Check for updates

Accepted Article

Genetic structure in Louisiana Iris species reveals patterns of recent and historical admixture

Alexander S. Zalmat[1], V. Alex Sotola[2], Chris C. Nice[1], and Noland H. Martin[1,3,*]

[1]*Department of Biology, Texas State University, San Marcos, Texas 78666-0202, USA.*

[2]*Department of Genetics, University of Georgia, Athens, Georgia 30602-7223, USA.*

[3]*Chair of the Population and Conservation Biology Program, Department of Biology, Texas State University, San Marcos, Texas 78666-0202, USA.*

Author for correspondence: nm14@txstate.edu*

Manuscript received:_________; revision accepted:________.

Genetic structure of Louisiana Iris species

Abstract

Premise: When divergent lineages come into secondary contact reproductive isolation may be incomplete, thus providing an opportunity to investigate how speciation is manifested in the genome. The Louisiana Irises (*Iris*, series *Hexagonae*) comprise a group of three or more ecologically and reproductively divergent lineages that can produce hybrids where they come into contact. In this study we sought to estimate standing genetic variation to understand the current distribution of population structure in the Louisiana Irises.

Methods: We used genotyping-by-sequencing techniques to sample the genomes of Louisiana Iris species across their ranges. Twenty populations were sampled (total n=632) across 11,249 loci. Population genetic data were assessed using ENTROPY and PCA models.

Results: We discovered evidence for interspecific gene flow in parts of the range and

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/ajb2.1758.

This article is protected by copyright. All rights reserved.

Accepted Article

revealed patterns of population structure at odds with widely accepted nominal taxonomy. Undescribed hybrid populations were discovered that were designated as belonging to the *I. brevicaulis* lineage. *Iris nelsonii* shared significant ancestry with only one of the purported parent species, *I. fulva,* evidence inconsistent with a hybrid origin.

Conclusions: This study provides several key findings important to the investigation of standing genetic variation in the Louisiana Iris species complex. *Iris brevicaulis* has a large amount of genetic diversity within it relative to the other nominal species. In addition, this study has discovered a previously unknown hybrid zone between *I. brevicaulis* and *I. hexagona* along the Texas coast. Finally, *I. nelsonii* does not appear to have mixed ancestry from three parental taxa as has been the longstanding hypothesis.

Key words: speciation, homoploid hybrid species, Single Nucleotide Polymorphisms (SNPs), hybrid zone, gene flow, plant evolution, *Iris hexagona, Iris fulva, Iris brevicaulis, Iris nelsonii*, species complex, Iridaceae

INTRODUCTION

Speciation has traditionally been thought of as occurring through the gradual accumulation of prezygotic or postzygotic isolating barriers that arise among genetically diverging lineages (Dobzhasnky, 1937; Mayr, 1942; Coyne and Orr, 2004). Because speciation is rarely instantaneous (but see Grant, 1981; Levin, 1983, 2002; Wood et al., 2009), early stages of population divergence often occur in the face of considerable gene flow via natural hybridization (Arnold et al., 2015). Natural hybridization and gene flow can, in some cases, certainly lead to reductions in diversity, potentially resulting in the fusion of lineages or genetic swamping of one taxon over the other (Allendorf et al., 2001; Edmands, 2007). Despite this, there is a general consensus that natural

This article is protected by copyright. All rights reserved.

Accepted Article

hybridization and gene flow can also act as a creative force resulting in increased genetic diversity within introgressed populations, adaptive introgression of alleles, and even the evolution of new reproductively isolated hybrid lineages (Arnold, 1997; Arnold et al., 2008; Anderson and Stebbins, 1954; Gompert and Buerkle, 2016; Martin et al., 2005, 2006; Rieseberg et al., 2003; Soltis and Soltis, 2009; Sotola et al., 2019; Taylor et al., 2009).

The growing body of evidence that divergence can occur in the face of gene flow (Gavrilets, 1999; Morales et al., 2017; Nosil, 2008; Pinho and Hey, 2010; Taylor et al., 2013; Wang et al., 2018) is seemingly difficult to reconcile under a strict interpretation of the biological species concept (Mayr, 1942), where a complete cessation of gene flow across the entirety of the genome is required in order for taxa to be appropriately considered as distinct species. However, Wu's (2001) genic view of speciation posits that entire genomes need not be protected from gene flow in order for divergence to take place. Rather, a small number of regions distributed throughout the genome, each of which is selected against in heterospecific genomic backgrounds, is potentially sufficient to allow for continued divergence, even if introgression occurs in other regions of the genome (Wu, 2001). Furthermore, allopatry, particularly when accompanied by divergent selection leading to local adaptation, is well understood to increase rates of divergence among populations (Coyne and Orr, 2004). However, complete allopatry is likely ephemeral throughout the process of divergence, and secondary contact among diverging taxa is likely commonplace (Martin et al., 2007, 2008; Sung et al., 2018; Sotola et al., 2019). By examining divergent and diverging lineages that have undergone secondary contact at varying timescales and geographic scales, the degree to which divergent

This article is protected by copyright. All rights reserved.

Accepted Article

selection and local adaptation directly result in reproductive isolation between these taxa can be explored (Hamilton et al., 2013; Melo et al., 2014; Holliday et al., 2016; Zhao et al., 2020). Thus, identifying the degree to which natural hybridization and gene flow occur at various temporal, genomic, and geographic scales is important for a more complete understanding of not only the long and short-term effects of gene exchange, but also the very evolutionary and ecological processes by which new species arise and persist in nature.

The Louisiana Iris species complex (*Iris* series *Hexagonae*) offers an excellent opportunity to investigate mechanisms which influence reproductive isolation leading to increased genetic divergence and subsequent speciation. This group provides a system in which to study the evolutionary importance of hybridization and introgression as all species are interfertile yet appear to maintain their evolutionary independence in the face of long-documented natural hybridization (Carney and Arnold, 1997; Taylor et al., 2013; Wesseling and Arnold, 2000). The taxonomic diversity within the Louisiana Iris complex was initially thought to be quite high with the earliest descriptions by Small and Alexander (1931) originally positing that the system consisted of over 80 distinct species occurring in southern Louisiana, USA. Later investigations revealed that this diversity of form was largely the result of naturally occurring hybrids that were formed primarily between three morphologically, reproductively, and ecologically distinct (though in much of their ranges geographically overlapping - Fig. 1) species; *Iris hexagona, I. fulva,* and *I. brevicaulis* (Viosca, 1935; Forster, 1937; Riley, 1938; Arnold et al., 1991; Hamlin et al., 2017). Hamlin et al. (2017) performed phylogenetic work on broadly sampled *I. fulva, I. brevicaulis*, and *I. hexagona* and proposed that *I. brevicaulis* and *I. fulva* were sister taxa.

This article is protected by copyright. All rights reserved.

Accepted Article

Some researchers have more recently suggested that *I. hexagona* may be more appropriately divided into additional geographically distinct species: *I. giganticaerulea, I. hexagona,* and *I. savannarum* (Meerow et al., 2011). An additional described species, *I. nelsonii*, occupies a restricted range within a single southern Louisiana swamp, occurring sympatrically with the three more widespread species. Randolph (1966) described *I. nelsonii* as being of likely homoploid hybrid origin between *I. fulva, I. hexagona,* and possibly *I. brevicaulis* on the basis of cytological and morphological evidence (Randolph et al., 1961; Randolph, 1966), and *I. nelsonii* has been widely referred to as a 'classic example' of a homoploid hybrid species, with the majority of the genome consisting of *I. fulva* and small introgressed regions coming from *I. brevicaulis* and *I. hexagona* (Coyne and Orr, 2004). Arnold et al. (1990), using a small number of allozyme markers, found that a minority of loci were shared by both *I. brevicaulis* and *I. nelsonii* individuals, though no such evidence of shared loci between *I. nelsonii* and *I. hexagona* was detected. Arnold (1993), using four RAPD markers that were diagnostic for *I. fulva, I. hexagona*, and *I. brevicaulis*, found that *I. nelsonii* shared both *I. hexagona* and *I. brevicaulis* alleles (in addition to mostly *I. fulva* alleles), leading him to conclude that a majority of the genome of *I. nelsonii* was largely derived from *I. fulva* with smaller contributions from *I. hexagona* and *I. brevicaulis* (Arnold et al., 1990; Arnold, 1993). Of note, none of the hypothesized introgressed alleles from *I. hexagona* and *I. brevicaulis* were revealed to be fixed in *I. nelsonii* in either of these molecular studies (Arnold et al., 1990; Arnold, 1993), and thus contemporary hybridization and introgression – as opposed to ancient homoploid hybridization - could not be ruled out as an explanation for these patterns.

In the current study we refer to four nominal taxa: *Iris brevicaulis, I. fulva, I.*

This article is protected by copyright. All rights reserved.

Accepted Article

hexagona, and *I. nelsonii* broadly distributed in Eastern North America (Figure 1), recognizing the taxonomic uncertainty described within "*I. hexagona*." Of the three widespread species, *I. brevicaulis* and *I. fulva* are more ecologically similar with respect to abiotic associations; both species largely occupy moist understory habitats throughout their broad ranges within the Mississippi River drainage (Viosca 1935). However, the flowers of both species are quite different, with *I. fulva* flowers being copper red with a hummingbird pollination syndrome, and *I. brevicaulis* flowers being bright blue with a bumblebee pollination syndrome (Viosca 1935; Wesselingh and Arnold, 2000; Bouck et al., 2007; Martin et al., 2008). *Iris hexagona* is a common species distributed along open coastal habitats (Figure 1). This species is much larger than either *I. brevicaulis* or *I. fulva*, and it shares a bee pollination syndrome with *I. brevicaulis* (Taylor et al. 2013). The reproductive phenology of *I. brevicaulis* occurs, on average, much later than any of the other species, and is likely one of the most important barriers to hybridization in sympatry (Cruzan and Arnold, 1994; Martin et al., 2007). *Iris nelsonii* is both genetically and morphologically most similar to *I. fulva* and also has morphological characteristics indicative of hummingbird pollination (Randolph, 1966; Arnold et al., 1990; Arnold 1993; Taylor et al., 2013). A number of reproductive isolating barriers including habitat isolation, phenological isolation, pollinator isolation (both behavioral and mechanical), interspecific pollen competition, and postzygotic isolation (in the form of both hybrid inviability and hybrid sterility) reduce the rate at which hybrids form in sympatry (reviewed by Arnold et al. 2010). Despite these identified barriers, current – and extensive - hybridization is widely documented between most of the taxa, especially in southern Louisiana (though no such hybridization between *I. brevicaulis* and *I. hexagona*

This article is protected by copyright. All rights reserved.

Accepted Article

has yet been identified (Carney and Arnold, 1997; Taylor et al., 2013; Wesseling and Arnold, 2000)).The extent to which historical and current gene flow has played a role in the broad genetic structuring of this Louisiana Iris species complex throughout the ranges of these taxa is still unknown.

In this study, we used a genotype by sequencing approach along with extensive geographic sampling of the four widely recognized Louisiana Iris species (*I. fulva, I. hexagona, I. brevicaulis*, and *I. nelsonii*) to address the following three objectives. The first objective was to assess the distribution of genomic variation within and between the nominal taxonomy of Louisiana Iris. Second, we set out to resolve the taxonomic confusion associated with *I. hexagona*, specifically targeting populations across the range of this coastal species. And third, we sought to readdress the purported hybrid origin of *I. nelsonii* with a robust genome-wide dataset two orders of magnitude larger than any other molecular study.

MATERIALS AND METHODS

Sample collection- Tissue samples were collected throughout the year of 2015 from individuals spanning locations across all four known ranges of the Louisiana Iris species (Fig. 1). Sample locations in southern Louisiana were previously identified by two of the authors (A. Zalmat and N. Martin) through extensive field work in the area over the past two decades. Collection locales outside of central Louisiana were identified in various herbarium accessions and discussions with *Iris* enthusiasts, though most often habitat modification required searching areas beyond where the original accessions and/or sightings were originally observed. Five to ten locations per species were sampled, except in the case of *I. nelsonii*, for which the only known location was sampled (Table

This article is protected by copyright. All rights reserved.

Accepted Article

1). At each location, some individual plants were flowering during the time of collection and these individuals could be readily identified to species. Louisiana Iris species can reproduce clonally, and these clones can readily be identified as they often grow in a circular pattern. Care was taken to ensure that the same genotype was not sampled twice, and no clonal replicates were detected after performing genetic analyses. In some cases, hybrid individuals were found and noted during the time of collection (Table 1). Collection sites containing flowering *I. hexagona* individuals included two sites in the Florida peninsula and several coastal sites in Louisiana and Texas. Collections of *I. brevicaulis* included two sites in the northern reaches of the range including a site in Illinois-labeled here as “bfi”- previously sampled by Hamlin and colleagues (Hamlin and Arnold, 2014), and several sites in coastal Texas and Louisiana. *Iris fulva* collections were made at five sites spanning a large portion of the range, including a site in Mississippi, labeled as “fcm” in Table 1, also sampled previously by Hamlin and colleagues (Hamlin and Arnold, 2014). *Iris nelsonii* individuals were sampled from their only known location in Abbeville, Louisiana, USA. Because hybridization in southern Louisiana has been well documented by previous researchers (and some of the co-authors here), such hybrid zones were not targeted in this study. However, a coastal marsh in Brazoria County, Texas was extensively sampled where there appeared to be an extensive contact zone between *I. brevicaulis* and *I. hexagona* with a number of phenotypically intermediate hybrid individuals, as no such *I. brevicaulis* X *I. hexagona* purported hybrid zone has been documented. Leaf tissue from 6-142 individuals per sampling locale was collected, placed in a coin envelope, and then dried on silica prior to DNA extraction.

DNA sequence generation, assembly, and variation-DNA was extracted from

This article is protected by copyright. All rights reserved.

Accepted Article

tissue samples using a standard CTAB extraction protocol in a 96-well plate format (Doyle 1991). A reduced representation library was generated for each individual following the methods of Parchman et al. (2012) and Gompert et al. (2014). DNA from each individual was digested with restriction enzymes EcoRI and MseI. Adaptors that included an 8 to 10 base pair oligonucleotide barcode for individual identification was then ligated to the generated fragments. These restriction ligation products were then amplified for two rounds of PCR using standard Illumina (San Diego, California, USA) primers. After PCR, the products were pooled across individuals and size selected for 300-400 bp length fragments using BluePippin technology. The final DNA library was then sequenced twice across two lanes at the University of Texas Genomic Sequencing and Analysis Facility (Austin, Texas, USA) on an Illumina HiSeq. 4000 platform. The resulting single end, 100bp sequence reads were then filtered to remove reads that belong to genomes used in the Illumina control libraries. The reads were then processed using custom scripts (available from the corresponding author) to remove barcodes and adaptor sequences in order to obtain final reads ranging from 84-86 bp in length. No reference genome is available for Iris species, so a *de novo* assembly the following the assembly strategy from dDocent was performed (Puritz et al., 2014). If reads were shared by fewer than four individuals in the dataset or were represented fewer than four times, they were removed. The resulting set of reads was then assembled using an 80% sequence similarity threshold using CD-hit software (Fu et al., 2012). From this, reference scaffolds were generated, and all sequence data were then assembled to the reference using Burrows Wheeler Aligner, BWA version 0.7.5a-r405 (Li and Durbin, 2009). Single nucleotide polymorphisms (SNPs) were then identified using SAMtools (ver. 0.1.19) and BCFtools

This article is protected by copyright. All rights reserved.

Accepted Article

(ver. 0.1.19.). In order for SNPs to be identified they had to be present in at least 50% of individuals in the data set and had to have a minimum of one read at that site. Genotype likelihood estimates were generated for the resulting 224,414 loci spanning 632 individuals. Allele frequency estimates were generated from the genotype likelihood estimates and loci with a global minor allele frequency of less than 0.01 were excluded. To reduce the effects of linkage disequilibrium amongst SNPs, one variable site was chosen per reference scaffold. The resulting dataset consisted of 11,249 loci, which were used in the population structure analysis.

Population structure and gene flow- To quantify the geographic distribution of genomic variation within and among the currently recognized species of Louisiana Iris, population genetic parameters were estimated using Entropy (Gompert et al., 2014), a hierarchical Bayesian model similar to the correlated allele frequencies admixture model in Structure (Pritchard et al., 2000). Entropy makes population genetic parameter estimates with no *a-priori* knowledge of sample localities, and accounts for variation in sequence depth and genotyping errors in a Bayesian framework. The user designates the assumed number of clusters (k), and the Entropy algorithm estimates parameters based on the designated number of clusters and the posterior probability of the allele frequencies for each k. We compared models assuming two to ten clusters (k = 2-10). For each model we iterated 100,000 MCMC steps, sampling for each parameter every 10 steps (thinning) and dropping the first 5,000 steps (burn-in). We ran two chains for each model of k. We did not assume that there was a "best" k solution, but rather we compared models across k to gain insight into different levels of population structure (Gilbert et al., 2012; Lawson et al., 2012; Meirmans, 2015; Janes et al., 2017; Driscoe et al., 2019; Sotola et al., 2019)

This article is protected by copyright. All rights reserved.

Accepted Article

as each level of k can provide biologically pertinent information about genomic structuring at different scales. The Gelman-Rubin diagnostic statistic was used to check chain convergence for each level of k. Runs above k=7 showed poor mixing and a lack of chain convergence via the Gelman-Rubin statistic; they were therefore not used for interpretation, which are provided in the supplemental information. To visualize and summarize the distribution of genomic variation, we used a Principal Components Analysis using the *prcomp* function in R. This PCA is performed on the matrix of genotype probabilities from Entropy with individuals being the rows and loci being columns in this matrix.

Genetic divergence was explored using Nei's G_{ST} (Nei, 1987), which was calculated using genotype probabilities estimated from Entropy. Genotype probabilities were averaged across all "biologically interpretable" runs of k (i.e., k=2 through k=7) and across all MCMC chains, which were then used to calculate allele frequencies in R (R Core Team, 2019). These allele frequencies were in turn used to calculate pairwise G_{ST} values. Pairwise comparisons included between each sampled population in order to show all patterns: between and within putative species, providing fine-scaled resolution to determine the partitioning of genetic divergence. In order to get representation of *I. hexagona* from the Matagorda Bay, TX area we removed individuals from the sampled contact zone that assigned at 98% or higher to the cluster representing *I. hexagona* in the k = 2 ENTROPY model. In our G_{ST} plot these individuals have been given the three-letter designation "hbt".

RESULTS

Sequence depth, assembly, and sampling- A total of 572,371,746 sequence reads

This article is protected by copyright. All rights reserved.

Accepted Article

was obtained and the *de novo* assembly produced 49,785 scaffolds onto which the rest of the reads were assembled. Across these scaffolds, 218,743 variable sites were discovered in the variant calling process. Once loci with a minor allele frequency of less than 0.01 were removed and one variable site per scaffold was chosen, a total of 11,249 SNPs were included in the dataset. The final dataset included 632 individuals from 20 locations and included representatives from four Louisiana Iris species ranges (Fig. 1; Table 1). Sample sizes of each locality ranged from 6 to 139 (Table 1) and the mean individual sequence depth (the average number of reads per locus per individual) was found to be 15.66 (SD = 2.09).

Population structure- Admixture proportions were calculated in ENTROPY for k=2 through k=10; however, only up to k=7 is presented. At and after k=8, there were no new biologically interpretable clusters and mixing was poor, thus we conclude that models above k=7 represent poor fits for the data. For k = 2, the model separated individuals designated as *I. hexagona* from all other sample localities (Fig. **2a**). The model at k = 2 also shows some admixture between *I. hexagona* individuals and individuals from the remaining cluster in the sympatric contact zone indicating ongoing gene flow (Fig. **2a**). For k = 3, the model adds a group that includes individuals designated as *I. fulva* and *I. nelsonii*, and there is admixture observed within the northern and Louisiana localities of *I. brevicaulis* from all three genetic clusters (Figure **2b**). All *I. nelsonii* and *I. fulva* individuals belong to the same cluster, indicating that there is more genetic variation within the *I. brevicaulis* lineage than there is between *I. nelsonii* and *I. fulva*. At k = 4 the model still does not differentiate between individuals designated as *I. nelsonii* and *I. fulva*, though two individuals pre-designated in the field as *I. fulva* and one

This article is protected by copyright. All rights reserved.

Accepted Article

pre-designated as *I. nelsonii* have mixed ancestry from *I. brevicaulis* in southern Louisiana. This is unsurprising as hybridization is common in the area (Carney and Arnold, 1997; Taylor et al., 2013; Wesseling and Arnold, 2000), and given the broad sympatry of the species in the area and nature of the sampling (many individuals were collected that were not flowering), it would be unusual to detect no hybrids in this region. Given that flowering phenology varies between these species (Taylor et al., 2009), and that collections of non-flowering individuals were made in this sympatric part of the range, it is probable that these individuals represent misidentifications. A new cluster is also formed at k = 4 containing the individuals designated as *I. brevicaulis* from the two localities in Louisiana (Fig. **2c**). The two northern *I. brevicaulis* sites remain admixed with ancestry from the three other clusters in the model. At k = 5, the admixed northern *I. brevicaulis* locality mentioned above in the k=4 model is now revealed as its own distinct cluster. *I. fulva* and *I. nelsonii* are still indistinguishable from each other (Fig. **2d**). At k=6, the observed variation within the *I. fulva* group is similar to that found in the *I. nelsonii* samples, although there are no pure individuals within either group (Fig. **2e**). At k= 7, the Florida population of *I. hexagona* splits out as its own genetic group (Fig. **2f**). At k=8 and above the model becomes a poor fit to the data, adding additional clusters that do not show high levels of assignment in any individuals in the dataset while still maintaining the major clusters resolved under the k = 5 model (see Appendices S2 and S3 in supplementary data with this article).

The first three principal components in the PCA ordination explained 57.95% of the genetic variation in this Iris group. The first PC explained 34.55%, PC2 explained 20.4%, and PC3 explained only 3% of the variation (Fig. 3, 4). Individuals of *I.*

This article is protected by copyright. All rights reserved.

Accepted Article

brevicaulis, *I. fulva*, *I. hexagona*, and *I. nelsonii* clustered into five clearly distinct groups along the three PC axes. Individuals designated as *I. brevicaulis* formed three clearly distinct clusters: the two northernmost *I. brevicaulis* localities formed one cluster, the two Louisiana *I. brevicaulis* localities forming a second, and the Texas *I. brevicaulis* localities forming the third. Individuals designated as *I. fulva* and *I. nelsonii* together formed a single cluster that was not resolved across any PC axis. Individuals pre-designated as *I. hexagona* formed a single cluster far removed from other clusters along the first PC axis. Individuals sampled from the Texas *I. hexagona/I. brevicaulis* contact zone mostly separated one of two clusters with some intermediate individuals spanning the first PC axis in between pure *I. brevicaulis* and *I. hexagona*, indicating hybridization and gene flow at the Texas locality. A very small number of individuals that were clearly misidentified in the field can be seen clustering with unrelated taxa (a number of individuals that were not flowering were collected for this study in clearly sympatric populations, and it is difficult to differentiate among the taxa using only leaf morphology).

Pairwise G_{ST} measurements within and across previously described taxonomic divisions (i.e., nominal “species”) revealed unexpected patterns of both within- and between-species genomic differentiation (Table 2). Measures of G_{ST} within widely collected samples of purported *I. fulva* (from IL/KY to MS/LA) were quite small (0.0097 +/- 0.0049, mean +/- 1 SD) and similar in magnitude to disparately sampled *I. hexagona* collections (FL to TX, 0.0141 +/- 0.0052). In contrast, the mean and standard deviation of intraspecific *Iris brevicaulis* measures of G_{ST} (0.0705 +/- 0.0525) were, however, an order of magnitude higher than those of intraspecific *I. fulva* and *I. hexagona* measures.

This article is protected by copyright. All rights reserved.

This high mean-G_{ST} among *I. brevicaulis* localities is reflected by the fact that three highly genetically differentiated lineages were uncovered that are currently encompassed by the *I. brevicaulis* taxon (Fig. **2c,d**). Measures of G_{ST} within Texas *I. brevicaulis* are small (0.0193 +/- 0.0076), as is G_{ST} measured between the two Louisiana populations of *I. brevicaulis* (0.0053). Recall that only a single Illinois population of *I. brevicaulis* was collected for this study, however, pairwise Texas x Louisiana *I. brevicaulis* G_{ST} measurements (0.1348 +/- 0.0086), Illinois x Louisiana (0.0971 +/- 0.0023), and Texas x Illinois (0.0717 +/- 0.0081) are much greater than the within-Texas and within-Louisiana *I. brevicaulis* G_{ST} measures, reflecting the high levels of intraspecific differentiation that has been uncovered within this taxon.

The large amounts of variation uncovered among geographically separated *I. brevicaulis* collections is especially stark when considering other interspecific pairwise G_{ST} comparisons. For example, Texas x Louisiana *I. brevicaulis* G_{ST} measures equal those of Texas *I. brevicaulis* x *I. hexagona* comparisons (0.1335 +/- 0.0151) and are over ⅓ higher than those of Texas *I. brevicaulis* x *I. fulva* comparisons (0.1004 +/- 0.0090). Interspecific measures of G_{ST} were highest between *I. fulva* and *I. hexagona* (0.1921 +/- 0.0215) and Louisiana *I. brevicaulis* and *I. hexagona* (0.2068 +/- 0.0167).

Iris nelsonii has been described only from a single swamp in southern Louisiana, therefore no inter-population measures of G_{ST} were possible in the single collection included in this study. *Iris nelsonii* measures of interspecific G_{ST} were high between populations of *I. hexagona* (0.1848 +/- 0.0228) and Texas, Louisiana, and Illinois *I. brevicaulis* (0.1600 +/- 0.0093, 0.0944 +/- 0.0071, and 0.1184 respectively - only a single Illinois *I. brevicaulis* x *I. nelsonii* comparison could be calculated, so no standard

Accepted Article

This article is protected by copyright. All rights reserved.

deviation is presented). Measures of G_{ST} between *I. nelsonii* x *I. fulva* were much lower, however (0.0114 +/- 0.0038) than for any other nominal taxon in this study, falling well within the range of intraspecific *I. fulva* measures of G_{ST} (again, 0.0097 +/- 0.0049, mean +/- 1 SD) calling into question the hypothesized evolutionary uniqueness of *I. nelsonii*.

DISCUSSION

This study sought to assess the standing genomic variation within and among the currently recognized taxa of the Louisiana Irises throughout a broad sample of much of the range of the species complex. There was strong evidence supporting the existence of three divergent lineages that agreed with the delineation of the species *I. fulva, I. hexagona,* and *I. brevicaulis*. A fourth named species, *I. nelsonii*, has long been hypothesized to be of homoploid hybrid origin (Arnold, 1993), however results from the current study found no evidence to support this hypothesis. Furthermore, our genome wide analysis did not distinguish *I. nelsonii* from *I. fulva,* not even at fine scales of genetic structure (see Appendix S1in supplementary data). This result is in strong contrast with previous work and previous assumptions that *I. nelsonii* is a separate species – much less a unique homoploid hybrid species (Arnold et al., 1990; Arnold, 1993; Arnold et al., 1991; Randolph, 1966). Additionally, this is the first study to find strong support for two separate lineages with mixed ancestry within what is currently described as *I. brevicaulis*. Populations of *I. brevicaulis* sampled in Texas exhibited no evidence of past hybridization, while those purported to be *I. brevicaulis* collected in Louisiana and Illinois were revealed to have more complicated evolutionary histories and might be two separate stabilized three-way hybrid lineages (involving *I. fulva, I. hexagona*, and the Texas *I. brevicaulis* lineages). However, other population processes

This article is protected by copyright. All rights reserved.

Accepted Article

Accepted Article

could lead to this pattern of apparent admixture in our clustering analysis (Lawson et al. 2018) and the hypothesis of independent hybrid lineages in Louisiana and Illinois will require further analysis and investigation. We also find that within the coastal *I. hexagona* group, we see only fine-scale population structure that corresponds with the range in which *I. savanarrum* is found in Florida at k=7 (Fig. **2f**). Finally, and perhaps unsurprisingly, given the long history of natural hybridization studies in the Louisiana Iris species complex, evidence of current admixture between the three supported nominal taxa was discovered as well. This is the first study, however, to document an extensive hybrid zone between ostensibly "pure" *I. brevicaulis* and *I. hexagona* in coastal Texas. Further studies are clearly warranted to determine the mechanism by which such ongoing hybridization is occurring at this (and potentially other as-of-yet uninvestigated) locales. Overall, the Louisiana Iris system continues to be a compelling case study of the evolutionary importance of natural hybridization - both current and historical.

Contemporary hybridization in Louisiana Iris- The importance of natural hybridization in the evolutionary process has long been documented in a diverse array of living organisms (Abbott et al., 2016; Arnold, 2016; Carney and Arnold, 1997; Seehausen, 2004). Indeed Viosca (1935), Anderson (1949) and Arnold (1997) have each written entire books, decades apart, devoted largely to natural hybridization with Louisiana Iris as the focal taxa exemplifying the process. In the current study we did not specifically target natural hybrid zones, as has been done by this research group in the past (e.g., Sung et al. 2018), but rather we sought to determine broader patterns of genetic variation that exist across the nominal taxa of Louisiana Iris. In many of the extensively sampled collection locales, a number of hybrid individuals were, perhaps not

This article is protected by copyright. All rights reserved.

Accepted Article

surprisingly, identified (Fig. 4). The newly discovered hybrid zone is clearly active with individual admixture proportions in these collection locales ranging from "100% pure" *I. brevicaulis* to "100% pure" *I. hexagona*, and a full range of intermediate genotypes in between (Fig. 2-bar plots showing intermediate assignments in the Texas contact zone). Such extensive hybridization between these two taxa has likely been missed by previous researchers and hybrid fanciers because the two species have not been studied extensively across the Texas coast. The two species involved in this hybrid zone share some similarities in flower morphology in terms of color and pollinator syndrome. However, they differ strikingly in terms of flowering stalk height and shape with *I. hexagona* growing tall straight stalks, and *I. brevicaulis* growing short "zig-zagging" stalks. Indeed, the hybrid individuals discovered in this study were flowering at the time of sampling and a stunning amount of variation in flowering stalk height was observed (A. Zalmat, personal observation). This is in stark contrast to hybrid zones involving the red-flowered *I. fulva* and either of the blue-flowered species *I. brevicaulis* or *I. hexagona*, where hybrid individuals can often be readily identified as having intermediate "purple" colors as well as even non-intermediate transgressive floral morphologies that do not resemble either parent (Ballerini et al., 2012; Bouck et al., 2007; Martin et al., 2008; Sung et al., 2018; Taylor et al., 2013). In the latter examples, it is believed that floral morphologies are one of the primary reproductive isolating barriers that reduce interspecific hybridization and prevent the localized amalgamation of the two species, and in fact a number of studies have been performed examining the genetic architecture of these floral differences as they relate to pollinator behavior (Wesseling and Arnold, 2000; Martin et al., 2008; Shaw et al., 2017). This hybrid zone provides interesting research opportunities to examine the

This article is protected by copyright. All rights reserved.

Accepted Article

maintenance of these two taxa where floral divergence is likely not the primary driver of reproductive isolation.

Historical admixture in the Louisiana Irises- The Louisiana Iris species, *I. nelsonii*, has been hailed as a classic example of homoploid hybrid speciation, even by authors who have long doubted the overall importance of the process in generating biodiversity (e.g., Coyne and Orr, 2004). *Iris nelsonii* was first hypothesized to be of hybrid origin based on chromosomal and morphological evidence (Randolph et al., 1961; Randolph, 1966), and this was later supported through molecular evidence via small isozyme (Arnold et al., 1990) and Random Amplified Polymorphic DNA (Arnold, 1993) datasets. Here we used a much larger genomic dataset, and while we do find evidence of some recent localized admixture between purported *I. nelsonii* individuals and *I. brevicaulis* and *I. hexagona*, which is consistent with the observations made by Arnold et al. (1990) and Arnold (1993)- we find no evidence of a hybrid origin for this species. If *I. nelsonii* were to have a signature of mixed ancestry from two or even three parental lineages, we should expect to see this reflected in our clustering analysis in the form of *I. nelsonii* individuals having mixed ancestry from the other purported taxa at lower values of k, however we did not observe this pattern (Figure 4). In fact, we find no evidence that the *I. nelsonii* individuals are any more genetically divergent from *I. fulva*, than different *I. fulva* populations are from each other (Table 2); we observed consistent and repeated assignments of the *I. nelsonii* individuals to the *I. fulva* cluster, regardless of the k examined. This absence of genetic structure within *I. fulva* populations (and between *I. nelsonii* and *I. fulva* populations) is likely due to gene flow reflecting pollen movement rather than seed dispersal – the seeds float and would primarily drift downstream in a

This article is protected by copyright. All rights reserved.

Accepted Article

southerly direction. The primary pollinator for this species is the Ruby-Throated hummingbird, *Archilochus colubris* (Courter et al. 2013). Further, there is no evidence of behavioral isolation with respect to pollinator choice between *I. fulva* and *I. nelsonii* (Martin and Taylor 2013).

Under a homoploid hybrid species hypothesis, we should expect *I. nelsonii* individuals to show intermediate clustering in the PCAs, but this was also not observed (Figure 3, 4, Appendix S1-see supplementary data with this article). Instead, we observed evidence in all cases that *I. nelsonii* is indistinguishable from any of the sampled populations of *I. fulva.* PCA analyses that examined the full dataset as well as one that sampled a subset of the data including only samples of the four purported species in Southern Louisiana for finer scale analysis, clustered *I. nelsonii* individuals within the *I. fulva* cluster. While we do acknowledge that some phenotypic differences do apparently exist between purported *I. fulva* and *I. nelsonii* species (Randolph et al., 1961; Randolph, 1966), these morphological differences are not largely reflected in a variety of common garden settings (Taylor et al 2011), suggesting much of the phenotypic differences observed among populations of purported *I. nelsonii* and *I. fulva* might be due to environmental variation. The findings from this extensive genomic dataset therefore cast much doubt not only on the homoploid hybrid origin of *I. nelsonii*, but on the species designation of this taxon as well.

Interestingly, we discovered individuals that were identified phenotypically as *I. brevicaulis* that showed the patterns of admixture we would expect to observe in a stabilized hybrid lineage (Figures 2-4). One population occurs in central Louisiana, and the other further north in Illinois. Both populations show admixture from multiple

This article is protected by copyright. All rights reserved.

parental groups at the lower Ks, yet also show strong assignment to their own clusters at higher Ks, indicating that there is some strong population structuring in these groups. Furthermore, in PCA these groups form clusters that intermediate to the other three major groups. Although we cannot confidently say that these groups represent reproductively isolated homoploid hybrid species, it is apparent from these data that these groups exhibit strong population structure that is built upon admixed genomes, a feature one might expect in a homoploid hybrid species. With this discovery, a more robust analysis of these hybrid genomes is warranted to assess the extent of stabilization and ongoing gene flow that may be occurring. We plan to address these questions in future work focusing more closely on the hybridization dynamics at play in these, and the other admixed populations discovered in this study.

The range-wide population genetic analyses of Louisiana Iris presented here revealed both expected and unexpected results. The current taxonomy of Louisiana Iris was largely confirmed for *I. brevicaulis, I. hexagona,* and *I. fulva.* However, there was no broad support for the infrequently-used *I. savanarrum* taxon, as it appears to be only slightly differentiated from *I. hexagona* individuals in the Northern and Western Gulf slope. Evidence of long-distance (up to 70km) pollen dispersal has been hypothesized for *I. hexagona*, which could potentially explain the reduced amount of genetic structuring in this study (Arnold et al. 1992). New findings of extensive contemporary hybridization between *I. hexagona* and *I. brevicaulis* was documented for the first time outside of Louisiana, and further genomic investigations are warranted to examine the genetic and ecological mechanisms by which these species are maintained in sympatry.

Surprisingly, *I. nelsonii* showed no genetic signature of homoploid hybrid origin.

This article is protected by copyright. All rights reserved.

Accepted Article

Further, this nominal taxon was also no more differentiated from *I. fulva* populations, than intraspecific populations of *I. fulva* were from each other. This is most surprising considering the long history of I. *nelsonii* being known as a "classic example" of homoploid hybrid speciation and considering it does have a slightly larger floral and leaf morphology than "typical" *I. fulva* populations. Further investigations of range-wide morphological variation in *I. fulva* are warranted to determine whether *I. nelsonii* is truly divergent, or if these results are within the standing genetic variation of *I. fulva*.

Paradoxically, while the results of this study clearly provide no support for *I. nelsonii* as a homoploid hybrid species, we found strong evidence that two different "*I. brevicaulis*-like" populations appear to be possible stable homoploid hybrids. While *I. nelsonii* can no longer be considered a "textbook example" of homoploid hybrid speciation, the Louisiana Iris system still offers unique and surprising opportunities to examine the evolutionary dynamics of contemporary and historic hybridization and will remain a "model organism" for such studies for a long time to come.

Acknowledgements

We thank the American Iris Society and the Society for Louisiana Irises for providing funding and support. We also thank all of the landowners for allowing us to sample the irises on their land. We would also like to thank Joy Sung, Keith Usry, and Aiko Amano, and Chris Kneupper for help in sampling and laboratory work. We would also like to thank the reviewers for the extremely helpful comments and suggestions that increased the academic quality of this manuscript.

Author Contribution

All authors have contributed to the content of the manuscript. A.S.Z. collected

This article is protected by copyright. All rights reserved.

Accepted Article

Accepted Article

samples, performed laboratory methods, and analyzed data. V.A.S provided analytical expertise, performed analyses and contributed to the production of graphics. N.H.M. provided laboratory space, supplies, and contributed heavily to the writing of the manuscript. C.C.N. provided theoretical and analytical expertise and contributed to DNA library prep and data analysis. All authors contributed to revision and writing of the manuscript.

Data Availability Statement

The data that support the findings of this study are openly available in the Dryad repository: https://datadryad.org/stash/dataset/doi:10.5061/dryad.jm63xsjbm

Supporting Information

Additional Supporting Information may be found online in the supporting information section at the end of the article:

- Appendix S1-PCA of southern Louisiana sample localities. *PCA plot including only individuals sampled from southern Louisiana. Green points represent I. hexagona samples, pink points indicate I. brevicaulis samples, red points represent I. fulva, and orange points indicate I. nelsonii.*
- Appendix S2-Barplots of K=8 and K=9. *Bar plots showing ENTROPY output for K=8 (A) and K=9 (B). No further substructure was detected at these Ks as compared with lower K values.*
- Appendix S3-Gelman-Rubin convergence diagnostic statistic table. *Table with Gelman-Rubin convergence diagnostic statistic for each run of K reported.*

References

This article is protected by copyright. All rights reserved.

Accepted Article

Abbott, R. J., N. H. Barton, and J. M. Good. 2016. Genomics of hybridization and its evolutionary consequences. *Molecular ecology* 25: 2325–2332.

Abbott, R. J., M. J. Hegarty, S. J. Hiscock, and A. C. Brennan. 2010. Homoploid hybrid speciation in action. *Taxon* 59: 1375–1386.

Allendorf, F. W., R. F. Leary, P. Spruell, and J. K. Wenburg. 2001. The problems with hybrids: setting conservation guidelines. *Trends in Ecology and Evolution* 16.

Anderson, E. 1949. Introgressive Hybridization. J. Wiley, New York.

Arnold, M. L. 2016. Anderson's and Stebbins' Prophecy Comes True: Genetic Exchange in Fluctuating Environments. *Systematic Botany* 41: 4–16.

Arnold, M. L. 1993. Iris nelsonii (Iridaceae): Origin and genetic composition of a homoploid hybrid species. *American Journal of Botany* 80: 577–583.

Arnold, M. L. 1997. Natural Hybridization and Evolution. Oxford University Press, Oxford, England.

Arnold, M. L., B. D. Bennett, and J. L. Hamrick. 1990. Allozyme variation in louisiana irises: A test for introgression and hybrid speciation. *Heredity* 65: 297–306.

Arnold, M. L., A. L. Brothers, J. A. P. Hamlin, S. J. Taylor, and N. H. Martin. 2015. Divergence with Gene Flow-What Humans and Other Mammals Got up to. Springer International Publishing.

Arnold, M. L., C. M. Buckner, and J. J. Robinson. 1991. Pollen-mediated introgression and hybrid speciation in Louisiana irises. *Proceedings of the National Academy of Sciences of the United States of America* 88: 1398–1402.

Arnold, M. L., J. J. Robinson, C. M. Buckner, and B. D. Bennet. 1992. Pollen dispersal and interspecific gene flow in Louisiana irises. *Heredity* 88: 399-404.

This article is protected by copyright. All rights reserved.

Accepted Article

Arnold, M. L., Y. Sapir, and N. H. Martin. 2008. Genetic Exchange and the Origin of Adaptations: Prokaryotes to Primates. *Philisophical Transactions of The Royal Society* 363: 2813–2820.

Arnold, M. L., S. Tang, S. J. Knapp, and N. H. Martin. 2010. Asymmetric introgressive hybridization among Louisiana Iris species. *Genes* 1: 9–22.

Ballerini, E. S., A. N. Brothers, S. Tang, S. J. Knapp, A. Bouck, S. J. Taylor, M. L. Arnold, and N. H. Martin. 2012. QTL mapping reveals the genetic architecture of loci affecting pre- and post-zygotic isolating barriers in Louisiana Iris. *BMC Plant Biology* 12: 1–12.

Bouck, A., S. R. Wessler, and M. L. Arnold. 2007. QTL analysis of floral traits in Louisiana Iris hybrids. *Evolution* 61: 2308–2319.

Carney, S. E., and M. L. Arnold. 1997. Differences in pollen-tube growth rate and reproductive isolation between Louisiana irises. *Journal of Heredity* 88: 545–549.

Courter, J. R., R. J. Johnson, C. W. Bridges, and K. G. Hubbard. 2013. Assessing migration of Ruby-throated Hummingbirds (*Archilochus colubris*) at broad spatial and temporal scales. *The Auk*, 130 (1), 107-117.

Coyne, J. A., and A. H. Orr. 2004. *Speciation*. Sinauer Associates, Sunderland, MA.

Cruzan, M. B., and M. L. Arnold. 1994 Assortative mating and natural selection in an Iris hybrid zone. *Evolution* 48**:** 1946–1958.

Dobzhasnky, T. 1937. Genetic Nature of Species Differences. *The American Naturalist* 71: 404–420.

Doyle, J. 1991. DNA protocols for plants. Molecular techniques in taxonomy (pp. 283–293). Berlin Heidelberg: Springer.

This article is protected by copyright. All rights reserved.

Accepted Article

Driscoe, A. L., C. C. Nice, R. W. Busbee, G. R. Hood, S. P. Egan, and J. R. Ott. 2019. Host plant associations and geography interact to shape diversification in a specialist insect herbivore. *Molecular Ecology* 28: 4197–4211.

E. Anderson and G. L. Stebbins, J. 1954. Hybridization as an Evolutionary Stimulus. *Evolution* 8: 378–388.

Edmands, S. 2007. Between a rock and a hard place: Evaluating the relative risks of inbreeding and outbreeding for conservation and management. *Molecular Ecology* 16: 463–475.

Forster, R. C. 1937. A cyto-taxonomic survey of the North American species of Iris. *Contributions from the Gray Herbarium* CXIX: 3–80.

Fu, L., B. Niu, Z. Zhu, S. Wu, and W. Li. 2012. CD-HIT: Accelerated for clustering the next-generation sequencing data. *Bioinformatics* 28: 3150–3152.

Gavrilets, S. 1999. A dynamical theory of speciation on holey adaptive landscapes. *American Naturalist* 154: 1–22.

Gilbert, K. K., A. Rose, D. G. Bock, M. T. Franklin, N. C. Kane, J.-S. Moore, B. T. Moyers, et al. 2012. Recommendations for utilizing and reporting population genetic analyses: the reproducibility of genetic clustering using the program STRUCTURE. *Molecular Ecology* 21: 4925–4930.

Gompert, Z., and C. A. Buerkle. 2016. What, if anything, are hybrids: enduring truths and challenges associated with population structure and gene flow. *Evolutionary Applications* 9: 909–923.

Gompert, Z., J. A. Fordyce, M. L. Forister, A. M. Shapiro, and C. C. Nice. 2006. Homoploid hybrid speciation in an extreme habitat. *Science* 314: 1923–1925.

This article is protected by copyright. All rights reserved.

Accepted Article

Gompert, Z., L. K. Lucas, C. A. Buerkle, M. L. Forister, J. A. Fordyce, and C. C. Nice. 2014. Admixture and the organization of genetic diversity in a butterfly species complex revealed through common and rare genetic variants. *Molecular Ecology* 23: 4555–4573.

Grant, V. 1981. Plant Speciation. Columbia University Press, New York.

Hamilton, J. A., C. Lexer, and S. N. Aitken. 2013. Differential introgression reveals candidate genes for selection across a spruce (Picea sitchensis × P. glauca) hybrid zone. *New Phytologist* 197: 927–938.

Hamlin, J. A., and M. L. Arnold. 2014. Determining population structure and hybridization for two iris species. *Ecology and Evolution* 4 (6): 743-755

Hamlin, J. A., T. J. Simmonds, and M. L. Arnold. 2017. Niche conservatism for ecological preference in the Louisiana iris species complex. Biological Journal of the Linnean Society 120 (1): 144-154.

Holliday, J. A., L. Zhou, R. Bawa, M. Zhang, and R. W. Oubida. 2016. Evidence for extensive parallelism but divergent genomic architexture of adaptation along altitudinal gradients in P. opulus trichocarpa. *New Phytologist* 209: 1240–1251.

Janes, J. K., J. M. Miller, J. R. Dupuis, R. M. Malenfant, J. C. Gorrell, C. I. Cullingham, and R. L. Andrew. 2017. The K = 2 conundrum. *Molecular Ecology* 26: 3594–3602.

Jiggins, C. D., C. Salazar, M. Linares, and J. Mavarez. 2008. Hybrid trait speciation and Heliconius butterflies. *Philosophical Transactions of the Royal Society B: Biological Sciences* 363: 3047–3054.

Lamichhaney, S., F. Han, M. T. Webster, L. Anderson, R. B. Grant, and P. R. Grant. 2018. Rapid hybrid speciation in Darwin's finches. *Science* 359: 172.

This article is protected by copyright. All rights reserved.

Accepted Article

Lawson, D. J., G. Hellenthal, S. Myers, and D. Falush. 2012. Inference of population structure using dense haplotype data. *PLoS Genetics* 8: 11–17.

Levin, D. A. 1983. Polyploidy and Novelty in Flowering Plants. *The American Naturalist* 122.

Levin, D. A. 2002. The Role of Chromosomal Change in Plant Evolution. Oxford University Press, Oxford, England.

Li, H., and R. Durbin. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25: 1754–1760.

Martin, N. H., A. C. Bouck, and M. L. Arnold. 2006. Detecting adaptive trait introgression between Iris fulva and I. brevicaulis in highly selective field conditions. *Genetics* 172: 2481–2489.

Martin, N. H., A. C. Bouck, and M. L. Arnold. 2005. Loci Affecting Long-Term Hybrid Survivorship in Louisiana Irises: Implications for Reproductive Isolation and Introgression. *Evolution* 59: 2116.

Martin, N. H., A. C. Bouck, and M. L. Arnold. 2007. The genetic architecture of reproductive isolation in Louisiana Irises: Flowering phenology. *Genetics* 175: 1803–1812.

Martin, N. H., Y. Sapir, and M. L. Arnold. 2008. The genetic architecture of reproductive isolation in Louisiana Irises: Pollination syndromes and pollinator preferences. *Evolution* 62: 740–752.

Martin, N. H., and S. J. Taylor. 2013. Floral preference, flower constancy, and pollen transfer efficiency of the ruby-throated hummingbird (*Archilochus colubris*) in mixed arrays of *Iris nelsonii* and *Iris fulva*." *Evolutionary Ecology Research* 15 (7):

This article is protected by copyright. All rights reserved.

783-792.

Mavárez, J., and M. Linares. 2008. Homoploid hybrid speciation in animals. *Molecular Ecology* 17: 4181–4185.

Mayr, E. 1942. Systematics and the Origin of Species. Columbia University Press, New York.

Meerow, A. W., M. Gideon, D. N. Kuhn, S. Mopper, and K. Nakamura. 2011. The genetic mosaic of Iris series Hexagonae in Florida: Inferences on the Holocene history of the Louisiana irises and anthropogenic effects on their distribution. *International Journal of Plant Sciences* 172: 1026–1052.

Meirmans, P. G. 2015. Seven common mistakes in population genetics and how to avoid them. *Molecular Ecology* 24: 3223–3231.

Melo, M. C., A. Grealy, B. Brittain, G. M. Walter, and D. Ortiz-Barrientos. 2014. Strong extrinsic reproductive isolation between parapatric populations of an Australian groundsel. *New Phytologist* 203: 323–334.

Morales, A. E., N. D. Jackson, T. A. Dewey, B. C. O'Meara, and B. C. Carstens. 2017. Speciation with gene flow in North American Myotis bats. *Systematic Biology* 66: 440–452.

Nei, M. 1987. Molecular evolutionary genetics. Columbia University Press, New York.

Nosil, P. 2008. Speciation with gene flow could be common. *Molecular Ecology* 17: 2103–2106.

Ottenburghs, J. 2018. Exploring the hybrid speciation continuum in birds. *Ecology and Evolution* 8: 13027–13034.

Parchman, T. L., Z. Gompert, J. Mudge, F. D. Schilkey, C. W. Benkman, and C. A.

This article is protected by copyright. All rights reserved.

Accepted Article

Accepted Article

Buerkle. 2012. Genome-wide association genetics of an adaptive trait in lodgepole pine. *Molecular Ecology* 21: 2991–3005.

Pinho, C., and J. Hey. 2010. Divergence with gene flow: Models and data. *Annual Review of Ecology, Evolution, and Systematics* 41: 215–230.

Pritchard, J. K., M. Stephens, and P. Donelly. 2000. Inference of population structure using multilocus genotype data. *Genetics* 155: 945–959.

Puritz, J. B., C. M. Hollenbeck, and J. R. Gold. 2014. dDocent: A RADseq, variant-calling pipeline designed for population genomics of non-model organisms. *PeerJ* 2014.

Randolph, L. F. 1966. Iris nelsonii, a new species of Louisiana Iris of hybrid origin. *Baileya* 14: 143–169.

Randolph, L. F., J. Mitra, and I. S. Nelson. 1961. Cytotaxonomic studies of Louisiana Irises. *International Journal of Plant Sciences* 123.

Rieseberg, L. H., O. Raymond, D. M. Rosenthal, Z. Lai, K. Livingstone, T. Nakazato, J. L. Durphy, et al. 2003. Major ecological transitions in wild sunflowers facilitated by hybridization. *Science* 301: 1211–1216.

Riley, H. P. 1938. A character analysis of colonies of Iris fulva, Iris hexagona var. giganticaerulea and natural hybrids. *American Journal of Botany* 25: 727–738.

Seehausen, O. 2004. Hybridization and adaptive radiation. *Trends in Ecology and Evolution* 19: 198–207.

Selz, O. M., R. Thommen, M. E. Maan, and O. Seehausen. 2014. Behavioural isolation may facilitate homoploid hybrid speciation in cichlid fish. *Journal of Evolutionary Biology* 27: 275–289.

This article is protected by copyright. All rights reserved.

Accepted Article

Shaw, J. P., S. J. Taylor, M. C. Dobson, and N. H. Martin. 2017. Pollinator isolation in Louisiana iris: Legitimacy and pollen transfer. *Evolutionary Ecology Research* 18: 429–441.

Small, J. K., and E. J. Alexander. 1931. Botanical interpretation of the iridaceous plants of the Gulf States. *Contributions from the New York Botanical Garden* 327: 325–357.

Soltis, P. S., and D. E. Soltis. 2009. The role of hybridization in plant speciation. *Annual Review of Plant Biology* 60: 561–588.

Sotola, A. V., D. S. Ruppel, T. H. Bonner, C. C. Nice, and N. H. Martin. 2019. Asymmetric introgression between fishes in the Red River basin of Texas is associated with variation in water quality. *Ecology and Evolution* 9: 2083–2095.

Sung, C. J., K. L. Bell, C. C. Nice, and N. H. Martin. 2018. Integrating Bayesian genomic cline analyses and association mapping of morphological and ecological traits to dissect reproductive isolation and introgression in a Louisiana Iris hybrid zone. *Molecular Ecology* 27: 959–978.

Taylor, S. J., M. L. Arnold, and N. H. Martin. 2009. The genetic architecture of reproductive isolation in louisiana irises: Hybrid fitness in nature. *Evolution* 63: 2581–2594.

Taylor, S. J., R. W. Willard, J. P. Shaw, M. C. Dobson, and N. H. Martin. 2011. Differential response of the homoploid hybrid species Iris nelsonii (Iridaceae) and its progenitors to abiotic habitat conditions. *American Journal of Botany*, 98(8): 1309-1316.

Taylor, S. J., L. D. Rojas, S. W. Ho, and N. H. Martin. 2013. Genomic collinearity and

This article is protected by copyright. All rights reserved.

Accepted Article

the genetic architecture of floral differences between the homoploid hybrid species Iris nelsonii and one of its progenitors, Iris hexagona. *Heredity* 110: 63–70.

Viosca, P. 1935. The irises of southeastern Louisiana: a taxonomic and ecological interpretation. *Bulletin of the American Iris Society*.

Wang, Y., D. Lu, Y. J. Chung, and S. Xu. 2018. Genetic structure, divergence and admixture of Han Chinese, Japanese and Korean populations. *Hereditas* 155: 19.

Wesseling, R. A., and M. L. Arnold. 2000. Pollinator behavior and the evolution of Louisiana Iris hybrid zones. *Journal of Evolutionary Biology* 13: 171–180.

Wood, T. E., N. Takebayashi, M. S. Barker, I. Mayrose, P. B. Greenspoon, and L. H. Rieseberg. 2009. The frequency of polyploid speciation in vascular plants. *Proceedings of the National Academy of Sciences of the United States of America* 106: 13875–13879.

Wu, C. I. 2001. The genic view of the process of speciation. *Journal of Evolutionary Biology* 14: 851–865.

Yakimowski, S. B., and L. H. Rieseberg. 2014. The role of homoploid hybridization in evolution: A century of studies synthesizing genetics and ecology. *American Journal of Botany* 101: 1247–1258.

Zhao, W., Y. Q. Sun, J. Pan, A. R. Sullivan, M. L. Arnold, J. F. Mao, and X. R. Wang. 2020. Effects of landscapes and range expansion on population structure and local adaptation. *New Phytologist* 228: 330–343.

This article is protected by copyright. All rights reserved.

TABLES

Table 1. Sample localities that were included in the study. Each location has a three letter ID, a species designation based on the phenotype of individuals sampled at the location, and the number of individuals collected from the locality.

Location	ID	Species	n	Latitude	Longitude
Fayette, Illinois	bfi	*brevicaulis*	16	38.927	-89.114
St. Landry, Louisiana	bsf	*brevicaulis*	28	30.547	-91.981
Iberia, Louisiana	bil	*brevicaulis*	16	29.978	-91.754
Brazos, Texas	bzt	*brevicaulis*	34	30.569	-96.202
Galveston, Texas	ugt	*brevicaulis*	6	29.509	-95.116
Fort Bend, Texas	uft	*brevicaulis*	9	29.379	-95.583
Matagorda, Texas	bbt	*brevicaulis*	7	28.915	-95.756
Brazoria, Texas	uht	*brevicaulis*	7	29.088	-95.575
Brazoria, Texas	ubt	*brevicaulis*	25	29.009	-95.486
Brazoria, Texas	hbz	*brev/hex contact zone*	139	28.882	-95.584
Brazoria, Texas	hbt	*hexagona*	9	28.881	-95.586
St. Mary, Louisiana	hml	*hexagona*	76	29.776	-91.774
Assumption, Louisiana	ual	*hexagona*	18	20.907	-91.985
Lee, Florida	uuf	*hexagona*	60	26.575	-81.822
Union, Illinois	fui	*fulva*	25	37.443	-89.396
Fulton, Kentucky	ffk	*fulva*	20	36.525	-89.315
Carroll, Mississippi	fcm	*fulva*	10	33.413	-90.155
St. Landry, Louisiana	fll	*fulva*	48	30.546	-91.864
St. Martin, Louisiana	fml	*fulva*	7	30.165	-91.814
Vermillion, Louisiana	nvl	*nelsonii*	72	29.902	-92.096
		Total =	632		

This article is protected by copyright. All rights reserved.

Accepted Article

Table 2. G_{ST} matrix showing pairwise sample locality level G_{ST} values. Colored cells represent comparisons within a species.

	bfi	bil	bsf	bzt	ubt	uft	ugt	uht	bbt	fcm	ffk	fll	fml	fui	nvl	hbt	hml	ual	uuf
bfi	0	0.095	0.099	0.064	0.061	0.069	0.078	0.08	0.077	0.119	0.115	0.122	0.132	0.112	0.118	0.184	0.171	0.149	0.185
bil		0	0.005	0.127	0.119	0.132	0.14	0.144	0.14	0.094	0.091	0.096	0.106	0.089	0.089	0.144	0.128	0.109	0.143
bsf			0	0.129	0.122	0.134	0.142	0.146	0.142	0.105	0.101	0.106	0.117	0.099	0.099	0.149	0.133	0.115	0.148
bzt				0	0.006	0.01	0.019	0.02	0.015	0.155	0.151	0.158	0.168	0.148	0.154	0.213	0.2	0.18	0.214
ubt					0	0.01	0.018	0.02	0.014	0.147	0.143	0.15	0.16	0.14	0.146	0.201	0.188	0.169	0.202
uft						0	0.024	0.026	0.019	0.16	0.156	0.162	0.172	0.152	0.158	0.218	0.204	0.185	0.218
ugt							0	0.034	0.026	0.168	0.163	0.17	0.181	0.16	0.166	0.225	0.211	0.191	0.225
uht								0	0.03	0.172	0.168	0.175	0.185	0.165	0.17	0.229	0.215	0.195	0.229
bbt									0	0.168	0.164	0.17	0.18	0.161	0.166	0.224	0.211	0.191	0.225
fcm										0	0.007	0.008	0.018	0.007	0.012	0.208	0.193	0.158	0.208
ffk											0	0.006	0.016	0.004	0.01	0.204	0.189	0.155	0.204
fll												0	0.01	0.006	0.008	0.209	0.194	0.159	0.209
fml													0	0.015	0.018	0.218	0.204	0.168	0.218
fui														0	0.009	0.202	0.187	0.153	0.202
nvl															0	0.201	0.186	0.152	0.2
hbt																0	0.01	0.015	0.021
hml																	0	0.007	0.013
ual																		0	0.019
uuf																			0

Color	Species
Blue	I. brevicaulis
Red	I. fulva
Yellow	I. nelsonii
Purple	I. hexagona

FIGURE LEGENDS

This article is protected by copyright. All rights reserved.

Accepted Article

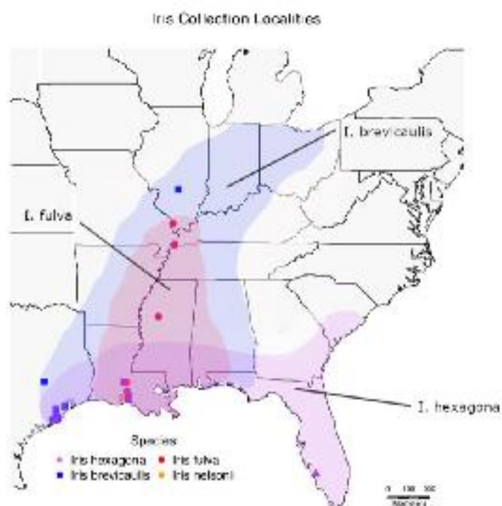
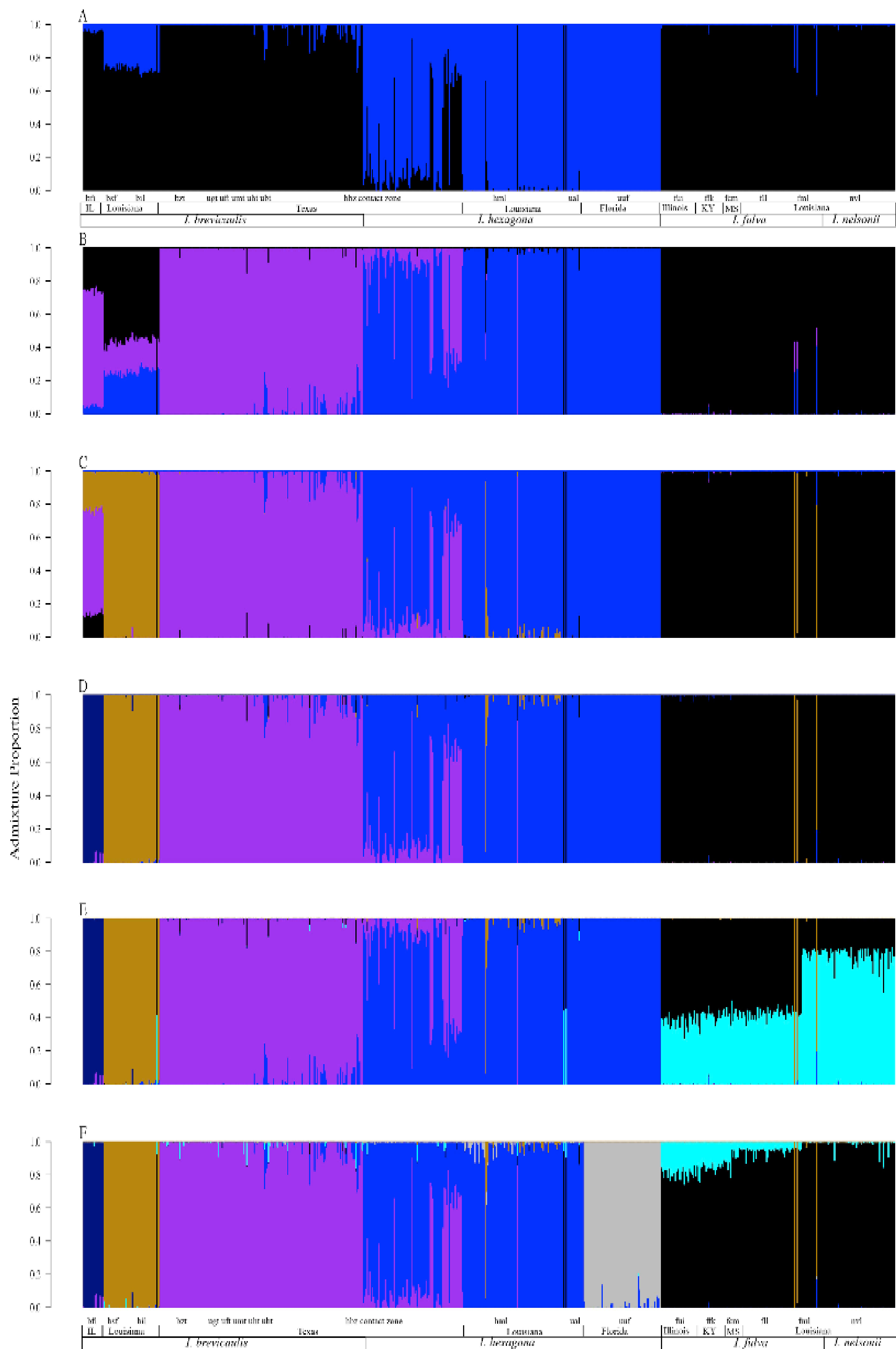


Figure 1. Map of collection localities in eastern USA. Species designations were based on phenotype of the sampled individuals at each locality upon collection. Ranges are overlaid onto sample localities which are depicted as triangles, squares, and circles depending on species.

This article is protected by copyright. All rights reserved.

Accepted Article

This article is protected by copyright. All rights reserved.



Accepted Article

This article is protected by copyright. All rights reserved.

Figure 2-Admixture proportions as estimated with ENTROPY. Each bar represents an individual and is colored to depict the individuals estimated admixture proportion inherited from k source populations. Starting with k=2 (A), we show each model output incrementally until k=7 (F), after which the model is a poor fit to the data. The species designation and location information are depicted below the plot for ease of reference.

Accepted Article

This article is protected by copyright. All rights reserved.

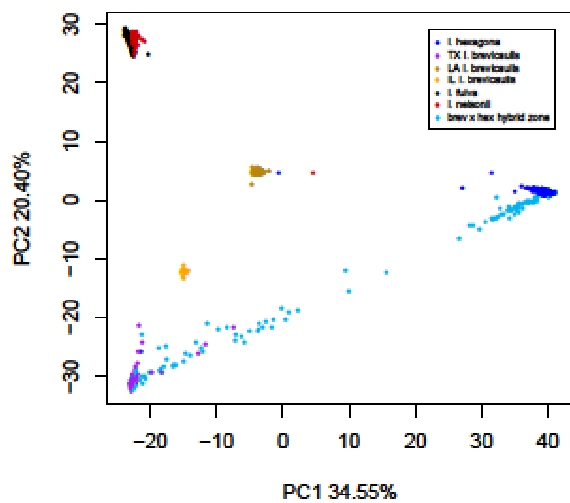


Figure 3. PCA plot of genotype likelihoods across individuals. The PC 1 axis is plotted against the PC 2 axis. Individuals are colored based on significant population structure findings as discovered in our ENTROPY analyses.

This article is protected by copyright. All rights reserved.

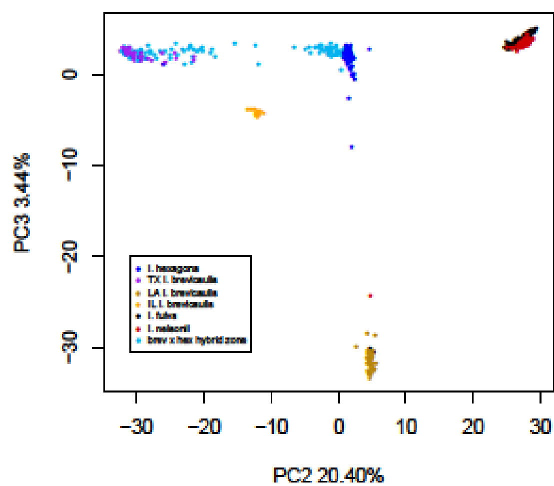


Figure 4. PCA plot of genotype likelihoods across individuals. The PC 2 axis is plotted against the PC 3 axis. Individuals are colored based on significant population structure findings as discovered in our ENTROPY analyses.

This article is protected by copyright. All rights reserved.

Accepted Article