Teleostei: Lutjanidae) in the Western Atlantic (WA) where mtDNA has delimited fewer species than initially documented. For over a century, two allopatric species of red snappers have been recognized on the basis of morphological and meristic traits, including number of scales in the lateral line (or scale counts in rows above and below the lateral line) and modal differences in anal-fin ray counts [23]. The northern red snapper, *Lutjanus campechanus* (Poey, 1860), is distributed along the US East coast and the Gulf of Mexico; the southern red snapper, *Lutjanus purpureus* (Poey, 1866), occurs in the Caribbean Sea and southwards through northeastern Brazil. Recent attempts to investigate population genetic structure and to evaluate the degree of similarity of red snappers using mtDNA sequences [24,25] failed, however, to discriminate the nominal species as independent evolutionary groups. These studies have ultimately suggested that the northern and southern red snappers constitute a single species (*L. campechanus*) that exhibits phenotypic variability throughout the WA. This decision has been recently adopted by several taxonomic authorities [26,27], carrying downstream repercussions for conservation and fisheries management. The conflicting morphological and mitochondrial evidence has raised a controversial case of species delimitation where an accurate taxonomic demarcation is of particular concern, as red snappers have been widely overfished and are currently listed as vulnerable by the IUCN Red List of Threatened Species [28].

Using genome-wide markers generated via RAD sequencing approaches, we test the discordance between the mtDNA- and morphology-based hypotheses that has led to a continuing conflict of species delimitation in WA red snappers. We show that southern and northern red snappers represent two independent evolutionary lineages that should be recognized as distinct species. We highlight the importance of using genomic approaches to reconcile complex species delimitation scenarios where different lines of evidence conflict.

2. Material and methods

(a) Sampling

We examined a total of 178 red snapper individuals (105 of *L. campechanus* and 73 of *L. purpureus*) collected from 15 locations across the WA (figure 1; electronic supplementary material, table S1). Georeferenced data are available for most sampling sites; for others, the approximate location was inferred by interpreting collecting site descriptions. We also attempted to acquire specimens that would fill the sampling gap through the Caribbean Islands or intermediate populations in Central America. Although we actively searched for over 2 years in two key Caribbean locations (Puerto Rico and San Andrés Island, Colombia), all surveys were unsuccessful (see details in electronic supplementary material, figure S1). Given the apparent scarcity of the species in the region, other Caribbean locations were also ineffectively probed for samples through networking efforts. To emphasize the low abundance of red snappers in many Caribbean localities, we generated a map with total records for both species using reports available from FishNet (www.fishnet2.net) and the Ocean Biogeographic Information System (OBIS) (www.iobis.org) (electronic supplementary material, figure S1).

(b) Molecular protocols, mitochondrial data and SNP genotyping

All individuals examined were sequenced using restriction-digest-associated DNA sequencing (RADseq) approaches by applying the double-digest (ddRADseq) protocol of Peterson *et al.* [29]. This technique allows for the low cost discovery and genotyping of thousands of genetic markers and is particularly useful for non-model organisms [30]. In order to compare the population structure using genome-wide RADseq markers to that obtained with mtDNA (e.g. as in previous studies [24,25]), a subset of 83 samples was barcoded using the mtDNA gene cytochrome-c oxidase subunit I (*COI*) following standard protocols [31] (electronic supplementary material, table S2). Additional mtDNA sequences for *COI* and *D-loop* were downloaded from available data on NCBI (electronic supplementary material, table S3). ddRADseq data were processed using several packages, including Stacks v1.49 [32], FastQC v0.11.5 (www.bioinformatics.babraham.ac.uk/projects/fastqc/), Tassel v5.2.43 [33], and VCFtools v0.1.15 [34]. Different combinations of assembly parameters were first tested on a subset of 30 samples (following [35]) in Stacks. Final locus assembly was performed using a minimum of five raw reads required to form a stack, and allowing a maximum of two mismatches between stacks and three mismatches between loci of different individuals. Loci with a minimum allele frequency of 0.05 and a maximum observed heterozygosity of 0.70 were further excluded as potential paralogues. The sensitivity of results to number of individuals and missing data was also evaluated by applying a variety of filters. Four datasets that contained between 21 431 and 55 795 loci were first selected based on loci present in multiple predetermined numbers of populations (p) and percentage of individuals (r) (p11r50, p12r50, p9r60 and p8r60). A second filter ('min. sites') was applied after removing individuals

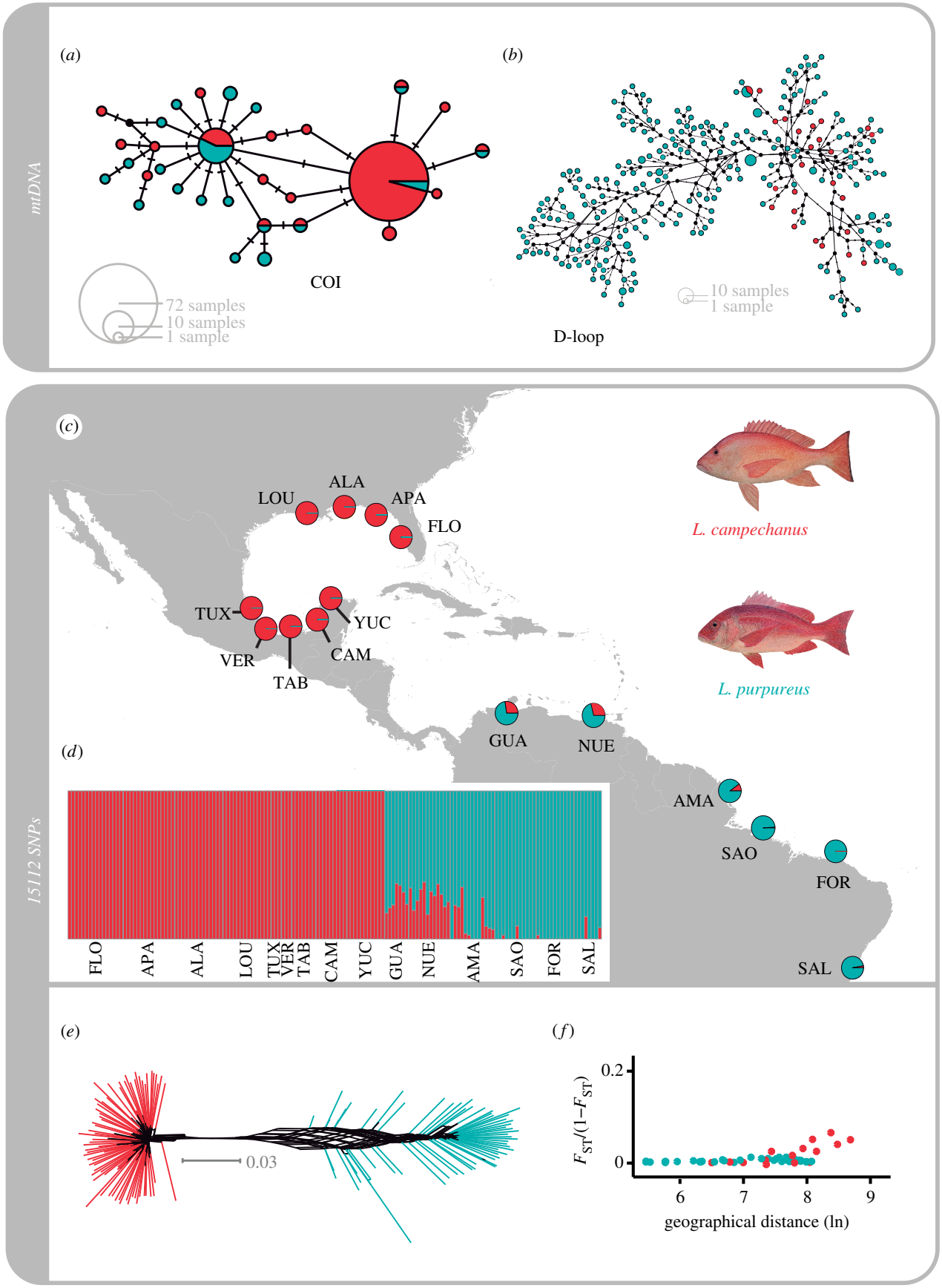

Figure 1. Genetic structure of WA red snappers. The northern red snapper (*Lutjanus campechanus*; red) and the southern red snapper (*Lutjanus purpureus*; green) are recognized as two separate species by conventional taxonomy on the basis of morphological characters. Consistent with previous studies, haplotype networks based on mitochondrial DNA sequences lack discriminatory power at the species level: (*a*) COI; (*b*) D-loop. However, a Bayesian structure analysis using 15 112 genome-wide SNPs identifies two main genetic clusters ($K = 2$) that are concordant with the traditional taxonomic delineations. Average admixture proportions were calculated for either (*c*) populations or (*d*) individuals (each structure bar representing the probability of assignment to each cluster). These results are congruent with (*e*) the estimated phylogenetic network (see additional trees in figure 2). (*f*) Correction of Mantel correlogram between Weir and Cockerham F_{ST} values versus least cost path geographical distances provide little support for a model of isolation by distance for intraspecific comparisons. Population information, descriptions, and abbreviations are given in electronic supplementary material, table S1.

with different thresholds for missing sites (0.75, 0.50, 0.25 and 0.05). These filters resulted in 20 datasets (electronic supplementary material, table S4), of which six were further selected according to the amount of missing data (9–46%), number of individuals present (44–155), number of SNPs (15 112–42 406), and number of populations (8–15). Additional details on molecular protocols for *de novo* assembly of RAD loci are given in the electronic supplementary material.

(c) Phylogenetic and coalescent analyses

A phylogenetic network was computed based on 15 112 SNPs with the Neighbor-Net algorithm in SplitsTree v4.14.6 (www.splitstree.org). Phylogenetic reconstruction was performed in a maximum-likelihood (ML) framework using the software RAxML v8 [36]. Trees were inferred for the six SNP datasets selected in the previous step. All invariant sites were removed from the matrices using the R package phrynomics (https://github.com/bbanbury/phrynomics/). Two alternative ML analyses were performed: one in which heterozygous alleles were collapsed using ambiguity (IUPAC) codes, and another using concatenated variants (extracted in order of appearance from the VCF file). No major differences were found between trees reconstructed from these variants and only the latter trees are reported. To account for acquisition biases inherent to SNP datasets [37], we used the GTR + I model with ascertainment bias correction (ASC) in RAxML. Nodal support was assessed in RAxML using 100 rapid bootstrap replicates. For the mitochondrial matrix, haplotype networks of 654 bp *COI* and 858 bp *D-loop* were constructed using the TCS Network available in PopArt [38]. The *COI* sequences were aligned using available references of *L. campechanus* and *L. purpureus* from GenBank (accession no. EU752115 and EU752118).

We also assessed the fit of the two alternative scenarios of species delimitation in WA red snappers in a coalescent framework. We used the Bayes factor delimitation (BFD*) method implemented for genome-wide SNP data [39] in the programs SNAPP v1.3.0 [40] and BEAST2 v2.4.1 [41]. To reduce computational burden, we first applied additional filters to the p12r50 dataset (with 15 112 SNPs from 15 populations; see above) by retaining both loci and individuals from each population with the lowest proportions of missing data. Three subsets with 58–108 individuals and 149–957 loci were assembled (see electronic supplementary material). To set up priors and MCMC runs, we carefully followed the guidelines outlined in the BFD* tutorial by A. Leaché (http://www.beast2.org/bfd/). Because the scenarios tested contained fewer than three species-tree nodes (e.g. for one species the leaf node is also the root node), we removed all tree operators from the analyses (R. Bouckaert 2019, personal communication). Finally, we compared the marginal likelihood estimates for the alternative scenarios using Bayes factors.

(d) Population structure analyses

Principal component analyses (PCAs) were first computed on allele frequencies using Tassel v5.2.43 [33]. The analyses for p12r50 and p8r60 matrices were performed with three different proportions of 'min. sites' (0.75, 0.25 and 0.5), and figures were plotted using the R [42] package ggplot [43]. The p12r50 matrix was selected for downstream analyses, as this dataset captures the genetic information of 155 individuals from all 15 populations while maximizing population discrimination. Next, the fastStructure v1.0 [44] package, a Bayesian clustering method, was used for inferring population structure. The number of population clusters was evaluated by running multiple values of K (1–18) using a logistic prior. The best-fitting model complexity was selected with the chooseK.py routine and the resulting K value was re-run through fastStructure 25 times with multiple random starting seeds to identify the five highest values of the log-marginal likelihood (LLBO). Final plots were constructed using disruct.py, available from fastStructure. Lastly, Weir and Cockerham F_{ST} values were estimated using the R package hierfstat [45] using 100 bootstrap replicates. Because large amounts of loci with missing data can deviate true values of summary statistics [46], we calculated F_{ST} values from two datasets: (i) p12r50-155, as previously selected; and (ii) p12r50-89, excluding three populations with few represented loci. F_{ST} values were plotted as heatmaps using the R package gplots [47].

(e) Isolation by distance

Limited dispersal capabilities in panmictic populations often result in a correlation between geographical distance and genetic differentiation among populations—a process termed isolation-by-distance (IBD) [48]. To test whether the red snapper populations follow a pattern of IBD, we performed a Mantel test using a correlation and a major axis correction [49,50] between Weir and Cockerham F_{ST} values among populations and their corresponding geographical distances (including 15 populations). Geographical distances were calculated via the least cost path (LCP) distance over seawater using the R package marmap [51]. We constrained the LCP to depth values between 10 and 190 m, which constitute the depth range of suitable habitat for red snappers [26]. Because these species can also disperse through oceanic currents during their pelagic larval phase, additional mantel tests were conducted using Euclidean geographical distances (computed with the R package adegenet [52]). Results of Mantel tests were not affected by the use of LCP or Euclidean distances; therefore, only the former results are reported (figure 1*f*).

(f) Hybridization

In order to test for ongoing hybridization between the two species, we used the R package gghybrid to estimate the hybrid index (HI)—a measure of genetic admixture within individuals [53]. The gghybrid package uses a Bayesian algorithm on bi-allelic genomic data to calculate the proportion of allele copies coming from parental reference sets [54] while applying a logit-logistic model for the genomic cline curve [54,55]. We ran HI estimations using 10 000 MCMC iterations and estimated posterior probability values after a 5000 iteration burnin. We selected the northernmost populations of *L. campechanus* (Florida and Apalachicola) and the southernmost populations of *L. purpureus* (Fortaleza and Salvador) as parental references in order to reduce the probability of gene exchange between major lineages. By selecting populations with a low probability of contact, the analysis focuses on loci that are highly differentiated in the parental reference populations.

3. Results

In agreement with previous studies [24,25], our mtDNA haplotype networks fail to delimit the nominal species as distinct haplogroups (figure 1*a*,*b*). The *COI* network shows an intermingling of *L. campechanus* and *L. purpureus* (figure 1*a*), whereas the *D-loop* network identifies one haplogroup formed solely by *L. purpureus* individuals and another where haplotypes of *L. campechanus* are nested within the *L. purpureus* populations (figure 1*b*). Similarly, the mtDNA *COI* tree lacks resolution and reveals no geographical segregation of individuals based on unique haplotypes (figure 2*a*). By contrast, trees inferred with genome-wide RADseq data (15 112–42 406 loci) consistently

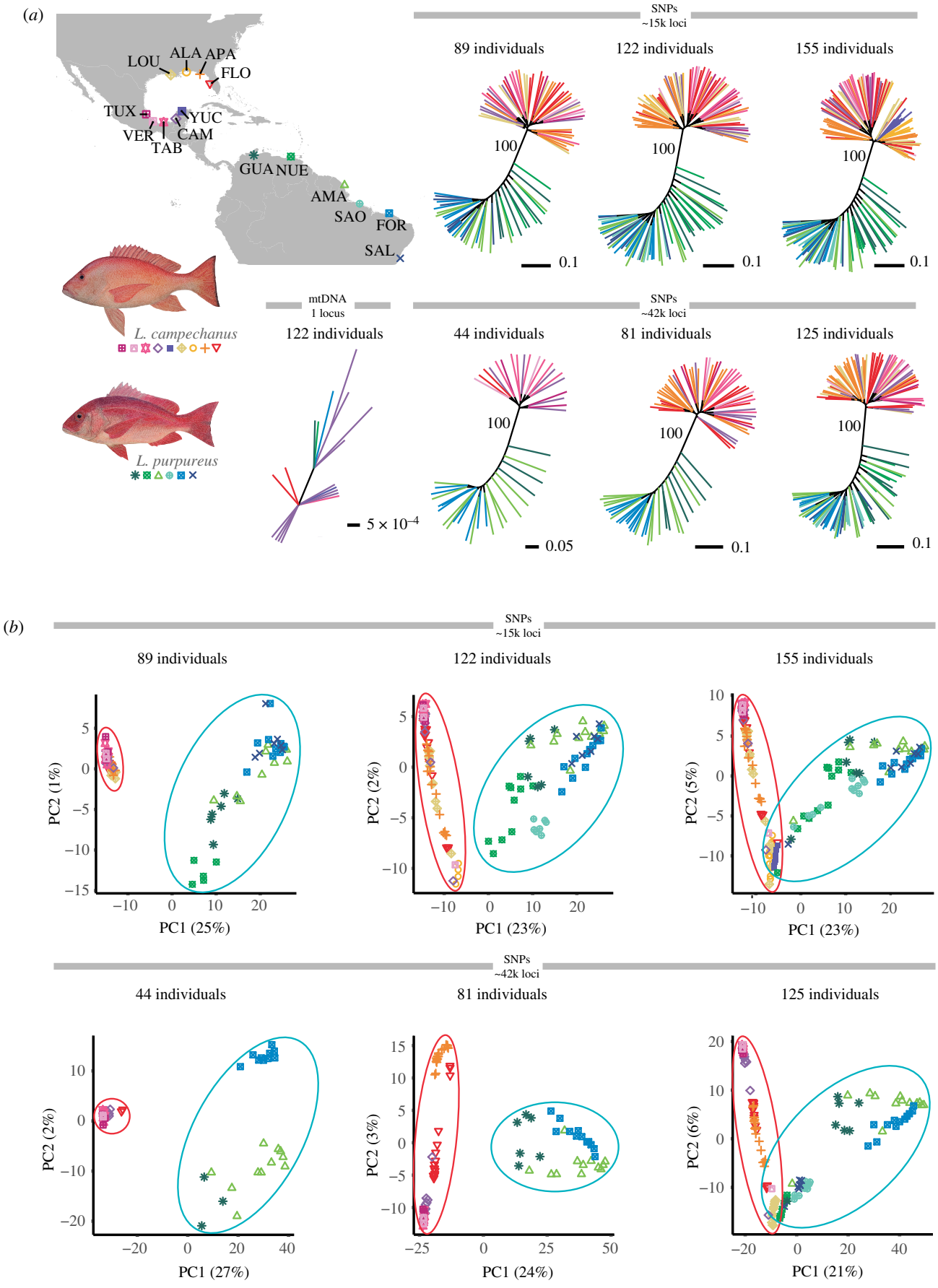

Figure 2. (*a*) Phylogenetic and (*b*) principal component analyses (PCAs). Phylogenetic trees and PCAs based on approximately 15 000–42 000 SNPs resolve two well-differentiated clusters that are congruent with the morphology-based hypothesis. These results are largely consistent regardless of the number of individuals or loci analysed. As in previous studies, a phylogenetic tree based on mtDNA (*a*; *COI* sequences) fails to identify genetic structure that aligns with the recognized species boundaries. Trees and PCAs plots are colour-coded according to their geographical location in map. Abbreviations in map correspond to locality information given in electronic supplementary material, table S1; see also electronic supplementary material, figure S3 for additional details on PCAs. (Online version in colour.)

resolve two divergent and well-supported reciprocally monophyletic groups (bootstrap support = 100%) that match the established species boundaries for *L. campechanus* and *L. purpureus* (figures 1*e* and 2*a*). There is no apparent pattern of geographical segregation within each clade, as individuals are not clustered in the SNP-based trees by populations/locations. Coalescent-based analyses using the BFD* method also provide overwhelming support in favour of the two species delimitation scenario (Bayes factors for two versus one species 2310.28–22 356.22; see details in electronic supplementary material, table S5). A list of diagnostic SNPs differentiating populations of *L. campechanus* from *L. purpureus*, which can be used for barcoding purposes, is given in the electronic supplementary material, table S6.

Population structure results based on fastStructure analyses of SNP data delimit the northern and southern lineages as separate units (figure 1*c*,*d*), with a best-fitting model supporting two meta-populations ($K = 2$). Although there have been recent concerns that structure analyses tend to be biased in favour of $K = 2$ [56], we note that our K scheme is consistent with the results inferred using multiple lines of evidence (figures 1*e* and 2; electronic supplementary material, table S5). In the PCAs of RAD loci, the first principal component accounts for 21–27% of the variation and is congruent with the separation of *L. campechanus* from *L. purpureus* (figure 2*b*; electronic supplementary material, figure S3). The second principal component represents 1–6% of the genetic variation, resulting in scattered populations on a cline that unveils fine-scale patterns of population structure according to geography (e.g. Veracruz-Tuxpan and Alabama-Louisiana define slope extremes of *L. campechanus*; the same is true for Guajira and Fortaleza in *L. purpureus*). While results using the 42 406 SNPs matrix (figure 2*b*; electronic supplementary material, figure S3d–f) show a much clearer species demarcation relative to the 15 112 SNPs matrix (figure 2*b*; electronic supplementary material, figure S3a–c), PCAs overall identify the same clustering patterns regardless of the number of SNPs analysed. Main variations on the observed genetic groups were influenced by the number of individuals contained ('min. sites' filter) in each PCA analysis, where PCAs generated with 'min. site 0.05' superimpose populations from both species that contained the highest amount of missing data (e.g. Yucatán; figure 2*b*; electronic supplementary material, figure S3a). These results emphasize that missing data can bias the results obtained with large RADseq datasets [57].

Weir and Cockerham F_{ST} values are substantially lower at intra- versus inter-specific levels (electronic supplementary material, figure S4). *Lutjanus campechanus* shows genetic differences between 0.0010 and 0.0119 in Tabasco-Veracruz and Veracruz-Apalachicola respectively, whereas *L. purpureus* presents a range of F_{ST} values from 0 in Fortaleza-Salvador and 0.0588 in Fortaleza-La Guajira. Negative and zero F_{ST} values were common across adjacent populations for both species, suggesting higher genetic differences at intra- rather than inter-population scales (i.e. individuals from adjacent locations may form a panmictic population [58,59]). By contrast, interspecific comparisons show substantially higher F_{ST} values, varying from 0.1240 among geographically closer populations (Veracruz-Nueva Esparta) to 0.3406 among the more distant comparisons (Fortaleza-Apalachicola).

Results of Mantel tests are marginally significant when the northern and southern lineages are each analysed in isolation ($p = 0.06$ for *L. campechanus*; $p = 0.02$ for *L. purpureus*; electronic supplementary material, figure S5a,b). No substantial association between genetic and geographical distances (IBD), however, is detected for Euclidean nor LCP distances, as the points do not form continuous linear plots (figure 1*f*; electronic supplementary material, figure S5c,d).

Admixture plots from fastStructure reveal introgression (figure 1*c*,*d*) at geographically neighbouring populations between the two species in northern South America. This result is further confirmed with the hybrid index estimated with gghybrid (electronic supplementary material, figure S6), which identifies admixture in geographically intermediate populations in Colombia (La Guajira), Venezuela (Nueva Esparta), and to some extent Brazil (Amapá). By contrast, none of the populations of *L. campechanus* reveal signs of ongoing introgression. Taken together, these results indicate a pattern of unidirectional interspecific introgression in *L. purpureus* from *L. campechanus*.

4. Discussion

(a) Species delimitation

Western Atlantic red snappers represent two commercially important species whose taxonomic status has been recently challenged. Described on the basis of morphological and meristic traits [23], *Lutjanus campechanus* and *L. purpureus* were recently considered to be conspecific based on assessments of genetic structure that examined mitochondrial DNA sequences and failed to delineate the formerly recognized species boundaries [24,25]. Despite finding concordant results with previous studies based on expanded mitochondrial *COI* and *D-loop* sequences (figures 1*a*,*b* and 2*a*), genome-wide analyses implementing SNP data (approx. 15 000–42 000) identify remarkable genetic divergences between the northern and southern red snappers. These results are supported by structure analyses (figure 1*c*,*d*,*f*), phylogenetic inferences (figures 1*e* and 2*a*), coalescent tests (electronic supplementary material, table S5), PCAs (figure 2*b*), and geographical patterns of population abundances (from FishNet and OBIS; electronic supplementary material, figure S1b), all of which are in agreement with the morphospecies delimitation and are largely robust to the number of individuals, SNPs, or missing data included in each of the data matrices analysed. The mito-nuclear discordance observed for WA red snappers can be the result of mitochondrial introgression or incipient sorting of mitochondrial haplotypes—the most likely biological sources of genealogical incongruence among recently diverged species [9]. Notably, a recent unpublished study that compared the otolith shape among different populations and species of WA red snappers using geometric morphometric approaches also identified well-differentiated and non-overlapping clusters that are consistent with the evolutionary units delineated here using genomic data [60].

While Mantel tests and correlogram analyses indicate that intraspecific populations separated by vast geographical distances have a smaller likelihood of gene flow, our analyses find tenuous support for a pattern of intraspecific IBD (figure 1*f*; electronic supplementary material, figure S5). We find evidence, however, of ongoing interspecific hybridization

(figure 1*c*,*d*) through circulating gene flow across geographically neighbouring locations in northern South America (figure 1*c*,*d*; electronic supplementary material, figure S6). Interspecific hybridization is not rare across sister species of marine fishes (e.g. *Haemulon maculicauda* and *H. flaviguttatum* [61]) and could also lead to mito-nuclear discordance via genetic introgression [18]. In the face of introgression, genetic structure is expected to reflect geographical patterns, particularly when sister species pairs become geographically isolated and subsequently come into secondary contact [17,61]. In this case, geographical patterns appear to support secondary admixture over incomplete lineage sorting (ILS), where ongoing hybridization leads to nuclear introgression [17,18].

Remarkably, the apparent direction of the aforementioned introgression (north to south) runs counter to the progression of marine currents in the Greater Caribbean (south to north). Although a northward route seems more plausible than the reverse, this is not entirely unexpected in light of the complex patterns of connectivity in the Greater Caribbean [62,63]. For instance, the progression of the lionfish invasion in the area has taken place southwards, from Florida to South America [64]. An alternative explanation to the observed pattern is that genetic structure reflects the maintenance of ancestral polymorphisms (ILS), possibly as a result of the recent divergence of the species. This interpretation, however, seems unfeasible considering that cline analyses (gghybrid) account for ancestral shared polymorphisms by focusing on loci that are highly differentiated in the parental reference populations. Another possibility involves secondary contact after divergence between the two species, at a time when *L. campechanus* co-occurred in the south, and either southern *L. campechanus* populations are now extinct or they have been diluted into a dominant *L. purpureus* genetic and demographic background. All these possibilities remain to be explored in greater depth using demographic and migration tests.

Although we were unable to examine samples from the western and northern Caribbean region (electronic supplementary material, figure S1), this does not necessarily represent a caveat in our study. The scarcity of records over time in the Bahamian, Eastern Caribbean, Greater Antilles, and Southwestern Caribbean marine ecoregions (regionalization according to [65]; electronic supplementary material, figure S1) suggest that the populations of red snappers are not completely established, possibly formed by vagrant individuals. This, in fact, reflects and actual gap in the connectivity of the two species that reinforces our observations and emphasizes a regional discontinuity pattern. Notwithstanding a worst-case scenario with well-established intermediate populations in the Caribbean and a smooth cline of admixture between the northern and southern lineages, the vast genomic divergences observed between these lineages provide strong evidence for the delimitation of two discrete taxonomic units. For instance, while low genetic differentiation values were estimated intraspecifically despite great geographical distances, the closest interspecific locations sampled feature high genetic divergences. Populations of *L. purpureus* from Nueva Esparta and São Salvador da Bahia are separated by an F_{ST} of 0.04 and an LCP of 4819 km; *L. campechanus* from Campeche and Florida have an F_{ST} of 0.003 and an LCP of 2953 km. These results are congruent with observed values proposing panmictic within-species populations [59,66]. Conversely, the corresponding interspecific values between Puerto de Tuxpan (northern red snapper) and La Guajira (southern red snapper) are 0.18 and 4847 km, respectively (electronic supplementary material, figure S4a).

It is important to note that we are not splitting species here based on genetic differences alone (e.g. [6]). Instead, we are testing the morphological and mitochondrial hypotheses in light of analyses based on thousands of genetic markers, with the former setting a century-long precedent on the validation of two species. In our present scheme, finding the cline between the two lineages would be difficult, as the lack of samples from intermediate locations precludes the determination of accurate geographical boundaries and the extent of the hybrid zone. Novel approaches allow the delimitation of species in the presence of gene flow [67]; however, these require gene trees as input, which is unfeasible using SNP data. Therefore, although we cannot confidently assert that these populations represent two valid species under the Biological Species Concept (BSC), they do represent two well-defined entities that match the Phylogenetic Species Concept (PSC)—the most commonly used criterion to delimit species in ichthyology [3]. Given that similar controversies exist about the specific taxonomic status of other living and extinct organisms (e.g. Neanderthal and Denisovan hominids [68]), debating whether these lineages fail to match particular aspects of species concepts can be a difficult and possibly futile endeavour. As Darwin notes: '. . .to discuss if they are rightly called species or varieties, before any definition of these terms has been generally accepted, is vainly to beat the air' [69, p. 49].

(b) Conservation implications

Red snappers represent some of the most economically important commercial and recreational fisheries in the Western Atlantic (WA), generating estimated annual revenues of over USD $27 million in the US alone [70]. Such fishing pressures have had an adverse effect on their populations, leading to heavily overfished stocks [71]. Delimitation of their species boundaries is imperative as the IUCN only lists the northern red snapper as Vulnerable (the southern red snapper has not been evaluated) [28]. Generally, an accurate evaluation of species—in particular commercially important and threatened species—represents the basic scientific knowledge required to determine conservation status that is assessed by international conservation organizations including the IUCN and the Food and Agriculture Organization of the United Nations (FAO), as natural species do not follow political delimitations. Even though the northern and southern stocks are managed by different legal entities, it is crucial to include genetic information as a baseline for planned stock enhancement [72] in multiple countries. Finally, correct delimitation of species also has implications for the enforcement of seafood mislabelling given that only *L. campechanus* has been traditionally allowed to use the US market name 'red snapper'; other species labelled as red snapper, including *L. purpureus*, are considered misbranded [73].

5. Conclusion

Our genome-wide analyses provided a strong signal of genetic differentiation among the northern and southern red snappers in the Western Atlantic, reconciling a long-standing

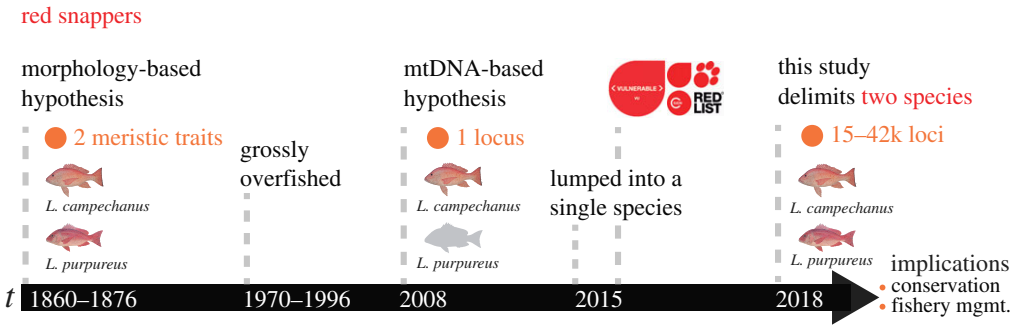

Figure 3. Timeline of the controversial species delimitation of WA red snappers and present resolution using genomic data. (Online version in colour.)

conflict of species delimitation between mtDNA and morphology (figure 3). These results highlight the importance of using powerful markers for addressing complex species delineation problems, particularly with organisms that rely on accurate recognition of species boundaries to inform their conservation status. We conclude that the two red snappers should be managed as separate taxonomic units. Our findings further emphasize the importance of implementing genomic approaches to settle species delimitation disagreements, where more conventional methods that lack discerning power may lead to the underestimation of biological diversity. These results ultimately align with observations from recent studies [15,16], which recommend that taxonomic decisions should strive to be conservative when based on single locus inferences, as those can be affected by incomplete lineage sorting or introgression in recent speciation events.

Data accessibility. Datasets available from the Dryad Digital Repository: https://doi.org/10.5061/dryad.sk61618 [74].

Authors' contributions. R.B.-R., G.G., R.S. and J.D. designed study; G.G., J.D., C.d.R.P.-M., R.S., L.L., O.D.-D., D.M., D.P., L.B., K.F., J.N.-B. and A.A. provided samples; C.d.R.P.-M. and R.B.-R. generated data; C.d.R.P.-M., R.B.-R., S.M.V.B. and A.M.-Y. ran analyses; C.d.R.P.-M., R.B.-R., S.M.V.B. and R.A.R.-V. drafted the manuscript; H.O.-Z. provided bioinformatics support; all authors contributed to the writing.

Competing interests. We declare we have no competing interests.

Funding. This project was funded by National Science Foundation grants NSF-DEB-1541491 and NSF-DEB-1457184, and the Fondo Institucional Para la Investigación (FIPI) of the University of Puerto Rico – Río Piedras grant 20FIP1470016 (R.B.-R.). Additional support was provided by a Consejo Nacional de Ciencia y Tecnología (CONACYT), doctorate scholarship 411629 (C.d.R.P.-M.), and the High Computing Performance Facility at UPR-RP (supported INBRE P20GM103475 grant to H.O.-Z.).

Acknowledgements. The authors are grateful to Y. Ortiz (Sequencing and Genotyping Facility at UPR-RP) for assistance in library preparation; to C. Loeffer, H. Flores-Quintana, E. Torres, F. del Moral, R. Lara and J. Sánchez for providing tissue samples; to K. Carpenter for useful comments on snapper taxonomy; to X. Velez-Zuazo for help with tree inference estimation and visualization; and to J.C. Bonilla for providing bioinformatics support. This study is contribution no. 485 of the Instituto de Estudios en Ciencias del Mar (Cecimar), Universidad Nacional de Colombia sede Caribe.

References

1. Thielsch A, Knell A, Mohammadyari A, Petrusek A, Schwenk K. 2017 Divergent clades or cryptic species? Mito-nuclear discordance in a *Daphnia* species complex. *BMC Evol. Biol.* **17**, 1–9. (doi:10.1186/s12862-017-1070-4)
2. Bickford D, Lohman DJ, Sodhi NS, Ng PKL, Meier R, Winker K, Ingram KK, Das I. 2007 Cryptic species as a window on diversity and conservation. *Trends Ecol. Evol.* **22**, 148–155. (doi:10.1016/j.tree.2006.11.004)
3. Victor BC. 2015 How many coral reef fish species are there? Cryptic diversity and the new molecular taxonomy. In *Ecology of fishes on coral reefs: the functioning of an ecosystem in a changing world* (ed. C Mora), pp. 76–87. Cambridge, UK: Cambridge University Press.
4. Pinzo JH, LaJeunesse TC. 2011 Species delimitation of common reef corals in the genus *Pocillopora* using nucleotide sequence phylogenies, population genetics and symbiosis ecology. *Mol. Ecol.* **20**, 311–325. (doi:10.1111/j.1365-294X.2010.04939.x)
5. Balasundaram SV, Engh IB, Skrede I, Kauserud H. 2015 How many DNA markers are needed to reveal cryptic fungal species? *Fungal Biol.* **119**, 940–945. (doi:10.1016/j.funbio.2015.07.006)
6. Fennessy J, Bidon T, Reuss F, Kumar V, Elkan P, Nilsson MA, Vamberger M, Fritz U, Janke A. 2016 Multi-locus analyses reveal four giraffe species instead of one. *Curr. Biol.* **26**, 2543–2549. (doi:10.1016/j.cub.2016.07.036)
7. Suchan T, Espíndola A, Rutschmann S, Emerson BC, Gori K, Dessimoz C, Arrigo N, Ronikier M, Alvarez N. 2017 Assessing the potential of RAD-sequencing to resolve phylogenetic relationships within species radiations: the fly genus *Chiastocheta* (Diptera: Anthomyiidae) as a case study. *Mol. Phylogenet. Evol.* **114**, 189–198. (doi:10.1016/j.ympev.2017.06.012)
8. Funk DJ, Omland KE. 2003 Species-level paraphyly and polyphyly: frequency, causes, and consequences, with insights from animal mitochondrial DNA. *Annu. Rev. Ecol. Evol. Syst.* **34**, 397–423. (doi:10.1146/annurev.ecolsys.34.011802.132421)
9. Mutanen M *et al.* 2016 Species-level para- and polyphyly in DNA barcode gene trees: strong operational bias in European Lepidoptera. *Syst. Biol.* **65**, 1024–1040. (doi:10.1093/sysbio/syw044)
10. Spinks PQ, Thomson RC, Shaffer BH. 2014 The advantages of going large: genome-wide SNPs clarify the complex population history and systematics of the threatened western pond turtle. *Mol. Ecol.* **23**, 2228–2241. (doi:10.1111/mec.12736)
11. Martin SH, van Belleghem SM. 2017 Exploring evolutionary relationships across the genome using topology weighting. *Genetics* **206**, 429–438. (doi:10.1534/genetics.116.194720)
12. Toffoli D, Hrbek T, Góes de Araújo ML, de Almeida MP, Charvet-Almeida P, Farias IP. 2008 A test of the utility of DNA barcoding in the radiation of the freshwater stingray genus *Potamotrygon* (Potamotrygonidae, Myliobatiformes). *Genet. Mol. Biol.* **31**, 324–336. (doi:10.1590/S1415-47572008000200028)
13. Roberts MA, Schwartz TS, Karl SA. 2004 Global population genetic structure and male-mediated gene flow in the green sea turtle (*Chelonia mydas*): analysis of microsatellite loci.

Genetics **166**, 1857–1870. (doi:10.1534/genetics.166.4.1857)

14. Milá B, Van Tassell JL, Calderón JA, Rüber L, Zardoya R. 2017 Cryptic lineage divergence in marine environments: genetic differentiation at multiple spatial and temporal scales in the widespread intertidal goby *Gobiosoma bosc*. *Ecol. Evol.* **7**, 5514–5523. (doi:10.1002/ece3.3161)
15. Corrigan S, Maisano P, Eddy C, Duffy C, Yang L, Li C, Bazinet AL, Mona S, Naylor GJP. 2017 Historical introgression drives pervasive mitochondrial admixture between two species of pelagic sharks. *Mol. Phylogenet. Evol.* **110**, 122–126. (doi:10.1016/j.ympev.2017.03.011)
16. Mateus CS, Stange M, Berner D, Roesti M, Quintella BR, Alves MJ, Almeida PR, Salzburger W. 2013 Strong genome-wide divergence between sympatric European river and brook lampreys. *Curr. Biol.* **23**, R649–R650. (doi:10.1016/j.cub.2013.06.026)
17. Weigand H, Weiss M, Cai H, Li Y, Yu L, Zhang C, Leese F. 2017 Deciphering the origin of mito-nuclear discordance in two sibling caddisfly species. *Mol. Ecol.* **26**, 5705–5715. (doi:10.1111/mec.14292)
18. Toews DPL, Brelsford A. 2012 The biogeography of mitochondrial and nuclear discordance in animals. *Mol. Ecol.* **21**, 3907–3930. (doi:10.1111/j.1365-294X.2012.05664.x)
19. Pie MR, Bornschein MR, Ribeiro LF, Faircloth BC, McCormack JE. 2017 Phylogenomic species delimitation in microendemic frogs of the Brazilian Atlantic Forest. *bioRxiv* **55**.
20. Martin SH *et al*. 2013 Genome-wide evidence for speciation with gene flow in *Heliconius* butterflies. *Genome Res.* **23**, 1817–1828. (doi:10.1101/gr.159426.113)
21. Van Belleghem SM, Vangestel C, De Wolf K, De Corte Z, Möst M, Rastas P, De Meester L, Hendrickx F. 2018 Evolution at two time frames: polymorphisms from an ancient singular divergence event fuel contemporary parallel evolution. *PLoS Comput. Biol.* **14**, 1–26. (doi:10.1371/journal.pgen.1007796)
22. Karanth KP. 2017 Species complex, species concepts and characterization of cryptic diversity: vignettes from Indian systems. *Curr. Sci.* **112**, 1320–1324.
23. Anderson WD. 2002 Lutjanidae. In *FAO species identification guides for fishery purposes: the living marine resources of the western central Atlantic* (ed. KE Carpenter), pp. 1479–1504. Rome, Italy: FAO.
24. Gomes G, Sampaio I, Schneider H. 2012 Population structure of *Lutjanus purpureus* (Lutjanidae—Perciformes) on the Brazilian coast: Further existence evidence of a single species of red snapper in the western Atlantic. *An. Acad. Bras. Cienc.* **84**, 979–999. (doi:10.1590/S0001-37652012000400013)
25. Gomes G, Schneider H, Vallinoto M, Santos S, Ortí G, Sampaio I. 2008 Can *Lutjanus purpureus* (South red snapper) be 'legally' considered a red snapper (*Lutjanus campechanus*)? *Genet. Mol. Biol.* **31**, 372–376. (doi:10.1590/S1415-47572008000200035)
26. Robertson DR, Van Tassell JL. 2015 Shorefishes of the Greater Caribbean online information system. Version 1.0 Smithsonian Tropical Research Institute, Balboa, Panama. See http://biogeodb.stri.si.edu/caribbean/en/pages (accessed 20 August 2005).
27. Fricke R, Eschmeyer WN, Van der Laan R (eds). 2019 Eschmeyer's catalog of fishes: genera, species, references. See http://researcharchive.calacademy.org/research/ichthyology/catalog/fishcatmain.asp (accessed 2 October 2019).
28. Anderson W, Claro R, Cowan J, Lindeman K, Padovani-Ferreira B, Rocha LA. 2015 *Lutjanus campechanus* (errata version published in 2017). The IUCN Red List of Threatened Species 2015: e.T194365A115334224.
29. Peterson BK, Weber JN, Kay EH, Fisher HS, Hoekstra HE. 2012 Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PLoS ONE* **7**, e37135. (doi:10.1371/journal.pone.0037135)
30. Andrews KR, Good JM, Miller MR, Luikart G, Hohenlohe PA. 2016 Harnessing the power of RADseq for ecological and evolutionary genomics. *Nat. Rev. Genet.* **17**, 81–92. (doi:10.1038/nrg.2015.28)
31. Weigt LA, Driskell AC, Baldwin CC, Ormos A. 2012 DNA barcoding fishes. In *DNA barcodes, Methods in Molecular Biology (Methods and Protocols)*, vol. 858 (eds W Kress, D Erickson), pp. 109–126. Totowa, NJ: Humana Press.
32. Catchen JM, Hohenlohe PA, Bassham S, Amores A, Cresko WA. 2013 Stacks: an analysis tool set for population genomics. *Mol. Ecol.* **22**, 3124–3140. (doi:10.1111/mec.12354)
33. Bradbury PJ, Zhang Z, Kroon DE, Casstevens TM, Ramdoss Y, Buckler ES. 2007 TASSEL: Software for association mapping of complex traits in diverse samples. *Bioinformatics* **23**, 2633–2635. (doi:10.1093/bioinformatics/btm308)
34. Danecek P *et al.* 2011 The variant call format and VCFtools. *Bioinformatics* **27**, 2156–2158. (doi:10.1093/bioinformatics/btr330)
35. Mastretta-Yanes A, Arrigo N, Alvarez N, Jorgensen TH, Piñeros D, Emerson BC. 2015 Restriction site-associated DNA sequencing, genotyping error estimation and de novo assembly optimization for population genetic inference. *Mol. Ecol.* **15**, 28–41. (doi:10.1111/1755-0998.12291)
36. Stamatakis A. 2014 RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312–1313. (doi:10.1093/bioinformatics/btu033)
37. Leaché AD, Oaks JR. 2017 The utility of single nucleotide polymorphism (SNP) data in phylogenetics. *Annu. Rev. Ecol. Evol. Syst.* **48**, 69–84. (doi:10.1146/annurev-ecolsys-110316-022645)
38. Clement M, Snell Q, Walker P, Posada D, Crandall K. 2002 TCS: Estimating Gene Genealogies. *Proc. 16th Int. Parallel Distrib. Process. Symp.* **2**, 184.
39. Leaché A, Fujita M, Minin V, Bouckaert R. 2014 Species delimitation using genome-wide SNP data. *Syst. Biol.* **63**, 534–542. (doi:10.1101/001172)
40. Bryant D, Bouckaert R, Felsenstein J, Rosenberg NA, RoyChoudhury A. 2012 Inferring species trees directly from biallelic genetic markers: bypassing gene trees in a full coalescent analysis. *Mol. Biol. Evol.* **29**, 1917–1932. (doi:10.1093/molbev/mss086)
41. Bouckaert R, Heled J, Kühnert D, Vaughan T, Wu CH, Xie D, Suchard MA, Rambaut A, Drummond AJ. 2014 BEAST 2: a software platform for Bayesian evolutionary analysis. *PLoS Comput. Biol.* **10**, 1–6. (doi:10.1371/journal.pcbi.1003537)
42. RStudio Team. 2015 RStudio: integrated development environment for R. See www.rstudio.com/products/RStudio.
43. Wickham H. 2009 *ggplot2: elegant graphics for data analysis*. New York, NY: Springer.
44. Raj A, Stephens M, Pritchard JK. 2014 FastSTRUCTURE: variational inference of population structure in large SNP data sets. *Genetics* **197**, 573–589. (doi:10.1534/genetics.114.164350)
45. Goudet J. 2005 HIERFSTAT, a package for R to compute and test hierarchical *F*-statistics. *Mol. Ecol. Notes* **2**, 184–186. (doi:10.1111/j.1471-8286.2004.00828.x)
46. Arnold B, Corbett-Detig RB, Hartl D, Bomblies K. 2013 RADseq underestimates diversity and introduces genealogical biases due to nonrandom haplotype sampling. *Mol. Ecol.* **22**, 3179–3190. (doi:10.1111/mec.12276)
47. Warnes GR. 2019 Gplots: various R programming tools for plotting data. 2.6.0 ed. See http://cran.r-project.org/web/packages/gplots/index.html.
48. Aguillon SM, Fitzpatrick JW, Bowman R, Schoech SJ, Clark AG, Coop G, Chen N. 2017 Deconstructing isolation-by-distance: the genomic consequences of limited dispersal. *PLoS Genet.* **13**, 1–27. (doi:10.1371/journal.pgen.1006911)
49. Diniz-filho JAF, Soares TN, Lima JS, Dobrovolski R, Landeiro VL, Pires M, Telles DC, Rangel TF, Bini LM. 2013 Mantel test in population genetics. *Genet. Mol. Biol.* **485**, 475–485. (doi:10.1590/S1415-47572013000400002)
50. Rousset F. 1997 Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. *Genetics* **145**, 1219–1228.
51. Pante E, Simon-Bouhet B. 2013 marmap: a package for importing, plotting and analyzing bathymetric and topographic data in R. *PLoS ONE* **8**, e73051. (doi:10.1371/journal.pone.0073051)
52. Jombart T. 2008 Adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics* **24**, 1403–1405. (doi:10.1093/bioinformatics/btn129)
53. Buerkle CA. 2005 Maximum-likelihood estimation of a hybrid index based on molecular markers. *Mol. Ecol. Notes* **5**, 684–687. (doi:10.1111/j.1471-8286.2005.01011.x)
54. Bailey RI. 2018 gghybrid: evolutionary analysis of hybrids and hybrid zones.
55. Fitzpatrick BM. 2013 Alternative forms for genomic clines. *Ecol. Evol.* **3**, 1951–1966. (doi:10.1002/ece3.609)

56. Janes JK, Miller JM, Dupuis JR, Malenfant M, Gorrell JC, Cullingham CI, Andrew RL. 2017 The K=2 conundrum. *Mol. Ecol.* **26**, 3594–3602. (doi:10.1111/mec.14187)
57. Leaché AD, Banbury BL, Felsenstein J, De Oca ANM, Stamatakis A. 2015 Short tree, long tree, right tree, wrong tree: new acquisition bias corrections for inferring SNP phylogenies. *Syst. Biol.* **64**, 1032–1047. (doi:10.1093/sysbio/syv053)
58. Wells RJD, Cowan JH, Fry B. 2008 Feeding ecology of red snapper *Lutjanus campechanus* in the northern Gulf of Mexico. *Mar. Ecol. Prog. Ser.* **361**, 213–225. (doi:10.3354/meps07425)
59. Silva R, Sampaio I, Schneider H, Gomes G. 2016 Lack of spatial subdivision for the snapper *Lutjanus purpureus* (Lutjanidae—Perciformes) from Southwest Atlantic based on multi-locus analyses. *PLoS ONE* **11**, e0161617. (doi:10.1371/journal.pone.0161617)
60. Marval-Rodríguez A, Renán-Galindo X, Montero-Muñoz J, Galindo-Cortés G, Jiménez-Baldillo M de L, Brulé T. Inter- and intraspecific differences of *Lutjanus campechanus* and *Lutjanus purpureus* in otolith shape. Presented at 71st Gulf and Caribbean Fisheries Institute Conference, The Royal Decameron Isleño Hotel, San Andres, Colombia, 5–9 November.
61. Bernal MA, Gaither MR, Simison WB, Rocha LA. 2017 Introgression and selection shaped the evolutionary history of sympatric sister-species of coral reef fishes (genus: *Haemulon*). *Mol. Ecol.* **26**, 639–652. (doi:10.1111/mec.13937)
62. Cowen RK, Paris CB, Srinivasan A. 2006 Scaling of connectivity in marine populations. *Science* **311**, 522–527. (doi:10.1126/science.1122039)
63. Cowen RK, Sponaugle S. 2009 Larval dispersal and marine population connectivity. *Ann. Rev. Mar. Sci.* **1**, 443–466. (doi:10.1146/annurev.marine.010908.163757)
64. Betancur-R R, Hines A, Acero PA, Ortí G, Wilbur AE, Freshwater DW. 2011 Reconstructing the lionfish invasion: insights into Greater Caribbean biogeography. *J. Biogeogr.* **38**, 1281–1293. (doi:10.1111/j.1365-2699.2011.02496.x)
65. Spalding MD *et al*. 2007 Marine ecoregions of the world: a bioregionalization of coastal and shelf areas. *Bioscience* **57**, 573–583.
66. Alva-Campbell Y, Floeter SR, Robertson DR, Bellwood DR, Bernardi G. 2010 Molecular phylogenetics and evolution of *Holacanthus* angelfishes (Pomacanthidae). *Mol. Phylogenet. Evol.* **56**, 456–461. (doi:10.1016/j.ympev.2010.02.014)
67. Jackson ND, Carstens BC, Morales AE, O'Meara BC. 2017 Species delimitation with gene flow. *Syst. Biol.* **66**, 799–812. (doi:10.1093/sysbio/syw117)
68. Gibbons A. 2011 A new view of the birth of *Homo sapiens*. *Science* **331**, 392–394. (doi:10.1126/science.331.6016.392)
69. Darwin C. 1859 *The origin of species*. London, UK: John Murray.
70. NOAA. 2016 Annual commercial landings by group. See https://www.st.nmfs.noaa.gov/st1/commercial/landings/annual_landings.html.
71. Marko PB, Lee SC, Rice AM, Gramling JM, Fitzhenry TM, McAlister JS, Harper GR, Moran AL. 2004 Mislabelling of a depleted reef fish. *Nature* **430**, 309–310. (doi:10.1038/430309b)
72. Garber AF, Tringali MD, Stuck KC. 2004 Population structure and variation in red snapper (*Lutjanus campechanus*) from the Gulf of Mexico and Atlantic Coast of Florida as determined from mitochondrial DNA control region sequence. *Mar. Biotechnol.* **6**, 175–185. (doi:10.1007/s10126-003-0023-7)
73. US Food and Drug Administration. 1980 Compliance Policy Guide 540.475 Snapper—Labeling. See https://www.fda.gov/ICECI/ComplianceManuals/CompliancePolicyGuidanceManual/ucm074504.htm.
74. Pedraza-Marrón C del R *et al*. 2019 Data from: Genomics overrules mitochondrial DNA, siding with morphology on a controversial case of species delimitation. Dryad Digital Repository. (doi:10.5061/dryad.sk61618)