

**Title: Rapid speciation via the evolution of pre-mating isolation  
in the Iberá Seedeater**

**Authors:** Sheela P. Turbek<sup>1\*</sup>, Melanie Browne<sup>2</sup>, Adrián S. Di Giacomo<sup>2</sup>, Cecilia Kopuchian<sup>2</sup>, Wesley M. Hochachka<sup>3</sup>, Cecilia Estalles<sup>4</sup>, Darío A. Lijtmaer<sup>4</sup>, Pablo L. Tubaro<sup>4</sup>, Luís Fábio Silveira<sup>5</sup>, Irby J. Lovette<sup>3,6</sup>, Rebecca J. Safran<sup>1</sup>, Scott A. Taylor<sup>1</sup>, and Leonardo Campagna<sup>3,6\*</sup>.

**Affiliations:**

<sup>1</sup>Department of Ecology and Evolutionary Biology, University of Colorado, Boulder, CO, USA.

<sup>2</sup>Centro de Ecología Aplicada del Litoral (CECOAL, CONICET), Corrientes, Argentina.

<sup>3</sup>Fuller Evolutionary Biology Program, Cornell Lab of Ornithology, Ithaca, NY, USA.

<sup>4</sup>Museo Argentino de Ciencias Naturales Bernardino Rivadavia (MACN, CONICET), Buenos Aires, Argentina.

<sup>5</sup>Museu de Zoologia da Universidade de São Paulo, São Paulo, Brazil.

<sup>6</sup>Department of Ecology and Evolutionary Biology, Cornell University, Ithaca, NY, USA.

\*Correspondence to: sheela.turbek@colorado.edu and lc736@cornell.edu

**Abstract:** Behavioral isolation can catalyze speciation and permit the slow accumulation of additional reproductive barriers between co-occurring organisms. We illustrate how this process occurs by examining the genomic and behavioral bases of pre-mating isolation between two bird species (*Sporophila hypoxantha* and the recently discovered *S. iberaensis*) that belong to the southern capuchino seedeaters, a recent, rapid radiation characterized by variation in male plumage coloration and song. Though these two species co-occur without obvious ecological barriers to reproduction, we document behaviorally-mediated species recognition and strong assortative mating associated with genomic regions underlying male plumage patterning. Plumage differentiation likely originated through the reassembly of standing genetic variation, indicating how novel sexual signals may quickly arise and maintain species boundaries.

**One Sentence Summary:** Novel mating signals restrict gene flow between co-occurring and closely related bird species.

**Main Text:** Organisms in the early stages of speciation provide an opportunity to understand the processes that govern reproductive isolation between taxa (1). Pre-mating isolation (*e.g.*,  
5 ecological or behavioral mechanisms that prevent individuals from interbreeding) is a powerful barrier that can separate sympatric species early in divergence (2–4). While post-mating barriers, such as genetic incompatibilities, take longer to accumulate than the time to speciation of many taxa (5, 6), learned or genetic preferences can diverge over shorter timescales and generate assortative mating (7–9), fueling rapid speciation and paving the way for the accumulation of  
10 additional reproductive barriers (3, 6, 10). Tracking mating decisions among wild populations early in speciation can improve our understanding of how behavioral isolation promotes divergence.

Southern capuchino seedeaters (*Sporophila*) are one of the most rapid avian radiations, showing remarkably low levels of ecological and genomic divergence (11, 12). Like Lake  
15 Victoria cichlids, where differences in male coloration promoted rapid diversification (13), the Neotropical southern capuchinos radiated within the last million years to form ten predominantly sympatric species that differ primarily in male plumage coloration and song (11, 12). Field experiments suggest that divergent male traits govern conspecific recognition and territorial defense (14). Nonetheless, viable hybrids between capuchino species are readily produced in the  
20 field (15) and captivity (16), suggesting a lack of genetic incompatibilities.

Here we take advantage of the identification of *S. iberensis* (the Iberá Seedeater), a newly described species from Iberá National Park, Argentina, where six other southern capuchinos co-occur during the breeding season (17), to study the importance of pre-mating

barriers early in speciation. *Sporophila iberensis* was first observed in October of 2001 (18), has a breeding range contained entirely within that of *S. hypoxantha* (Fig. 1A-B, Fig. S1), and breeds primarily in the northern portion of the Iberá wetlands (in the 111,000 hectare San Nicolás Reserve), where both species hold neighboring territories. Unlike its congeners, *S. iberensis* is increasing in local abundance (Fig. 1C, Fig. S1C). The species' small breeding range, combined with the fact that this region was unexplored from an ornithological perspective until the beginning of the 21st century due to a lack of public roads, suggests that *S. iberensis* likely already existed in the area and went undescribed (Supplementary Text). This is consistent with other southern capuchino species that have small and restricted breeding ranges (e.g., *S. melanogaster* and *S. nigrorufa* (Fig. S1A) (17)) and another more distantly related species in this taxonomically challenging genus (*S. beltoni*), which has a limited breeding range and was only recently identified in South America (19).

Throughout two breeding seasons, we located and monitored 128 nests of *S. hypoxantha* and *S. iberensis*, the only two southern capuchinos observed successfully breeding in the San Nicolás Reserve in Iberá National Park during the study (Fig. 1D-E; Table S1). We collected samples for genomic analyses from 80 nestlings and 126 adults and performed behavioral experiments of these two species to examine 1) the role of assortative mating in the maintenance of species boundaries, 2) the phenotypic traits underlying species recognition, 3) the genomic basis of such traits, and 4) the origin of these genomic variants.

## Results

### *S. hypoxantha* and *S. iberensis* show low genomic differentiation

We examined the degree of genomic divergence between these species using shotgun short-read whole-genome sequences from 16 individuals of *S. hypoxantha* and 21 individuals of *S. iberaensis* (20 males and 17 females; Table S2), identifying ~13.3 million single nucleotide polymorphisms (SNPs). The 42 SNPs showing the highest differentiation ( $F_{ST} > 0.85$ ,  $\max = 0.94$ ) were concentrated in three relatively narrow (30-50 kb) divergence peaks, which were located on separate chromosomes (1, 11, and sex chromosome Z; Fig. 1F) and exhibited increased absolute sequence divergence ( $D_{XY}$ ; Fig. S2). As among other capuchinos (11), *S. hypoxantha* and *S. iberaensis* are characterized by extremely low genomic differentiation (mean  $F_{ST} = 0.006 \pm 0.059$  SD) and no mitochondrial divergence (Fig. S3). However, individuals clustered by species in a genome-wide principal component analysis (PCA; Fig. S4). Individuals also clustered by species in separate PCAs performed with the SNPs from each peak (Fig. 1G-I), but only one peak completely differentiated the two species (scaffold 430; Fig. 1H). Despite being located on different chromosomes, the three regions showed high values of linkage disequilibrium within and among the peaks (Fig. S5), indicating their co-inheritance.

The lack of fixed differences (*i.e.*,  $F_{ST} = 1$ ) between *S. iberaensis* and *S. hypoxantha* among the 42 highly differentiated SNPs identified in our genome-wide  $F_{ST}$  analysis motivated us to search for the extent of shared variants between the species in the divergence peaks. First, we PCR-amplified and Sanger sequenced a ~700 bp region that included 15 of the 64 SNPs with  $F_{ST}$  greater than 0.79 within the peak on scaffold 430 (mean  $F_{ST} = 0.872 \pm 0.028$  SD), allowing us to assess the genetic variation in this region for a sample of 202 individuals. We observed 21 out of 200 *S. iberaensis* haplotypes that grouped with those of *S. hypoxantha* (Fig. 2). While each species had a common haplotype, a few *S. iberaensis* individuals carried the *S. hypoxantha* haplotype, and a small proportion of haplotypes appeared to be intermediate. Some of the

intermediate haplotypes belonged to *S. iberaensis* individuals and clustered with *S. hypoxantha*, yet we also observed intermediate haplotypes in *S. hypoxantha* that did not cluster with *S. iberaensis* (Fig. 2). We obtained similar results when conducting this haplotype-based analysis on the 37 individuals with whole-genome sequencing data for all the variants found in the peak on scaffold 430 (Fig. S6) and the SNPs showing the highest level of differentiation within the peaks on scaffolds 430 (Fig. S7) and 257 (Fig. S8, Supplementary Text). Taken together, these findings are consistent with *S. hypoxantha* variants segregating within *S. iberaensis* at the sites showing the highest differentiation between both species (and to a lesser extent in the reverse direction), either due to incomplete lineage sorting or past events of hybridization.

### **Divergence peaks contain plumage coloration genes**

We identified 12 genes within the divergence peaks (Fig. 1G-I; Table 1). Two peaks (scaffolds 257 and 430) contained genes known to be involved in melanic coloration (*TYRP1*, *OCA2*, and *HERC2*; Fig. 1G-I; Table 1) (20). Most highly differentiated SNPs (98%) were located in non-coding regions (Table S3), which may contain cis-regulatory elements that generate phenotypic variation (11). Although genes of small effect located outside of the divergence peaks could contribute to phenotypic differentiation, only 1.4% of SNPs in the genome had moderate  $F_{ST}$  values ( $F_{ST} > 0.2$ ; Fig. S9), suggesting that high differentiation is largely confined to the  $F_{ST}$  peaks.

### ***S. hypoxantha* and *S. iberaensis* mate assortatively in sympatry**

*S. hypoxantha* and *S. iberaensis* females do not show clear morphological characters that allow their identification to species. However, given that birds can detect wavelengths in the ultraviolet

range (300-400 nm) that are not perceived by humans (21), we examined the extent to which females of *S. hypoxantha* and *S. iberaensis* overlap in plumage coloration from an avian visual perspective. Benites *et al.* (22) detected coloration differences among females of four capuchino species that could be perceived by birds; however, we found that a large percentage of the convex hulls encompassing females of *S. iberaensis* in tetrahedral color space (a model of avian vision) were contained within those of *S. hypoxantha*, and the species largely overlapped in coloration across the avian visual spectrum (Fig 3). Therefore, we used the divergent genomic regions to identify females to species and quantify assortative mating. Paired males and females ( $N = 17$  pairs) clustered together in a tree based on whole-genome data (Fig. 4A) and shared the same  $F_{ST}$  peaks (Fig. S10), indicating a lack of hybrid pairs. We expanded this analysis to pairs for which we lacked whole-genome data ( $N = 23$ ) by using double-digest restriction site-associated DNA (ddRAD) sequencing to genotype all sampled males, females, and nestlings ( $N = 206$ ) at ~61.5 thousand SNPs. Despite the extremely shallow genomic differentiation between these species, individuals clustered into two groups in a PCA, matching the phenotype of the male attending each nest (Fig. 4B). This signal was derived mainly from the cumulative effect of SNPs with low  $F_{ST}$  values, as the ddRAD data only contained 28 SNPs that fell within the  $F_{ST}$  peaks identified from the whole-genome data (Fig. S11A) and showed the same pattern when those SNPs were excluded from the PCA (Fig. S11B).

Because mating outside of the social pair bond is common in birds (23), we also used 281 highly informative ddRAD loci to evaluate patterns of paternity. While the rate of extra-pair mating was very high ( $> 52\%$ ; 35/67 offspring with known social fathers), all extra-pair offspring that matched candidate fathers in the dataset were sired by males of the same species as their social father ( $N = 18$ ; Table S4). In addition, both social ( $N = 40$ ) and genetic pairs ( $N = 27$ )

clustered by species based on genomic PC1 score (Fig. 4C), indicating that assortative mating is maintained via both social and extra-pair mating.

### Species discrimination is based on plumage and song traits

5 *Sporophila hypoxantha* and *S. iberaensis* mate assortatively despite holding neighboring territories during the breeding season (Fig. 1E), breeding synchronously (Fig. S12), and foraging together on the same grasses (24). In addition to male plumage patterning, capuchinos differ in song (Fig. S13), a culturally transmitted trait acquired primarily through social learning in songbirds, though there is a genetic component of early song discrimination (25). Therefore, differences in male plumage patterning and song, rather than temporal or spatial barriers to reproduction, likely mediate mate choice and prevent interbreeding through genetic and/or imprinting mechanisms (4, 9, 14).

To test the roles of divergent plumage patterning and song in species recognition and pre-mating isolation, we presented territorial males of *S. hypoxantha* ( $N = 40$ ) and *S. iberaensis* ( $N =$  15 36) with all combinations of conspecific and heterospecific capuchino song and plumage (using song playback and artificial mounts, see Fig. 5A-C, Fig. S14), as well as that of a sympatric and ecologically similar heterospecific control (*S. collaris*), and assessed their behavioral responses. Across 240 trials (24 per treatment/species), we recorded aggressive behaviors and generated a response intensity score using PCA (Fig. S15). Each species responded most aggressively to the combination of conspecific song and plumage, exhibited intermediate responses to the treatments with mismatched traits, and largely ignored the heterospecific capuchino traits and those of the control species (Fig. 5D-E, Fig. S16, Table S5). Generalized linear mixed models confirmed that 20 both song and plumage are used to recognize sexual competitors, with significant effects on the

intensity of the males' response (song/plumage:  $P < 0.0001$ ) and attack behavior (song:  $P = 0.005$ , plumage:  $P = 0.012$ ) in both species (Table 2).

### **Existing mutations in novel combinations underlie the plumage phenotype of *S. iberensis***

5 To investigate the origin of the novel *S. iberensis* plumage phenotype, we examined genomic differentiation across the broader capuchino radiation (~28.2 million SNPs across 127 individuals from 12 species). We generated phylogenies using maximum likelihood for the entire genome and the regions containing divergence peaks. The whole-genome tree showed patterns consistent with recent speciation (Fig. 6A, Fig. S17), such as a lack of species-level monophyly possibly due to hybridization and incomplete lineage sorting, a result that was further supported 10 by demographic modeling (Supplementary Text, Fig. S18). Despite this phylogenetic uncertainty, *S. iberensis* formed a clade, as did most individuals from other species with restricted ranges (see *S. melanogaster* in green and *S. nigrorufa* in yellow; Fig. 6A, Fig. S17). In contrast, *S. iberensis* did not form a species-specific clade in the phylogenies derived from the regions containing divergence peaks (see arrows in Fig. 6B-C, Fig. S19), unlike most *S.* 15 *melanogaster* (indicated with a circle in Fig. 6B) and *S. ruficollis* (indicated with a circle in Fig. 6C) individuals. Although multiple species shared variants with *S. iberensis* at the individual divergence peaks (e.g., *S. ruficollis* in peak 257 and five other species in peak 430; Fig. 6B-C, Fig. S19), the particular combination found only in *S. iberensis* distinguished it from other 20 capuchinos (note that no other capuchino shares variants with *S. iberensis* at both divergence peaks; Fig. 6B-C, Fig. S20). This result implies that the *S. iberensis* phenotype likely arose through the reshuffling of standing genetic variation that already existed within the other

southern capuchinos, providing a mechanism for rapid speciation without the long period required for relevant mutations to arise *de novo* (26, 27).

## Discussion

5 Our findings point to pre-mating isolation through assortative mate choice, based on genetically inherited (plumage color) and culturally inherited traits (song, but see (25)), as a primary mechanism promoting divergence between these co-occurring capuchino species. While we never observed hybrid pairs during this study, selection against intermediate traits (plumage patterns or songs) or mismatched plumage and song traits in hybrids could further strengthen  
10 assortative mating through reinforcement (28). Most divergence peaks in capuchinos (11) and one of the three peaks between *S. iberensis* and *S. hypoxantha* are located on the sex chromosome Z. Loci on sex chromosomes are thought to have a disproportionate effect on hybrid fitness (large-Z effect (29)) and may have played a predominant role in the evolution of the southern capuchino radiation. Functional studies of specific variants in these divergent  
15 genomic regions will help clarify how novel allele associations could lead to different plumage phenotypes. Though the ultimate fate of the incipient *S. iberensis* species remains uncertain, our findings illustrate how phenotypically differentiated lineages can form and rapidly become reproductively isolated from co-occurring, syntopic species (30, 31). Our results suggest that the reshuffling of standing genetic variation can generate novel phenotypes that are targeted by  
20 sexual selection. Assortative mating based on these traits may maintain species boundaries early in speciation while subsequent reproductive barriers accumulate.

**Fig. 1. Geographic context and genomic characterization of the study species.** (A) Plumage phenotype and (B) breeding distribution of *S. hypoxantha* (red) (17) and *S. iberaensis* (blue circles are observations in the eBird database). The arrow indicates the study site location. (C) Increase in reporting rate probability for *S. iberaensis* in the eBird database (dotted lines are 95% prediction intervals for the estimated probabilities). (D) Typical breeding habitat. (E) Spatial distribution of 49/128 nests of *S. hypoxantha* and *S. iberaensis* found during the study. (F) Pattern of genomic differentiation between individuals of *S. hypoxantha* ( $N = 16$ ) and *S. iberaensis* ( $N = 21$ ). Divergence peaks are labeled according to their scaffold and corresponding chromosome in the zebra finch assembly. The plot contains the 733 largest scaffolds. (G-I) Genomic locations of individual SNPs with  $F_{ST} > 0.85$  on scaffold (G) 257, (H) 430, and (I) 762. Genes within 50-kb of these SNPs are depicted with arrows drawn to scale, with those involved in coloration highlighted in red. The insets show PCAs of the SNPs under the peaks.

**Fig. 2. Clustering of haplotypes obtained from the region of highest differentiation on scaffold 430.**

Phased genotypes of males, females, and nestlings of *S. hypoxantha* and *S. iberaensis* ( $N = 202$ ) for 15 highly divergent SNPs located in the peak on scaffold 430 generated from either whole-genome or Sanger sequence data (~700 bp). Each row represents a single chromosome, and each individual is represented twice in the tree. The four nucleotides are color-coded as indicated in the upper left corner. *S. hypoxantha* individuals and the majority of *S. iberaensis* birds have species-specific haplotypes. However, 17/100 (17%) *S. iberaensis* birds possessed one haplotype that clustered with *S. hypoxantha* and two *S. iberaensis* individuals (2%) clustered with *S. hypoxantha* on the basis of both haplotypes. The most common haplotype

for each species is indicated at the bottom of the two main clusters. For graphical clarity, identical copies of each of these common haplotypes were omitted from the tree.

**Fig. 3. Females of *S. hypoxantha* and *S. iberaensis* overlap in plumage coloration.**

5 Phenotype, degree of overlap, and reflectance patterns across the avian visual spectrum for the (A) crown, (B) throat, (C) belly, and (D) rump of *S. hypoxantha* ( $N = 22$ ) and *S. iberaensis* ( $N = 20$ ) females. The lines indicate mean reflectance for each group, the shaded areas depict the standard deviation, and the arrows indicate the location of each measured plumage patch. The gray polygons in the insets show the extent of overlap in tetrahedral color space between the two  
 10 species for the crown (43%), throat (72%), belly (58%) and rump (83%).

**Fig. 4. No evidence of hybridization through social or extra-pair mating.** (A) Whole-genome coalescent tree showing the relationship between males and females of *S. hypoxantha* and *S. iberaensis* ( $N = 37$ ). (B) PCA from double-digest restriction site-associated DNA (ddRAD) sequencing data depicting genomic differentiation between males, females, and nestlings of *S. hypoxantha* and *S. iberaensis* ( $N = 206$ ). (C) Genomic PC1 scores of males and females of *S. hypoxantha* and *S. iberaensis* for social ( $N = 40$ ) and genetic pairs ( $N = 27$ ; *i.e.*, pairs that fertilized within-pair or extra-pair offspring). For all plots, females were classified based on the phenotype of their social mate.  
 15  
 20

**Fig. 5. Territorial males of both species respond most aggressively to conspecific song and plumage.** Artificial mounts of (A) *S. iberaensis*, (B) *S. hypoxantha*, and (C) *S. collaris* (the control) alongside breeding males of the two capuchino species. Two mounts were created per

species for use in the behavioral experiment. **(D-E)** Behavioral response intensity (PC1) of territorial males of *S. hypoxantha* and *S. iberaensis*, respectively, to combinations of conspecific (CON), heterospecific capuchino (HET), and control (CONTROL) song and plumage. Different letters indicate statistical significance between treatment groups (Tukey HSD; adjusted  $P < 0.05$ ,  $N = 120$  per species).

**Fig. 6. *S. iberaensis* is monophyletic but shares variants at divergence peaks with other capuchinos.** Capuchino phylogeny inferred using maximum likelihood based on **(A)** whole-genome data and SNPs from the peaks on **(B)** scaffold 257 and **(C)** scaffold 430. The black square bracket and arrows indicate clades containing all or most *S. iberaensis* individuals and the black circles indicate clades of other species with species-specific variants at the peaks of differentiation. Outgroups are shown above the dashed line in the legend.

**Table 1. Regions of elevated genomic differentiation between the two species.**

Scaffold	Chr	Peak size (kb)	Highest $F_{ST}$ (over 5-kb window)	No. SNPs with $F_{ST} > 0.85$	No. genes (known function)	Coloration genes	Coloration gene function
257	Z	50	0.533	4	2 (2)	<i>TYRP1</i>	Enzyme involved in the production of melanin (32–35)
430	1	45	0.485	36	5 (2)	<i>OCA2</i> , <i>HERC2</i>	<i>OCA2</i> : Melanosomal transmembrane protein (36–38), <i>HERC2</i> : Contains a regulatory sequence that controls <i>OCA2</i> expression (39)
762	11	30	0.435	2	5 (3)	-	-

**Table 2. Species discrimination is based on both plumage and song.** Generalized linear mixed model results examining the behavioral responses of territorial males of *S. hypoxantha* and *S. iberaensis* to mount presentations and song playbacks when the heterospecific control (*S.*

*collaris*) trials were excluded. Significant results ( $P < 0.05$ ) are highlighted in bold. Plumage (conspecific vs. heterospecific) and song (conspecific vs. heterospecific) had a significant effect on response intensity regardless of whether outliers (observations outside 1.5 \* interquartile range) in each treatment group were included ( $N = 192$ ) or excluded ( $N = 179$ ; plumage:  $P < 0.0001$ , song:  $P < 0.0001$ ), while the species of the focal male did not affect behavioral response. We detected an additional significant interaction between song and plumage on response intensity ( $P = 0.03$ ) when outliers were removed, which could indicate a synergistic effect when both traits belong to the same species.

Response intensity (PC1) <sup>1</sup>	Estimate	Std. error	t value	P value
Intercept	-1.63	0.29	-5.70	<b>0.0007</b>
Species	-0.16	0.26	-0.62	0.535
Plumage	0.93	0.18	5.04	<b>&lt;0.0001</b>
Song	1.41	0.19	7.58	<b>&lt;0.0001</b>
Plumage × Song	0.24	0.26	0.92	0.360

Attack behavior <sup>2</sup>	Estimate	Std. error	z value	P value
Intercept	0.11	0.78	0.14	0.889
Species	1.19	1.00	1.20	0.232
Plumage	-2.76	1.10	-2.52	<b>0.012</b>
Song	-3.19	1.14	-2.79	<b>0.005</b>
Plumage × Song	-20.38	12802.17	-0.002	0.999

<sup>1</sup>Model included male ID (SD = 0.89; 95% CI of SD = 0.70-1.12) and female presence (SD = 0.26; 95% CI of SD = 0.03 – 1.21) as random effects.

<sup>2</sup>Model included male ID (SD = 2.90; 95% CI of SD = 1.34-6.28) as a random effect.

## Materials and methods

### Field methods

We have carried out extensive field work in Iberá National Park (Argentina) since 2007, encountering individuals of all seven southern capuchino seedeaters that breed in the region. In the San Nicolás Reserve (28° 07' 41.4" S, 57° 26' 04.7" W), where this study took place, our group has conducted studies on the breeding ecology of capuchinos since 2014. During the study, *S. hypoxantha* and *S. iberaensis* were the only two southern capuchino species observed

successfully breeding in San Nicolás. From November 2016 – January 2017 and October –  
December 2018, we located and monitored 128 nests of the two species (*S. hypoxantha*:  $N = 65$ ;  
*S. iberaensis*:  $N = 63$ ). We collected blood samples from the brachial veins of 126 adults and 77  
nestlings (*S. hypoxantha*:  $N = 40$  adult males, 23 adult females, 40 nestlings; *S. iberaensis*: 42  
5 adult males, 21 adult females, 37 nestlings), as well as tissue samples from two unhatched eggs  
of *S. hypoxantha* and one unhatched egg of *S. iberaensis* (Table S1). In addition, we collected  
feather samples from four plumage patches across the body (crown, throat, belly, and rump) of  
individuals of *S. hypoxantha* ( $N = 46$  males, 22 females) and *S. iberaensis* ( $N = 41$  males, 20  
females) to examine plumage coloration (described in the *Feather coloration* section). Males  
10 were attracted with playback and captured with mist nets during the nest construction, egg  
laying, and nestling provisioning stages, while females were captured at the nest during nestling  
provisioning. We measured and banded each individual with a numbered aluminum band and  
unique combination of colored leg bands prior to release. Blood samples were stored in lysis  
buffer and DNA was extracted with DNeasy blood and tissue kits (Qiagen, CA, USA) for all  
15 subsequent genomic analyses. From October – December 2019, we carried out an additional  
behavioral experiment in the San Nicolás Reserve (described in the *Behavioral experiment*  
section) to test the importance of song and plumage coloration in species recognition and pre-  
mating isolation between *S. hypoxantha* and *S. iberaensis*.

Capuchino seedeaters are austral migrants that breed in the Iberá wetlands and migrate  
20 northwards during the non-breeding season (40). Over the course of the breeding season in 2018,  
we re-sighted nine individuals (16%) of *S. hypoxantha* ( $N = 4$ ) and *S. iberaensis* ( $N = 5$ ) of the 56  
adults that were banded in 2016, two breeding seasons prior. In addition, in 2019, we re-sighted  
22 banded males (26%) of *S. hypoxantha* ( $N = 12$ ) and *S. iberaensis* ( $N = 10$ ) of the 86 males

banded from 2016-2018. Almost all of the re-sighted males were holding territories in the same geographic area of the study site as in previous breeding seasons. This relatively low recapture rate, but high philopatry, may be attributed to low inter-annual survival, and demonstrates high turnover in the individuals that are present at the breeding site across different years. The combination of a low recapture rate and the fact that females are indistinguishable to the human eye makes it difficult to quantify assortative mating by directly tracking the mating decisions of banded individuals across years (see *Assortative mating* section).

### ***Whole-genome resequencing and variant discovery***

We generated shotgun short-read whole-genome sequences for 37 individuals of *S. hypoxantha* ( $N = 8$  males, 8 females) and *S. iberaensis* ( $N = 12$  males, 9 females). Whole-genome resequencing generated over 860 million paired-end reads with a length of 151 bp, producing an expected per-individual coverage ranging between  $3.9\times$  and  $10.7\times$  (median:  $5.4\times$ ; Table S2).

We evaluated the quality of individual libraries with FastQC (version 0.11.7) and used AdapterRemoval (version 2.1.7) to trim adapter sequences, filter by quality, and merge overlapping paired-end reads (41, 42). The filtered data were aligned to a previously assembled reference genome of *S. hypoxantha* (11) using the ‘very-sensitive-local’ option in Bowtie2 (version 2.3.4), and alignment statistics were subsequently obtained using Qualimap (version 2.2.1) (43, 44). A high percentage of reads aligned to the reference genome (*S. hypoxantha*: mean =  $98.6 \pm 0.2\%$  SD, *S. iberaensis*: mean =  $98.6 \pm 0.1\%$  SD), and the average depth of coverage following filtering and alignment was  $5.6\times$  per sample (range:  $3.8\times$ - $10.3\times$ ; Table S2).

We used SAMtools (version 1.7) to convert SAM to BAM files and sort and index the data (45). We then marked PCR duplicates with Picard Tools (version 2.17.10) and used

HaplotypeCaller in GATK (version 3.8.0) to perform SNP variant discovery and genotyping (46, 47). The following hard filtering parameters were used to exclude variants in GATK:  $QD < 2.0$ ,  $FS > 60.0$ ,  $MQ < 20.0$ , and  $ReadPosRankSum < -8.0$ . We additionally used VCFtools (version 0.1.13) to filter out variants that had a minor allele frequency of less than 8% (retaining alleles present in at least three homozygous individuals), a mean depth of coverage lower than 2 or greater than 50, more than 20% missing data, or were not biallelic (48). This pipeline produced 13,254,970 SNPs, and the average percent of missing data was 5% per individual.

### ***Population genomics***

To search for regions of elevated differentiation between *S. hypoxantha* and *S. iberaensis*, we computed average  $F_{ST}$  values for non-overlapping 5-kilobase (kb) windows, as well as individual SNPs, using VCFtools. Peaks of divergence were identified as 5-kb windows with elevated genomic divergence that contained at least one individual SNP with  $F_{ST}$  greater than 0.85. This criterion focused on the strongest putative targets of divergent selection, though it may have excluded regions under selection that contained genes of small effect. We identified three peaks of divergence between the two species by building Manhattan plots and conducting principal component analyses (PCAs) of the genomic data with the packages *qqman* and *SNPRelate* in R version 3.5.2 (49–51). In addition, we estimated  $D_{XY}$ , an absolute measure of divergence, over non-overlapping 5-kb windows for the three peaks of divergence using the custom script `popgenWindows.py` ([https://github.com/simonhmartin/genomics\\_general](https://github.com/simonhmartin/genomics_general)). We explored patterns of linkage disequilibrium in these regions by calculating the  $r^2$  statistic in plink version 1.9 (52).

The reference genome was assembled to contigs from short-read shotgun and mate-pair libraries, and subsequently assembled to scaffolds using long-read data from Pacific Biosciences

sequencing (11). Although some of these scaffolds are large, they are not assembled to chromosome level. To assign scaffolds with divergence peaks to chromosomes, we aligned them to the zebra finch assembly (*Taeniopygia\_guttata*-3.2.4) using the Satsuma synteny model from Satsuma version 3.1 (53). We assigned each scaffold to the chromosome with the top hit and examined the results with MizBee (54). Finally, we referred to the annotated *S. hypoxantha* genome in (11) to compile a list of genes within 50 kb of each divergence peak and searched for these annotations of interest in the UniProt (<https://www.uniprot.org/>) and Human Gene databases (<https://www.genecards.org>) to identify genes in the regions of elevated differentiation. The gene *OCA2* is adjacent to *HERC2* in the zebra finch, yet was not annotated in the *S. hypoxantha* reference genome. We located the *OCA2* coordinates in our reference genome by aligning the zebra finch mRNA (XM\_032749285.1) using BLAST (55). To search for areas that could play an important role in regulating the expression of *OCA2* (*i.e.*, cis-regulatory elements), we assessed the level of conservation of the intergenic region between *OCA2* and *HERC2*, the area of the genome showing the highest differentiation between *S. iberaensis* and *S. hypoxantha*, with respect to more distantly related birds by using the Bird PhastCons track (56) from the UCSC (University of California Santa Cruz) genome browser (57). Bird PhastCons scores are derived from a multigenome alignment of the budgerigar (*Melopsittacus undulatus*), zebra finch (*Taeniopygia guttata*), chicken (*Gallus gallus*), and turkey (*Meleagris gallopavo*) genomes, and represent the probability that a nucleotide belongs to a conserved element (ranging from 0 to 1). Areas that are highly conserved among distantly related species may contain regulatory elements that are important for controlling gene expression. We aligned the ~37 kb of sequence to the medium ground finch (*Geospiza fortis*) reference genome (geoFor1; to which the PhastCons scores were mapped) using BLAT (58) with a 96.4% identity.

In addition, we assembled full mitochondrial genomes from the filtered whole-genome sequences belonging to *S. hypoxantha* and *S. iberaensis* individuals with MITObim 1.9.1 (59), using the “quick” option and up to 40 iterations with the full mitochondrial genome from *Geospiza magnirostris* as a template (GenBank number NC\_039770.1). We aligned the 37 individual sequences with an average length of 16,562 bp in Geneious version 10.1.3 (60) and subsequently constructed an unrooted statistical parsimony network using PopART 1.7 (61). In addition, we used the same methodology to generate a network based on the recovered COI DNA barcodes, which are frequently used for species identification (62–64).

To generate phylogenetic hypotheses defined by the variants within the three divergence peaks in the context of the entire capuchino radiation, and compare these relationships to those in a whole-genome phylogeny of all capuchino species, we increased our genomic sampling to the ten southern capuchino species plus two outgroups. We combined the 37 whole-genome sequences obtained from *S. hypoxantha* and *S. iberaensis* with 72 additional individuals from nine capuchino species previously sequenced and published by Campagna *et al.* (11), 12 new individuals from *S. ruficollis* sequenced on a lane of Illumina NextSeq 500 (paired-end, 151 bp), and six additional individuals sequenced on a lane of Illumina NextSeq 500 (mid-output mode, paired-end, 151 bp). All additional sequencing was performed at the Cornell University Biotechnology Resource Center (BRC). We assembled a VCF file (as described above in the *Whole-genome resequencing and variant discovery* section) with a total of 127 individuals (28 *S. hypoxantha*, 21 *S. iberaensis*, 15 *S. ruficollis*, 12 *S. pileata*, 12 *S. palustris*, 12 *S. melanogaster*, 12 *S. nigrorufa*, 4 *S. bouvreuil*, 4 *S. hypochroma*, 3 *S. cinnamomea*, and 2 individuals each of *S. minuta* and *S. castaneiventris* as outgroups). We applied the same hard filters as described above and subsequently retained variants that were present in 80% of all individuals, had a depth of

coverage between 4 and 50, and a minor allele count of at least four. This combined dataset contained 32,993,511 SNPs after filtering. We explored the relationships among individuals and species by performing a PCA in *SNPRelate* and used VCFtools to create three additional files with subsets of SNPs from the regions defined by each divergence peak (1760 SNPs for the peak on scaffold 430, 1040 SNPs for the peak on scaffold 257, and 13 SNPs for the peak on scaffold 762). We used RAxML version 8.2.4 (65) to produce maximum likelihood phylogenies from the variants of each of the three divergence peaks, implementing the “ASC\_GTRGAMMA” model and the Lewis correction for ascertainment bias. For the more computationally demanding whole-genome phylogeny, we used RAxML-ng version 0.9.0 (66) and the “GTR+G+ASC\_LEWIS” model. RAxML-ng used the 28.2 million SNPs that had the minor allele in homozygosity in at least one individual (from the total of ~33 million variants in the dataset). This analysis ran for approximately 1600 clock hours on all 64 cores of a computer with 512 Gb of RAM. Despite not converging on a single best phylogeny, an inspection of trees from the final search rounds showed very little variation, with only minor changes at the tips of the tree. We therefore generated a smaller dataset by applying more stringent filtering parameters (85% of individuals present at a locus and a minimum minor allele frequency of 10%), which retained 6,283,771 SNPs. We ran RAxML-ng on this dataset under the same conditions as described above, except that we used a parsimony tree as a starting point. This strategy converged on a single best tree, and both datasets (28.2 million and 6.3 million SNPs) produced comparable topologies.

### ***Double-digest restriction site-associated DNA sequencing***

To determine species identity for individuals without whole-genome data, assign paternity, and analyze patterns of assortative mating, we sequenced 206 individuals (126 adults, 77 nestlings, and 3 unhatched eggs from 23 nests of *S. hypoxantha* and 20 nests of *S. iberaensis*) in two separate sequencing runs following the double-digest restriction site-associated DNA (ddRAD) sequencing protocol detailed in (67).

We used FASTX-Toolkit ([http://hannonlab.cshl.edu/fastx\\_toolkit/](http://hannonlab.cshl.edu/fastx_toolkit/)) to trim the 3' end of all reads to a length of 97 bp (FASTX Trimmer) and eliminated sequences (FASTX Quality Filter) if at least one base had a Phred score below 10 (90% call accuracy) or more than 5% of the bases had a score below 20 (99.9% call accuracy). We aligned the ddRAD data to the reference genome of *S. hypoxantha* using the 'sensitive' option in Bowtie2 (version 2.3.5) and sorted and indexed the data with SAMtools (version 1.9). We then used the *gstacks* and *populations* modules of Stacks (version 2.3) to call variants and remove loci that were present in fewer than 80% of individuals (68). The effective per-sample coverage was  $31.8 \times \pm 15.1 \times$  (mean  $\pm$  SD). This pipeline produced a VCF file containing 61,484 SNPs across the 206 individuals.

### ***Sanger sequencing of a region within the divergence peak on scaffold 430***

To investigate the genomic architecture of phenotypic differences between the species in more detail, we developed a pair of primers (forward: 5'-ATTGCTGGTGTCTCCTTATTGA-3'; reverse: 5'-ATGTCCCTTTGGCTGTCTG-3') to sequence a ~700-bp region on scaffold 430 (11,028,673 to 11,029,376 bp) that contained 12 highly divergent SNPs ( $F_{ST} > 0.85$ ) and an additional three SNPs with  $F_{ST} > 0.79$ . We amplified the divergent region via PCR for 165 individuals ( $N = 87$  adults, 77 nestlings, and one unhatched egg) with GoTaq colorless master mix (Promega, WI, USA) and the following thermal cycle profile: 3 min at 95 °C, followed by

25 cycles of 30 sec at 95 °C, 30 sec at 61 °C, and one min at 72 °C, and finally 5 min at 72 °C.

The PCR product was Sanger sequenced in both directions with the same primers used for amplification at the Cornell University BRC.

We used Unipro UGENE version 1.32.0 to trim primers and edit the Sanger sequences (69) and combined the information from these sequences with variants obtained through whole-genome sequencing to determine the genotypes of 202 individuals at 15 SNPs that showed high levels of differentiation within the peak on scaffold 430. We calculated  $F_{ST}$  values at each site using VCFtools and subsequently phased and imputed missing data (~5.2% or 158 out of 3060 genotypes, with a mean probability of  $0.996 \pm 0.0036$  SD) using BEAGLE version 3.3.2 (70). This resulted in 404 haplotypes, two per individual. We explored the relationships between individuals at these sites by calculating a distance matrix in the R package *vegan* (71) and plotting it with the function *phylo.heatmap()* from the R package *phytools* (72). We also compared these results to three similar plots derived from the *S. iberaensis* ( $N = 21$ ) and *S. hypoxantha* ( $N = 16$ ) individuals for which we had whole-genome sequencing data. We produced one plot for all SNPs found in the divergence peak on scaffold 430, one plot for the 64 SNPs with  $F_{ST} > 0.79$  found in the same region, and a third plot for the 13 SNPs with  $F_{ST} > 0.79$  found in the divergence peak on scaffold 257. We used an  $F_{ST}$  cutoff of 0.79, as the segment selected for PCR amplification included SNPs with this level of divergence.

### ***Assortative mating***

We analyzed patterns of social pairing from the whole-genome data ( $N = 17$  social pairs) by first creating a tree of individuals using SVDquartets, implemented in PAUP\*. SVDquartets is a coalescent-based method that compares possible quartet topologies for a set of four taxa,

selecting the topology with the lowest score (73). In addition, we used the R package *qqman* to create Manhattan plots comparing the level of differentiation between species for males and females independently in each of the three divergence peaks. We evaluated whether individuals that formed a social pair grouped together on the tree, as expected if assortative mating contributes to reproductive isolation between *S. hypoxantha* and *S. iberaensis*, and showed elevated levels of differentiation in the same genomic regions.

For all social pairs ( $N = 40$ ), including those with whole-genome data, we used the ddRAD pipeline to assign individuals to species and calculate the number of observed conspecific and heterospecific pairings. Specifically, we conducted a PCA of the genomic data using the *SNPRelate* package in R and evaluated if 1) individuals clustered by species in a PCA, and 2) socially or genetically determined male-female pairs (see *Paternity analysis* section) grouped together by species based on their diagnostic genomic PC1 scores.

### ***Paternity analysis***

We further filtered the VCF file from the ddRAD pipeline using the *populations* module of Stacks (version 2.3) to remove loci that had a minor allele frequency of less than 0.25, an observed heterozygosity greater than 0.7, or were present in fewer than 95% of individuals. We restricted the analysis to the first SNP per locus and used VCFtools to remove loci that were not in Hardy-Weinberg equilibrium or had a mean depth of coverage below 20. This pipeline produced a VCF file containing 281 highly informative loci across the 206 individuals that we used for paternity analysis.

Following filtering, we converted the VCF file to a format compatible with CERVUS 3.0.7, which takes a likelihood approach to assign paternity from SNP data (74). CERVUS

calculates the natural logarithm of the likelihood ratio (LOD score) for each potential pairing by comparing offspring genotypes to the genotypes of candidate parents and random individuals in the population. The LOD score thus estimates the relative likelihood that a sampled offspring was sired by a candidate father rather than a random male in the population. In addition, the program conducts a simulated parentage analysis using population allele frequencies and the proportion of candidate parents sampled in the dataset to calculate the critical differences in LOD scores necessary to assign paternity with either 80% or 95% confidence.

To determine critical LOD scores, we simulated paternity assignments for 100,000 offspring (the recommended number) using the following parameters: 122 candidate males, 67% of candidate males sampled, and the default of 1% of loci mistyped. We approximated the proportion of candidate males sampled by estimating the number of males of both species that held neighboring territories to the sampled males but were never caught. The proportion of typed loci for the simulation was 0.972. As known mothers (confirmed by catching females at the nest) were sampled for 88% of offspring, we included known mothers in the analysis and evaluated CERVUS assignments using trio LOD scores, which take into account potential genotyping errors and the genotypes of known mothers when assigning paternity. Our total sample included 82 candidate males. We accepted assignments if the number of mismatches between the assigned male and his offspring was less than or equal to the maximum observed number of mismatches between a mother and her known offspring (as in (67); max = 8, < 3% of 281 loci; Table S4). We assigned 51 of 77 nestlings (66%) to a candidate father with 95% confidence.

### ***Feather coloration***

We collected feathers from four plumage patches across the body (crown, throat, belly, and rump) from 68 individuals of *S. hypoxantha* (46 males and 22 females) and 61 individuals of *S. iberaensis* (41 males and 20 females) to examine plumage coloration. We stacked 10-15 feathers from each plumage patch on a non-reflective background surface (Flock Paper, Edmund Optics) to mimic their placement on the body of the bird. Reflectance data were generated relative to a white standard (WS-1-SL, Ocean Optics) and a dark standard (all light omitted) with an Ocean Optics Flame spectrometer connected to an Ocean Optics PX-2 pulsed xenon light source. We used the OceanView software package (version 1.6.7, Ocean Optics) to record the reflectance data, averaging 20 scans per measurement. For each plumage patch, we took three measurements per individual and averaged the measurements prior to subsequent analysis. We used the R package *pavo* to compare the reflectance curves and degree of overlap in tetrahedral color space for each plumage patch between individuals of *S. hypoxantha* and *S. iberaensis* (75).

### ***Behavioral experiment***

From October – December 2019, we located males of *S. hypoxantha* ( $N = 40$ ) and *S. iberaensis* ( $N = 36$ ) that were actively singing on their territories in the San Nicolás Reserve and carried out a behavioral experiment in which we presented them with the following five treatments: (1) conspecific mount and song, (2) heterospecific capuchino mount and conspecific song, (3) conspecific mount and heterospecific capuchino song, (4) heterospecific capuchino mount and heterospecific capuchino song, and (5) heterospecific control mount and song. While *S. hypoxantha* and *S. iberaensis* form social pairs from October-November, our paternity data indicate that extra-pair mating continues throughout December in the San Nicolás Reserve. Male responses to mount presentation and song playback are often used to infer the importance of pre-

mating isolation between divergent taxa (76–80), as numerous studies have found that the traits used by males to recognize sexual competitors are also employed in female mate choice (81–83). For the heterospecific control, we followed the methods of (14) and used *S. collaris*, which is closely related to our focal species but not a capuchino seedeater (14, 84). *S. collaris* breeds in sympatry and occupies a very similar ecological niche to *S. hypoxantha* and *S. iberensis* (85, 86). The heterospecific control treatment thus attempts to discriminate between aggressive responses to ecological and sexual competitors, given that all three species are grassland birds that feed on the seeds of tall grasses, such as *Paspalum durifolium* (Poaceae) and *Andropogon lateralis* (Poaceae), which dominate the landscape in the San Nicolás Reserve (24, 87). In particular, an elevated response to conspecific traits relative to the stimuli of the heterospecific capuchino and control would indicate that 1) capuchinos recognize members of their own species as sexual competitors, and 2) the conspecific traits that elicit an elevated response are involved in male-male competition and potentially female choice (81–83). In contrast, a similarly aggressive response to conspecific and heterospecific capuchino stimuli would suggest that capuchinos do not discriminate between *S. hypoxantha* and *S. iberensis*, recognizing males of both species as sexual and/or ecological competitors. Finally, an aggressive response to the control *S. collaris* stimuli would suggest that this more distantly related species, which is not a sexual competitor, elicits a response because it is recognized as an ecological competitor.

We recorded the geographic coordinates of each trial and tested focal males with as many treatments as possible (up to five treatments) by returning to the same geographic location multiple times. Trials performed with the same focal males were separated by at least one day, and the order in which treatments were presented was randomized. In addition, we randomized the order in which stimuli were presented across trials and ensured that the mounts (two of *S.*

*hypoxantha*, two of *S. iberensis*, and two of *S. collaris*) and playback files (10 of *S. hypoxantha*, 10 of *S. iberensis*, and five of *S. collaris*) were presented an equal number of times. Sixteen of the 76 focal males (21%) were color-banded from our field work in previous years. In addition, capuchino seedeaters exhibit a considerable degree of intraspecific variation in plumage coloration, which is likely associated with age, and only sing within their territories. Thus, by returning to the same location where a male was previously observed singing, and using color bands or plumage to identify individuals, we could be confident that the same individual was tested in subsequent trials.

During each trial, we located the focal male and set up the mount approximately 1-2 m off the ground < 35m from the focal male on a thin pole near vegetation suitable for perching. We hid a compact speaker (JBL Flip 5) in the vegetation under the mount and connected the speaker to a phone through Bluetooth to start the playback recordings. Each trial lasted a total of five minutes (the duration of the playback file), with the same observer (always SPT, for consistency in scoring behavioral responses) standing 20 meters away from the mount. We generated video recordings of each trial with a DSLR camera (Canon EOS 7D) and dictated vocalizations and behaviors into the camera during the trials. We recorded the following behavioral responses: the number of flights and amount of time spent at various distances from the mount, the number of attacks and amount of time spent attacking the mount, and the amount of time spent singing by each focal male, using the 2m pole on which the mount was placed to estimate distance from the mount. While females do not assist with territorial defense in capuchino seedeaters, we noted whether or not a female was observed during each trial in case female presence influences male response to territorial intrusion. Females were observed in 53 trials (22%). In total, we presented 32 males (16 *S. hypoxantha* and 16 *S. iberensis*) with all five

treatments and 44 males (24 *S. hypoxantha* and 20 *S. iberaensis*) with fewer than five treatments, for a total of 240 trials (24 trials per treatment for each species).

We ran a PCA on the correlation matrix of the behavioral response variables using the R package *vegan* to reduce the dimensionality of the behavioral data. The PCA identified three axes of variation (eigenvalues > 1) that collectively explained 79% of the variation in behavioral responses (PC1: 43%, PC2: 20%, PC3: 16%). All input variables associated with male aggression (*e.g.*, number of flights near the mount, proportion of time spent near the mount, and number of attacks at the mount) loaded positively on PC1, while proportion of time spent singing and proportion of time spent over six meters from the mount loaded negatively on PC1 (Fig. S15C), indicating that PC1 represented a reasonable overall summary of aggression. We therefore extracted PC1 to generate a response intensity score for each trial. In addition, we classified all trials as displaying attack behavior ('1') or not ('0'), with attack behavior defined as either swooping at or making direct contact with the mount, to examine a direct indicator of aggression. We carried out parallel analyses with response intensity (*i.e.*, PC1) and attack behavior as dependent variables using R version 3.5.2 (51) and fit generalized linear mixed models (GLMMs) with the R packages *lme4* (for linear mixed models) and *glmmTMB* (for mixed logistic regression models) to analyze the responses of territory owners to mount presentation and song playback (88, 89).

Because territorial males typically did not respond to the heterospecific control (*S. collaris*) stimuli, we ran separate analyses with and without the control trials (as in (90)). Excluding the control trials, we first ran GLMMs examining the effects of species, plumage, and song (fixed effects) on 1) response intensity (PC1) and 2) attack behavior (whether or not the mount was attacked at any point during the trial). We used a mixed logistic regression model

with a binomial distribution and logit link function to model attack behavior, which had a binary outcome (0 or 1). Preliminary models included treatment order, male ID, female presence (0 or 1), mount ID, and playback ID as random effects, with mount ID and playback ID nested within each plumage and song type, respectively, in order to control for repeated measures from individuals and mount/playback exemplar effects. We calculated a 95% confidence interval around the estimated standard deviation explained by the random effects using the *confint()* function from the R stats package. We excluded random effects from the model if the lower end of the confidence interval reached zero, indicating that the effect did not account for variation in the model (e.g., treatment order, mount ID, and playback ID), retaining male ID in the model of attack behavior and male ID and female presence in the model of response intensity.

In addition, we included the heterospecific control trials to run a GLMM for each focal species that tested the effect of treatment group on response intensity (PC1), incorporating male ID and female presence as random effects. We used the R package *emmeans* to run post-hoc pairwise comparisons between treatment groups using Tukey's honestly significant difference (HSD) test (91). Again, treatment order, mount ID, playback ID, and female presence (in the case of *S. hypoxantha*) had 95% confidence intervals that reached zero when included as random effects in the preliminary models and were therefore excluded from the final analyses. We verified the assumptions of the linear mixed models by generating Q-Q plots and plotting the residuals versus the fitted values.

### ***Abundance estimates***

To estimate whether *S. iberaensis* has increased in abundance across its breeding range since the species' first records in 2001 (18), we downloaded eBird data through January 2020 from the

February 2020 release of the eBird Basic Dataset (92). The downloaded dataset contained all available data for the control used in the behavioral experiment (*S. collaris*) and the seven capuchino species (*S. iberensis*, *S. hypoxantha*, *S. cinnamomea*, *S. palustris*, *S. pileata*, *S. ruficollis*, and *S. hypochroma*) that breed in Iberá National Park (17), as well as the Sampling Event Data (needed to infer non-detection records). eBird is an online database where scientists, researchers, and amateur naturalists can upload avian observations (93). We filtered the data for each species using the R package *auk* (94) to exclude records from incomplete checklists (*i.e.*, checklists in which some identified species were not reported) and observations that fell outside of a bounding box around the area encompassing all observations of *S. iberensis* in Argentina and Paraguay (the central range of *S. iberensis*), retaining only a single checklist from each set of non-independent (“shared”) checklists. We then inferred non-detection records (*i.e.*, “zero-filled the data”) using the *auk* package to create presence/absence data for each species. To more precisely define the spatial area of interest, we converted the presence-only data from *S. iberensis* into a spatial object and generated a convex hull polygon around the distribution of *S. iberensis* records using the function *gbuffer()* in the R package *rgeos* (95), adding a buffer of 1 map unit (with data in a South America Albers Equal Area Conic projection) outside of these locations; we only retained records of observations that fell within this polygon of interest. We then placed temporal restrictions on the remaining data, only retaining records from October-February, when capuchino seedeaters are present on the breeding grounds, and records beginning with the austral summer that spanned the years 2013-2014, the first summer for which multiple observations of *S. iberensis* existed in the eBird database.

After processing the data, we examined whether there have been any systematic changes in the reporting rates of the seven capuchino seedeater species over the past decade in order to

assess whether *S. iberaensis* has increased in prevalence relative to other capuchino species. We modeled changes in prevalence by fitting generalized additive models (GAMs; using R package *mgcv* (96)), in which the probability of reporting of the focal species was modeled as a function of the calendar year at the end of each austral summer. GAMs are able to identify arbitrary, continuous patterns of change through time, rather than forcing specific patterns onto the data. We used additional smoothing terms to account for variation in observation effort as described by the following variables: observation date, distance traveled, and duration of the observation period. We compared the GAM results for each capuchino species to determine whether 1) there was a significant change in the reporting probability of each species over time and 2) whether reporting probability consistently increased from 2014-2019.

### References and Notes:

1. R. Butlin *et al.*, What do we need to know about speciation? *Trends Ecol. Evol.* **27**, 27–39 (2012).
2. J. M. Sobel, M. A. Streisfeld, Strong premating reproductive isolation drives incipient speciation in *Mimulus aurantiacus*. *Evolution.* **69**, 447–461 (2015).
3. A. C. R. Lackey, J. W. Boughman, Evolution of reproductive isolation in stickleback fish. *Evolution.* **71**, 357–372 (2017).
4. J. A. C. Uy, D. E. Irwin, M. S. Webster, Behavioral Isolation and Incipient Speciation in Birds. *Annu. Rev. Ecol. Evol. Syst.* **49**, 1–24 (2018).
5. T. D. Price, M. M. Bouvier, The evolution of F1 postzygotic incompatibilities in birds. *Evolution.* **56**, 2083–2089 (2002).

6. J. A. Coyne, H. A. Orr, Patterns of speciation in *Drosophila*. *Evolution*. **43**, 362–381 (1989).
7. M. Xu, K. L. Shaw, Genetic coupling of signal and preference facilitates sexual isolation during rapid speciation. *Proc. R. Soc. B Biol. Sci.* **286**, 20191607 (2019).
- 5 8. O. Seehausen *et al.*, Speciation through sensory drive in cichlid fish. *Nature*. **455**, 620–626 (2008).
9. P. R. Grant, B. R. Grant, Role of sexual imprinting in assortative mating and premating isolation in Darwin’s finches. *Proc. Natl. Acad. Sci.* **115**, E10879–E10887 (2018).
10. T. C. Mendelson, Sexual isolation evolves faster than hybrid inviability in a diverse and sexually dimorphic genus of fish (Percidae: Etheostoma). *Evolution*. **57**, 317–327 (2003).
- 10 11. L. Campagna *et al.*, Repeated divergent selection on pigmentation genes in a rapid finch radiation. *Sci. Adv.* **3**, e1602404 (2017).
12. L. Campagna *et al.*, Rapid phenotypic evolution during incipient speciation in a continental avian radiation. *Proc. R. Soc. B Biol. Sci.* **279**, 1847–1856 (2011).
- 15 13. O. M. Selz, M. E. R. Pierotti, M. E. Maan, C. Schmid, O. Seehausen, Female preference for male color is necessary and sufficient for assortative mating in 2 cichlid sister species. *Behav. Ecol.* **25**, 612–626 (2014).
14. P. Benites, L. Campagna, P. L. Tubaro, Song-based species discrimination in a rapid Neotropical radiation of grassland seedeaters. *J. Avian Biol.* **46**, 55–62 (2015).
- 20 15. C. A. B. Medolago, M. C. Costa, L. F. Silveira, M. R. Francisco, Hybridization between two recently diverged Neotropical passerines: The Pearly-bellied Seedeater *Sporophila pileata*, and the Copper Seedeater *S. bouvreuil* (Aves, Passeriformes, Thraupidae). *PLoS One*. **15**, e0229714 (2020).

16. L. Campagna, P. Rodriguez, J. C. Mazzulla, Transgressive phenotypes and evidence of weak postzygotic isolation in F1 hybrids between closely related capuchino seedeaters. *PLoS One*. **13**, e0199113 (2018).
17. Birdlife International, IUCN Red List for birds (2020), (available at <http://www.birdlife.org>).
18. A. S. Di Giacomo, C. Kopuchian, Una nueva especie de capuchino (*Sporophila*: Thraupidae) de los Esteros del Iberá, Corrientes, Argentina. *Nuestras Aves*. **61**, 3–5 (2016).
19. M. Repenning, C. S. Fontana, A New Species of Gray Seedeater (Emberizidae: *Sporophila*) from Upland Grasslands of Southern Brazil. *Auk*. **130**, 791–803 (2013).
20. M. Abolins-Abols *et al.*, Differential gene regulation underlies variation in melanic plumage coloration in the dark-eyed junco (*Junco hyemalis*). *Mol. Ecol.* **27**, 4501–4515 (2018).
21. I. C. Cuthill *et al.*, Ultraviolet Vision in Birds. *Adv. Study Behav.* **29**, 159–214 (2000).
22. P. Benites, M. D. Eaton, D. A. Lijtmaer, S. C. Loughheed, P. L. Tubaro, Analysis from avian visual perspective reveals plumage colour differences among females of capuchino seedeaters (*Sporophila*). *J. Avian Biol.* **41**, 597–602 (2010).
23. S. C. Griffith, I. P. F. Owens, K. A. Thuman, Extra pair paternity in birds: a review of interspecific variation and adaptive function. *Mol. Ecol.* **11**, 2195–2212 (2002).
24. S. P. Turbek, M. Browne, C. Pasian, A. S. Di Giacomo, First nest description of the Iberá Seedeater (*Sporophila iberaensis*). *Wilson J. Ornithol.* **131**, 156–160 (2019).
25. D. Wheatcroft, A. Qvarnström, Genetic divergence of early song discrimination between

two young songbird species. *Nat. Ecol. Evol.* **1**, 192 (2017).

26. D. A. Marques, J. I. Meier, O. Seehausen, A Combinatorial View on Speciation and Adaptive Radiation. *Trends Ecol. Evol.* **34**, 531–544 (2019).

27. J. I. Meier, D. A. Marques, C. E. Wagner, L. Excoffier, O. Seehausen, Genomics of Parallel Ecological Speciation in Lake Victoria Cichlids. *Mol. Biol. Evol.* **35**, 1489–1506 (2018).

28. D. J. Howard, in *Hybrid zones and the evolutionary process*, R. G. Harrison, Ed. (Oxford University Press, New York, 1993), pp. 46–69.

29. H. Ellegren, Genomic evidence for a large-Z effect. *Proc. R. Soc. B Biol. Sci.* **276**, 361–366 (2009).

30. S. Lamichhaney *et al.*, Rapid hybrid speciation in Darwin’s finches. *Science.* **359**, 224–228 (2018).

31. E. B. Rosenblum *et al.*, Goldilocks Meets Santa Rosalia: An Ephemeral Speciation Model Explains Patterns of Diversification Across Time Scales. *Evol. Biol.* **39**, 255–261 (2012).

32. E. T. Domyan *et al.*, Epistatic and Combinatorial Effects of Pigmentary Gene Mutations in the Domestic Pigeon. *Curr. Biol.* **24**, 459–464 (2014).

33. E. E. Kenny *et al.*, Melanesian Blond Hair Is Caused by an Amino Acid Change in TYRP1. *Science.* **336**, 554–554 (2012).

34. J. Li *et al.*, A missense mutation in TYRP1 causes the chocolate plumage color in chicken and alters melanosome structure. *Pigment Cell Melanoma Res.* **32**, 381–390 (2019).

35. L. A. Lyons, I. T. Foe, H. C. Rah, R. A. Grahn, Chocolate coated cats: TYRP1 mutations

for brown color in domestic cats. *Mamm. Genome*. **16**, 356–366 (2005).

36. R. A. Sturm, T. N. Frudakis, Eye colour: portals into pigmentation genes and ancestry. *Trends Genet.* **20**, 327–332 (2004).

37. M. Visser, M. Kayser, F. Grosveld, R.-J. Palstra, Genetic variation in regulatory DNA elements: the case of OCA2 transcriptional regulation. *Pigment Cell Melanoma Res.* **27**, 169–177 (2014).

38. H. Klaassen, Y. Wang, K. Adamski, N. Rohner, J. E. Kowalko, CRISPR mutagenesis confirms the role of oca2 in melanin pigmentation in *Astyanax mexicanus*. *Dev. Biol.* **441**, 313–318 (2018).

39. M. Visser, M. Kayser, R.-J. Palstra, HERC2 rs12913832 modulates human pigmentation by attenuating chromatin-loop formation between a long-range enhancer and the OCA2 promoter. *Genome Res.* **22**, 446–455 (2012).

40. J. M. C. da Silva, Seasonal movements and conservation of seedeaters of the genus *Sporophila* in South America. *Stud. Avian Biol.* **19**, 272–280 (1999).

41. S. Andrews, FastQC: a quality control tool for high throughput sequence data (2010), (available at <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>).

42. M. Schubert, S. Lindgreen, L. Orlando, AdapterRemoval v2: rapid adapter trimming, identification, and read merging. *BMC Res. Notes.* **9**, 88 (2016).

43. B. Langmead, S. L. Salzberg, Fast gapped-read alignment with Bowtie 2. *Nat. Methods.* **9**, 357 (2012).

44. F. García-Alcalde *et al.*, Qualimap: evaluating next-generation sequencing alignment data. *Bioinformatics.* **28**, 2678–2679 (2012).

45. L. Heng *et al.*, The Sequence Alignment/Map format and SAMtools. *Bioinformatics*. **25**, 2078–2079 (2009).
46. Broad Institute, Picard Toolkit. (2018), (available at <http://broadinstitute.github.io/picard>).
47. A. McKenna *et al.*, The Genome Analysis Toolkit: a MapReduce framework for analyzing  
5 next-generation DNA sequencing data. *Genome Res.* **20**, 1297–1303 (2010).
48. P. Danecek *et al.*, The variant call format and VCFtools. *Bioinformatics*. **27**, 2156–2158 (2011).
49. S. Turner, qqman: an R package for visualizing GWAS results using Q-Q and manhattan plots. *J. Open Source Softw.* **3**, 731 (2018).
- 10 50. X. Zheng *et al.*, A high-performance computing toolset for relatedness and principal component analysis of SNP data. *Bioinformatics*. **28**, 3326–3328 (2012).
51. R Core Team, R: A language and environment for statistical computing (2018), (available at <http://www.r-project.org>).
52. S. Purcell *et al.*, PLINK: A Tool Set for Whole-Genome Association and Population-  
15 Based Linkage Analyses. *Am. J. Hum. Genet.* **81**, 559–575 (2007).
53. M. G. Grabherr *et al.*, Genome-wide synteny through highly sensitive sequence alignment: Satsuma. *Bioinformatics*. **26**, 1145–1151 (2010).
54. M. Meyer, T. Munzner, H. Pfister, MizBee: a multiscale synteny browser. *IEEE Trans. Vis. Comput. Graph.* **15**, 897–904 (2009).
- 20 55. S. F. Altschul, W. Gish, W. Miller, E. W. Myers, D. J. Lipman, Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410 (1990).

56. A. Siepel *et al.*, Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. *Genome Res.* **15**, 1034–1050 (2005).
57. W. J. Kent *et al.*, The Human Genome Browser at UCSC. *Genome Res.* **12**, 996–1006 (2002).
- 5 58. W. J. Kent, BLAT—The BLAST-Like Alignment Tool. *Genome Res.* **12**, 656–664 (2002).
59. C. Hahn, L. Bachmann, B. Chevreur, Reconstructing mitochondrial genomes directly from genomic next-generation sequencing reads—a baiting and iterative mapping approach. *Nucleic Acids Res.* **41**, e129–e129 (2013).
- 10 60. M. Kearse *et al.*, Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics.* **28**, 1647–1649 (2012).
61. J. W. Leigh, D. Bryant, POPART: full-feature software for haplotype network construction. *Methods Ecol. Evol.* **6**, 1110–1116 (2015).
- 15 62. P. D. N. Hebert, S. Ratnasingham, J. R. de Waard, Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. *Proc. R. Soc. London. Ser. B Biol. Sci.* **270**, S96–S99 (2003).
63. K. C. R. Kerr, D. A. Lijtmaer, A. S. Barreira, P. D. N. Hebert, P. L. Tubaro, Probing Evolutionary Patterns in Neotropical Birds through DNA Barcodes. *PLoS One.* **4**, e4379 (2009).
- 20 64. L. Campagna *et al.*, DNA barcodes provide new evidence of a recent radiation in the genus *Sporophila* (Aves: Passeriformes). *Mol. Ecol. Resour.* **10**, 449–458 (2010).
65. A. Stamatakis, RAxML version 8: a tool for phylogenetic analysis and post-analysis of

large phylogenies. *Bioinformatics*. **30**, 1312–1313 (2014).

66. A. M. Kozlov, D. Darriba, T. Flouri, B. Morel, A. Stamatakis, RAxML-NG: a fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference.

*Bioinformatics*. **35**, 4453–4455 (2019).

5 67. D. J. Thrasher, B. G. Butcher, L. Campagna, M. S. Webster, I. J. Lovette, Double-digest RAD sequencing outperforms microsatellite loci at assigning paternity and estimating relatedness: A proof of concept in a highly promiscuous bird. *Mol. Ecol. Resour.* **18**, 953–965 (2018).

68. J. Catchen, P. A. Hohenlohe, S. Bassham, A. Amores, W. A. Cresko, Stacks: an analysis  
10 tool set for population genomics. *Mol. Ecol.* **22**, 3124–3140 (2013).

69. K. Okonechnikov, O. Golosova, M. Fursov, the UGENE team, Unipro UGENE: a unified bioinformatics toolkit. *Bioinformatics*. **28**, 1166–1167 (2012).

70. S. R. Browning, B. L. Browning, Rapid and Accurate Haplotype Phasing and Missing-Data Inference for Whole-Genome Association Studies By Use of Localized Haplotype  
15 Clustering. *Am. J. Hum. Genet.* **81**, 1084–1097 (2007).

71. J. Oksanen *et al.*, Vegan: community ecology package. R package version 1.17-4 (2010).

72. L. J. Revell, phytools: an R package for phylogenetic comparative biology (and other things). *Methods Ecol. Evol.* **3**, 217–223 (2012).

73. J. Chifman, L. Kubatko, Quartet Inference from SNP Data Under the Coalescent Model.  
20 *Bioinformatics*. **30**, 3317–3324 (2014).

74. S. T. Kalinowski, M. L. Taper, T. C. Marshall, Revising how the computer program cervus accommodates genotyping error increases success in paternity assignment. *Mol.*

*Ecol.* **16**, 1099–1106 (2007).

75. R. Maia, C. M. Eliason, P.-P. Bitton, S. M. Doucet, M. D. Shawkey, pavo: an R package for the analysis, visualization and organization of spectral data. *Methods Ecol. Evol.* **4**, 906–913 (2013).
- 5 76. D. E. Irwin, S. Bensch, T. D. Price, Speciation in a ring. *Nature*. **409**, 333–337 (2001).
77. B. R. Grant, P. R. Grant, Lack of premating isolation at the base of a phylogenetic tree. *Am. Nat.* **160**, 1–19 (2002).
78. B. R. Grant, P. R. Grant, Simulating secondary contact in allopatric speciation: an empirical test of premating isolation. *Biol. J. Linn. Soc.* **76**, 545–556 (2002).
- 10 79. C. N. Balakrishnan, M. D. Sorenson, Song discrimination suggests premating isolation among sympatric indigobird species and host races. *Behav. Ecol.* **17**, 473–478 (2006).
80. J. A. C. Uy, R. G. Moyle, C. E. Filardi, Z. A. Cheviron, Difference in plumage color used in species recognition between incipient species is linked to a single amino acid substitution in the melanocortin-1 receptor. *Am. Nat.* **174**, 244–254 (2009).
- 15 81. M. C. Baker, Response of Male Indigo and Lazuli Buntings and Their Hybrids To Song Playback in Allopatric and Sympatric Populations. *Behaviour*. **119**, 225–242 (1991).
82. M. C. Baker, A. E. M. Baker, Reproductive behavior of female buntings: isolating mechanisms in a hybridizing pair of species. *Evolution*. **44**, 332–338 (1990).
83. M. A. Patten, J. T. Rotenberry, M. Zuk, Habitat selection, acoustic adaptation, and the evolution of reproductive isolation. *Evolution*. **58**, 2144–2155 (2004).
- 20 84. N. A. Mason, K. J. Burns, Molecular phylogenetics of the Neotropical seedeaters and

seed-finches (*Sporophila*, *Oryzoborus*, *Dolospingus*). *Ornitol Neotrop.* **24**, 139–155 (2013).

85. A. S. Di Giacomo, M. S. Abril, Áreas importantes para la conservación de las aves en la Argentina: sitios prioritarios para la conservación de la biodiversidad (Aves Argentinas/Asociación Ornitológica del Plata, 2005).

86. A. G. Di Giacomo, S. F. Krapovickas, Historia natural y paisaje de la Reserva El Bagual, provincia de Formosa, Argentina. *Temas Nat. y Conserv.* **4**, 1–592 (2005).

87. J. del Hoyo, A. Elliott, J. Sargatal, D. A. Christie, G. Kirwan, Handbook of the Birds of the World Alive. *Lynx Edicions, Barcelona* (2020), (available at <http://www.hbw.com/>).

88. D. Bates, M. Mächler, B. Bolker, S. Walker, Fitting Linear Mixed-Effects Models Using lme4. *J. Stat. Software.* **1**, (2015).

89. M. E. Brooks *et al.*, glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. *R J.* **9**, 378–400 (2017).

90. J. A. C. Uy, R. G. Moyle, C. E. Filardi, Plumage and song differences mediate species recognition between incipient flycatcher species of the Solomon Islands. *Evolution.* **63**, 153–164 (2009).

91. R. Lenth, H. Singmann, J. Love, Emmeans: Estimated marginal means, aka least-squares means. R package version (2018).

92. eBird Basic Dataset. Version: EBD\_relJan-2020. Cornell Lab of Ornithology, Ithaca, New York. Jan 2020.

93. B. L. Sullivan *et al.*, The eBird enterprise: An integrated approach to development and application of citizen science. *Biol. Conserv.* **169**, 31–40 (2014).

94. M. Strimas-Mackey, E. Miller, W. Hochachka, auk: eBird Data Extraction and Processing with AWK. R package version (2017).
95. R. Bivand, C. Rundel, rgeos: interface to geometry engine-open source (GEOS). R package version 0.3-23 (2017).
- 5 96. S. N. Wood, Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *J. R. Stat. Soc. Ser. B (Statistical Methodol.* **73**, 3–36 (2011).
97. N. Rohland, D. Reich, Cost-effective, high-throughput DNA sequencing libraries for multiplexed target capture. *Genome Res.* **22**, 939–946 (2012).
- 10 98. I. Gronau, M. J. Hubisz, B. Gulko, C. G. Danko, A. Siepel, Bayesian inference of ancient human demography from individual genome sequences. *Nat. Genet.* **43**, 1031 (2011).
99. L. Campagna, I. Gronau, L. F. Silveira, A. Siepel, I. J. Lovette, Distinguishing noise from signal in patterns of genomic divergence in a highly polymorphic avian radiation. *Mol. Ecol.* **24**, 4238–4251 (2015).
- 15 100. M. Plummer, N. Best, K. Cowles, K. Vines, CODA: convergence diagnosis and output analysis for MCMC. *R news.* **6**, 7–11 (2006).
101. L. Smeds, A. Qvarnström, H. Ellegren, Direct estimate of the rate of germline mutation in a bird. *Genome Res.* **26**, 1211–1218 (2016).
102. L. Campagna, K. G. McCracken, I. J. Lovette, Gradual evolution towards flightlessness in  
20 steamer ducks. *Evolution.* **73**, 1916–1926 (2019).
103. L. L. Short, Aves nuevas o poco comunes de Corrientes. *Rev. Mus. Argent. Cienc. Nat. “Bernardino Rivadavia” Inst. Nac. Invest. Cienc. Nat. Zool.* **9**, 283–309 (1971).

104. N. J. Collar, *Threatened birds of the Americas* (Smithsonian Institution Press in cooperation with International Council for Bird Preservation, 1992).
105. A. R. Giraud *et al.*, Aves de los esteros del Iberá (Corrientes, Argentina): Patrones de diversidad, historia natural y perspectivas de conservación. *Fauna del Iberá. Editor. Univ. la Univ. Nac. del Noreste, Corrientes* (2003).
106. R. M. Fraga, The avifauna of Estancia San Juan Poriahú, Iberá Marshes, Argentina: checklist and some natural history notes. *Cotinga*. **16**, 81–86 (2001).
107. L. Campagna, L. F. Silveira, P. L. Tubaro, S. C. Loughheed, Identifying the Sister Species to the Rapid Capuchino Seedeater Radiation (Passeriformes: *Sporophila*). *Auk*. **130**, 645–655 (2013).
108. J. W. Poelstra, N. Vijay, M. P. Hoepfner, J. B. W. Wolf, Transcriptomics of colour patterning and coloration shifts in crows. *Mol. Ecol.* **24**, 4617–4628 (2015).

**Acknowledgments:** We thank Constanza Pasian, Bárbara Rodríguez-Quintana, and Bruna Rodrigues do Amaral for field assistance, Ethan Turbek for generating the mounts, and feedback from the Taylor, Safran, and Lovette lab groups. **Funding:** grants from AGA, NGS, APS, ASN, WOS, BOU, AOS, SSB, CU-Boulder, NSF GRFP (SPT), NSF (DEB-1555769) (IJL), FAPESP-2017/23548-2 and CNPq-308337/2019-0 (LFS), ANPCyT and CONICET (ASDG and DAL), and NGC (MB). **Author contributions:** Designed project: SPT, RJS, LC. Performed fieldwork: SPT, MB, ASDG, CK. Obtained genomic data: SPT, LC, IJL, MCE, DAL, PLT, LFS. Analyzed genomic data: SPT, LC. Performed and analyzed behavioral experiments: SPT. Analyzed eBird data: WMH. Supervised research: RJS, SAT, LC. Wrote manuscript: SPT, LC with edits from all

authors. **Competing interests:** None. **Data and materials availability:** Dryad  
(doi:10.5061/dryad.cjsxksn4h), NCBI (BioProject-PRJNA382416).

**Supplementary Materials:**

Materials and Methods

5 Supplementary Text

Figures S1-S20

Tables S1-S5

References (97-108)



## Supplementary Materials for

Rapid speciation via the evolution of pre-mating isolation in the Iberá Seedeater

Sheela P. Turbek, Melanie Browne, Adrián S. Di Giacomo, Cecilia Kopuchian, Wesley M. Hochachka, Cecilia Estalles, Darío A. Lijtmaer, Pablo L. Tubaro, Luís Fábio Silveira, Irby J. Lovette, Rebecca J. Safran, Scott A. Taylor, and Leonardo Campagna

Correspondence to: [sheela.turbek@colorado.edu](mailto:sheela.turbek@colorado.edu) and [lc736@cornell.edu](mailto:lc736@cornell.edu)

### **This PDF file includes:**

Materials and Methods  
Supplementary Text  
Figures S1-S20  
Tables S1-S5

### **Table of Contents**

- 1. Materials and Methods**
  - 1.1. Field methods
  - 1.2. Whole-genome resequencing and variant discovery
  - 1.3. Double-digest restriction site-associated DNA sequencing
  - 1.4. Demographic modeling
  - 1.5. Behavioral experiment
  
- 2. Supplementary Text**
  - 2.1. Natural history of *S. iberensis* and history of the San Nicolás Reserve and museum collections in Iberá National Park
  - 2.2. Haplotype phasing from the divergence peaks on scaffolds 257 and 430
  - 2.3. Demographic modeling and relationships between *S. iberensis* and other southern capuchinos
  
- 3. Supplementary Figures**

- Fig. S1. Geographic distribution of the southern capuchino seedeaters and their prevalence in Iberá National Park.
- Fig. S2. Pattern of absolute genomic differentiation across divergence peaks.
- Fig. S3. Mitochondrial unrooted statistical parsimony networks of *S. hypoxantha* and *S. iberaensis*.
- Fig. S4. Whole-genome PCA showing genomic differentiation between *S. hypoxantha* and *S. iberaensis*.
- Fig. S5. Pattern of linkage disequilibrium across divergence peaks.
- Fig. S6. Clustering of haplotypes from all SNPs within the peak on scaffold 430.
- Fig. S7. Clustering of haplotypes from the SNPs showing the highest  $F_{ST}$  in the peak on scaffold 430.
- Fig. S8. Clustering of haplotypes from the SNPs showing the highest  $F_{ST}$  in the peak on scaffold 257.
- Fig. S9. Frequency of  $F_{ST}$  values for individual SNPs in the whole-genome and ddRAD datasets.
- Fig. S10. Manhattan plot from the whole-genome data showing genomic differentiation across divergence peaks for males and females of *S. hypoxantha* and *S. iberaensis*.
- Fig. S11. Manhattan plot and PCA from the ddRAD data showing genomic differentiation between *S. hypoxantha* and *S. iberaensis*.
- Fig. S12. Number and timing of *S. hypoxantha* and *S. iberaensis* nests found across two breeding seasons.
- Fig. S13. Example sonograms derived from *S. hypoxantha* and *S. iberaensis* male songs.
- Fig. S14. Plumage coloration of *S. hypoxantha* and *S. iberaensis* vs. the mounts used in the behavioral experiment.
- Fig. S15. PCA of the response variables from the behavioral experiment.
- Fig. S16. Boxplot showing male responses to conspecific, heterospecific, and control treatments.
- Fig. S17. Whole-genome phylogenetic tree of ten capuchinos and two outgroup species.
- Fig. S18. Demographic reconstruction for southern capuchino seedeaters.
- Fig. S19. Phylogenetic tree of ten capuchinos and two outgroup species for the peak on scaffold 762.
- Fig. S20. Manhattan plots from whole-genome data showing genomic differentiation between *S. iberaensis* and other capuchino species.

#### 4. Supplementary Tables

- Table S1. Details on the samples used in the study.
- Table S2. Sequencing statistics for the whole-genome data.
- Table S3. Position of highly divergent SNPs ( $F_{ST} > 0.85$ ) relative to identified genes.
- Table S4. Parentage assignments from the ddRAD data.

- Table S5. Generalized linear mixed model results analyzing the effect of treatment group on behavioral responses of *S. hypoxantha* and *S. iberensis*.

## **1 Materials and Methods**

### **1.1 Field methods**

Research was approved by the University of Colorado Institutional Animal Care and Use Committee (IACUC) protocols no. 2498 and 2683 and conducted in accordance with USFWS permit MB12129A-10, USDA permit 112702, and permission from the Ministerio de Ambiente y Desarrollo Sostenible de la Nación (Argentina), the Administración de Parques Nacionales (Argentina), and the Dirección de Parques y Reservas (Province of Corrientes, Argentina).

### **1.2 Whole-genome resequencing and variant discovery**

Individually barcoded libraries were prepared with 200 ng of each sample and an insert size of 550 bp following the TruSeq Nano DNA library preparation kit protocol (Illumina). The 37 libraries were then pooled into two groups of 19 and 18 samples using concentrations of adapter-ligated DNA determined through digital polymerase chain reaction (PCR). We sequenced the pooled libraries on two Illumina NextSeq 500 lanes at the Cornell University Biotechnology Resource Center (BRC). Whole-genome resequencing generated over 860 million paired-end reads with a length of 151 bp, producing an expected per-individual coverage ranging between 3.9× and 10.7× (median: 5.4×; Table S2).

### **1.3 Double-digest restriction site-associated DNA sequencing**

We digested 100-500 ng of DNA from each individual with SbfI and MspI (NEB) and ligated one of 20 P1 adapters (containing unique inline barcodes) and a P2-MspI adapter to each sample.

The samples were pooled together in groups of 20, purified with 1.5× volumes of magnetic beads made with Sera-Mag Magnetic Speed-beads (FisherSci), as detailed in (97), and size-selected for 450-600 bp fragments using BluePippin (Sage Science) at the Cornell University BRC. We added the Illumina index group sequences and sequencing adapters following size selection by performing 11 PCR cycles with Phusion DNA Polymerase (NEB). The products of the reactions were purified with 0.7 volumes of magnetic beads, and pooled in equimolar ratios to generate two libraries for sequencing. Each library was sequenced on one lane of an Illumina HiSeq 2500 at the BRC to generate 101-bp single-end reads.

#### **1.4 Demographic modeling**

We generated demographic inferences using G-PhoCS (Generalized Phylogenetic Coalescent Sampler) version 1.3 (98), which allowed us to co-estimate effective population sizes, splitting times, and bi-directional migration rates (*i.e.*, introgression). We decided to conduct this analysis using ddRAD data, which provides two phased haplotypes per individual for a sub-sample of loci across the genome. This strategy allowed us to combine the ddRAD data generated for *S. iberaensis* and *S. hypoxantha* in this study ( $N = 206$  individuals), with previously generated ddRAD data for 70 individuals belonging to six southern capuchino species from (99) (*S. bouvreuil*, *S. hypoxantha*, *S. melanogaster*, *S. pileata*, *S. ruficollis*, and *S. palustris*). We first trimmed the 3' end of the sequences from (99) to a length of 90 bp (from an original 132 bp after filtering and demultiplexing) to match the locus size generated in this study for *S. iberaensis* and *S. hypoxantha*. We subsequently aligned these data to the *S. hypoxantha* reference genome and assembled ddRAD loci in Stacks as described in the *Double-digest restriction site-associated DNA sequencing* section. We discarded individuals with more than 10% missing data and

exported loci present in at least 90% of all individuals, which allowed us to retain a total of 5,852 variant and invariant (*i.e.*, without SNPs) loci for 256 individuals, containing a total of 54,698 SNPs. This reference-based assembly allowed us to recover ~2,000 more loci than what was previously used for demographic modeling by Campagna *et al.* (99). We explored the relationships among individuals and species, and compared the patterns to those observed from whole-genome sequencing data, by performing PCAs in *SNPRelate*. Prior to conducting the G-PhoCS analysis, we removed the invariant enzyme cut site from the 5' end of all loci.

The G-PhoCS analysis is based on the phylogenetic relationships among the species included in the demographic reconstruction. Because these relationships cannot be fully resolved within the rapid radiation of southern capuchinos (*e.g.*, see Fig. 6A), we devised a strategy that compared species by pairs (see Fig. S6 in Campagna *et al.* (99)). We conducted a total of 16 G-PhoCS runs, including all pairwise combinations of six southern capuchino species and *S. bouvreuil* as an outgroup. We also included one model in which *S. iberaensis* was sister to the remaining five species combined. This strategy had the advantage of including a smaller number of parameters than if all species had been included in a single model. Because of the computationally intensive nature of this analysis, each species was represented by 10 individuals, sampling one haplotype per locus per individual. In addition, we only included males in the analysis so that the sampling of the sex chromosome Z would be equivalent to that of autosomes. For the model that combined all species as sister to *S. iberaensis*, we sampled two haplotypes per species (total of 10 haplotypes as in all other models). Each pairwise model estimated 13 demographic parameters: three current and two ancestral population sizes, two splitting times, and six migration rates (*i.e.*, both directions for every pairwise combination among three populations: two ingroup taxa and one outgroup). In the models that included *S. iberaensis*, we

did not allow gene flow between this species and the outgroup *S. bouvreuil*, as these species are allopatric. We ran the multi-threaded version of the program for 750,000 iterations, discarding the initial 150,000 iterations as burn-in, and left the remaining parameters as default. Once each run was completed, we checked that the traces from the different parameter estimates were stationary by visualizing them using the *coda* package in R (100). We converted the median and 95% Bayesian credible intervals for each parameter from mutation scale to generations or individuals using an approximate mutation rate estimate of  $10^{-9}$  per bp per generation (101), as described in (98). We expressed migration (*i.e.*, introgression from one species into another) as the number of individual migrants per generation, calculated as indicated in (102), which is independent of our assumption of mutation rate. We note that the mutation rate is a rough approximation, and therefore focus the interpretation of the results on relative comparisons between estimates of different parameters, which are not influenced by our assumptions of mutation rate.

## 1.5 Behavioral experiment

To create song stimuli for the behavioral experiment, we recorded the long-range advertisement songs of ten individuals of *S. hypoxantha* and *S. iberaensis*, respectively, in the San Nicolás Reserve during the 2018 breeding season using a Marantz PMD 661 digital recorder at a 48 kHz sampling rate and 16 bit-depth PCM, combined with a Sennheiser ME66 shotgun microphone. For the control files, we downloaded five high-quality tracks of *S. collaris* (catalog numbers: XC47158, XC51947, XC73787, XC108310 and XC108311) that were recorded in the neighboring province of Santa Fe, Argentina from 2006-2012 and archived on Xeno-canto (<https://www.xeno-canto.org/>). To generate each playback file, we selected 10-15 of the cleanest

songs from each recording, inserted a standard interval of five seconds of silence between them, which is representative of natural singing behavior in capuchinos, and looped over the songs to create a five-minute track. We used Ocenaudio 3.7.9 to standardize the maximum amplitude of all playback files to -3dB and included five seconds of silence at the beginning of each track.

For the plumage stimuli, we created six artificial mounts from hand-painted 3D printed models (two representing *S. hypoxantha*, two representing *S. iberaensis*, and two representing *S. collaris*). We generated a base model of a mounted specimen of a male black-headed grosbeak (*Pheucticus melanocephalus*) at the Denver Museum of Nature and Science using a technique known as photogrammetry, which involves constructing a 3D model from a series of overlapping photographs taken from numerous angles around the specimen. We imported the photographs into the basic version of Agisoft Photoscan, which we used to process the photographs, assemble the model, and generate a high-resolution point cloud. The point cloud was exported in PLY format and loaded into CloudCompare to generate a high-resolution mesh. We then imported the mesh into MeshLab to simplify the model down to one million polygons, used the digital sculpting tool ZBrush to modify the bill and fill in missing parts on the tail and underside of the bird's body, and loaded the model into the sculpting program 3DCoat to make it 'watertight' (*i.e.*, a completely solid object ready for 3D printing). We printed a smaller version of the mount (~9 cm from bill to tail) to represent the two southern capuchino species, as *S. hypoxantha* and *S. iberaensis* are morphometrically indistinguishable, and a larger version of the mount (~12 cm from bill to tail) to represent *S. collaris*, which is considerably larger in size (87). The models were printed with a Wanhao Duplicator 5S 3D printer and heavily sanded and painted with an acrylic medium to fill in the uneven surface of the printed models. We used Vallejo hobby

acrylic paints to paint the models and sealed them with a matte varnish (Fig. 5A-C). Finally, we glued magnets to the base of the models in order to mount them on a pole in the field.

To determine whether the mounts accurately represented male coloration, we referred to standardized photographs and feather samples that were collected from four plumage patches across the body (crown, throat, belly, and rump) of 46 breeding males of *S. hypoxantha* and 41 breeding males of *S. iberaensis* from 2016-2018. Reflectance data for each plumage patch on the live birds were generated following the methods in the *Feather coloration* section, and the colors of the hand-painted models were visually compared to the spectrophotometer readings of *S. hypoxantha* and *S. iberaensis* breeding males to ensure accurate coloration (Fig. S14). Both capuchino species have crown feathers that reflect wavelengths in the near ultraviolet (300-400 nm), which could not be incorporated into the models. Nonetheless, live birds of *S. hypoxantha* and *S. iberaensis* do not differ in crown coloration. As a result, failing to incorporate UV reflectance into the models should have the same effect on the behavioral responses of both species. Each of the six mounts (including the two heterospecific controls) elicited attack behavior from conspecific males over the course of the study, which suggests that the mounts were sufficiently realistic representations of the three species.

## **2 Supplementary Text**

### **2.1 Natural history of *S. iberaensis* and history of the San Nicolás Reserve and museum collections in Iberá National Park**

Until recently, the fauna of Iberá National Park was primarily represented in museum collections from localities in the south (*e.g.*, near Colonia Carlos Pellegrini; 28° 32' 10" S, 57° 10' 25" W). Ornithological expeditions, including those by William H. Partridge in the 1960s, and Julio R.

Contreras in the 1990s (specimens deposited in the Museo Argentino de Ciencias Naturales “Bernardino Rivadavia” in Argentina and museums in the United States), generally focused on other areas of the province of Corrientes. Indeed, many capuchinos that are common in Iberá National Park first appeared in collections in the 1970s (103), and common grassland birds from Corrientes are poorly represented relative to birds from other regions (104). Giraudo *et al.* (105) and Fraga (106) explored several areas in the northern portion of the Esteros del Iberá wetlands (Estancia El Tránsito, 28° 29' S, 57° 40' W; Estancia San Nicolás, 28° 8' S, 57° 26' W; Loreto and Estancia San Juan Poriahú; 27° 42' S, 57° 11' W), but did not report *S. iberaensis* or *S. hypoxantha*, now the two most abundant capuchino species in San Nicolás. The San Nicolás Reserve served as a cattle farm until 2009, when it was converted into a protected area, possibly contributing to grassland recovery and the population increase of *S. iberaensis*.

## 2.2 Haplotype phasing from the divergence peaks on scaffolds 257 and 430

We phased the variants found within the peak on scaffold 430 to obtain the haplotypes observed among the 37 individuals (16 *S. hypoxantha* and 21 *S. iberaensis*). This region contained 157 SNPs with an  $F_{ST}$  value above 0.5 (mean of  $0.3209 \pm 0.2817$  SD for a total of 625 SNPs), allowing most haplotypes from each species to cluster together (Fig. S6; with the exception of one *S. hypoxantha* haplotype from a male with typical plumage). However, when we conducted the same analysis on the 64 SNPs showing the highest level of differentiation within the peak, we observed two haplotypes of *S. iberaensis* that clustered with those of *S. hypoxantha* (Fig. S7). The equivalent plot from the peak on scaffold 257 (derived from 13 SNPs) showed six *S. iberaensis* haplotypes that clustered with those of *S. hypoxantha* (Fig. S8). One *S. iberaensis* male had a mismatched haplotype on both peaks, yet possessed *S. iberaensis* breeding plumage.

The different haplotype-based analyses from these peak regions are consistent with incomplete lineage sorting or low levels of hybridization. This result is also supported by the extensive sharing of mitochondrial DNA between *S. iberensis* and *S. hypoxantha* (Fig. S3). We note, however, that we did not observe hybridization directly in the pairs sampled in this study, nor individuals that could clearly be assigned as early generation hybrids. In addition, these findings indicate potential redundancy in the genes contributing to phenotypic differences across peaks and suggest that the genotypes that lead to *S. iberensis* plumage may be dominant over those that produce *S. hypoxantha* plumage, as *S. iberensis* males with mixed genotypes in the variants showing the highest level of differentiation still possess *S. iberensis* plumage. A better understanding of the causal sites responsible for both plumage patterns is needed to shed light on how *TYRP1* and *OCA2* act together to shape these phenotypes.

### **2.3 Demographic modeling and relationships between *S. iberensis* and other southern capuchinos**

Even though *S. iberensis* is broadly sympatric with other southern capuchinos in Iberá National Park, and breeds side by side with *S. hypoxantha* in the San Nicolás Reserve, *S. iberensis* individuals form their own clade in a phylogenetic tree (Fig. 6A) and can be clearly differentiated from other capuchino seedeaters in PCAs based on SNPs from either whole-genome or ddRAD sequencing data (Fig. S18A-B). While the *S. iberensis* clade is sister to a clade composed mostly of *S. hypoxantha* individuals (Fig. 6A), the relationship between *S. iberensis* and other species in the radiation remains unclear. To better understand the evolutionary history of *S. iberensis*, we conducted demographic analyses in G-PhoCS, which allowed us to infer splitting times, effective population sizes, and migration rates between *S.*

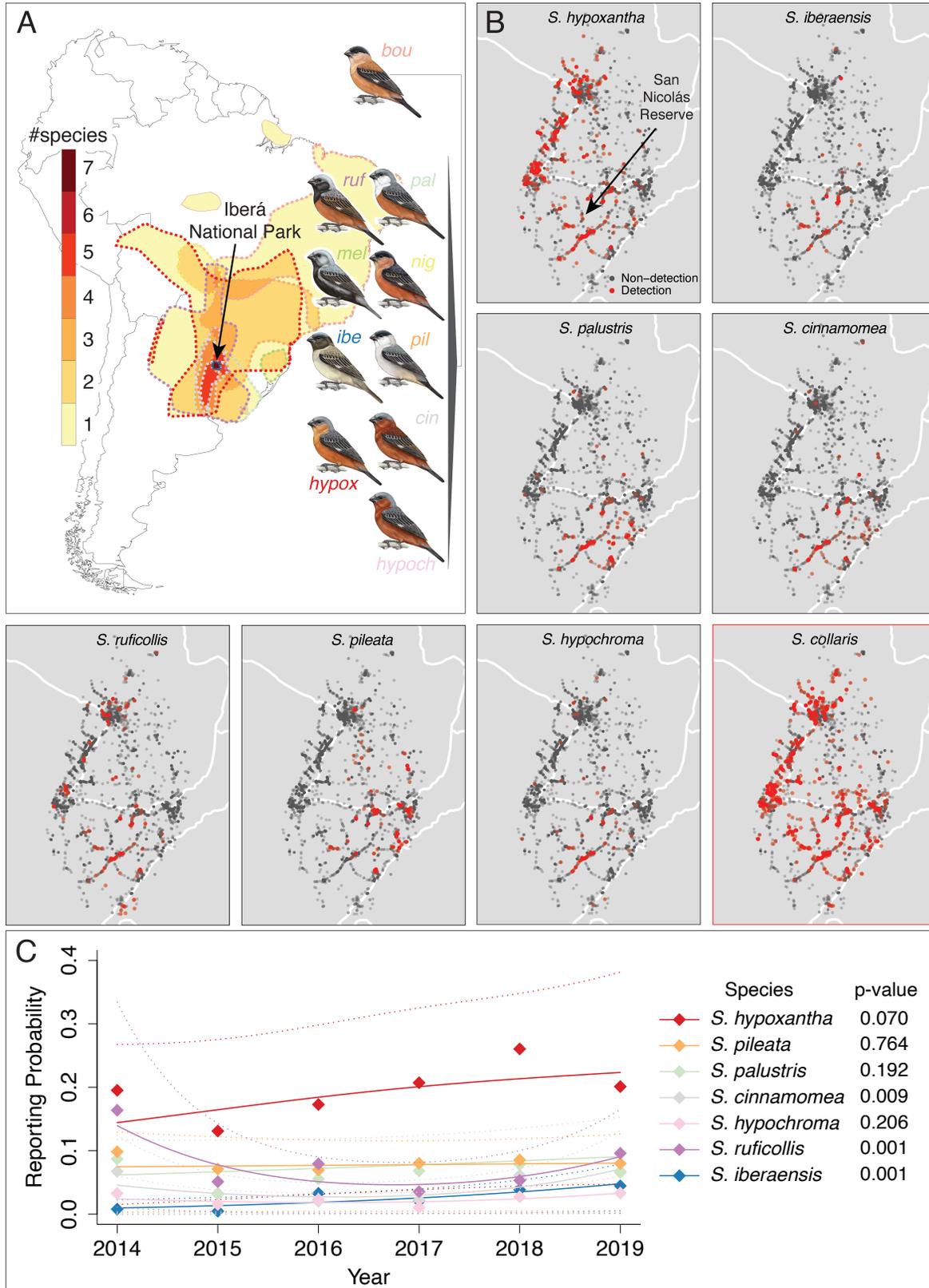
*iberaensis* and five other southern capuchinos (*S. hypoxantha*, *S. melanogaster*, *S. pileata*, *S. ruficollis* and *S. palustris*). We also compared the parameters estimated for *S. iberaensis* mentioned above to equivalent parameters inferred between the latter southern capuchino species alone (*i.e.*, without the inclusion of *S. iberaensis*). We conducted a total of 16 G-PhoCS runs, each including one of the 15 pairwise combinations of six capuchino species as sister to each other and *S. bouvreuil* as an outgroup (107). The 16th G-PhoCS run compared *S. iberaensis* to the five remaining southern capuchinos combined, with *S. bouvreuil* as an outgroup (Fig. S18C).

The results from the 16 G-PhoCS runs could be grouped into two general patterns. The majority of runs (11 out of 16) identified a divergence time within the ingroup comparable to that between the ingroup and *S. bouvreuil*, and inferred a smaller ancestral effective population size than the current effective population sizes of the ingroup taxa (Fig. S18D). These models identified migration between the three current populations, including from and into *S. iberaensis*. The second pattern (Fig. S18E) included 5 models (*S. hypoxantha* vs. *S. iberaensis*, *S. ruficollis* vs. *S. iberaensis*, *S. hypoxantha* vs. *S. melanogaster*, *S. palustris* vs. *S. pileata*, and *S. palustris* vs. *S. ruficollis*) that identified a very recent split between the ingroup taxa and an older split between the ingroup and the outgroup. These models estimated a large ancestral effective population size and smaller current effective population sizes for the ingroup, and detected gene flow only between the ingroup and the outgroup. In addition, these models include the two comparisons that show the lowest mean genome-wide  $F_{ST}$  with respect to *S. iberaensis*: *S. hypoxantha* ( $0.0055 \pm 0.0137$  SD), and *S. ruficollis* ( $0.0037 \pm 0.0184$  SD; Fig. S20).

We note that these estimates are based on a rough assumption of mutation rate and do not include all species interacting in the model at the same time. Therefore, the results must be interpreted with caution. However, they contribute to our understanding of the evolutionary

history of *S. iberaensis* by showing that, although this species was recently discovered, it has similar demographic parameters to those observed among the other southern capuchinos. For example, the most recent identified split between two capuchinos was estimated at ~1500 generations, between *S. iberaensis* and *S. hypoxantha*. However, the 95% credible interval of this estimate overlapped with the inferred splitting time between *S. hypoxantha* and *S. melanogaster*. We also observed levels of gene flow between *S. iberaensis* and other species (*S. melanogaster*, *S. pileata*, and *S. palustris*) that are similar to those inferred between other capuchinos, indicating low levels of hybridization during speciation. These results suggest that *S. iberaensis* could have diverged recently from *S. hypoxantha*, and it is possible that the process occurred with the introgression of specific genomic regions from other taxa.

### 3 Supplementary Figures



**Fig. S1. Geographic distribution of the southern capuchino seedeaters and their prevalence in Iberá National Park. (A)** Cartoon showing the range overlap of southern capuchino

seedeaters, modified from (11). Species are abbreviated using the first few letters of each name.

**(B)** Location of eBird reports across the breeding range of *S. iberaensis* for the seven capuchino species (rectangles outlined in gray) that breed in Iberá National Park, as well as *S. collaris*

(lower right corner; rectangle outlined in red), the control in the behavioral experiment. The red circles indicate locations where each focal species was observed, and the gray symbols denote

areas where eBird checklists were submitted but the focal species was not detected for years in

which the focal species was reported in the eBird database. Across the limited breeding range of

*S. iberaensis*, *S. hypoxantha* is the most common and *S. iberaensis* is one of the least common

capuchino species. *S. collaris*, while not a member of the southern capuchino radiation, is

common in Iberá National Park and breeds in sympatry with *S. iberaensis* and *S. hypoxantha*.

**(C)** Change in reporting probability in the eBird database from 2014-2019 for the capuchino

species that breed in Iberá National Park. The points depict reporting probability per year and the

solid lines indicate the fitted results of generalized additive models (GAMs) that model the effect of year on reporting probability while controlling for the confounding effects of observation date,

distance traveled, and duration of the observation period. The dotted lines show the 95%

prediction intervals for the estimated probabilities. The p-values from the GAMs are the

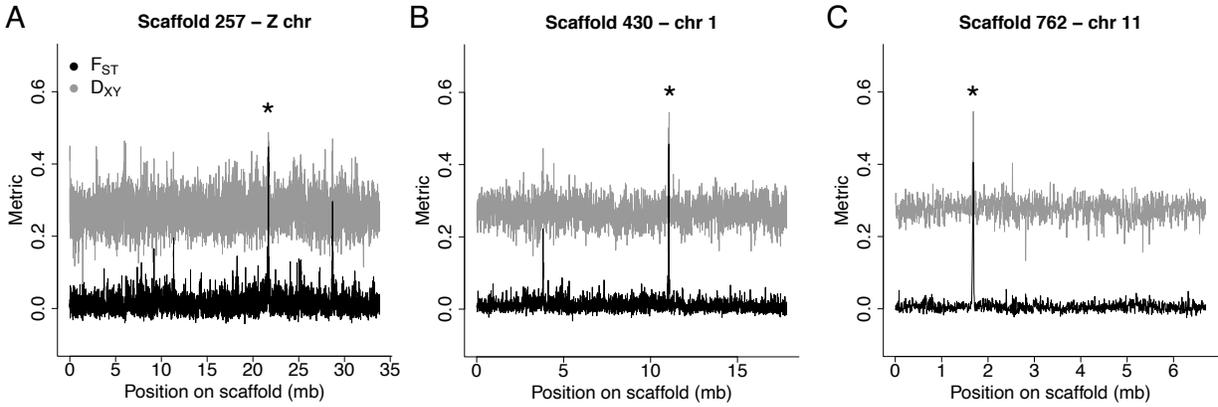
probabilities that there have been systematic changes in the relationship between year and

reporting probability for each capuchino species. *S. iberaensis* has increased in local prevalence

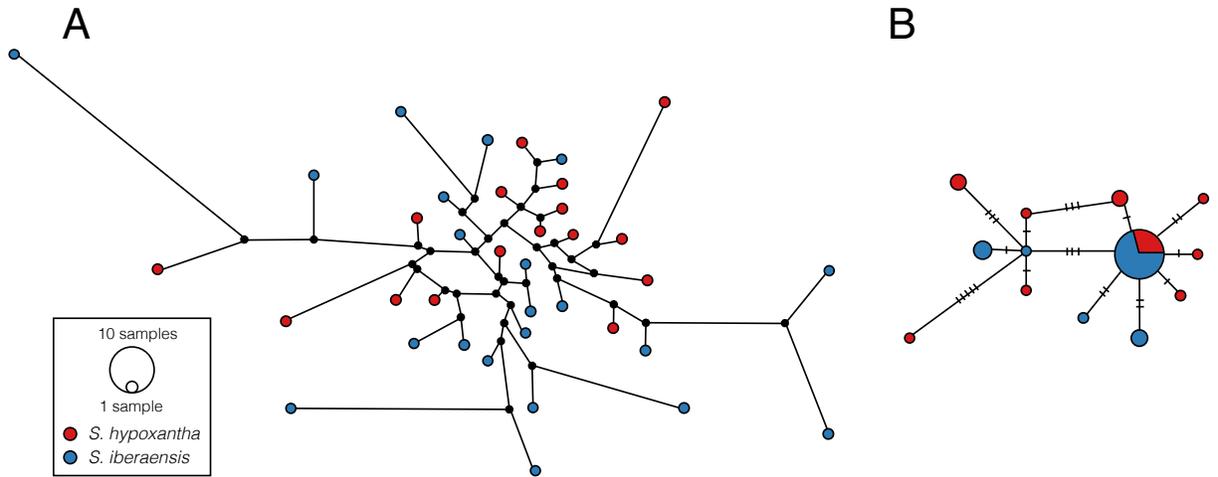
to surpass that of *S. hypochroma* across its limited breeding range. *S. ruficollis* and *S.*

*cinnamomea* also showed systematic changes in reporting probability, but did not consistently

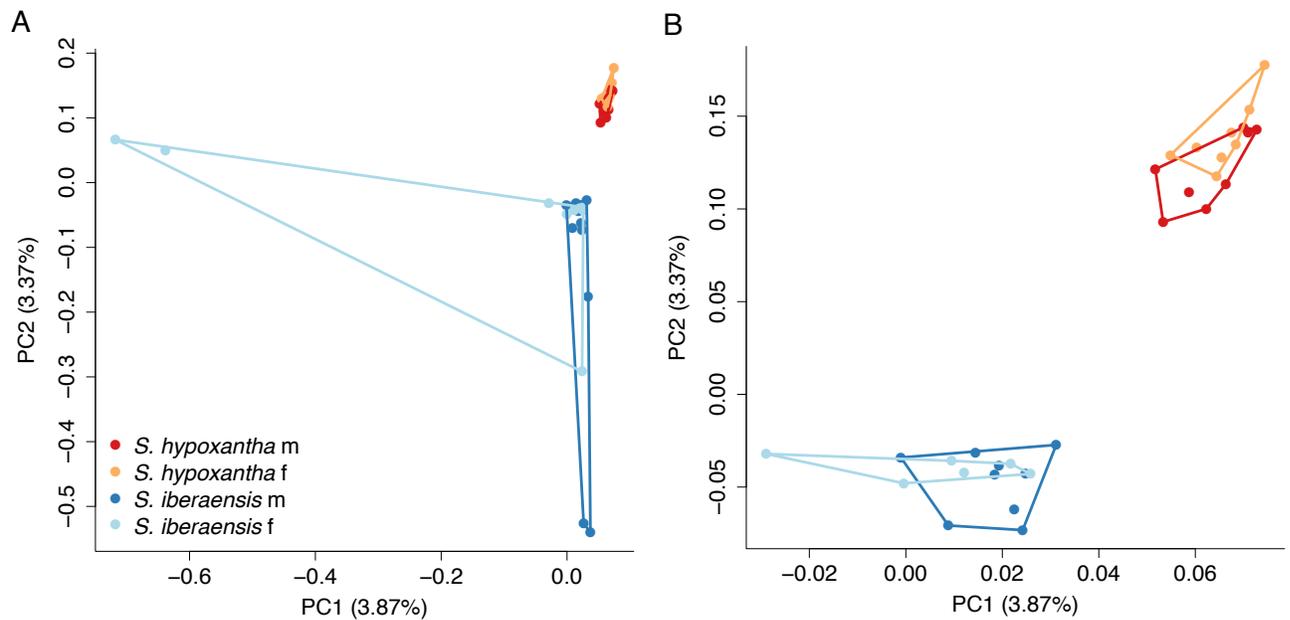
increase in prevalence over time.



**Fig. S2.  $F_{ST}$  peaks were characterized by increased absolute differentiation between *S. hypoxantha* and *S. iberaensis*.** Plots comparing patterns of relative genomic differentiation ( $F_{ST}$ ) and absolute genomic differentiation ( $D_{XY}$ ) over 5-kb windows across (A) scaffold 257, (B) scaffold 430, and (C) scaffold 762 ( $N = 37$ ). Asterisks indicate the divergence peaks that contained at least one individual SNP with  $F_{ST}$  greater than 0.85 (as in Fig. 1F).



**Fig. S3. Mitochondrial unrooted statistical parsimony networks for *S. hypoxantha* and *S. iberensis*.** (A) Haplotype network based on a 17,004 bp alignment of the mitochondrial genome. (B) Haplotype network based on 697 bp of the mitochondrial COI DNA barcode region commonly used for species identification. The length of the branches connecting haplotypes is proportional to the number of nucleotide differences between them, which are indicated by the number of line marks in each branch (omitted for simplicity in the case of the full mitochondrial network). Black dots in (A) indicate hypothetical, unobserved haplotypes. The two capuchino species show no differentiation in mitochondrial DNA with either approach.



**Fig. S4. Pattern of genomic differentiation between *S. hypoxantha* and *S. iberaensis*.**

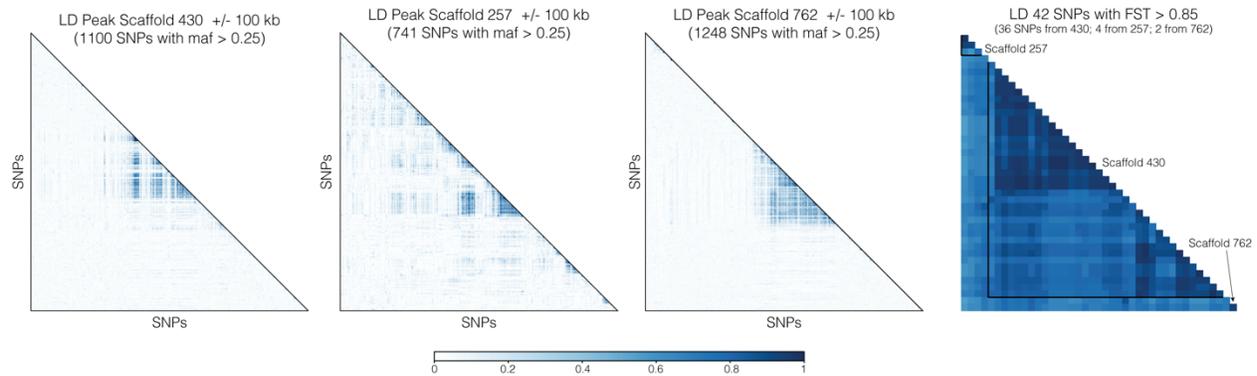
Principal component analysis (PCA) generated with whole-genome data (~13.3 million SNPs) depicting the degree of genomic differentiation between (A) all males and females of *S.*

*hypoxantha* and *S. iberaensis* ( $N = 37$ ) and (B) individuals with a low relatedness statistic ( $N =$

31). Three pairs of *S. iberaensis* individuals, which appear as outliers in (A), were more closely

related to each other than the others (unadjusted  $A_{jk}$  relatedness statistic: B009595 vs. B009507 = 0.46; B009502 vs. B009504 = 0.19; B009508 vs. B009542 = 0.07). Female capuchinos were

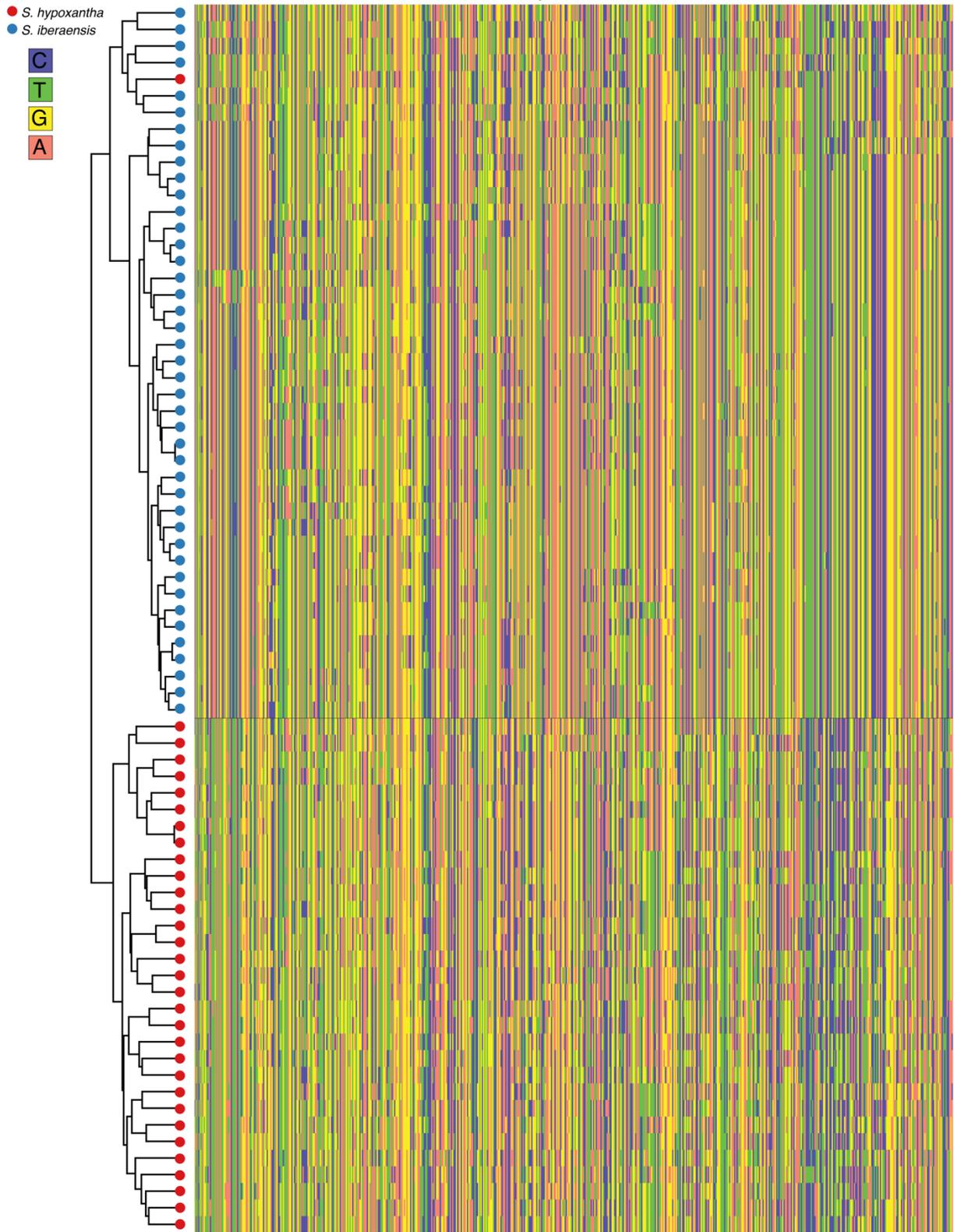
classified based on the phenotype of their social mate.



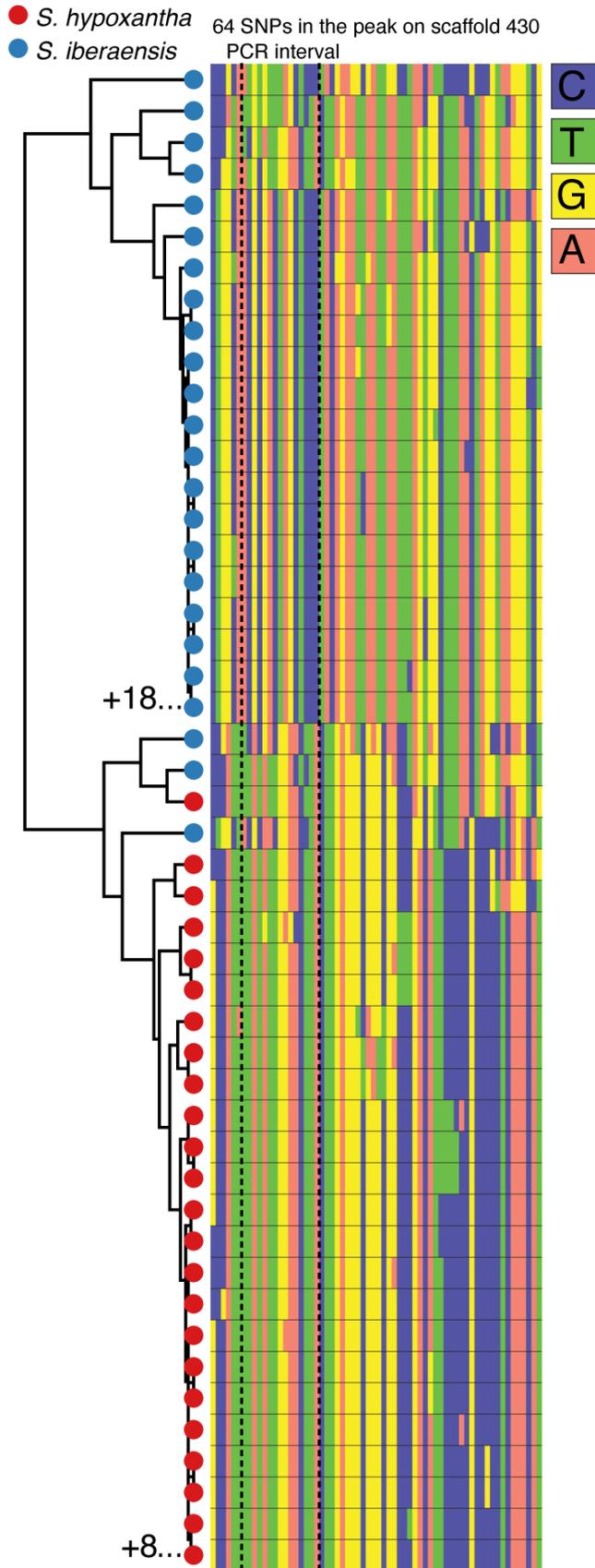
**Fig. S5. Pattern of linkage disequilibrium (LD).**

LD measured by the  $r^2$  statistic across the peak on scaffold 430, 257, and 762, as well as among all possible combinations of the 42 highly differentiated SNPs with  $F_{ST} > 0.85$  within and among each of the three peaks. LD between SNPs located within each peak was high (mean: 0.82, range: 0.59-1 for the peak on scaffold 257; mean: 0.82, range 0.56-1 for the peak on scaffold 430; and 0.94 for the comparison of the two SNPs on the peak on scaffold 762). The values of LD among comparisons of SNPs from different peaks (which were located on separate chromosomes) were also elevated, with a mean of 0.69 (0.56-0.88). Inter-chromosomal LD was somewhat lower but comparable when calculated separately for *S. hypoxantha* (mean: 0.26) and *S. iberaensis* (mean: 0.28). However, the two species tend to have different alleles at each of the 42 highly differentiated SNPs within the three divergence peaks (see haplotype plots below).

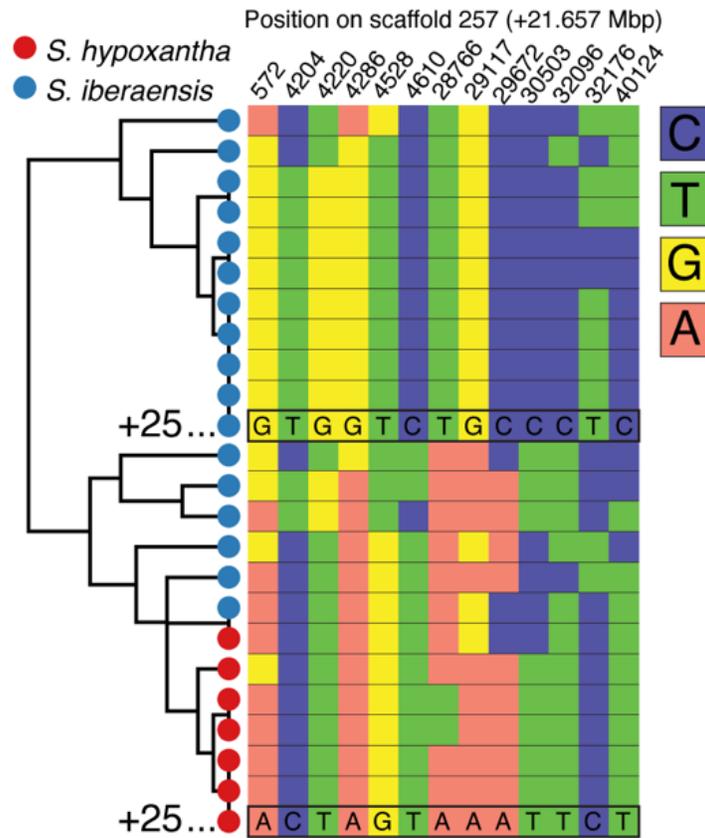
SNPs from the peak on scaffold 430



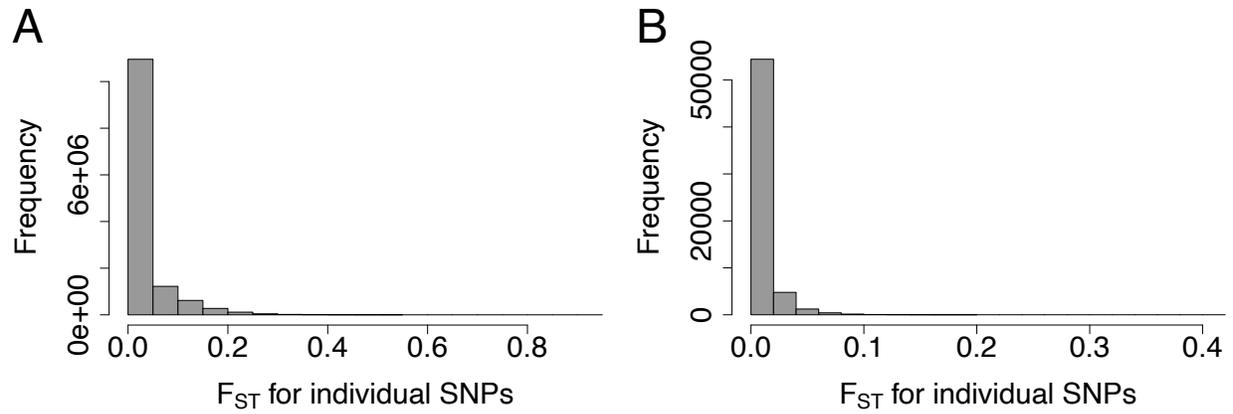
**Fig. S6. Nearly all individuals cluster by species in a haplotype tree derived from the SNPs within the peak on scaffold 430.** Each row represents a single chromosome, and each individual is represented twice in the tree. The analysis was conducted with the individuals for which we had whole-genome data (*S. hypoxantha*:  $N = 16$ ; *S. iberaensis*:  $N = 21$ ). The four nucleotides are color-coded as indicated in the upper left corner. One haplotype from *S. hypoxantha* male B009525 grouped with those of *S. iberaensis*.



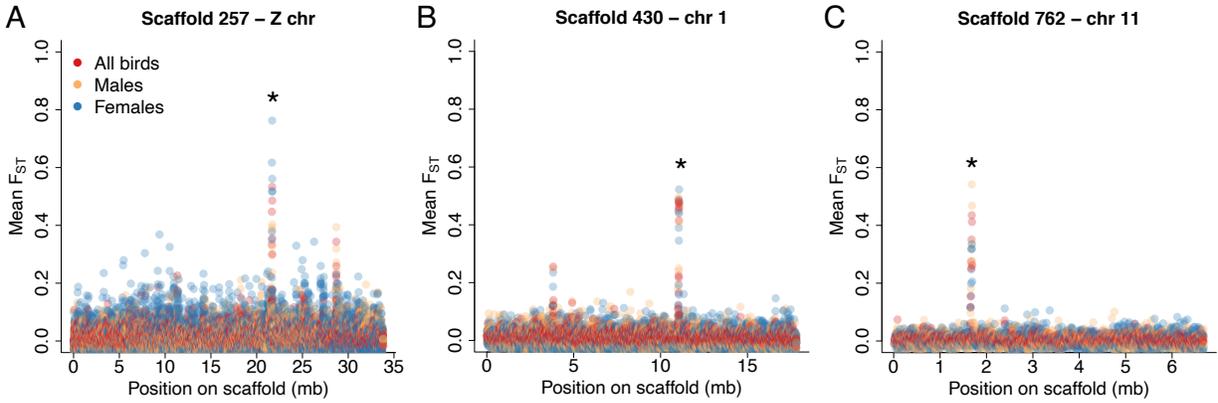
**Fig. S7. Clustering of haplotypes from the SNPs showing the highest  $F_{ST}$  in the peak on scaffold 430.** Most haplotypes (two per individual: *S. hypoxantha*,  $N = 32$ ; *S. iberensis*,  $N = 42$ ) cluster by species, except for three haplotypes derived from *S. iberensis* that group with those of *S. hypoxantha* (one haplotype from male B009552, one from male B009504, and one from female B009507). The analysis included the 36 SNPs with  $F_{ST} > 0.85$  in this region, plus an additional 28 SNPs with  $F_{ST} > 0.79$ . We included these additional SNPs as the segment selected for PCR amplification (indicated by vertical dashed lines) included three SNPs with this level of divergence. For graphical clarity, 18 identical copies of the most common *S. iberensis* haplotype and 8 identical copies of the most common *S. hypoxantha* haplotype were omitted from the tree. Other details as in Fig. S6.



**Fig. S8. Clustering of haplotypes from the SNPs showing the highest  $F_{ST}$  in the peak on scaffold 257.** The haplotypes from *S. hypoxantha* ( $N = 32$ ) and *S. iberaensis* ( $N = 42$ ) individuals cluster by species, except for six *S. iberaensis* haplotypes that group with those of *S. hypoxantha*. The mismatched haplotypes belong to male B009508 (both haplotypes), male B009561 (both haplotypes), B009552 (one haplotype), and B009583 (one haplotype). Note that male B009552 also possesses a mismatched haplotype in the peak on scaffold 430 (Fig. S7), yet has typical *S. iberaensis* plumage. This analysis included the 13 SNPs with  $F_{ST} > 0.79$  in this region. The most common haplotype for each species is indicated at the bottom of the two main clades, and 25 identical copies of these haplotypes from each species were omitted for graphical clarity. Other details as in Fig. S6.

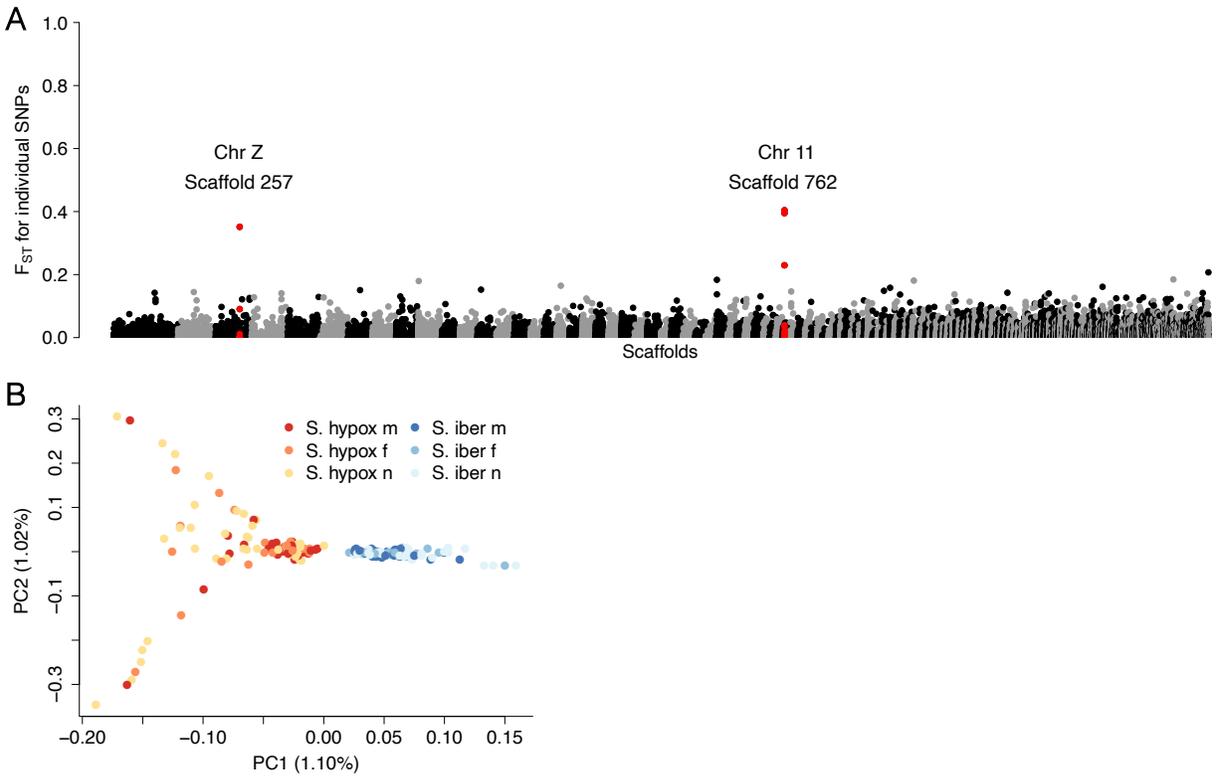


**Fig. 9. Few SNPs with  $F_{ST}$  greater than 0.2 were present in the whole-genome and ddRAD datasets.** Histograms showing the frequency of  $F_{ST}$  values for individual SNPs in the (A) whole-genome and (B) ddRAD data. Only 1.4% of SNPs in the whole-genome data and 0.008% of SNPs in the ddRAD data had moderate  $F_{ST}$  values ( $> 0.2$ ).

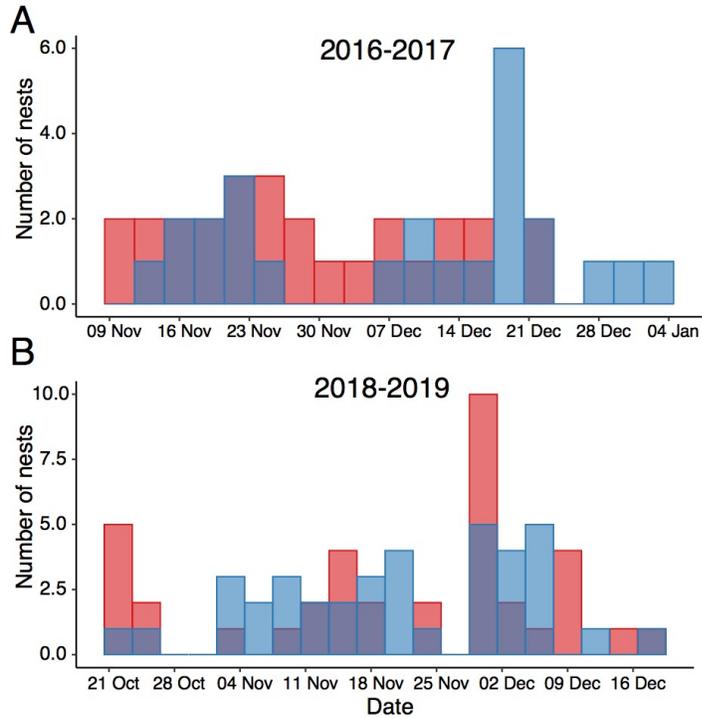


**Fig. S10. Females and males exhibit elevated divergence in the same genomic regions.**

Manhattan plots showing regions of elevated genomic differentiation ( $F_{ST}$ ) over 5-kb windows on (A) scaffold 257, (B) scaffold 430, and (C) scaffold 762 for all individuals ( $N = 37$ ), males ( $N = 20$ ), and females ( $N = 17$ ) of *S. hypoxantha* and *S. iberaensis*. Asterisks indicate the divergence peaks that contained at least one individual SNP with  $F_{ST}$  greater than 0.85 when all individuals were compared (as in Fig. 1F). Female capuchinos were classified based on the phenotype of their social mate.

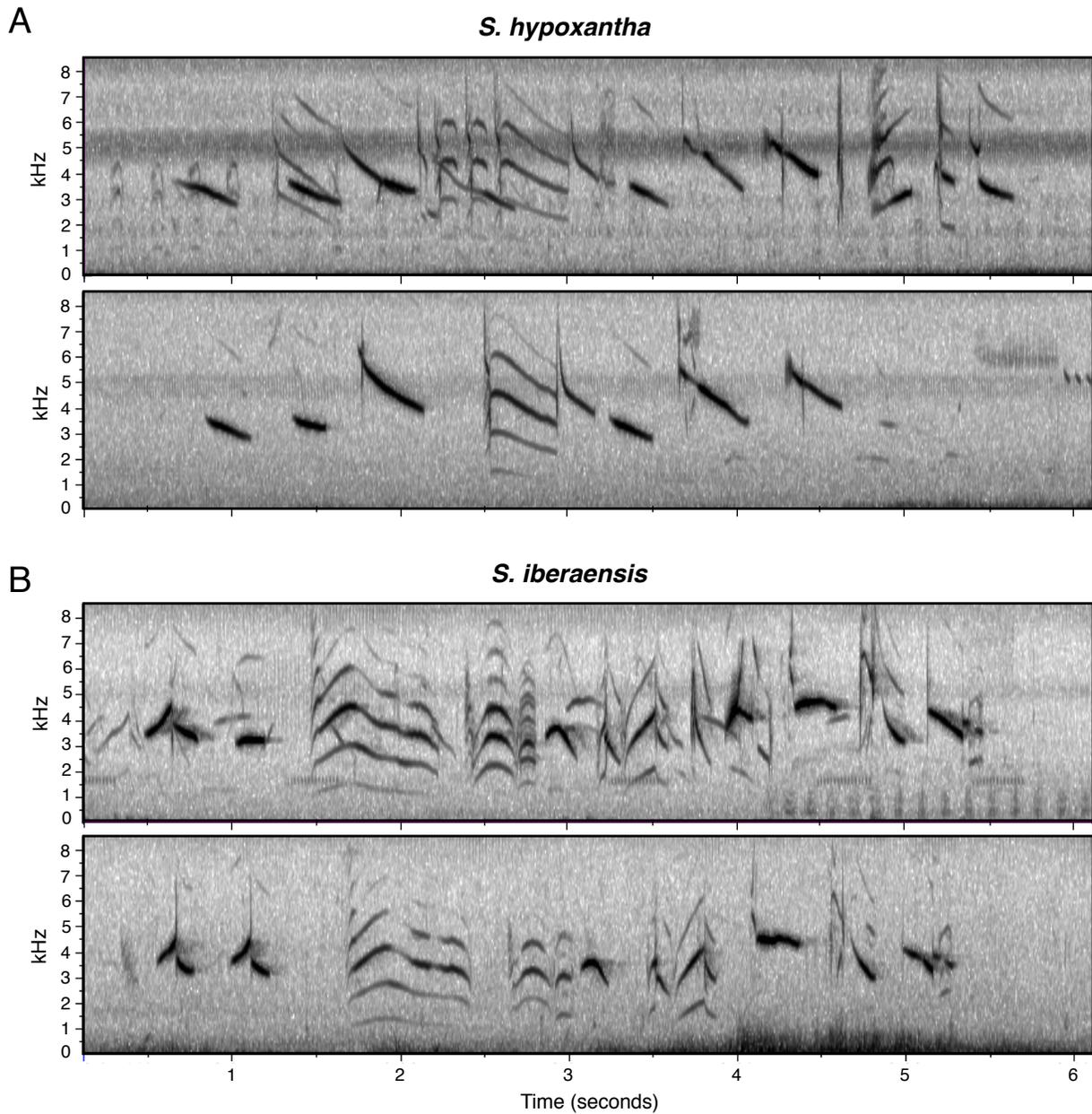


**Fig. S11. Few SNPs in the ddRAD data fell within the  $F_{ST}$  peaks found in the whole-genome data. (A)** Manhattan plot generated with ddRAD data showing the pattern of genomic differentiation between *S. hypoxantha* and *S. iberaensis* ( $N = 206$ ). The 28 SNPs that fell within  $F_{ST}$  peaks identified from the whole-genome data are highlighted in red and labeled according to their scaffold and corresponding chromosome in the zebra finch assembly. The plot contains 571 scaffolds. **(B)** PCA generated with ddRAD data depicting the pattern of genomic differentiation between males, females, and nestlings of *S. hypoxantha* and *S. iberaensis* when the 28 SNPs that fell within  $F_{ST}$  peaks identified from the whole-genome data were removed ( $N = 206$  individuals). The ddRAD data distinguished the two species even when the SNPs that fell within  $F_{ST}$  peaks were excluded. The two species are separated in PC1, while *S. hypoxantha* shows additional variation in PC2.



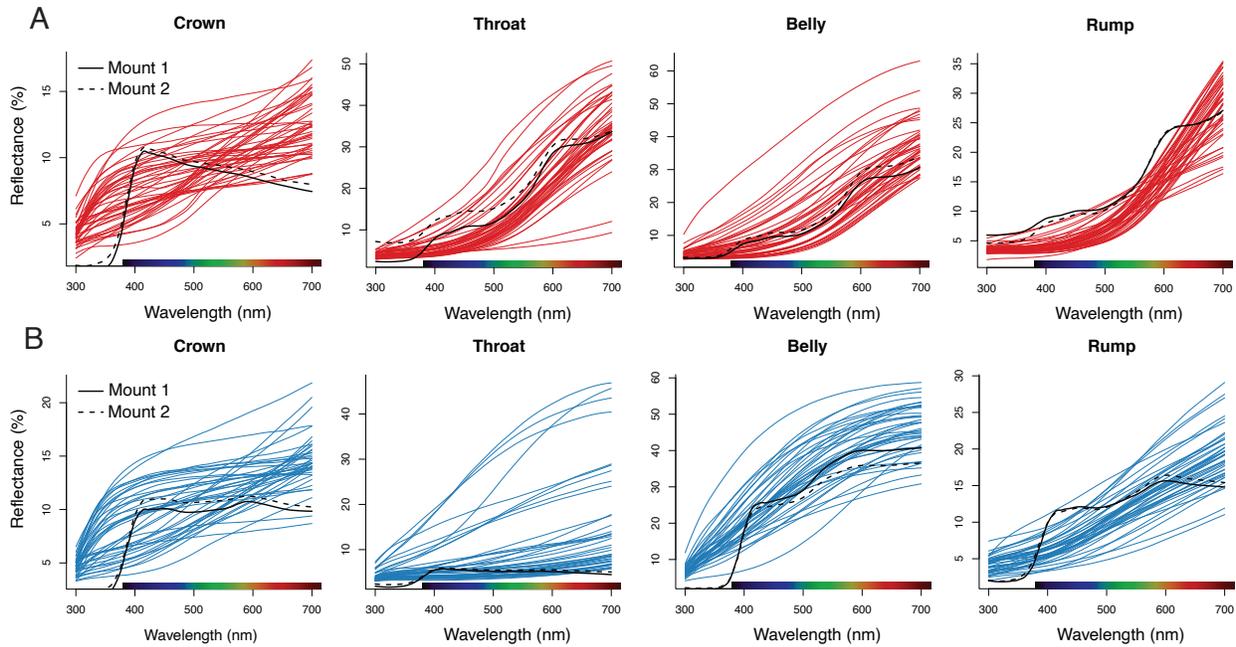
**Fig. S12. *S. hypoxantha* and *S. iberaensis* breed synchronously in Iberá National Park.**

Histograms showing the number of nests of *S. hypoxantha* (red) and *S. iberaensis* (blue) found in the San Nicolás Reserve over the course of two breeding seasons: **(A)** 2016-2017 ( $N = 52$ ) and **(B)** 2018-2019 ( $N = 76$ ). Overlapping bars are shown in purple.

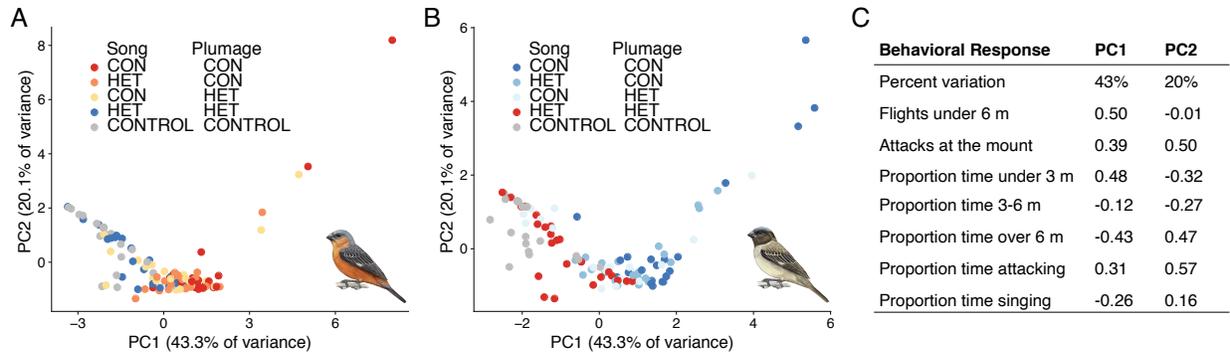


**Fig. S13. Differences in male song between *S. hypoxantha* and *S. iberensis*.**

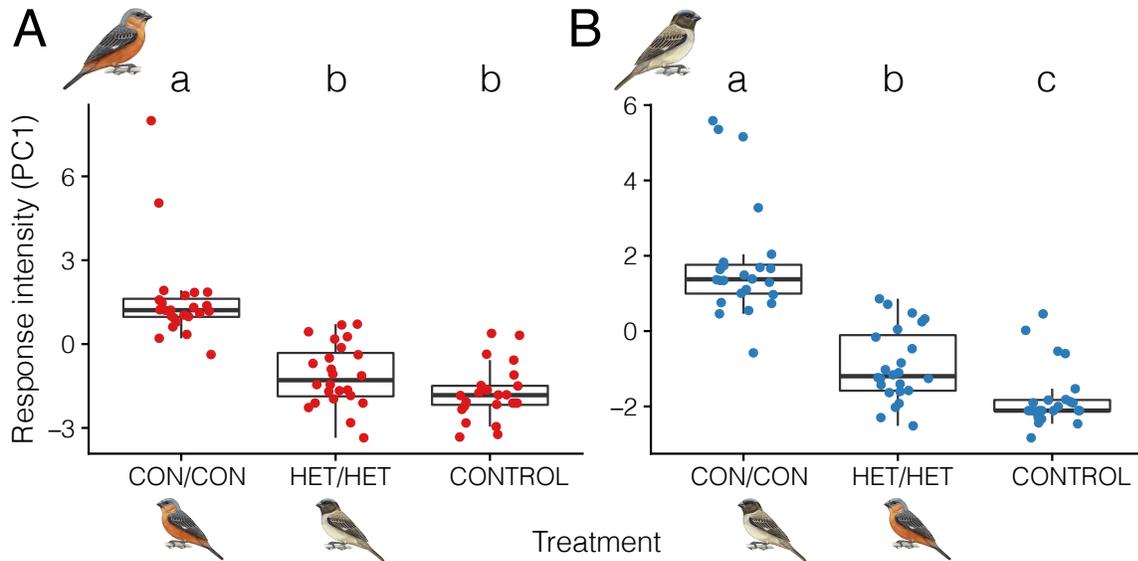
Representative spectrograms showing the typical song of two (A) *S. hypoxantha* and (B) *S. iberensis* males.



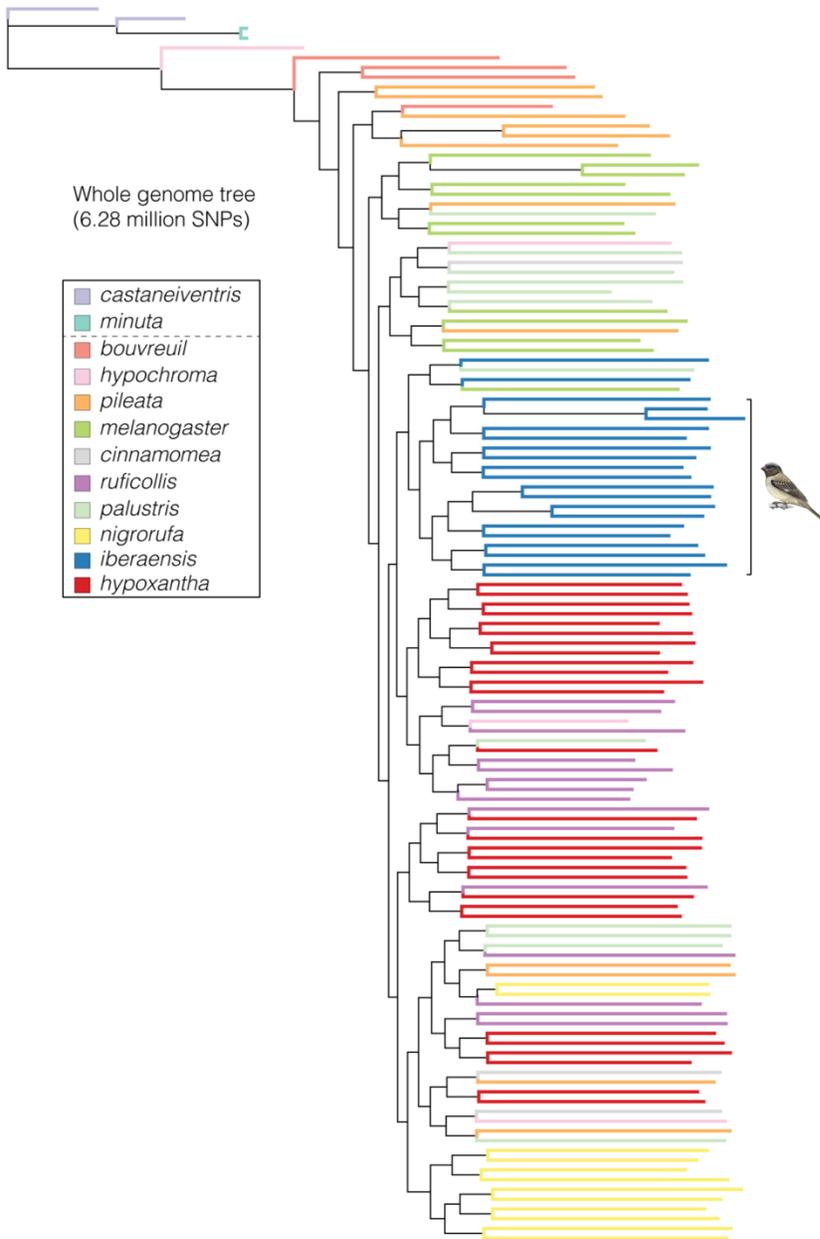
**Fig. S14. Comparison between the plumage coloration of breeding males and the mounts used in the behavioral experiment.** Reflectance patterns across the avian visual spectrum for the crown, throat, belly, and rump of the two mounts used for each species (solid black and dashed lines) in the behavioral experiment and breeding males of **(A)** *S. hypoxantha* ( $N = 46$ ) and **(B)** *S. iberaensis* ( $N = 41$ ).



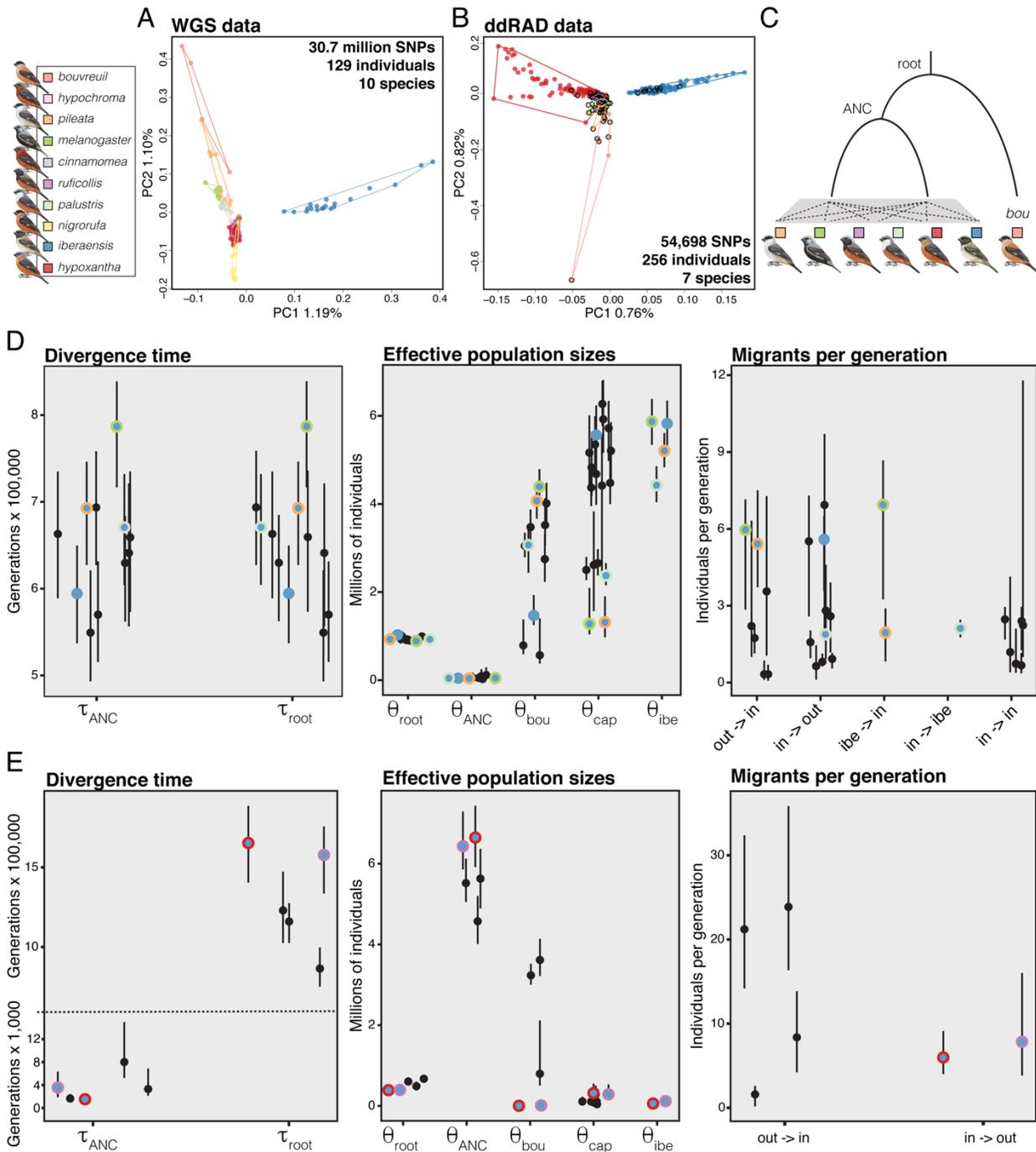
**Fig. S15. Principal component analysis (PCA) of the correlation matrix of the behavioral response variables.** PC axes for territorial males of **(A)** *S. hypoxantha* ( $N = 120$ ) and **(B)** *S. iberaensis* ( $N = 120$ ) when presented with various combinations of conspecific (CON), heterospecific (HET), and *S. collaris* (CONTROL) song and plumage. Points are colored by treatment. **(C)** Behavioral responses and their factor loadings for the first two principal components. Positive values of PC1 were associated with more flights within six meters of the mount, more attacks at the mount, and a greater proportion of time spent within three meters of mount.



**Fig. S16. Territorial males of (A) *S. hypoxantha* and (B) *S. iberaensis* responded more aggressively to conspecific song and plumage (CON/CON) than that of the heterospecific capuchino (HET/HET) or control (*S. collaris*; CONTROL). Points are colored by species and positive values of PC1 were associated with more flights within six meters of the mount, more attacks at the mount, and a greater proportion of time spent within three meters of mount ( $N = 24$  trials per treatment). We ran generalized linear mixed models examining the effect of treatment group on response intensity when the mismatched treatments were excluded. Different letters indicate statistical significance between treatment groups (Tukey HSD method; adjusted  $P < 0.05$ ). For *S. hypoxantha*, we detected significant differences between the CON/CON and HET/HET treatments ( $P < 0.0001$ ) and the CON/CON and CONTROL treatments ( $P < 0.0001$ ). For *S. iberaensis*, we detected significant differences between all treatment groups (CON/CON vs. HET/HET:  $P < 0.0001$ , CON/CON vs. CONTROL:  $P < 0.0001$ , HET/HET vs. CONTROL:  $P = 0.008$ ). When outliers (observations outside  $1.5 \times$  interquartile range; 7 for *S. hypoxantha* and 10 for *S. iberaensis*) were removed, we detected an additional significant difference in the response of *S. hypoxantha* to the HET/HET and CONTROL treatments ( $P = 0.003$ ).**



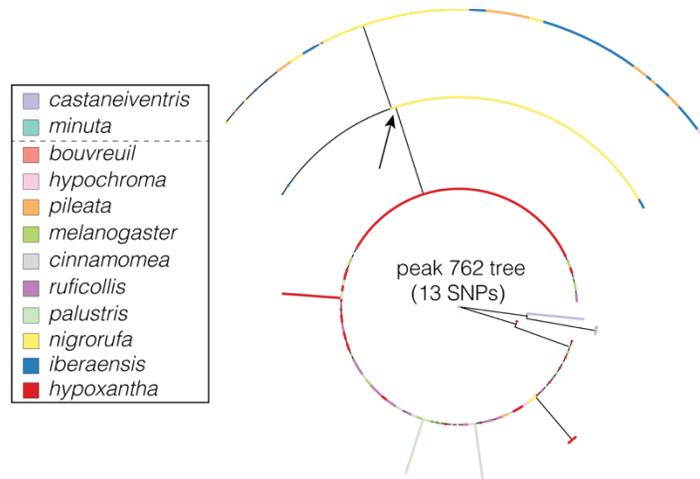
**Fig. S17. Phylogenetic tree inferred using maximum likelihood depicting the relationship between ten capuchino species and two outgroups based on whole-genome data.** This dataset retained 6,283,771 SNPs that were present in at least 85% of individuals and had a minor allele frequency of at least 10%. The black square bracket to the right of the tips in the tree indicates the clade containing the majority of *S. iberaensis* individuals. Outgroups are indicated above the dashed line in the legend.



**Fig. S18. Demographic reconstructions for southern capuchino seedeaters.**

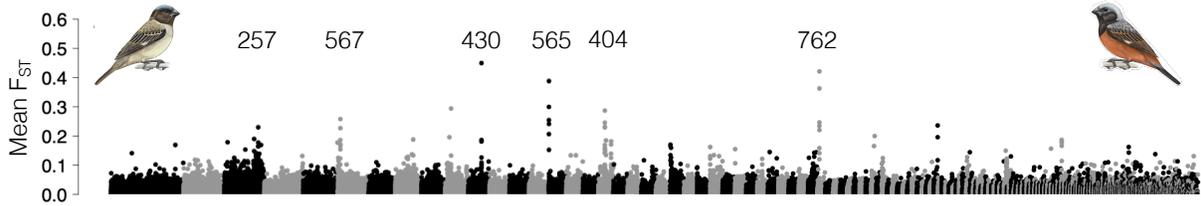
Clustering patterns in PCAs for capuchino individuals based on (A) whole-genome sequencing (WGS) data with the individuals and species shown in Fig. 6, and (B) ddRAD data. For the PCA based on WGS data, we excluded the outgroups *S. minuta* and *S. castaneiventris*, one outlier *S.*

*hypochroma* individual, and three other individuals showing high levels of relatedness with respect to another member of their species (one *S. iberensis*, one *S. pileata*, and one *S. melanogaster*). The individuals used for demographic modeling are shown in (B) as circles with a black outline. (C) Tree representing the ancestral and current populations used for demographic reconstructions. The results from the 16 G-PhoCS runs could be grouped into two general patterns. (D) The first pattern, which included most runs, identified similar divergence times within the ingroup and between the ingroup and *S. bouvreuil*, inferred a smaller ancestral effective population size than the current effective population sizes of the ingroup taxa, and detected gene flow among the three populations. (E) The second pattern included five runs that identified a much more recent split between the ingroup taxa, estimated smaller current effective population sizes for the ingroup, and only detected gene flow with the outgroup. For each run, the posterior median value is shown with a circle and the 95% credible interval with a vertical black line. Blue circles represent estimates from models that included *S. iberensis*, with the second ingroup species indicated by the color of the outline. The solid blue circles represent the comparison of *S. iberensis* vs. all other ingroup species combined. *S. bouvreuil* is abbreviated as bou, capuchinos as cap, and *S. iberensis* as ibe. Note the split y-axis for the divergence time plot in (E). Only estimates of migration with 95% credible intervals that do not overlap with zero are shown. The two highest migration estimates from the ingroup to the outgroup in (E) are not shown in the plot due to their large 95% credible interval (ranging from 61 to 343 and 6 to 97, respectively). The split between *S. iberensis* and *S. hypoxantha*, estimated at 1500 generations, was the most recent split identified between two capuchinos.

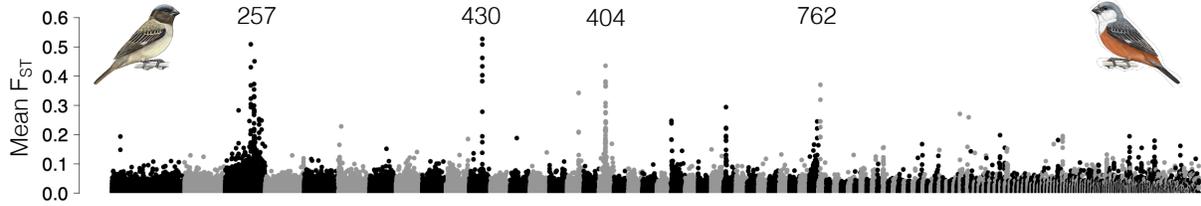


**Fig. S19. Phylogenetic tree inferred using maximum likelihood depicting the relationship between ten capuchino species and two outgroups based on the 13 SNPs that fell within the  $F_{ST}$  peak on scaffold 762.** The tree was generated using whole-genome data and the arrow indicates the clade that contains all individuals of *S. iberensis*, as well as many other species.

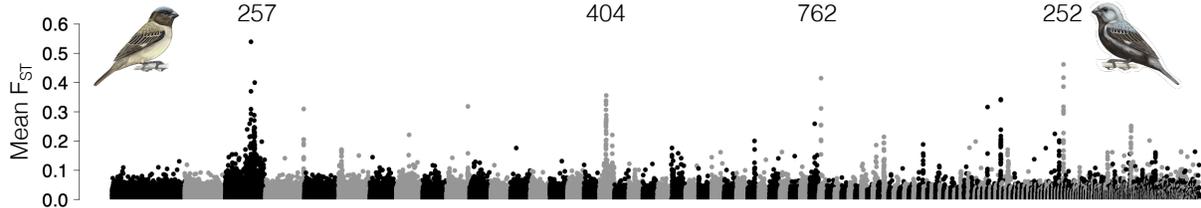
*S. iberensis* vs. *S. ruficollis*:  $F_{ST} = 0.0037 \pm 0.0184$



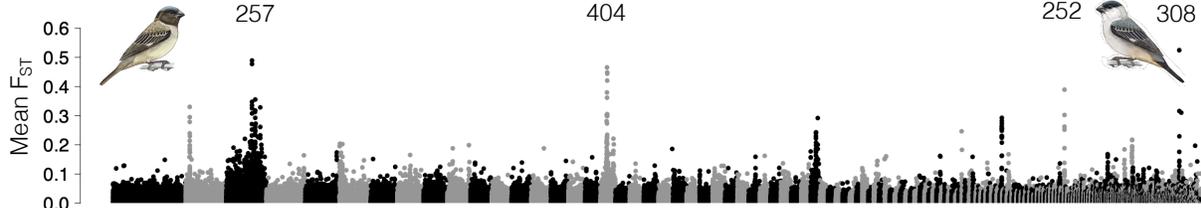
*S. iberensis* vs. *S. palustris*:  $F_{ST} = 0.0055 \pm 0.0202$



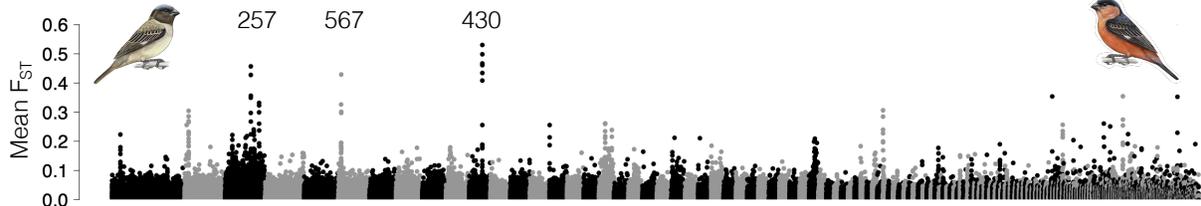
*S. iberensis* vs. *S. melanogaster*:  $F_{ST} = 0.0080 \pm 0.0208$



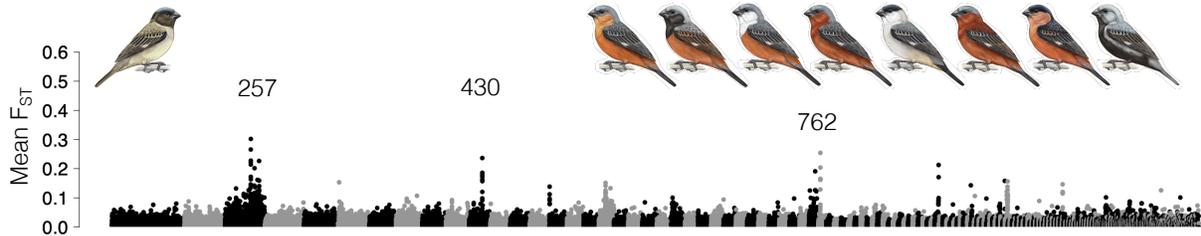
*S. iberensis* vs. *S. pileata*:  $F_{ST} = 0.0092 \pm 0.0221$



*S. iberensis* vs. *S. nigrorufa*:  $F_{ST} = 0.0115 \pm 0.0256$



*S. iberensis* vs. all capuchinos



Scaffolds ordered by size

**Fig. S20. Manhattan plots from the whole-genome data showing regions of elevated genomic differentiation ( $F_{ST}$ ) over 5-kb windows between *S. iberaensis* and other southern capuchino species.** Genome-wide  $F_{ST}$  (mean  $\pm$  SD) is indicated for each comparison and the main peaks are labeled with their corresponding scaffold, which matched the following chromosomes in the zebra finch assembly: Chr 1 (scaffold 430), Chr 2 (scaffold 567), Chr 4 (scaffold 565), Chr 11 (scaffold 762), Chr 20 (scaffold 252), and Chr Z (scaffolds 257, 404, and 308). The plots contain the 381 largest scaffolds. The final comparison examined differentiation between *S. iberaensis* and eight other capuchino species combined ( $N = 21$  *S. iberaensis*, 28 *S. hypoxantha*, 15 *S. ruficollis*, 12 *S. pileata*, 12 *S. palustris*, 12 *S. melanogaster*, 12 *S. nigrorufa*, 3 *S. hypochroma*, and 3 *S. cinnamomea*). The additional divergence peaks contain *SLC45A2* (scaffold 404) and *ASIP* (scaffold 252), two genes involved in the melanogenesis pathway (108).

#### 4 Supplementary Tables

**Table S1. Individuals sampled for the study and the type of genetic data obtained for each individual.** We generated whole-genome sequencing (WGS) data for 37 individuals of *S. hypoxantha* ( $N = 8$  males, 8 females) and *S. iberaensis* ( $N = 12$  males, 9 females) and double-digest restriction site-associated DNA (ddRAD) sequencing data for 206 individuals of *S. hypoxantha* ( $N = 40$  males, 23 females, 40 nestlings, and 2 unhatched eggs) and *S. iberaensis* ( $N = 42$  males, 21 females, 37 nestlings, and 1 unhatched egg). The Sanger column indicates the individuals for which we obtained Sanger sequences for a region in the peak on scaffold 430 (Fig. 2).

Species	Sex	Band No.	Year	WGS	ddRAD	Sanger
<i>S. hypoxantha</i>	Male	B009522	2016	Y	Y	N
<i>S. hypoxantha</i>	Male	B009524	2016	Y	Y	N
<i>S. hypoxantha</i>	Male	B009525	2016	Y	Y	N
<i>S. hypoxantha</i>	Male	B009526	2016	Y	Y	N
<i>S. hypoxantha</i>	Male	B009530	2016	Y	Y	N
<i>S. hypoxantha</i>	Male	B009554	2016	Y	Y	N
<i>S. hypoxantha</i>	Male	B009556	2016	Y	Y	N
<i>S. hypoxantha</i>	Male	B009574	2016	Y	Y	N
<i>S. hypoxantha</i>	Female	B009503	2016	Y	Y	N
<i>S. hypoxantha</i>	Female	B009523	2016	Y	Y	N
<i>S. hypoxantha</i>	Female	B009529	2016	Y	Y	N
<i>S. hypoxantha</i>	Female	B009555	2016	Y	Y	N
<i>S. hypoxantha</i>	Female	B009557	2016	Y	Y	N
<i>S. hypoxantha</i>	Female	B009570	2016	Y	Y	N
<i>S. hypoxantha</i>	Female	B009573	2016	Y	Y	N
<i>S. hypoxantha</i>	Female	B009580	2016	Y	Y	N
<i>S. iberaensis</i>	Male	B009502	2016	Y	Y	N
<i>S. iberaensis</i>	Male	B009504	2016	Y	Y	N
<i>S. iberaensis</i>	Male	B009508	2016	Y	Y	N
<i>S. iberaensis</i>	Male	B009528	2016	Y	Y	N
<i>S. iberaensis</i>	Male	B009540	2016	Y	Y	N

<i>S. iberaensis</i>	Male	B009543	2016	Y	Y	N
<i>S. iberaensis</i>	Male	B009552	2016	Y	Y	N
<i>S. iberaensis</i>	Male	B009560	2016	Y	Y	N
<i>S. iberaensis</i>	Male	B009561	2016	Y	Y	N
<i>S. iberaensis</i>	Male	B009562	2016	Y	Y	N
<i>S. iberaensis</i>	Male	B009583	2016	Y	Y	N
<i>S. iberaensis</i>	Male	B009592	2016	Y	Y	N
<i>S. iberaensis</i>	Female	B009501	2016	Y	Y	N
<i>S. iberaensis</i>	Female	B009505	2016	Y	Y	N
<i>S. iberaensis</i>	Female	B009507	2016	Y	Y	N
<i>S. iberaensis</i>	Female	B009527	2016	Y	Y	N
<i>S. iberaensis</i>	Female	B009542	2016	Y	Y	N
<i>S. iberaensis</i>	Female	B009581	2016	Y	Y	N
<i>S. iberaensis</i>	Female	B009587	2016	Y	Y	N
<i>S. iberaensis</i>	Female	B009590	2016	Y	Y	N
<i>S. iberaensis</i>	Female	B009595	2016	Y	Y	N
<i>S. hypoxantha</i>	Unhatched Egg	UH_Egg_1	2016	N	Y	N
<i>S. hypoxantha</i>	Nestling	B009514	2016	N	Y	Y
<i>S. hypoxantha</i>	Nestling	B009515	2016	N	Y	Y
<i>S. hypoxantha</i>	Nestling	B009516	2016	N	Y	Y
<i>S. hypoxantha</i>	Nestling	B009517	2016	N	Y	Y
<i>S. hypoxantha</i>	Nestling	B009518	2016	N	Y	Y
<i>S. hypoxantha</i>	Nestling	B009519	2016	N	Y	Y
<i>S. hypoxantha</i>	Nestling	B009531	2016	N	Y	Y
<i>S. hypoxantha</i>	Nestling	B009532	2016	N	Y	Y
<i>S. hypoxantha</i>	Nestling	B009537	2016	N	Y	Y
<i>S. hypoxantha</i>	Nestling	B009538	2016	N	Y	Y
<i>S. hypoxantha</i>	Nestling	B009567	2016	N	Y	Y
<i>S. hypoxantha</i>	Nestling	B009569	2016	N	Y	Y
<i>S. hypoxantha</i>	Nestling	B009571	2016	N	Y	Y
<i>S. hypoxantha</i>	Nestling	B009572	2016	N	Y	Y
<i>S. hypoxantha</i>	Nestling	B009575	2016	N	Y	Y
<i>S. hypoxantha</i>	Nestling	B009576	2016	N	Y	Y
<i>S. hypoxantha</i>	Nestling	B009578	2016	N	Y	Y
<i>S. hypoxantha</i>	Nestling	B009579	2016	N	Y	Y
<i>S. hypoxantha</i>	Nestling	B009584	2016	N	Y	Y
<i>S. hypoxantha</i>	Nestling	B009585	2016	N	Y	Y
<i>S. hypoxantha</i>	Male	B009533	2016	N	Y	Y
<i>S. hypoxantha</i>	Male	B009534	2016	N	Y	Y

<i>S. hypoxantha</i>	Male	B009541	2016	N	Y	Y
<i>S. hypoxantha</i>	Male	B009553	2016	N	Y	Y
<i>S. hypoxantha</i>	Male	B009559	2016	N	Y	Y
<i>S. hypoxantha</i>	Male	B009591	2016	N	Y	Y
<i>S. hypoxantha</i>	Female	B009506	2016	N	Y	Y
<i>S. hypoxantha</i>	Female	B009509	2016	N	Y	Y
<i>S. iberaensis</i>	Nestling	B009510	2016	N	Y	Y
<i>S. iberaensis</i>	Nestling	B009511	2016	N	Y	Y
<i>S. iberaensis</i>	Nestling	B009512	2016	N	Y	Y
<i>S. iberaensis</i>	Nestling	B009513	2016	N	Y	Y
<i>S. iberaensis</i>	Nestling	B009535	2016	N	Y	Y
<i>S. iberaensis</i>	Nestling	B009536	2016	N	Y	Y
<i>S. iberaensis</i>	Nestling	B009544	2016	N	Y	Y
<i>S. iberaensis</i>	Nestling	B009545	2016	N	Y	Y
<i>S. iberaensis</i>	Nestling	B009546	2016	N	Y	Y
<i>S. iberaensis</i>	Nestling	B009547	2016	N	Y	Y
<i>S. iberaensis</i>	Nestling	B009548	2016	N	Y	Y
<i>S. iberaensis</i>	Nestling	B009549	2016	N	Y	Y
<i>S. iberaensis</i>	Nestling	B009551	2016	N	Y	Y
<i>S. iberaensis</i>	Nestling	B009588	2016	N	Y	Y
<i>S. iberaensis</i>	Nestling	B009589	2016	N	Y	Y
<i>S. iberaensis</i>	Nestling	B009593	2016	N	Y	Y
<i>S. iberaensis</i>	Nestling	B009594	2016	N	Y	Y
<i>S. iberaensis</i>	Nestling	B009596	2016	N	Y	Y
<i>S. iberaensis</i>	Nestling	B009597	2016	N	Y	Y
<i>S. iberaensis</i>	Male	B009539	2016	N	Y	Y
<i>S. iberaensis</i>	Male	B009550	2016	N	Y	Y
<i>S. iberaensis</i>	Male	B009563	2016	N	Y	Y
<i>S. iberaensis</i>	Male	B009564	2016	N	Y	Y
<i>S. iberaensis</i>	Male	B009565	2016	N	Y	Y
<i>S. iberaensis</i>	Male	B009566	2016	N	Y	Y
<i>S. iberaensis</i>	Male	B009582	2016	N	Y	Y
<i>S. iberaensis</i>	Female	B009586	2016	N	Y	Y
<i>S. hypoxantha</i>	Unhatched Egg	UH_Egg_2	2018	N	Y	Y
<i>S. hypoxantha</i>	Nestling	B009063	2018	N	Y	Y
<i>S. hypoxantha</i>	Nestling	B009064	2018	N	Y	Y
<i>S. hypoxantha</i>	Nestling	B009079	2018	N	Y	Y
<i>S. hypoxantha</i>	Nestling	B009080	2018	N	Y	Y
<i>S. hypoxantha</i>	Nestling	B009108	2018	N	Y	Y

<i>S. hypoxantha</i>	Nestling	B009109	2018	N	Y	Y
<i>S. hypoxantha</i>	Nestling	B009110	2018	N	Y	Y
<i>S. hypoxantha</i>	Nestling	B009111	2018	N	Y	Y
<i>S. hypoxantha</i>	Nestling	B009112	2018	N	Y	Y
<i>S. hypoxantha</i>	Nestling	B009607	2018	N	Y	Y
<i>S. hypoxantha</i>	Nestling	B009608	2018	N	Y	Y
<i>S. hypoxantha</i>	Nestling	B009609	2018	N	Y	Y
<i>S. hypoxantha</i>	Nestling	B009610	2018	N	Y	Y
<i>S. hypoxantha</i>	Nestling	B009613	2018	N	Y	Y
<i>S. hypoxantha</i>	Nestling	B009614	2018	N	Y	Y
<i>S. hypoxantha</i>	Nestling	B009622	2018	N	Y	Y
<i>S. hypoxantha</i>	Nestling	B009623	2018	N	Y	Y
<i>S. hypoxantha</i>	Nestling	B009625	2018	N	Y	Y
<i>S. hypoxantha</i>	Nestling	B009626	2018	N	Y	Y
<i>S. hypoxantha</i>	Nestling	Nestling_1	2018	N	Y	Y
<i>S. hypoxantha</i>	Male	B009027	2018	N	Y	Y
<i>S. hypoxantha</i>	Male	B009028	2018	N	Y	Y
<i>S. hypoxantha</i>	Male	B009030	2018	N	Y	Y
<i>S. hypoxantha</i>	Male	B009031	2018	N	Y	Y
<i>S. hypoxantha</i>	Male	B009032	2018	N	Y	Y
<i>S. hypoxantha</i>	Male	B009033	2018	N	Y	Y
<i>S. hypoxantha</i>	Male	B009035	2018	N	Y	Y
<i>S. hypoxantha</i>	Male	B009055	2018	N	Y	Y
<i>S. hypoxantha</i>	Male	B009056	2018	N	Y	Y
<i>S. hypoxantha</i>	Male	B009057	2018	N	Y	Y
<i>S. hypoxantha</i>	Male	B009059	2018	N	Y	Y
<i>S. hypoxantha</i>	Male	B009060	2018	N	Y	Y
<i>S. hypoxantha</i>	Male	B009067	2018	N	Y	Y
<i>S. hypoxantha</i>	Male	B009077	2018	N	Y	Y
<i>S. hypoxantha</i>	Male	B009078	2018	N	Y	N
<i>S. hypoxantha</i>	Male	B009107	2018	N	Y	Y
<i>S. hypoxantha</i>	Male	B009116	2018	N	Y	Y
<i>S. hypoxantha</i>	Male	B009117	2018	N	Y	Y
<i>S. hypoxantha</i>	Male	B009118	2018	N	Y	Y
<i>S. hypoxantha</i>	Male	B009119	2018	N	Y	Y
<i>S. hypoxantha</i>	Male	B009122	2018	N	Y	Y
<i>S. hypoxantha</i>	Male	B009133	2018	N	Y	N
<i>S. hypoxantha</i>	Male	B009134	2018	N	Y	Y
<i>S. hypoxantha</i>	Male	B009601	2018	N	Y	Y

<i>S. hypoxantha</i>	Male	B009612	2018	N	Y	Y
<i>S. hypoxantha</i>	Male	B009615	2018	N	Y	Y
<i>S. hypoxantha</i>	Female	B009050	2018	N	Y	Y
<i>S. hypoxantha</i>	Female	B009051	2018	N	Y	Y
<i>S. hypoxantha</i>	Female	B009053	2018	N	Y	Y
<i>S. hypoxantha</i>	Female	B009066	2018	N	Y	Y
<i>S. hypoxantha</i>	Female	B009072	2018	N	Y	Y
<i>S. hypoxantha</i>	Female	B009073	2018	N	Y	Y
<i>S. hypoxantha</i>	Female	B009081	2018	N	Y	Y
<i>S. hypoxantha</i>	Female	B009520	2018	N	Y	Y
<i>S. hypoxantha</i>	Female	B009606	2018	N	Y	Y
<i>S. hypoxantha</i>	Female	B009616	2018	N	Y	Y
<i>S. hypoxantha</i>	Female	B009617	2018	N	Y	Y
<i>S. hypoxantha</i>	Female	B009618	2018	N	Y	Y
<i>S. hypoxantha</i>	Female	B009621	2018	N	Y	Y
<i>S. iberaensis</i>	Unhatched Egg	UH_Egg_3	2018	N	Y	N
<i>S. iberaensis</i>	Nestling	B009036	2018	N	Y	Y
<i>S. iberaensis</i>	Nestling	B009037	2018	N	Y	Y
<i>S. iberaensis</i>	Nestling	B009038	2018	N	Y	Y
<i>S. iberaensis</i>	Nestling	B009061	2018	N	Y	Y
<i>S. iberaensis</i>	Nestling	B009062	2018	N	Y	Y
<i>S. iberaensis</i>	Nestling	B009106	2018	N	Y	Y
<i>S. iberaensis</i>	Nestling	B009123	2018	N	Y	Y
<i>S. iberaensis</i>	Nestling	B009124	2018	N	Y	Y
<i>S. iberaensis</i>	Nestling	B009130	2018	N	Y	Y
<i>S. iberaensis</i>	Nestling	B009131	2018	N	Y	Y
<i>S. iberaensis</i>	Nestling	B009603	2018	N	Y	Y
<i>S. iberaensis</i>	Nestling	B009604	2018	N	Y	Y
<i>S. iberaensis</i>	Nestling	B009605	2018	N	Y	Y
<i>S. iberaensis</i>	Nestling	B009627	2018	N	Y	Y
<i>S. iberaensis</i>	Nestling	B009628	2018	N	Y	Y
<i>S. iberaensis</i>	Nestling	B009634	2018	N	Y	Y
<i>S. iberaensis</i>	Nestling	B009635	2018	N	Y	Y
<i>S. iberaensis</i>	Nestling	Nestling_2	2018	N	Y	Y
<i>S. iberaensis</i>	Male	B009029	2018	N	Y	Y
<i>S. iberaensis</i>	Male	B009040	2018	N	Y	Y
<i>S. iberaensis</i>	Male	B009041	2018	N	Y	Y
<i>S. iberaensis</i>	Male	B009042	2018	N	Y	Y
<i>S. iberaensis</i>	Male	B009043	2018	N	Y	Y

<i>S. iberensis</i>	Male	B009046	2018	N	Y	Y
<i>S. iberensis</i>	Male	B009047	2018	N	Y	Y
<i>S. iberensis</i>	Male	B009049	2018	N	Y	Y
<i>S. iberensis</i>	Male	B009054	2018	N	Y	Y
<i>S. iberensis</i>	Male	B009058	2018	N	Y	Y
<i>S. iberensis</i>	Male	B009074	2018	N	Y	Y
<i>S. iberensis</i>	Male	B009076	2018	N	Y	Y
<i>S. iberensis</i>	Male	B009113	2018	N	Y	Y
<i>S. iberensis</i>	Male	B009114	2018	N	Y	Y
<i>S. iberensis</i>	Male	B009115	2018	N	Y	Y
<i>S. iberensis</i>	Male	B009121	2018	N	Y	Y
<i>S. iberensis</i>	Male	B009128	2018	N	Y	Y
<i>S. iberensis</i>	Male	B009132	2018	N	Y	Y
<i>S. iberensis</i>	Male	B009139	2018	N	Y	Y
<i>S. iberensis</i>	Male	B009141	2018	N	Y	Y
<i>S. iberensis</i>	Male	B009602	2018	N	Y	Y
<i>S. iberensis</i>	Male	B009620	2018	N	Y	Y
<i>S. iberensis</i>	Male	B009630	2018	N	Y	Y
<i>S. iberensis</i>	Female	B009052	2018	N	Y	Y
<i>S. iberensis</i>	Female	B009083	2018	N	Y	Y
<i>S. iberensis</i>	Female	B009085	2018	N	Y	Y
<i>S. iberensis</i>	Female	B009086	2018	N	Y	Y
<i>S. iberensis</i>	Female	B009120	2018	N	Y	Y
<i>S. iberensis</i>	Female	B009127	2018	N	Y	Y
<i>S. iberensis</i>	Female	B009135	2018	N	Y	Y
<i>S. iberensis</i>	Female	B009140	2018	N	Y	Y
<i>S. iberensis</i>	Female	B009619	2018	N	Y	Y
<i>S. iberensis</i>	Female	B009629	2018	N	Y	Y
<i>S. iberensis</i>	Female	B009633	2018	N	Y	Y

**Table S2. Sequencing statistics for the whole-genome data.**

Band No.	Species	Raw Reads (M)	Expected Coverage <sup>1</sup>	Retained Reads (M)	% Discarded (QC)	Expected Coverage <sup>2</sup>	% Alignment	Depth of Coverage <sup>3</sup>	% Missing in VCF
B009522	<i>S. hypoxantha</i>	45.3	5.7	42.9	5.2	5.7	98.8	5.5	3.8
B009524	<i>S. hypoxantha</i>	46.2	5.8	42.8	7.4	5.6	98.7	5.6	3.7
B009525	<i>S. hypoxantha</i>	57.1	7.2	51.8	9.2	6.8	98.7	6.9	2.0
B009526	<i>S. hypoxantha</i>	57.2	7.2	50.8	11.2	6.7	98.9	6.9	1.8
B009530	<i>S. hypoxantha</i>	42.3	5.3	40.0	5.3	5.3	98.7	5.1	4.9
B009554	<i>S. hypoxantha</i>	38.0	4.8	34.8	8.4	4.6	98.4	4.6	7.5
B009556	<i>S. hypoxantha</i>	45.0	5.7	40.9	9.1	5.4	98.5	5.4	4.6
B009574	<i>S. hypoxantha</i>	59.1	7.4	54.2	8.3	7.1	98.5	7.1	2.2
B009503	<i>S. hypoxantha</i>	38.5	4.8	35.8	6.9	4.7	98.5	4.6	7.2
B009523	<i>S. hypoxantha</i>	45.2	5.7	41.3	8.7	5.4	98.6	5.4	4.1
B009529	<i>S. hypoxantha</i>	53.2	6.7	50.3	5.5	6.6	98.6	6.4	2.6
B009555	<i>S. hypoxantha</i>	43.2	5.4	39.0	9.8	5.1	98.3	5.1	5.4
B009557	<i>S. hypoxantha</i>	41.4	5.2	38.4	7.3	5.1	98.5	5.0	5.9
B009570	<i>S. hypoxantha</i>	47.5	6.0	44.3	6.8	5.8	98.3	5.7	4.4
B009573	<i>S. hypoxantha</i>	48.8	6.1	46.2	5.4	6.1	98.5	5.9	4.1
B009580	<i>S. hypoxantha</i>	55.8	7.0	52.5	5.9	6.9	98.4	6.7	2.9
B009502	<i>S. iberensis</i>	43.7	5.5	41.4	5.1	5.5	98.7	5.3	4.6
B009504	<i>S. iberensis</i>	47.2	5.9	44.5	5.7	5.9	98.7	5.7	3.6
B009508	<i>S. iberensis</i>	42.1	5.3	40.2	4.6	5.3	98.7	5.1	5.0
B009528	<i>S. iberensis</i>	37.5	4.7	35.6	5.1	4.7	98.7	4.6	6.7
B009540	<i>S. iberensis</i>	32.0	4.0	30.2	5.5	4.0	98.8	3.9	9.5
B009543	<i>S. iberensis</i>	48.0	6.0	45.5	5.3	6.0	98.6	5.8	3.4
B009552	<i>S. iberensis</i>	39.0	4.9	36.3	6.9	4.8	98.5	4.7	6.8
B009560	<i>S. iberensis</i>	40.0	5.0	37.8	5.5	5.0	98.6	4.9	6.6
B009561	<i>S. iberensis</i>	49.9	6.3	47.3	5.2	6.2	98.5	6.0	4.0
B009562	<i>S. iberensis</i>	31.3	3.9	29.3	6.3	3.9	98.3	3.8	12.3
B009583	<i>S. iberensis</i>	57.5	7.2	53.5	6.9	7.0	98.5	6.9	2.4
B009592	<i>S. iberensis</i>	46.1	5.8	41.1	10.7	5.4	98.6	5.5	4.2
B009501	<i>S. iberensis</i>	43.1	5.4	40.4	6.3	5.3	98.8	5.2	4.5
B009505	<i>S. iberensis</i>	48.7	6.1	45.6	6.4	6.0	98.7	5.9	3.3
B009507	<i>S. iberensis</i>	40.3	5.1	37.7	6.5	5.0	98.6	4.9	5.8
B009527	<i>S. iberensis</i>	46.6	5.9	44.2	5.3	5.8	98.5	5.6	4.0
B009542	<i>S. iberensis</i>	44.1	5.6	41.6	5.8	5.5	98.6	5.3	4.5
B009581	<i>S. iberensis</i>	41.9	5.3	39.5	5.6	5.2	98.5	5.1	6.1

B009587	<i>S. iberensis</i>	49.8	6.3	45.1	9.5	5.9	98.5	6.0	3.6
B009590	<i>S. iberensis</i>	44.1	5.6	38.3	13.2	5.0	98.4	5.2	5.2
B009595	<i>S. iberensis</i>	85.2	10.7	80.0	6.1	10.5	98.4	10.3	0.8

<sup>1</sup> Computed as the number of raw 151 bp reads divided by a genome size of ~1.2 Gbp.

<sup>2</sup> Computed as the number of reads retained after filtering multiplied by their average size, and then divided by a genome size of ~1.2 Gbp.

<sup>3</sup> Calculated from the bam files after alignment.

**Table S3. Position of highly divergent SNPs ( $F_{ST} > 0.85$ ) relative to identified genes.**

The ~37 kb of sequence between *OCA2* and *HERC2* produced a mean PhastCons conservation score of  $0.048 \pm 0.015$  SD. However, these values ranged from 0 to up to 0.997, with defined peaks and 215 positions showing a conservation score of 0.9 or higher. It is likely that these regions, which are conserved among distantly related species, contain cis-regulatory elements that are necessary to control the expression of *OCA2*. The highly divergent SNPs were located between 6,496 and 8,711 kb upstream of *OCA2* and from 30,506 kb downstream of *HERC2* to an intron within that gene. There was one highly differentiated SNP 17,815 kb upstream of *TYRP1*.

Scaffold	Position of highly divergent SNP	$F_{ST}$ value	Location relative to identified genes
430	11027082	0.852	Upstream of <i>OCA2</i> , Downstream of <i>HERC2</i>
430	11027706	0.857	Upstream of <i>OCA2</i> , Downstream of <i>HERC2</i>
430	11028530	0.915	Upstream of <i>OCA2</i> , Downstream of <i>HERC2</i>
430	11028725	0.874	Upstream of <i>OCA2</i> , Downstream of <i>HERC2</i>
430	11028788	0.887	Upstream of <i>OCA2</i> , Downstream of <i>HERC2</i>
430	11028845	0.884	Upstream of <i>OCA2</i> , Downstream of <i>HERC2</i>
430	11028855	0.941	Upstream of <i>OCA2</i> , Downstream of <i>HERC2</i>
430	11028917	0.884	Upstream of <i>OCA2</i> , Downstream of <i>HERC2</i>
430	11028925	0.882	Upstream of <i>OCA2</i> , Downstream of <i>HERC2</i>
430	11028946	0.858	Upstream of <i>OCA2</i> , Downstream of <i>HERC2</i>
430	11028951	0.858	Upstream of <i>OCA2</i> , Downstream of <i>HERC2</i>
430	11028961	0.858	Upstream of <i>OCA2</i> , Downstream of <i>HERC2</i>
430	11029143	0.855	Upstream of <i>OCA2</i> , Downstream of <i>HERC2</i>
430	11029157	0.883	Upstream of <i>OCA2</i> , Downstream of <i>HERC2</i>
430	11029206	0.916	Upstream of <i>OCA2</i> , Downstream of <i>HERC2</i>
430	11029270	0.885	Upstream of <i>OCA2</i> , Downstream of <i>HERC2</i>
430	11029304	0.885	Upstream of <i>OCA2</i> , Downstream of <i>HERC2</i>
430	11030000	0.862	Upstream of <i>OCA2</i> , Downstream of <i>HERC2</i>
430	11030054	0.884	Upstream of <i>OCA2</i> , Downstream of <i>HERC2</i>
430	11030110	0.884	Upstream of <i>OCA2</i> , Downstream of <i>HERC2</i>
430	11043140	0.879	Upstream of <i>OCA2</i> , Downstream of <i>HERC2</i>
430	11044894	0.878	Upstream of <i>OCA2</i> , Downstream of <i>HERC2</i>
430	11054240	0.933	Upstream of <i>OCA2</i> , Downstream of <i>HERC2</i>

430	11054328	0.885	Upstream of <i>OCA2</i> , Downstream of <i>HERC2</i>
430	11058206	0.858	Upstream of <i>OCA2</i> , Downstream of <i>HERC2</i>
430	11058495	0.875	Exon of <i>HERC2</i>
430	11058847	0.880	Intron of <i>HERC2</i>
430	11059368	0.873	Intron of <i>HERC2</i>
430	11060287	0.874	Intron of <i>HERC2</i>
430	11060381	0.906	Intron of <i>HERC2</i>
430	11060579	0.850	Intron of <i>HERC2</i>
430	11064841	0.850	Intron of <i>HERC2</i>
430	11064842	0.850	Intron of <i>HERC2</i>
430	11067576	0.879	Intron of <i>HERC2</i>
430	11067689	0.852	Intron of <i>HERC2</i>
430	11067987	0.883	Intron of <i>HERC2</i>
257	21661286	0.868	Intron of <i>TYRP1</i>
257	21661528	0.884	Intron of <i>TYRP1</i>
257	21661610	0.862	Intron of <i>TYRP1</i>
257	21687503	0.864	Upstream of <i>TYRP1</i>
762	1684725	0.878	Intron of <i>GTP2</i>
762	1685092	0.862	Intron of <i>GTP2</i>

**Table S4. Parentage assignments from the ddRAD data.**

<b>Nestling Band No.</b>	<b>Social Father Band No.</b>	<b>Social Father Species</b>	<b>Candidate Father Band No.</b>	<b>Candidate Father Species</b>	<b>No. Mismatching Loci (out of 281)</b>	<b>Extra-pair?</b>
B009036	B009128	<i>S. iberaensis</i>	B009563	<i>S. iberaensis</i>	7 (2.5%)	Y
B009037	B009121	<i>S. iberaensis</i>	B009058	<i>S. iberaensis</i>	0 (0%)	Y
B009038	B009121	<i>S. iberaensis</i>	B009058	<i>S. iberaensis</i>	2 (0.7%)	Y
B009061	B009046	<i>S. iberaensis</i>	Unassigned	Unassigned	23 (8.2%)	Y
B009062	B009046	<i>S. iberaensis</i>	B009046	<i>S. iberaensis</i>	5 (1.8%)	N
B009063	B009030	<i>S. hypoxantha</i>	B009030	<i>S. hypoxantha</i>	8 (2.8%)	N
B009064	B009030	<i>S. hypoxantha</i>	B009030	<i>S. hypoxantha</i>	2 (0.7%)	N
B009079	B009067	<i>S. hypoxantha</i>	B009030	<i>S. hypoxantha</i>	5 (1.8%)	Y
B009080	B009067	<i>S. hypoxantha</i>	B009030	<i>S. hypoxantha</i>	3 (1.1%)	Y
B009106	B009141	<i>S. iberaensis</i>	B009141	<i>S. iberaensis</i>	2 (0.7%)	N
B009108	B009117	<i>S. hypoxantha</i>	B009116	<i>S. hypoxantha</i>	7 (2.5%)	Y
B009109	B009117	<i>S. hypoxantha</i>	B009116	<i>S. hypoxantha</i>	4 (1.4%)	Y
B009110	B009117	<i>S. hypoxantha</i>	B009117	<i>S. hypoxantha</i>	2 (0.7%)	N
B009111	B009116	<i>S. hypoxantha</i>	B009116	<i>S. hypoxantha</i>	3 (1.1%)	N
B009112	B009116	<i>S. hypoxantha</i>	B009116	<i>S. hypoxantha</i>	5 (1.8%)	N
B009123	B009115	<i>S. iberaensis</i>	B009115	<i>S. iberaensis</i>	3 (1.1%)	N
B009124	B009115	<i>S. iberaensis</i>	B009115	<i>S. iberaensis</i>	3 (1.1%)	N
B009130	B009113	<i>S. iberaensis</i>	B009113	<i>S. iberaensis</i>	2 (0.7%)	N
B009131	B009113	<i>S. iberaensis</i>	B009528	<i>S. iberaensis</i>	1 (0.4%)	Y
B009510	B009508	<i>S. iberaensis</i>	B009528	<i>S. iberaensis</i>	6 (2.1%)	Y
B009511	B009508	<i>S. iberaensis</i>	B009528	<i>S. iberaensis</i>	6 (2.1%)	Y
B009512	B009504	<i>S. iberaensis</i>	Unassigned	Unassigned	22 (7.8%)	Y
B009513	B009504	<i>S. iberaensis</i>	B009504	<i>S. iberaensis</i>	0 (0%)	N
B009514	B009522	<i>S. hypoxantha</i>	B009522	<i>S. hypoxantha</i>	5 (1.8%)	N
B009515	B009522	<i>S. hypoxantha</i>	B009522	<i>S. hypoxantha</i>	3 (1.1%)	N
B009516	Unbanded	<i>S. hypoxantha</i>	Unassigned	Unassigned	23 (8.2%)	Unknown
B009517	Unbanded	<i>S. hypoxantha</i>	Unassigned	Unassigned	24 (8.5%)	Unknown
B009518	Unbanded	<i>S. hypoxantha</i>	Unassigned	Unassigned	22 (7.8%)	Unknown
B009519	Unbanded	<i>S. hypoxantha</i>	Unassigned	Unassigned	20 (7.1%)	Unknown
B009531	Unbanded	<i>S. hypoxantha</i>	Unassigned	Unassigned	20 (7.1%)	Unknown
B009532	Unbanded	<i>S. hypoxantha</i>	Unassigned	Unassigned	22 (7.8%)	Unknown
B009535	B009528	<i>S. iberaensis</i>	B009528	<i>S. iberaensis</i>	2 (0.7%)	N
B009536	B009528	<i>S. iberaensis</i>	B009508	<i>S. iberaensis</i>	2 (0.7%)	Y
B009537	B009524	<i>S. hypoxantha</i>	B009524	<i>S. hypoxantha</i>	5 (1.8%)	N

B009538	B009524	<i>S. hypoxantha</i>	B009524	<i>S. hypoxantha</i>	0 (0%)	N
B009544	B009539	<i>S. iberaensis</i>	B009583	<i>S. iberaensis</i>	2 (0.7%)	Y
B009545	B009539	<i>S. iberaensis</i>	Unassigned	Unassigned	18 (6.4%)	Y
B009546	B009540	<i>S. iberaensis</i>	B009539	<i>S. iberaensis</i>	5 (1.8%)	Y
B009547	B009540	<i>S. iberaensis</i>	B009539	<i>S. iberaensis</i>	4 (1.4%)	Y
B009548	B009550	<i>S. iberaensis</i>	B009550	<i>S. iberaensis</i>	2 (0.7%)	N
B009549	B009550	<i>S. iberaensis</i>	B009550	<i>S. iberaensis</i>	4 (1.4%)	N
B009551	Unbanded	<i>S. iberaensis</i>	Unassigned	Unassigned	17 (6.0%)	Unknown
B009567	B009556	<i>S. hypoxantha</i>	Unassigned	Unassigned	16 (5.7%)	Y
B009569	B009556	<i>S. hypoxantha</i>	B009556	<i>S. hypoxantha</i>	5 (1.8%)	N
B009571	B009526	<i>S. hypoxantha</i>	B009553	<i>S. hypoxantha</i>	7 (2.5%)	Y
B009572	B009526	<i>S. hypoxantha</i>	B009553	<i>S. hypoxantha</i>	6 (2.1%)	Y
B009575	B009522	<i>S. hypoxantha</i>	Unassigned	Unassigned	19 (6.8%)	Y
B009576	B009522	<i>S. hypoxantha</i>	B009522	<i>S. hypoxantha</i>	3 (1.1%)	N
B009578	B009530	<i>S. hypoxantha</i>	B009530	<i>S. hypoxantha</i>	5 (1.8%)	N
B009579	Unbanded	<i>S. hypoxantha</i>	B009541	<i>S. hypoxantha</i>	3 (1.1%)	Unknown
B009584	B009554	<i>S. hypoxantha</i>	B009554	<i>S. hypoxantha</i>	2 (0.7%)	N
B009585	B009554	<i>S. hypoxantha</i>	B009522	<i>S. hypoxantha</i>	3 (1.1%)	Y
B009588	B009562	<i>S. iberaensis</i>	Unassigned	Unassigned	26 (9.3%)	Y
B009589	B009562	<i>S. iberaensis</i>	Unassigned	Unassigned	21 (7.5%)	Y
B009593	B009552	<i>S. iberaensis</i>	Unassigned	Unassigned	20 (7.1%)	Y
B009594	B009552	<i>S. iberaensis</i>	Unassigned	Unassigned	17 (6.0%)	Y
B009596	B009583	<i>S. iberaensis</i>	B009583	<i>S. iberaensis</i>	0 (0%)	N
B009597	B009583	<i>S. iberaensis</i>	B009583	<i>S. iberaensis</i>	0 (0%)	N
B009603	B009042	<i>S. iberaensis</i>	B009042	<i>S. iberaensis</i>	2 (0.7%)	N
B009604	B009042	<i>S. iberaensis</i>	Unassigned	Unassigned	27 (9.6%)	Y
B009605	B009042	<i>S. iberaensis</i>	Unassigned	Unassigned	24 (8.5%)	Y
B009607	B009055	<i>S. hypoxantha</i>	Unassigned	Unassigned	16 (5.7%)	Y
B009608	B009055	<i>S. hypoxantha</i>	B009055	<i>S. hypoxantha</i>	4 (1.4%)	N
B009609	B009601	<i>S. hypoxantha</i>	Unassigned	Unassigned	20 (7.1%)	Y
B009610	B009601	<i>S. hypoxantha</i>	Unassigned	Unassigned	20 (7.1%)	Y
B009613	B009612	<i>S. hypoxantha</i>	B009612	<i>S. hypoxantha</i>	4 (1.4%)	N
B009614	B009612	<i>S. hypoxantha</i>	B009612	<i>S. hypoxantha</i>	5 (1.8%)	N
B009622	B009060	<i>S. hypoxantha</i>	B009060	<i>S. hypoxantha</i>	2 (0.7%)	N
B009623	B009060	<i>S. hypoxantha</i>	Unassigned	Unassigned	23 (8.2%)	Y
B009625	B009118	<i>S. hypoxantha</i>	B009118	<i>S. hypoxantha</i>	3 (1.1%)	N
B009626	B009118	<i>S. hypoxantha</i>	Unassigned	Unassigned	27 (9.6%)	Y
B009627	B009620	<i>S. iberaensis</i>	B009620	<i>S. iberaensis</i>	5 (1.8%)	N
B009628	B009620	<i>S. iberaensis</i>	B009508	<i>S. iberaensis</i>	3 (1.1%)	Y

B009634	Unbanded	<i>S. iberaensis</i>	Unassigned	Unassigned	19 (6.8%)	Unknown
B009635	Unbanded	<i>S. iberaensis</i>	Unassigned	Unassigned	21 (7.5%)	Unknown
Nestling_1	B009119	<i>S. hypoxantha</i>	Unassigned	Unassigned	20 (7.1%)	Y
Nestling_2	B009128	<i>S. iberaensis</i>	B009128	<i>S. iberaensis</i>	2 (0.7%)	N
UH_Egg_1	Unbanded	<i>S. hypoxantha</i>	Unassigned	Unassigned	0 (0%)	Unknown
UH_Egg_2	Unbanded	<i>S. hypoxantha</i>	Unassigned	Unassigned	24 (8.5%)	Unknown
UH_Egg_3	Unbanded	<i>S. iberaensis</i>	Unassigned	Unassigned	19 (6.8%)	Unknown

**Table S5. Post-hoc pairwise comparisons (Tukey HSD method) for generalized linear mixed models examining the effect of treatment group on behavioral response intensity**

**(PC1).** Territorial males of *S. hypoxantha* (top) and *S. iberaensis* (bottom) were presented with combinations of conspecific (CON), heterospecific (HET), and *S. collaris* (CONTROL) song and plumage. Significant results (adjusted  $P < 0.05$ ) are highlighted in bold. Treatment group had a significant effect on response intensity in both *S. hypoxantha* ( $P < 0.0001$ ,  $N = 120$ ) and *S. iberaensis* ( $P < 0.0001$ ,  $N = 120$ ) males. Asterisks indicate treatments groups that were not significantly different when outliers (observations outside 1.5 \* interquartile range; *S. hypoxantha*: 11 outliers, *S. iberaensis*: 11 outliers) were removed.\*

<i>S. hypoxantha</i>				
Treatment contrast <sup>1</sup>	Estimate	S.E.	df	P value
CONTROL – CON Plumage/CON Song	-3.10	0.28	90.7	<b>&lt;0.0001</b>
CONTROL – CON Plumage/HET Song	-1.80	0.27	85.6	<b>&lt;0.0001</b>
CONTROL – HET Plumage/CON Song	-2.17	0.28	90.7	<b>&lt;0.0001</b>
CONTROL – HET Plumage/HET Song	-0.60	0.28	89.0	0.2078
CON Plumage/CON Song – CON Plumage/HET Song	1.30	0.27	85.0	<b>0.0001</b>
CON Plumage/CON Song – HET Plumage/CON Song*	0.93	0.27	87.9	<b>0.0082</b>
CON Plumage/CON Song – HET Plumage/HET Song	2.51	0.28	89.9	<b>&lt;0.0001</b>
CON Plumage/HET Song – HET Plumage/CON Song	-0.37	0.27	85.0	0.6544
CON Plumage/HET Song – HET Plumage/HET Song	1.21	0.27	83.6	<b>0.0002</b>
HET Plumage/CON Song – HET Plumage/HET Song	1.57	0.28	89.9	<b>&lt;0.0001</b>
<i>S. iberaensis</i>				
Treatment contrast <sup>2</sup>	Estimate	S.E.	df	P value
CONTROL – HET Plumage/HET Song	-0.84	0.26	86.6	<b>0.0136</b>
CONTROL – HET Plumage/CON Song	-2.57	0.26	87.8	<b>&lt;0.0001</b>
CONTROL – CON Plumage/HET Song	-1.99	0.26	91.4	<b>&lt;0.0001</b>
CONTROL – CON Plumage/CON Song	-3.47	0.27	92.5	<b>&lt;0.0001</b>
CON Plumage/CON Song – CON Plumage/HET Song	-1.48	0.26	90.4	<b>&lt;0.0001</b>
CON Plumage/CON Song – HET Plumage/CON Song*	-0.90	0.26	86.8	<b>0.0068</b>
CON Plumage/CON Song – HET Plumage/HET Song	-2.62	0.26	90.4	<b>&lt;0.0001</b>
CON Plumage/HET Song – HET Plumage/CON Song	0.58	0.26	87.9	0.1745
CON Plumage/HET Song – HET Plumage/HET Song	-1.15	0.26	89.1	<b>0.0003</b>
HET Plumage/CON Song – HET Plumage/HET Song	-1.73	0.26	86.5	<b>&lt;0.0001</b>

<sup>1</sup>Model included male ID (SD = 0.90; 95% CI of SD = 0.64-1.21) as a random effect.

<sup>2</sup>Model included male ID (SD = 0.72; 95% CI of SD = 0.50-1.02) and female presence (SD = 0.46; 95% CI of SD = 0.09-1.49) as random effects.

\*Although we did not detect a significant difference between the CON Plumage/CON Song treatment and the HET Plumage/CON Song treatment in both species, suggesting that song may play a more important role than plumage in territorial interactions, this result was only obtained after excluding outliers and further research is necessary to evaluate the relative importance of these two divergent traits.