Introgression between *Sphyrapicus nuchalis* and *S. varius* sapsuckers in a hybrid zone in west-central Alberta

Journal:	*Journal of Avian Biology*
	Research Article
Keywords:	hybrid zone, introgression, red-naped sapsucker, yellow-bellied sapsucker

Abstract:	Studying species interactions at hybrid zones allows biologists to understand the forces that promote speciation. Hybridization among *Sphyrapicus nuchalis*, *S. varius*, and *S. ruber* has long been acknowledged, and hybrid zones between *S. nuchalis/S. ruber* and *S. varius/S. ruber* have been characterized with both genetic and genomic data. Using a combination of next-generation Restriction Site-Associated DNA sequencing (RAD-Seq) and traditional genetic methods, we examined patterns of introgression in the poorly characterized *S. nuchalis/S. varius* contact zone; the two most similar species in the complex, though they are not each other's closest relatives. We found high introgression rates, with several early and many advanced generation hybrids along a 275 km stretch of Rocky Mountain foothill, pointing to a well-established hybrid zone with hybrid individuals backcrossing with individuals from the parental species and each other. Plumage colouration in the hybrid zone was a relatively poor indicator of parental or hybrid status, which could be attributed to the possible involvement of few large effect genes.

Abstract

Studying species interactions at hybrid zones allows biologists to understand the forces that promote speciation. Hybridization among *Sphyrapicus nuchalis*, *S. varius*, and *S. ruber* has long been acknowledged, and hybrid zones between *S. nuchalis/S. ruber* and *S. varius*/*S. ruber* have been characterized with both genetic and genomic data. Using a combination of next-generation Restriction Site-Associated DNA sequencing (RAD-Seq) and traditional genetic methods, we examined patterns of introgression in the poorly characterized *S. nuchalis/S. varius* contact zone; the two most similar species in the complex, though they are not each other's closest relatives. We found high introgression rates, with several early and many advanced generation hybrids along a 275 km stretch of Rocky Mountain foothill, pointing to a well-established hybrid zone with hybrid individuals backcrossing with individuals from the parental species and each other. Plumage colouration in the hybrid zone was a relatively poor indicator of parental or hybrid status, which could be attributed to the possible involvement of few large effect genes.

Keywords: Hybrid zone; introgression; Genotyping-by-Sequencing; red-naped sapsucker; yellow-bellied sapsucker

Introduction

Describing the process by which new species arise (speciation) is central to comprehending avian biodiversity, and speciation can be best studied by examining populations at intermediate stages of species divergence (Grant & Grant 1992), where interbreeding (hybridization) between individuals from different populations is still possible. Studying hybridization among such forms often reveals the extent of reproductive isolation between these populations and levels of divergence. Hybridization is richest and most informative in hybrid zones where the breeding ranges of two adjacent species overlap (Hewitt 1988, Short 1972), which are often called "natural laboratories" of speciation because they present many different combinations of parental alleles, much like biologists might create in a laboratory (Hewitt 1988). The populations may collapse into a hybrid swarm, they may remain distinct despite introgression, and they may be ephemeral or stable, recurring or novel, distributed broadly or restricted to a narrow zone (Grant & Grant 1992, Moore 1977, Short 1972). Describing hybrid zone dynamics helps us understand the evolutionary histories of species and the processes (isolating mechanisms) that keep them apart and predict their evolutionary trajectories (Grant & Grant 1992).

Sphyrapicus varius, *S. nuchalis*, and *S. ruber* are three species of North American woodpeckers that hybridize in sympatry. Hybridization within this species complex has long been of interest to biologists, who have studied the forces maintaining several hybrid zones and changes within them (Billerman 2016, Billerman et al. 2016, 2019, Howell 1952, Johnson & Johnson 1985, Natola & Burg 2018, Scott et al. 1976, Seneviratne et al. 2012, 2016, Trombino 1998). *Sphyrapicus ruber*/*S. nuchalis* and *S. ruber/S. varius* hybrid zones have been studied both behaviourally and genetically (Billerman et al. 2016, 2019, Cicero & Johnson 1995, Grossen et

al. 2016, Johnson & Johnson 1985, Johnson & Zink 1983, Natola & Burg 2018, Seneviratne et al. 2016, 2012). Though these species hybridize, studies of mate choice in two hybrid zones show the forms mate assortatively, supporting their designations as separate species (Johnson & Johnson 1985, Seneviratne et al. 2012, 2016). Possible putative reproductive barriers between these species include different migratory strategies, breeding phenology, habitat preferences, and assortative mating (Billerman et al. 2016, Johnson & Johnson 1985, Natola & Burg 2018, Trombino 1998).

A *S. nuchalis/S. varius* hybrid zone exists in west-central Alberta, but it has not been well studied using genetic methods. Recent behavioural work (Jocelyn Hudon, unpubl. obs.) and some genetic data (Natola & Burg 2018) document hybridization within this hybrid zone, but we know little about the extent of recent or past hybridization or genetic introgression between the two species. Since the *S. varius/S. ruber* and *S. nuchalis/S. ruber* hybrid zones are well characterized, describing the *S. varius/S. nuchalis* hybrid zone stands to greatly improve our understanding of reproductive interactions among these three species and the isolating barriers that assist their differentiation, notably plumage. *S. nuchalis* and *S. varius* are of particular interest because, of the three species of sapsuckers noted above, they are the two most alike phenotypically with only slight differences in the amount of red plumage on the nape, throat, and malar, and white patterning on the back. This similarity belies their more distant genetic relationship (*S. nuchalis and S. ruber* are sister taxa; Johnson and Zink 1983, Cicero and Johnson 1995). By comparison, *S. ruber*'s entirely red head, chin, and breast contrast sharply with the variegated look of *S. nuchalis* and *S. varius*, and might help us tease out the role of phenotypic differences in keeping maintaining the species in this complex.

Here we characterize genetically an area where the ranges of *S. nuchalis* and *S. varius* overlap and the two forms hybridize in order to delineate the geographical extent of the hybrid zone, quantify introgression rates and characteristics of the contact zone, and compare early vs late generation hybrids. Using traditional and Restriction Site-Associated DNA sequencing (RAD-Seq) methods allows us to examine introgression patterns both across the genome and across the geographic range, which provides a broader context to investigate patterns of hybridization between *S. nuchalis* and *S. varius*.

Methods

Sample acquisition

We collected DNA samples from specimens in museum collections and birds caught with mist nets during the breeding season. We collected wild-caught samples from May to July to reduce the number of migrants caught. Birds were called in with playbacks and caught using 12 m mist nets. We took a small (<50 µL) sample of blood from the brachial vein, banded the birds, and took morphometric measurements and photographs. All birds were released on site and blood samples were stored in 99% ethanol. Museum specimens were selected from birds collected within the last 20 years during the breeding season to ensure data reflected contemporary patterns, however, we acknowledge this may contain multiple generations (Supplemental Material 1). Included in these museum specimens is material collected at the conclusion of a 15-year study of reproductive interactions between *S. varius* and *S. nuchalis* along the foothills of western Alberta between Strachan (53.4005˚ N, -117.8685˚ W) and the Porcupine Hills (49.0512˚ N, -113.9115˚ W), AB (Jocelyn Hudon, unpubl. obs.). Overall, the

phenotypic make-up of sapsuckers in the hybrid zone changed little over the 15 years of study of the hybrid zone (but see Geographic, genomic, and genetic patterns of introgression below).

We examined sapsuckers for the extent of red on the nape, the degree of invasion of the malar stripe and cheek patch by red in males, the amount of red on the throat of females, and the amount and distribution of white markings on the back, visible characteristics that differ noticeably between the two species of sapsuckers and are intermediate in the hybrids. With the sapsuckers from the hybrid zone, we scored the first two characteristics (three between the two sexes) on a scale of 0 to 3, in increments of half points, 0 corresponding to the conditions observed on phenotypically pure *S. varius*, and the third characteristic in the same fashion on a scale of 0 to 2, for a maximum of 8 points for a phenotypically pure *S. nuchalis* (Supplemental Material 2). We classified individuals with any evidence of intermediate characters as hybrids.

RAD-Seq methods

RAD-Seq DNA extraction, processing

We extracted DNA using a standard phenol-chloroform extraction procedure and sent 63 samples (34 *S. nuchalis*, 18 *S. varius*, and 11 hybrids) to Cornell University's Institute for Genomic Diversity (IGD) for Genotyping-by-Sequencing (GBS) following Elshire et al. (2011) with the restriction enzyme *PstI*. We sent an additional 84 samples (42 *S. nuchalis*, 23 *S. varius*, and 19 hybrids) to the Genomic Sequencing and Analysis Facility (GSAF) at the University of Texas for double digest RADseq (ddRAD) using the restriction enzyme pair *NlaIII* and *MluCI* following Peterson et al. (2012). See Supplemental Materials 3 for population locations and sample sizes. Ten individuals were sent to both facilities for a total of 137 unique samples, 57 collected within AB and 76 from outside AB.

SNP calling

Due to the lack of a reference genome for *Sphyrapicus*, we performed *de novo* SNP calling. For the GBS dataset, we used the GBS UNEAK analysis pipeline version 3.0, which is an extension of the JAVA program TASSEL (Bradbury et al. 2007), to filter reads and call SNPs. Quality filtering removed any reads with incorrect, missing, or multiple restriction cut sites or barcodes. Reads were aligned into identical sequence tags with a threshold of a minimum of 10 reads per tag for inclusion in the SNP calling process, the error tolerance rate set to 0.03, and a minimum minor allele frequency (MAF) of 0.05 for pairwise alignment identification of SNPs. We kept a maximum of one SNP per read to minimize linkage disequilibrium in downstream analyses, and only SNPs that were present in more than 40% of individuals were retained. We also filtered out individuals that had fewer than 10 % called genotypes.

We filtered, demultiplexed, and cleaned the ddRAD data using the STACKS v 2.5 process_radtags pipeline (J. Catchen et al., 2013; J. M. Catchen et al., 2011). We used the denovo_map STACKS pipeline to identify SNPs *de novo* with the number of reads required to create a stack set to 10. We used similar filtering parameters to the GBS dataset in STACKS by filtering for a MAF of 0.05 and only keeping SNPs present in 40 % of individuals. We then used PLINK v 1.07 (Purcell et al. 2007) to filter the ddRAD samples for individuals with no less than 10% called genotypes.

Genomic analyses of hybrid zone individuals

Pairwise F_{ST} was calculated using the software ARLEQUIN v 3.5.2.2 (Excoffier & Lischer 2010) for both the GBS and ddRAD datasets. We calculated two sets of pairwise F_{ST}

values for the two different datasets, one using only those individuals from allopatric populations and one for only phenotypically pure species within the hybrid zone. PCAs were made for both datasets in R v 4.0.3 (R Development Core Team 2020) in accordance with scripts from Irwin et al. (2016).

To determine the proportion of ancestry from each species in individuals from the hybrid zone, we used the program ADMIXTURE v 1.2.3 (Alexander et al., 2009). As a result, we ran each ADMIXTURE analysis for K = 1-5, using a quasi-Newton algorithm for accelerated convergence (Zhou et al., 2011) and a 5-fold cross-validation. We used the default block relaxation algorithm to perform point estimation stopping the analyses when the change in the log-likelihood of point estimations between iterations increased by <0.0001. We determined the number of clusters that best fit the data by the K value with the lowest cross-validation error. We ran ADMIXTURE on the two datasets separately as they both contain different loci and different individuals. After the analyses, we compared the results using t-tests of each species classification group and paired t-tests on samples sequenced using both RAD-Seq methods and found they were not significantly different (*S. nuchalis*, $p = 0.37$; *S. varius,* $p = 0.74$; hybrid, $p = 0.07$; samples sequenced with both technologies, $p = 0.94$). We therefore combined the results to view the complete dataset together. We used the resulting ancestry coefficient values to define individuals as either *S. nuchalis* ($Q \leq 0.1$), *S. varius* ($Q \geq 0.9$), or admixed ($0.1 < Q < 0.9$) in accordance with Billerman et al. (2019). For ADMIXTURE plots, we removed five samples lacking geographic data and plotted all AB individuals in ascending latitudinal order.

Genomic structure of the hybrid zone

153 To determine the proportion of F1 and advanced generation hybrids within AB, we
154 compared the hybrid index (HINDEX) score (the allele frequencies of the individual compared
155 to the allele frequencies of *a priori* defined parental populations) to the interspecific
156 heterozygosity score (mean number of loci that are heterozygous for parental alleles) for each
157 individual within AB using the R package INTROGRESS v 1.22 (Gompert & Buerkle, 2010;
158 Gompert & Buerkle, 2009; R Development Core Team, 2013) for both GBS and ddRAD datasets
159 individually. We compared the HINDEX results using t-tests as in ADMIXTURE analyses and
160 found they were not significantly different (HINDEX: *S. nuchalis*, $p = 0.13$; *S. varius,* $p = 0.16$;
161 hybrid, $p = 0.08$; samples sequenced with both technologies, $p = 0.95$; interspecific
162 heterozygosity: *S. nuchalis*, $p = 0.06$; *S. varius,* $p = 0.11$; hybrid, $p = 0.15$; samples sequenced
163 with both technologies, $p = 0.56$) and combined the two datasets for further analysis. We defined
164 parental individuals within AB as those with a hybrid index of either ≤ 0.1 (*S. nuchalis*) or ≥ 0.9
165 (*S. varius*).
166 We plotted the interspecific heterozygosity against the hybrid index for each individual to
167 distinguish pure individuals from F1 hybrids or advanced generation backcrosses. We used *S.*
168 *varius* and *S. nuchalis* individuals from allopatric populations with high assignment of ancestry
169 to their respective species in ADMIXTURE as *a priori* parental populations in the
170 INTROGRESS analysis. To determine the genetic composition of individuals within the hybrid
171 zone, we analyzed all individuals located within AB regardless of their phenotypic identification.
172
173 *Traditional genetic marker methods*
174 DNA extraction, amplification, and sequencing

We selected a total of 206 samples from populations outside and within Alberta from each species and 45 individuals designated as phenotypic hybrids (Supplemental Materials 4). We extracted total genomic DNA from blood samples using a modified Chelex extraction (Walsh et al., 1991, Burg and Croxall 2001). Following extraction, we stored all samples at -20˚C. We Sanger sequenced an array of genetic loci looking to find SNPs or indels that differentiated between species and found none that were diagnostic. Three, α-enolase, glyceraldehyde, and an anonymous nuclear marker, showed strong species-specific structure. We screened for polymorphisms at these loci across the species' shared and allopatric ranges.

We amplified a 370 bp segment of the α-enolase (Enol) nuclear gene in 17 individuals total from *S. nuchalis* and *S. varius* using the Enol8L731 and Enol9H912 primers (Supplemental Material 5). The thermal cycling profile was one cycle of 120 s at 94 ˚C, 45 s at 54 ˚C, 60 s at 72 ˚C; 37 cycles of 30 s at 94 ˚C, 45 s at 54 ˚C, 60 s at 72 ˚C; one cycle of 300 s at 72 ˚C and 20 s at 4 ˚C. The 25 µL PCR reaction contained 5x Green GoTaq® Flexi buffer (Promega), 0.2 mM dNTP, 1 mM $MgCl_2$, 0.4 µM primers Enol8L731 and Enol9H912, 0.5 U GoTaq® Flexi polymerase, and genomic DNA.

We amplified and sequenced a 450 bp region of the glyceraldehyde gene (GAPD) in a total of 11 individuals of both species using primers GAPD11L890 and GAPD12H950 (Supplemental Material 5). The thermal cycling profile was similar to that used to amplify Enol, but with a 60 ˚C annealing temperature, and 0.8 mM $MgCl_2$.

We sequenced 760 bp of an anonymous region (ANM) in 21 individuals from both species using primers TP1F4 and TP1R5 (Supplemental Material 5). The thermal cycling profile was similar to the profile used in Enol, but with a 48 ˚C annealing temperature, a 105 s extension time, 0.8 mM dNTP, and 2.5 mM $MgCl_2$.

198 We sent successfully amplified samples to NanuQ sequencing service at McGill
199 University, Montreal, Quebec for sequencing. We aligned sequences using MEGA v. 6 (Tamura
200 et al. 2011).
201
202 SNP screening
203 We used the aligned Enol, GAPD, and anonymous nuclear marker sequences to identify
204 SNPs. We detected a C/T SNP 213 bp from the 3' end of the Enol8L731 primer. The GAPD
205 sequences contained a 4 bp insertion/deletion that was associated with a CTC/ATG multiple
206 nucleotide polymorphism (hereafter referred to as a SNP for simplicity) 118 bp from the 3' end
207 of the GAPD11L890 primer. Sequences from the anonymous nuclear marker revealed a C/T
208 SNP 84 bp from the 3' end of the TP1F4 primer. We designed new primers for Enol and the
209 anonymous nuclear marker to anneal near the SNP to introduce a restriction enzyme cut sites
210 associated with the SNPs to allow screening with restriction enzymes. The Enol SapLM13
211 primer changed the original GTGG sequence to GCGG to cut with *NlaIII*. The new ANM primer
212 (TP1 SapRM13) changed the original CCG**T**TAA sequence to CCG**G**TAA to cut with *MspI*. We
213 added M13 tags to the 5' end either to allow screening on an acrylamide gel (Enol) or to increase
214 size differences of digested products on an agarose gel (GAPD, anonymous nuclear marker).
215 We used a standard PCR protocol with a 10 μL reaction containing 0.1 mM dNTP,
216 0.4 μM primers, 0.25 U GoTaq® Flexi polymerase, genomic DNA, and varying amounts of
217 $MgCl_2$, 5x GoTaq® Flexi buffer (Promega), and the addition or omission of 0.04 μM fluorescent
218 M13 tag (Supplemental Material 6). Thermal cycling protocols were the same as GAPD, with
219 different annealing temperatures (Supplemental Material 6). We digested PCR products with

restriction enzymes (Supplemental Material 6) and a 1x buffer for a minimum of 3 hours at 37 ˚C.

We ran digested Enol products on a 6% acrylamide gel on the LI-COR 4300 DNA Analyzer and scored individuals as homozygous for T (187 bp) or C (41 bp and 146 bp), or heterozygous (41 bp, 146 bp, and 187 bp) for 187 individuals. We ran the GAPD and the anonymous nuclear marker restriction digests on a 3% agarose gel. We scored 174 individuals for the GAPD locus as CTC/deletion (469 bp), ATG/insertion (149 bp and 320 bp), or heterozygous (149 bp, 320 bp, and 469 bp). We scored the products at the anonymous nuclear marker as homozygous C (41 bp and 103 bp) or T (144 bp), or heterozygous (41 bp, 103 bp, and 144 bp) for 206 individuals.

To determine statistical significance of SNP variation between *S. nuchalis* and *S. varius*, between each species and hybrids, and between zones of allopatry and sympatry within each species, we used Fisher's exact tests.

We ran genotype data at all three loci through STRUCTURE v. 2.3.4 (Pritchard et al. 2000) for ancestry assignment. We used a burn in of 10,000, MCMC length of 150,000, the loc priors setting and ran ten iterations for each K = 2-4. We used the lowest log-likelihood values to select optimal K for the data, and averaged Q values. We used the resulting ancestry coefficient values to define individuals as either *S. nuchalis* ($Q \leq 0.1$), *S. varius* ($Q \geq 0.9$), or admixed ($0.1 < Q < 0.9$) in accordance with Billerman et al. (2019).

Results

RAD-Seq

After filtering, the GBS dataset contained 133,191,542 good barcoded reads which made up 119,371 reciprocal tag pairs and resulted in a total of 3,860 SNPs shared between *S. varius* and *S. nuchalis* while the ddRAD dataset contained 316,037,253 good barcoded reads which made up 80,819 stacks and resulted in a total of 205 such SNPs. Thirteen samples were removed from the ddRAD dataset due to missing data resulting in 71 samples (18 *S. varius*, 35 *S. nuchalis*, and 18 hybrids, see Supplemental Material 3). The pairwise F_{ST} for the GBS dataset was 0.26 ($p < 0.05$) using only allopatric individuals and 0.16 ($p < 0.05$) using only phenotypically "pure" individuals from within AB. The pairwise F_{ST} for the ddRAD dataset was 0.26 ($p < 0.05$). After filtering too few pure *S. varius* from within the hybrid zone remained in the ddRAD dataset to calculate a pairwise F_{ST} of only phenotypically "pure" individuals from within AB. The PCA differentiated *S. varius* and *S. nuchalis* along PC1 with two mostly separate clusters and hybrid individuals positioned between them along the first principal component (Figure 1). However, some AB samples classified as *S. nuchalis* or *S. varius* grouped with a mostly hybrid cluster along the centre of PC1 in the GBS dataset and many phenotypically hybrid samples were clustered with the genotypically parental groups on either end of PC1 using both RAD-Seq datasets. ADMIXTURE identified K = 2 as having the lowest cross-validation error, and therefore the optimal number of clusters for both the ddRAD and GBS datasets.

The ADMIXTURE plot differentiated individuals outside of AB into two distinct clusters – one for *S. varius* and one for *S. nuchalis* (Figure 2a). Individuals identified phenotypically as *S. nuchalis* within AB were more admixed than both *S. nuchalis* individuals outside the zone and *S. varius* within AB. *Sphyrapicus varius* within AB showed little or no admixture, with the exception of one individual with 37% *S. nuchalis* ancestry (Figure 2a).

The hybrid individuals within AB had a mix of ancestry from both species. The phenotypic hybrid score of birds in the hybrid zone correlated loosely with their ancestry coefficient (Q values) for both RAD-Seq ($r = 0.78$) and traditional genetic marker data ($r = 0.60$; Figure 3). However, as in the PCAs, many birds that were phenotypically classed as hybrids showed genotypic scores typical of parental birds. The amount of admixture varied throughout AB, with individuals located further north having a higher proportion of *S. varius* ancestry and those to the south, a greater proportion of *S. nuchalis* ancestry (Figure 2a). Individuals with ancestry coefficient (Q) values ≥ 0.1 begin occurring at about 49.8˚ N and admixture is evident until about 52.3˚ N where no more individuals with $Q \leq 0.9$ are found. This corresponds to a hybrid zone approximately 275 km long (Figure 4). Within AB, 26% of birds were classified as genomic *S. nuchalis* ($Q \leq 0.1$), 22% as genomic *S. varius* ($Q \geq 0.9$), and 52% as genomic admixed ($0.1 < Q < 0.9$). Females were not underrepresented in the hybrid zone as compared to the allopatric zone.

Individuals within AB exhibited low overall interspecific heterozygosity ($0.11 - 0.41$; Figure 5). None of the samples in AB had an interspecific heterozygosity of zero, indicating all AB samples have heterozygous loci from both species (Figure 5). The maximum hybrid index value observed in the hybrid zone was > 0.999 (a hybrid), and the minimum hybrid index value was < 0.001 (a *S. nuchalis*). Eleven of the 21 individuals phenotypically identified as *S. nuchalis* in AB were identified as genotypic hybrids (HINDEX 0.1- 0.9), three of the 29 phenotypic hybrids were genotypic *S nuchalis* (HINDEX ≤ 0.1) and two were genotypic *S. varius* (HINDEX ≥ 0.9), and eight phenotypic *S. varius* within AB were genotypic hybrids (HINDEX 0.1-0.9). The ancestry of hybrids within AB was not skewed towards one species or the other, but instead varied across the hybrid index between the two species, similar to the results with ADMIXTURE

(Figures 1, 2 and 5). We found no F1 generation hybrids (HINDEX = 0.5, interspecific heterozygosity = 1.0; Figure 5), but this may be because shared variation between the species prevents early generation birds from having an interspecific heterozygosity equal to 1. The higher than expected interspecific heterozygosity for pure individuals within the hybrid zone and the prevalence of samples with low interspecific heterozygosity is indicative of widespread allele sharing and advanced generation hybrids.

Traditional genetic markers

Enol showed nearly diagnostic SNPs, with a T allele being characteristic of *S. nuchalis* (107 of 126 alleles) and a C allele being more common in *S. varius* (121 of 162 alleles; $p < 0.0001$; Table 1, Figure 6). *Sphyrapicus nuchalis* had a significantly higher proportion of GAPD insertions (106 of 118 alleles) than *S. varius* (86 of 138 alleles; $p < 0.0001$; Table 1a,b), and *S. nuchalis* tended to have proportionally fewer T alleles (85 of 128) at the anonymous nuclear marker than *S. varius* (169 of 194 alleles; $p < 0.0001$) (Table 1a,b). The two species showed significant differences at all three loci when considering only parental populations outside central Alberta (Enol, GAPD, anonymous nuclear marker, $p < 0.0001$; Table 1b).

Birds identified as *S. nuchalis* within AB had significantly different allele frequencies from both *S. varius* in AB and phenotypic hybrids at all three loci ($p \leq 0.002$) (Table 1b). However, *S. varius* had no loci with allele frequencies that differed significantly from hybrids ($p \geq 0.3307$).

Log likelihood values identified K = 2 as optimal for STRUCTURE. The clusters generally described the two species well, with most individuals in central AB expressing shared ancestry (Figure 2b). In AB, the STRUCTURE plot shows admixture extending through the

northern extent of our transect, revealing more extensive introgression in phenotypically *S. varius* individuals than in phenotypically *S. nuchalis* individuals. Individuals with ancestry coefficient values ≥ 0.1 begin occurring at about 49.8˚ N and remain throughout the rest of the transect (up to 60.0˚ N), representing a transect of more than 1130 km. Within AB, 21% of birds were classified as genetic *S. nuchalis* (Q ≤ 0.1), 0% as genetic *S. varius* (Q ≥ 0.9), and 89% as genetically admixed (0.1 < Q < 0.9).

Of 44 individuals examined using both RAD-Seq and traditional methods, Q values assigned clusters similarly in 27 samples and differently in 17 (Supplemental Material 7). Of the 17, ten had Q values of >0.8 to the same cluster.

Discussion

Traditional markers vs RAD-Seq

We provide the results for traditional genetic markers for a methodological comparison with RAD-Seq because it is generally financially and practically more accessible to researchers. The traditional genetic data differentiated admixed and non-admixed populations (Figure 2b), and corroborated the RAD-Seq findings of strong admixture between the two species along a north-south gradient of western AB. However, RAD-Seq data documented less admixture of northern *S. varius* compared to the traditional methods. The discrepancy between the two methods is likely due to the lower resolution inherent in using three loci compared to thousands of loci with RAD-Seq. Because of this general limitation of the traditional loci approach, we refer primarily to the RAD-Seq results throughout the discussion.

Rates of hybridization and introgression

Our data record high levels of introgression and a continuum from early to advanced generation hybrids. Most individuals within west-central AB were genetically admixed (52% RAD-Seq, 89% traditional loci) with intermediate Q scores, compared to 44-70% individuals in a *S. ruber/S. nuchalis* hybrid zone in northern California and Oregon (Billerman et al. 2019), 15% in the *S. ruber/S. nuchalis* hybrid zone in central British Columbia (Seneviratne et al. 2016) and 3% in the *S. ruber/S. varius* hybrid zone in British Columbia (Seneviratne et al. 2016). The majority of admixed individuals appear to be advanced generation hybrids (intermediate HINDEX scores, low interspecific heterozygosity). The presence of many advanced generation hybrids with low interspecific heterozygosity and a range of HINDEX scores indicates a well-established hybrid zone with many hybrid individuals backcrossing with individuals from the parental species and each other. Unlike northern California and Oregon, in west-central Alberta many of these individuals had mixed phenotypic characteristics (i.e. hybrids; Figure 3).

Geographic, genomic, and genetic patterns of introgression

Our data document genetic differentiation among allopatric populations of *S. nuchalis* and *S. varius* that largely disappears within the hybrid zone. This is supported by the strong genetic differentiation of populations outside AB from individuals within AB (Figure 2), and the low differentiation between individuals phenotypically resembling the parent species in AB and hybrids (Figure 5). Admixture of individuals within and near the hybrid zone occurs largely along a north-south gradient, in part due to limited suitable habitat west and east of the area of contact. *Sphyrapicus nuchalis* largely occurs in the south through low elevation passes across the Rockies, whereas *S. varius* reaches the contact zone through a narrow strip of forest between the Rocky Mountains and the Canadian prairies that is connected to the boreal forest to the north.

Our sampling throughout approximately 1200 km of AB shows genomically admixed individuals present over approximately 275 km along the foothills east of the Rocky Mountains (Figure 4).

Sphyrapicus nuchalis and *S. varius* have more similar habitat preferences than either species does to *S. ruber.* Both favor secondary growth deciduous and mixed forests and specialize on similar plant species including aspen (*Populus spp.*) and birch (*Betula spp.*) trees (Walters et al., 2002a, 2002b). However, *S. nuchalis* generally breeds in higher elevation areas and *S. varius* is more tightly linked to riparian and early succession habitats (Walters et al., 2002a, 2002b). In our study area (hybrid zone) both forms breed primarily along east-west riparian corridors populated by large aspens. Similar habitat preferences increase the likelihood for the parental species to occupy the same area, and sympatry of the forms will increase the likelihood of admixture. Furthermore, the hybrid zone is also located along a transitional environment between montane habitat to the south and the boreal forest to the north, where the habitats preferred by the two species blend (Downing and Pettapiece 2006). It is possible that in this blended environment intermediate phenotypes could be at an advantage over the parental phenotypes (the bounded hybrid superiority hypothesis of Moore (1977)). Additionally, hybrid zones made up of mostly late generation hybrids could permit selection to create novel allelic combinations adapted to the intermediate environment (Hamilton & Aitken, 2013; Milne & Abbott, 2008; Pinheiro et al., 2016). However, in the absence of data on hybrid fitness and relative abundance, we are unable to directly address this hypothesis. We note that the hybrid zone appears to be moving, *S. varius* and its characteristics, shifting south slightly over the 15 years of study of reproductive interactions between the two forms in western Alberta (Jocelyn Hudon, unpubl. obs.). Because a bounded hybrid superiority model depends on hybrids being more fit in a particular intermediate habitat, our moving hybrid zone may be inconsistent with

379 this model unless the intermediate habitat is shifting south as well. A southward shift in the
380 location of the hybrid zone could help explain why the allele frequencies of traditional marker
381 loci did not differ between AB *S. varius* (as opposed to allopatric *S. varius*) and the hybrids. The
382 "tail end" of *S. nuchalis* traditional marker genotypes found in AB *S. varius* may reflect a
383 remnant of the hybrid zone's southern shift, though this pattern is not reflected in the RAD-Seq
384 data.

385 In this part of Alberta where *S. varius* and *S. nuchalis* meet, the mixed forests that both
386 species require are largely restricted to riparian corridors in low-lying areas separated by an
387 extensive coniferous cover on the adjacent hills and in one spot the drier Bow River valley,
388 resulting in an important drop in density of sapsuckers compared to that in more distant areas
389 where suitable habitat is more readily available. Such density troughs create conditions that can
390 slow the movement of, or even trap, hybrid zones (Barton & Hewitt 1985, Hewitt 1988).
391 Immigration from larger populations to the north and/or to the south may contribute individuals
392 to these zones, but limited opportunities to mate with conspecifics in "density traps" with many
393 hybrids may further promote hybridization (Hewitt 1988, McCracken et al. 2013).

394 This part of Alberta is the meeting place of other species or strongly differentiated
395 subspecies of birds that also hybridize there, including the Yellow-shafted and Red-shafted
396 Northern Flicker (*Colaptes auratus auratus* and *C. a. cafer*, Wiebe 2000), the Oregon and Slate-
397 coloured Dark-eyed Junco (*Junco hiemalis oreganus* and *J. h. hiemalis*, Friis et al. 2016), and the
398 Audubon's and Myrtle forms of the Yellow-rumped Warbler (*Dendroica coronata audubonii* and
399 *D. c. coronata*, Brelsford and Irwin 2008), although none have been genetically characterized in
400 this area. Additional avian taxon pairs hybridize farther north in the Peace River region east of
401 the Rocky Mountains, including Mourning and MacGillivray's warblers (*Oporornis tolmiei* and

O. philadelphia, Irwin et al. 2009) and Winter and Pacific Wrens (*Troglodytes hiemalis* and *T. pacificus*, Toews and Irwin 2008). Toews and Irwin (2008) have also noted low wren densities in that contact zone. This broader context highlights the influence regional geographical features and/or habitat have on the location and outcomes of these species' contact. It would be interesting to know if any of the other species pairs' hybrid zones have experienced recent coincident southern shifts.

Plumage insights

All pre-genomic species classifications in this study were made based on plumage. *Sphyrapicus nuchalis* and *S. varius* are the more difficult sapsuckers to differentiate by plumage, but these classifications were carefully made using an eight-point system developed by J. Hudon (Supplemental Materials 2). Despite meticulous classification, our phenotypic and genotypic classifications were not always concordant. Nineteen AB sapsuckers had genotypic (Q values) and phenotypic classifications that did not accord, with eight phenotypic hybrids classified as genotypic parental types, and 11 phenotypic parental birds were genetically admixed (Figures 2a, 3a). Similar patterns were seen in other analyses (Figures 1, 5). This contrasts with the finding that plumage is generally a reliable predictor of ancestry in *S. nuchalis* and *S. ruber* (Billerman et al. 2019). This could be due to the plumage differences between *S. nuchalis* and *S. varius* being controlled by a few genes of large effect. Judging from the largely independent segregation of the three characteristics that we scored in the hybrids, at least three genomic regions must be involved. Howell (1952) noted the different sapsucker species differ primarily in the presence or absence of red colouration at the tip of black and white feathers on the head, neck, and upper breast. There are other examples where colour differences between closely related species are

425 explained by few genes or small genomic regions (e.g., Toews et al. 2016, Wang et al. 2020,
426 Aguillon et al. 2021). This means that species designations based on plumage should be taken
427 with care, and analyses should be based upon genetic classifications and additional data
428 wherever possible.

429 Plumage is often identified as a potentially important isolating barrier preventing the
430 production of a hybrid swarm and genetic homogenization in sapsuckers (Billerman et al. 2019,
431 Johnson & Zink 1983, Seneviratne et al. 2016). Though assortative mating has been observed or
432 inferred in both *S. nuchalis/S. ruber* (Billerman et al. 2019, Johnson & Johnson 1985) and *S.*
433 *varius/S. ruber* (Seneviratne et al. 2012, 2016) hybrid zones, no data has yet been published on
434 whether *S. nuchalis* and *S. varius* mate assortatively based on plumage phenotypes. In
435 Kananaskis Country, about halfway down the hybrid zone, the distribution of phenotypic hybrid
436 scores in 2000 did not differ significantly from a unimodal normal distribution, and pairings at
437 36 nests were no different from what would be expected from random pairings amongst available
438 phenotypes (Jocelyn Hudon, unpubl. obs.), suggesting little assortative mating. We note that they
439 are the two forms that look most alike, though they are not each other's closest relatives
440 (Johnson and Zink 1983, Cicero and Johnson 1995). Indeed, hybridization between the two
441 species does not appear to be prevented by their strong genetic differentiation – F_{ST} between *S.*
442 *nuchalis* and *S. varius* (0.26 this study, 0.23 Grossen et al. 2016) is large compared to either
443 species and *S. ruber* (0.18 *ruber/varius*, 0.06 *ruber/nuchalis*, Grossen et al. 2016). This raises an
444 intriguing question: how can assortative mating limit introgression if mate choice is based on a
445 trait that is not closely associated with ancestry, a phenomenon recently described by Semenov et
446 al. (2017) in white wagtails (*Motacilla alba*). If genotype and phenotype are not highly
447 correlated, assortative mating based on plumage may not necessarily lead to assortative mating at

most loci, and what would appear to be an isolating barrier may actually permit widespread introgression.

S. nuchalis/S. varius in the sapsucker hybridization complex

We have shown that hybridization between *S. nuchalis*/*S. varius* occurs in a somewhat stable hybrid zone 275 km in length in western Alberta, thus relatively wide compared to the 65 km for a S. *ruber/S. nuchalis* hybrid zone in southern British Columbia and 84 km for S. *ruber/S. varius* hybrid zone in northern British Columbia (775 and 400 km transects surveyed respectively; Seneviratne et al. 2016). Hybridization rates were quite high, and a variety of hybrid classes existed suggesting backcrossing occurs frequently. These are somewhat surprising results given the strong differentiation between *S. nuchalis* and *S. varius*. It could be that the greater phenotypic similarities in *S. nuchalis/S. varius* facilitate more introgression than would be expected from their high genomic differentiation and more distant evolutionary history.

Plumage does not appear to be a reliable indicator of genotype among *S. nuchalis* and *S. varius* individuals, despite high plumage/genotype correlations in another sapsucker pair (Billerman et al. 2019). The contrasts amongst this and other well documented sapsucker hybrid zones are intriguing for such closely related species and may offer a unique opportunity to try to decouple isolating barriers and hybridization rates in a future study integrating all three hybrid zone combinations.

Acknowledgements

We would like to acknowledge the assistance of R. Adams, B. Brinkman, K. Dohms, G. M. Erickson, B. Graham, C. Kaluthota, L. Lait, C. Macfarlane, P. Pulgarin, C. Scobie, and C.

Welke in the field. We also thank the following museums for supplying samples: Royal Alberta Museum, Canadian Museum of Nature, Field Museum, Museum of Southwest Biology, and Royal Saskatchewan Museum. Funding has been provided by Natural Sciences and Engineering Research Council of Canada Discovery Grant (TMB), Alberta Innovates (TMB), and the North American Bluebird Society (LN).

Data Accessibility

Data used in this manuscript will be made available on Dryad upon acceptance.

Author Contributions

JH and TB initiated the study. JH collected and scored the phenotype of sapsuckers from the hybrid zone. AC processed and analyzed the RAD-Seq data, while LN processed and analyzed the traditional genetic markers data. LN synthesized and organized the results. All authors contributed to the project development, analyses, and writing.

485 **References**

486 Aguillon, S.M., Walsh, J., & Lovette, I.J. (2021). Extensive hybridization reveals multiple
487 coloration genes underlying a complex plumage phenotype. Proc. R. Soc. B, 288,
488 20201805.
489 Alexander, D. H., Novembre, J., & Lange, K. (2009). Fast model-based estimation of ancestry in
490 unrelated individuals. Genome Research, 19(9), 1655–1664.
491 Barton, N. H., & Hewitt, G.M. 1985. Analysis of hybrid zones. Annual Review of Ecology and
492 Systematics 16, 113–148.
493 Billerman, S. M. (2016). An Integrative Approach to Understanding Hybrid Zone Movement and
494 Introgression in *Sphyrapicus* Woodpeckers. PhD Thesis. University of Wyoming.
495 Billerman, S. M., Cicero, C., Bowie, R. C. K., & Carling, M. D. (2019). Phenotypic and genetic
496 introgression across a moving woodpecker hybrid zone. Molecular Ecology, 28(7), 1692–
497 1708.
498 Billerman, S. M., Murphy, M. A., & Carling, M. D. (2016). Changing climate mediates
499 sapsucker (Aves : *Sphyrapicus*) hybrid zone movement. Ecology and Evolution, 6(22),
500 7976–7990.
501 Bradbury, P. J., Zhang, Z., Kroon, D. E., Casstevens, T. M., Ramdoss, Y., & Buckler, E. S.
502 (2007). TASSEL: Software for association mapping of complex traits in diverse samples.
503 Bioinformatics, 23(19), 2633–2635.
504 Brelsford, A. & Irwin, D. E. (2009). Incipient speciation despite little assortative mating: the
505 yellow-rumped warbler hybrid zone. Evolution, 63(12), 3050-3060.
506 Burg, T.M., & Croxall, J.P. (2001). Global relationships amongst black-browed and grey-headed
507 albatrosses: analysis of population structure using mitochondrial DNA and microsatellites.

508 Molecular Ecology, 10(11), 2647–2660.
509 Catchen, J., Hohenlohe, P. A., Bassham, S., Amores, A., & Cresko, W. A. (2013). Stacks: An
510 analysis tool set for population genomics. Molecular Ecology, 22(11), 3124–3140.
511 Catchen, J. M., Amores, A., Hohenlohe, P., Cresko, W., Postlethwait, J. H., & De Koning, D.-J.
512 (2011). Stacks: Building and genotyping loci *de novo* from short-read sequences. G3, 1(3),
513 171–182.
514 Cicero, C., & Johnson, N. K. (1995). Speciation in sapsuckers (*Sphyrapicus*): III. Mitochondrial-
515 DNA sequence divergence at the cytochrome-b locus. Auk, 112(3), 547–553.
516 Downing, D.J., & Pettapiece, W.W. (2006). Natural Regions Committee 2006. Natural Regions
517 and Subregions of Alberta. Government of Alberta. Pub. No. T/852.
518 Excoffier, L. & Lischer, H. E. L. (2010). Arlequin suite ver 3.5: a new series of programs to
519 perform population genetics analyses under Linux and Windows. Molecular Ecology
520 Resources, 10(3), 564-567.
521 Friesen, V. L., Congdon, B. C., Walsh, H. E., & Birt, T. P. (1997). Intron variation in marbled
522 murrelets detected using analyses of single-stranded conformational polymorphisms.
523 Molecular Ecology, 6(11), 1047–1058.
524 Friis, G., Aleixandre, P., Rodríguez-Estrella, R, Navarro-Sigüenza, A.G., & Milá, B. (2016).
525 Rapid postglacial diversification and long-term stasis within the songbird genus *Junco*:
526 phylogeographic and phylogenomic evidence. Molecular Ecology, 25(24), 6175-6195.
527 Gompert, Z., & Buerkle, C. A. (2009). A powerful regression-based method for admixture
528 mapping of isolation across the genome of hybrids. Molecular Ecology, 18(6), 1207–1224.
529 Gompert, Z., & Buerkle, C. A. (2010). INTROGRESS: A software package for mapping
530 components of isolation in hybrids. Molecular Ecology Resources, 10(2), 378–384.

Grant, P. R., & Grant, B. R. (1992). Hybridization of bird species. Science, 256(5054), 193–197.

Grossen, C., Seneviratne, S. S., Croll, D., & Irwin, D. E. (2016). Strong reproductive isolation and narrow genomic tracts of differentiation among three woodpecker species in secondary contact. Molecular Ecology, 25(17), 4247–4266.

Hamilton, J. A., & Aitken, S. N. (2013). Genetic and morphological structure of a spruce hybrid (*Picea sitchensis x P. glauca*) zone along a climatic gradient. American Journal of Botany, 100(8), 1651–1662.

Hewitt, G. M. (1988). Hybrid zones - natural laboratories for evolutionary studies. Trends in Ecology & Evolution, 3(7), 158–167.

Howell, T. R. (1952). Natural history and differentiation in the yellow-bellied sapsucker. Condor, 54(5), 237–282.

Irwin, D. E., Brelsford, A., Toews, D. P. L., MacDonald, C., & Phinney, M. (2009). Extensive hybridization in a contact zone between MacGillivray's warblers *Oporornis tolmiei* and mourning warblers *O. philadelphia* detected using molecular and morphological analyses. Journal of Avian Biology, 40(5), 539-552.

Irwin, D.E., Alcaide, M., Delmore, K.E., Irwin, J.H., & Owens, G. (2016). Recurrent selection explains parallel evolution of genomic regions of high relative but low absolute differentiation in a ring species. Molecular Ecology, 25(18), 4488–4507.

Johnson, N. K., & Johnson, C. B. (1985). Speciation in sapsuckers (*Sphyrapicus*): II. Sympatry, hybridization, and mate preference in *S. ruber daggetti* and *S. nuchalis*. Auk, 102(1), 1–15.

Johnson, N. K., & Zink, R. M. (1983). Speciation in sapsuckers (*Sphyrapicus*): I. Genetic differentiation. Auk, 100(4), 871–884.

McCracken, K. G., Wilson, R. E., & Martin, A. R. (2013). Gene flow and hybridization between

554 numerically imbalanced populations of two duck species on the subantarctic island of South
555 Georgia. PLoS ONE, 8(12), 1–8.
556 Milne, R. I., & Abbott, R. J. (2008). Reproductive isolation among two interfertile
557 *Rhododendron* species: Low frequency of post-F1 hybrid genotypes in alpine hybrid zones.
558 Molecular Ecology, 17(4), 1108–1121.
559 Moore, W. S. (1977). An evaluation of narrow hybrid zones in vertebrates. The Quarterly
560 Review of Biology, 52(3), 263–277.
561 Nadeau, N. J., Mundy, N. I., Gourichon, D., & Minvielle, F. (2007). Association of a single-
562 nucleotide substitution in TYRP1 with roux in Japanese quail (*Coturnix japonica*). Animal
563 Genetics, 38(6), 609–613.
564 Natola, L., & Burg, T. M. (2018). Population genetics and speciation of yellow-bellied, red-
565 naped, and red-breasted sapsuckers (*Sphyrapicus varius*, *S. nuchalis*, and *S. ruber*). Journal
566 of Heredity, 109(6), 663–674.
567 Peterson, B. K., Weber, J. N., Kay, E. H., Fisher, H. S., & Hoekstra, H. E. (2012). Double digest
568 RADseq: An inexpensive method for *de novo* SNP discovery and genotyping in model and
569 non-model species. PLoS ONE, 7(5), e37135.
570 Pinheiro, F., Gouveia, T. M. Z. de M. E, Cozzolino, S., & Cafasso, D. (2016). Strong but
571 permeable barriers to gene exchange between sister species of *Epidendrum*. American
572 Journal of Botany, 103(8), 1472–1482.
573 Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using
574 multilocus genotype data. Genetics, 155(2), 945–959.
575 Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., Maller, J.,
576 Sklar, P., de Bakker, P. I. W., Daly, M. J., & Sham, P. C. (2007). PLINK: A tool set for

whole-genome association and population-based linkage analyses. American Journal of Human Genetics, 81(3), 559–575.

Scott, D. M., Ankney, C. D., & Jarosch, C. H. (1976). Sapsucker hybridization in British Columbia: changes in 25 years. Condor, 78(2), 253–257.

Semenov, G. A., Smith, C. C. R., Scordato, E. S. C., Khaydarov, D. R., Kane, N. C., & Safran, R. J. (2017). Effects of assortative mate choice on the genomic and morphological structure of a hybrid zone between two bird subspecies. Molecular Ecology, 26(22), 6430–6444.

Seneviratne, S. S., Davidson, P., Martin, K., & Irwin, D. E. (2016). Low levels of hybridization across two contact zones among three species of woodpeckers (*Sphyrapicus* sapsuckers). Journal of Avian Biology, 47(6), 887–898.

Seneviratne, S. S., Toews, D. P. L., Brelsford, A., & Irwin, D. E. (2012). Concordance of genetic and phenotypic characters across a sapsucker hybrid zone. Journal of Avian Biology, 43(2), 119–130.

Short, L. L. (1972). Hybridization, taxonomy and avian evolution. Missouri Botanical Garden Press, 59(3), 447–453.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., & Kumar, S. (2011). MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Molecular Biology and Evolution, 28(10), 2731–2739.

Team, R. D. C. (2011). R: a language and environment for statistical computing. R Foundation for Statistical Computing.

Toews, D.P.L., & Irwin, D. E. (2008). Cryptic speciation in a Holarctic passerine revealed by genetic and bioacoustic analyses. Molecular Ecology, 17(11), 2691-2705.

Toews, D.P.L., Taylor, S.A., Vallender, R., Brelsford, A., Butcher, B.G., Messer, P.W., Lovette,

I.J. (2016). Plumage genes and little else distinguish the genomes of hybridizing Warblers. Current Biol, 26(17), 1-6.

Trombino, C. L. (1998). Species interactions in the hybrid zone between red-breasted (*Sphyrapicus ruber*) and red-naped (*Sphyrapicus nuchalis*) sapsuckers: Fitness consequences, reproductive character displacement, and nest site selection. Northern Illinois University.

Walsh, P. S., Metzger, D. A., & Higuchi, R. (1991). Chelex-100 as a medium for simple extraction of DNA for PCR-Based typing from forensic material. BioTechniques, 10(4), 506–513.

Walters, E. L., Miller, E. H., & Lowther, P. E. (2002a). Red-breasted sapsucker (*Sphyrapicus ruber*) and red-naped sapsucker (*Sphyrapicus nuchalis*). Birds of North America, *663*, 1–32.

Walters, E. L., Miller, E. H., & Lowther, P. E. (2002b). Yellow-bellied sapsucker. Birds of North America, 662, 2–24.

Wang, S., Rohwer, S., de Zwaan, D. E., Toews, D. P. L., Lovette, I. J., Mackenzie, J., & Irwin, D. (2020). Selection on a small genomic region underpins differentiation in multiple color traits between two warbler species. Evol Letters, 4(6), 502-515.

Wiebe, K.L. (2000). Assortative mating by color in a population of hybrid northern flickers. Auk, 117(2), 518-522.

Zhou, H., Alexander, D., & Lange, K. (2011). A quasi-Newton acceleration for high-dimensional optimization algorithms. Statistics and Computing, 21(2), 261–273.

Tables.

Table 1 SNP assignments by locus and population (a). Sample sizes recorded as number of alleles (n). Comparisons among species, hybrids, and populations within Alberta (b). P-values from Fisher's exact tests on the left, bold indicates comparisons are significant ($p < 0.05$), values to the right of the p-values indicate number of alleles (n).

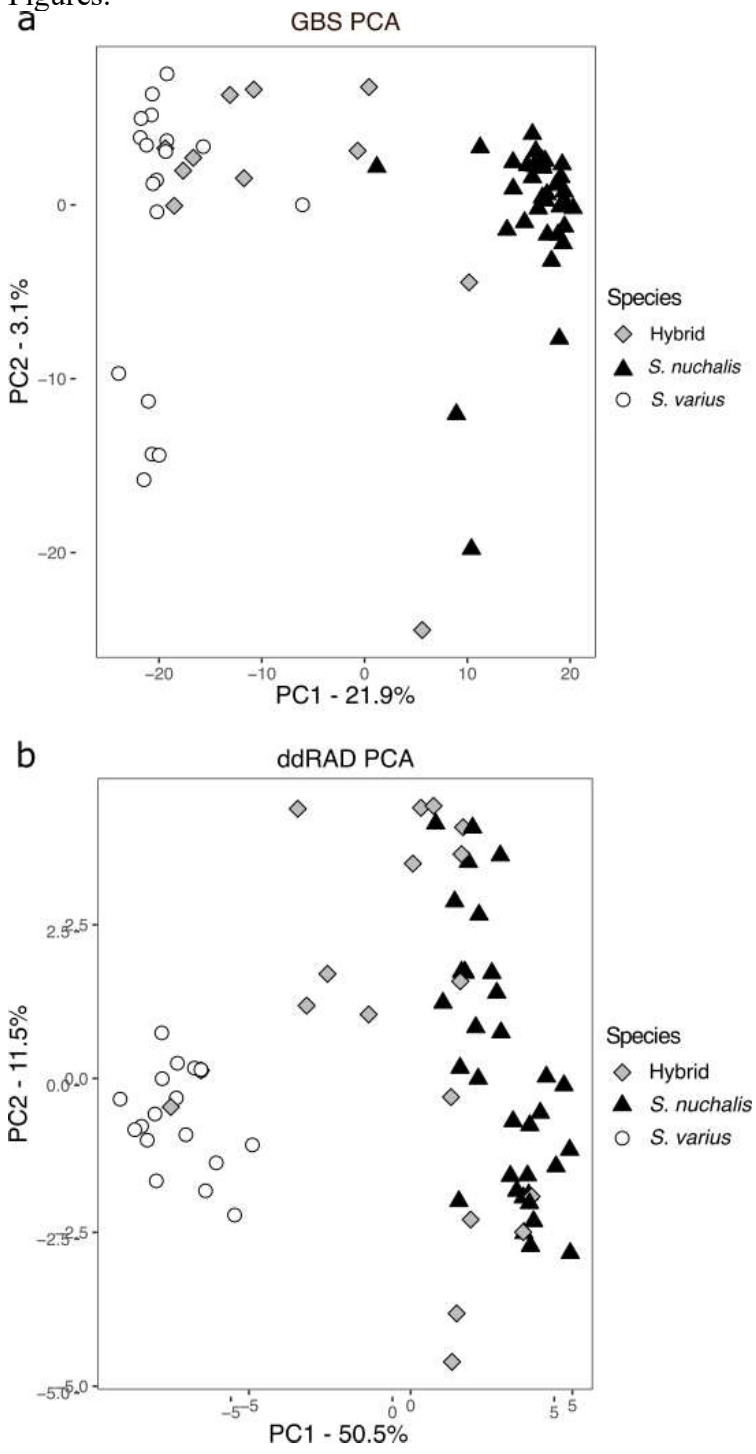
a.

	Enol			GAPD			anonymous nuclear marker		
	T	C	n	insertion	deletion	n	T	C	n
S. nuchalis NM	28	2	30	24	6	30	23	7	30
S. nuchalis ID	20	0	20	19	1	20	13	7	20
S. nuchalis AB	59	17	76	63	5	68	49	29	78
Hybrid	32	54	86	55	35	90	80	10	90
S. varius AB	34	50	84	45	31	76	91	21	112
S. varius SK	3	25	28	18	10	28	27	1	28
S. varius IL	1	25	26	21	9	30	29	1	30
S. varius NSNB	3	21	24	2	2	4	22	2	24
Total	180	194	374	247	99	346	334	78	412

b.

	Enol		GAPD		anonymous nuclear marker	
S. nuchalis x S. varius	**p < 0.0001**	288	**p < 0.0001**	256	**p < 0.0001**	322
AB *S. nuchalis x* AB *S. varius*	**p < 0.0001**	160	**p < 0.0001**	144	**p = 0.002**	190
AB *S. nuchalis x* hybrid	**p < 0.0001**	162	**p < 0.0001**	158	**p < 0.001**	168
AB *S. varius x* hybrid	0.6406	170	0.4271	166	0.3307	202

629 Figures.



630
631
632 Figure 1. Principal Component Analysis (PCA) of GBS dataset (a) and ddRAD dataset (b). PC1
633 distinguishes *S. nuchalis* from *S. varius* samples, hybrids are intermediate.
634

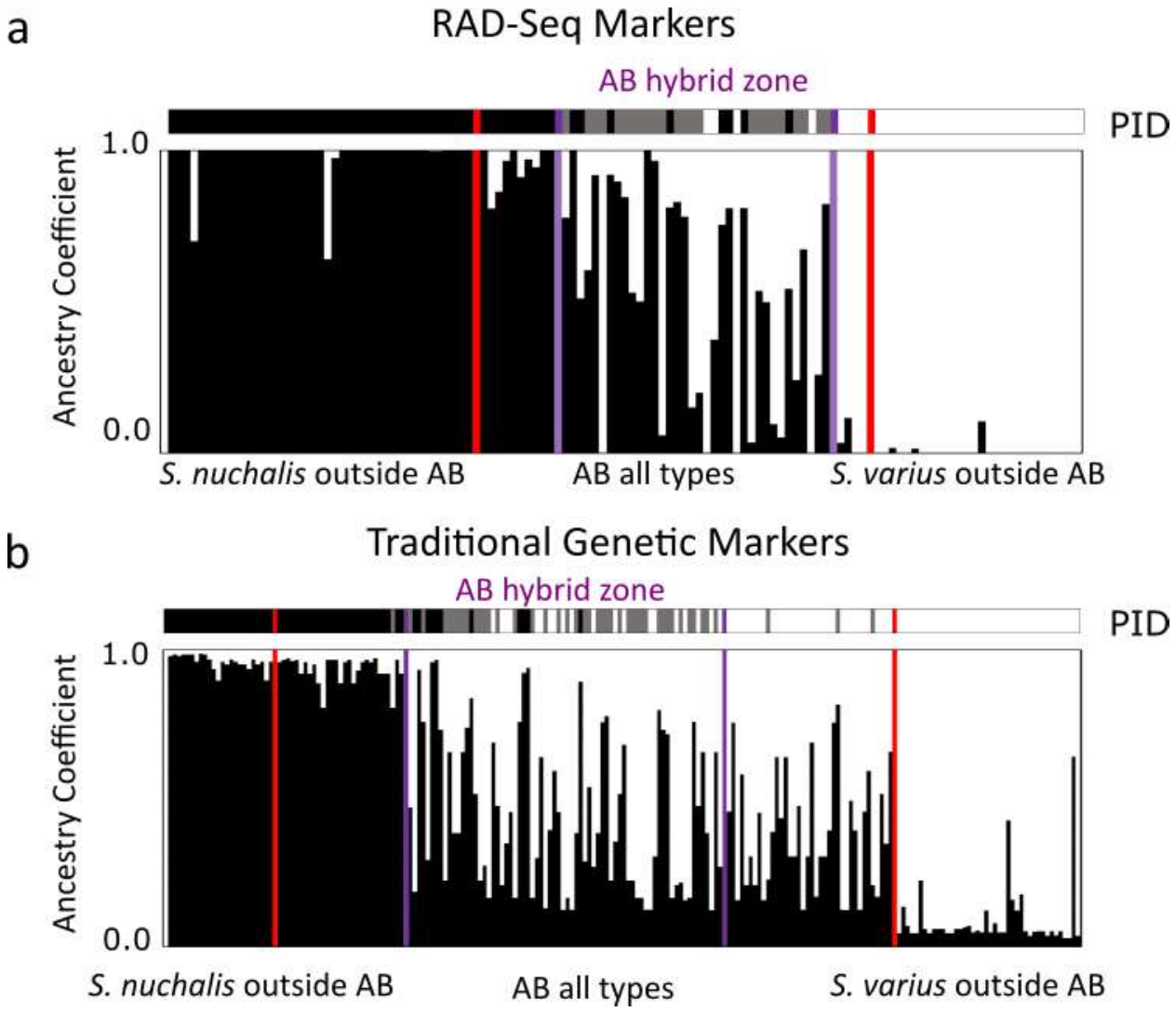


Figure 2. ADMIXTURE plot of the sapsuckers using RAD-Seq data (a) and STRUCTURE plot using traditional genetic markers (b) showing *S. nuchalis* ancestry (black) and *S. varius* ancestry (white). PID columns show phenotypic species identification for each individual in the STRUCTURE and ADMIXTURE plots, with black representing *S. nuchalis*, white for *S. varius*, and grey for hybrids. Red lines separate AB birds from birds sampled outside AB. The individuals within AB are organized south to north based on latitude. Violet separates birds trapped within the 275 km hybrid zone identified using RAD-Seq Q values. See Figure 4 and Supplemental Materials 3 and 4 for additional details on locations.

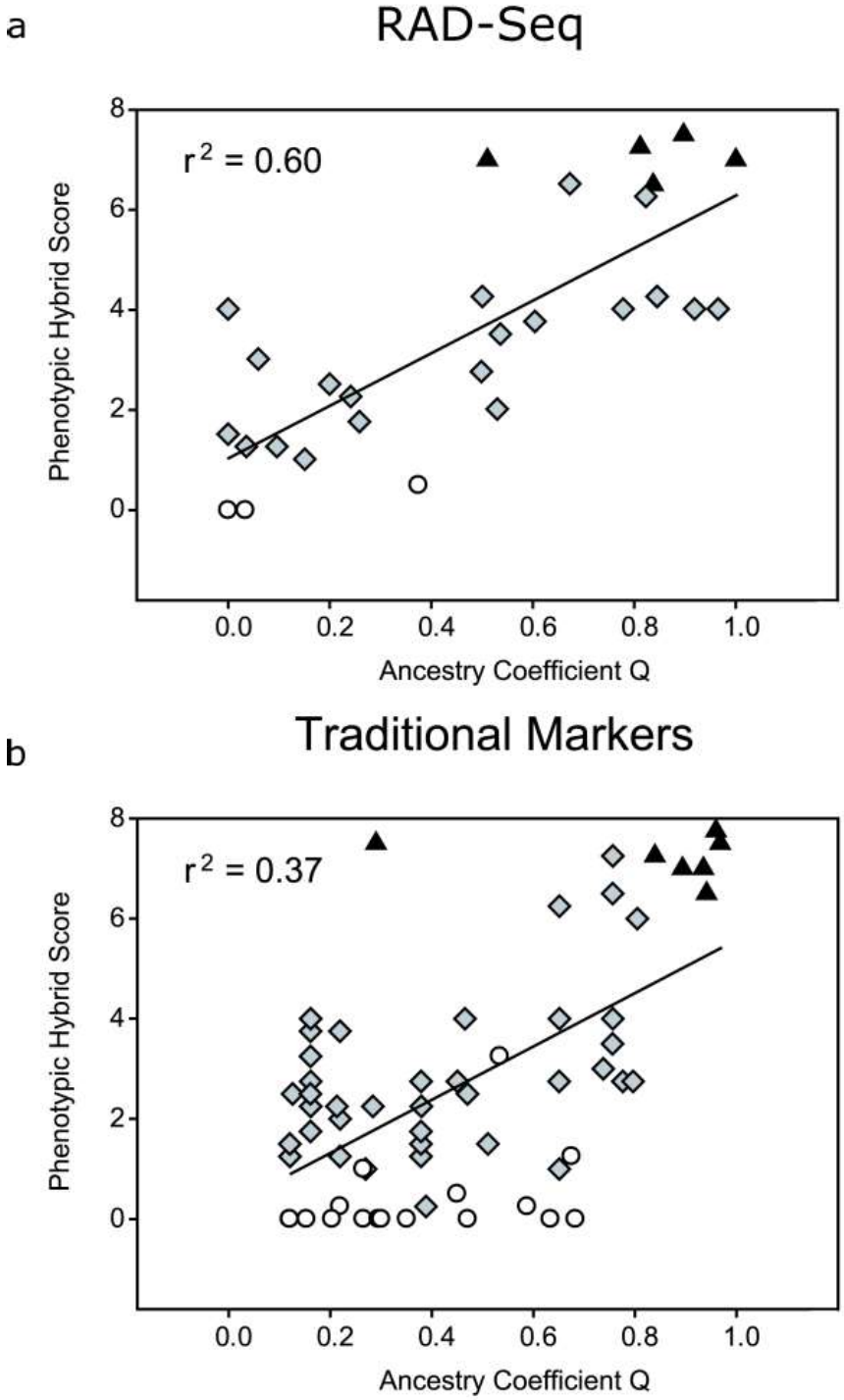
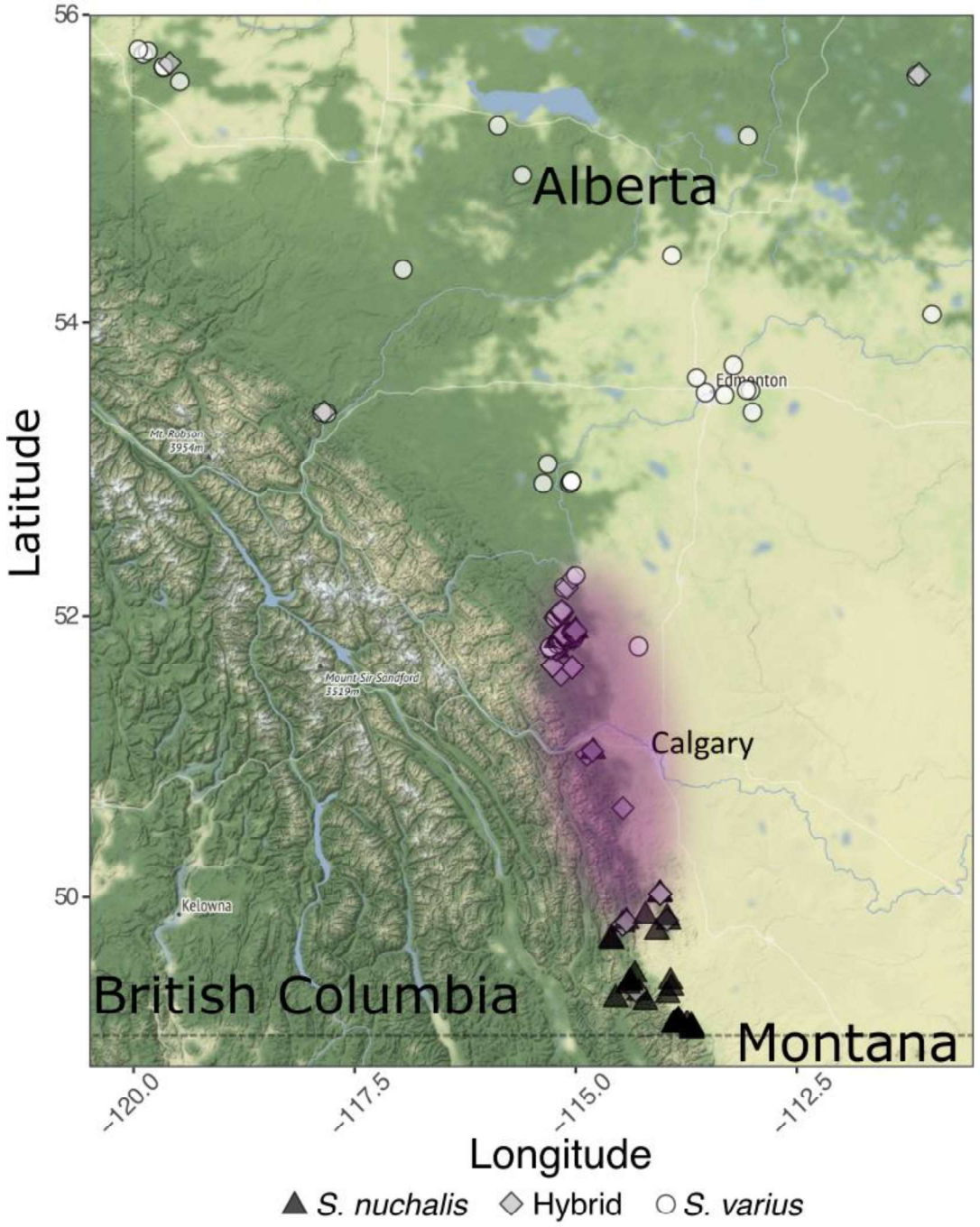


Figure 3. Relationship between the phenotypic hybrid scores of sapsuckers in the hybrid zone in west-central Alberta and the birds' ancestry coefficient (Q) as determined using RAD-Seq (a) or traditional genetic markers (b) data. Black triangles represent *S. nuchalis*, white circles represent *S. varius*, and grey diamonds are hybrids.

651



652
653
654 Figure 4. Sampling locations of individuals phenotypically identified as *S. nuchalis* (black
655 triangles), *S. varius* (white circles), and hybrids (grey diamonds). Violet indicates location of
656 hybrid zone identified using RAD-Seq loci.

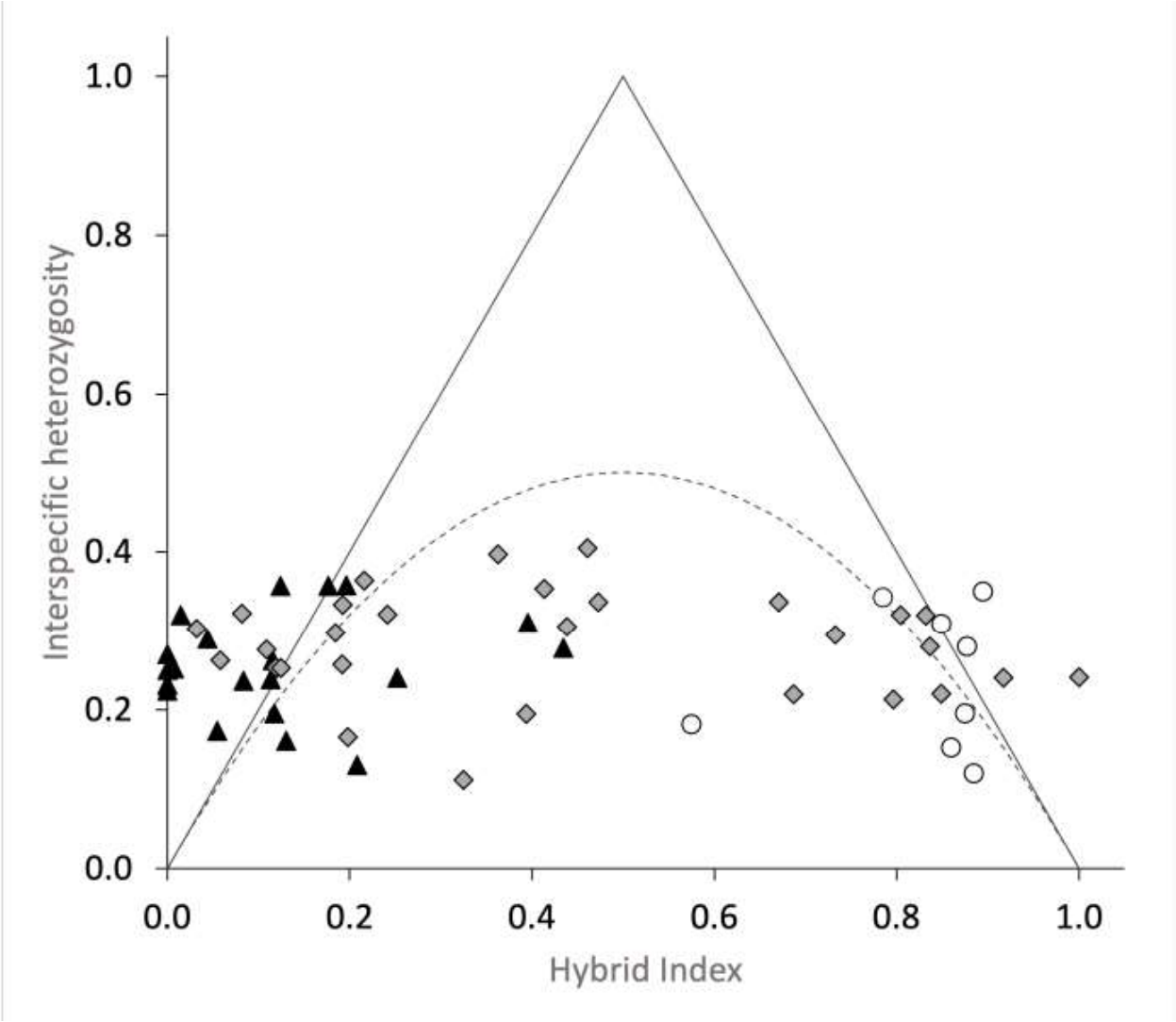


Figure 5. Interspecific heterozygosity and hybrid index (HINDEX) scores of each individual in AB for both GBS and ddRAD datasets combined. Symbols identify the forms sampled based on phenotype of the birds: pure *S. nuchalis* (black triangle), pure *S. varius* (white circles), and hybrids (grey diamonds). Individuals located on the solid line are back-crossed individuals. The dotted line indicates the expected distribution for individuals if mating was random throughout the zone.

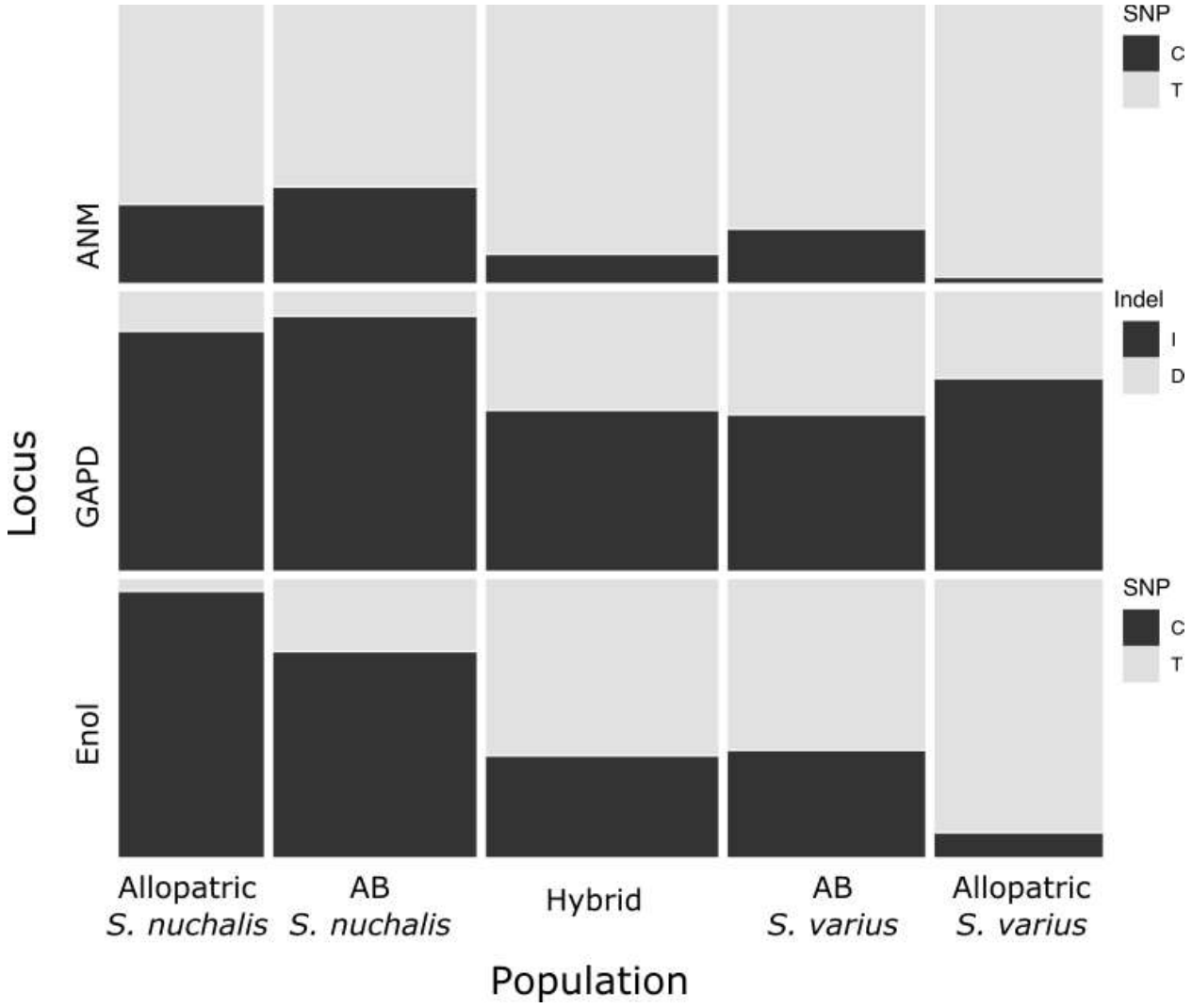


Figure 6. SNP assignments for Enol, GAPD, and anonymous nuclear marker for each population sampled. Sample sizes in Table 1.

	Enol			GAPD	
	T	C	n	insertion	deletion
S. nuchalis NM	28	2	30	24	6
S. nuchalis ID	20	0	20	19	1
S. nuchalis AB	59	17	76	63	5
Hybrid	32	54	86	55	35
S. varius AB	34	50	84	45	31
S. varius SK	3	25	28	18	10
S. varius IL	1	25	26	21	9
S. varius NSNB	3	21	24	2	2
Total	180	194	374	247	99

	anonymous nuclear marker		
n	T	C	n
30	23	7	30
20	13	7	20
68	49	29	78
90	80	10	90
76	91	21	112
28	27	1	28
30	29	1	30
4	22	2	24
346	334	78	412

	Enol		GAPD	
S. nuchalis x S. varius	**p < 0.0001**	288	**p < 0.0001**	258
AB *S. nuchalis x* AB *S. varius*	**p < 0.0001**	160	**p < 0.0001**	146
AB *S. nuchalis x* hybrid	**p < 0.0001**	160	**p < 0.0001**	156
AB *S. varius x* hybrid	0.9078	172	0.4271	166

anonymous nuclear marker	
$p < 0.0001$	322
$p = 0.002$	190
$p < 0.001$	166
0.3307	204

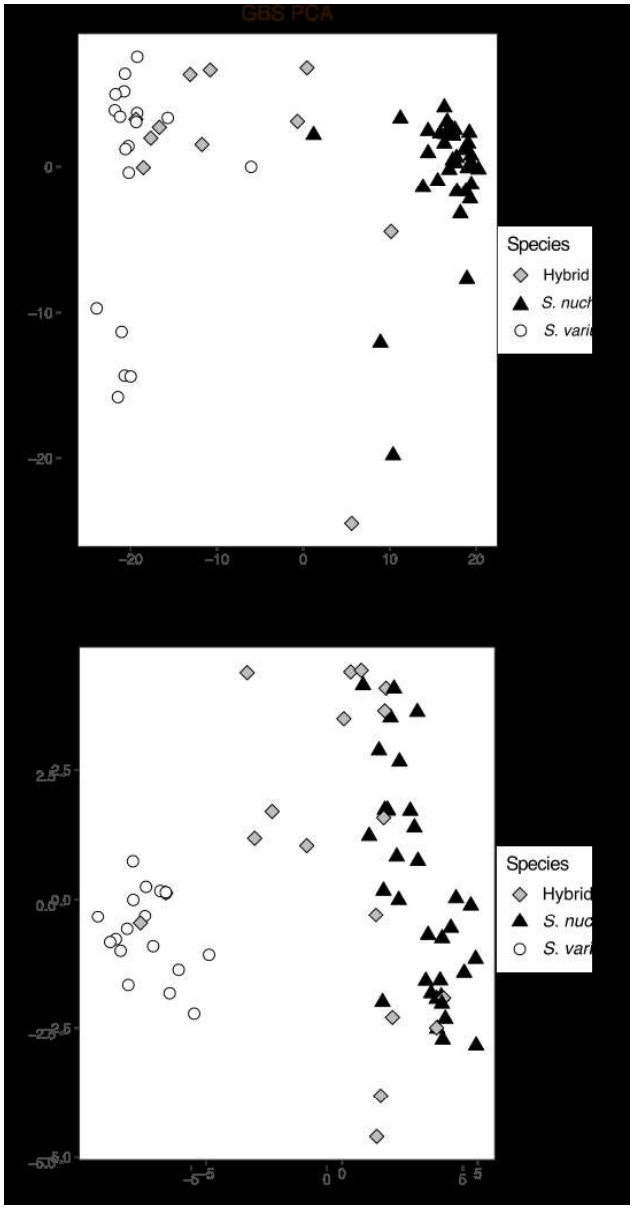


Figure 1. Principal Component Analysis (PCA) of GBS dataset (a) and ddRAD dataset (b). PC1 distinguishes *S. nuchalis* from *S. varius* samples, hybrids are intermediate.

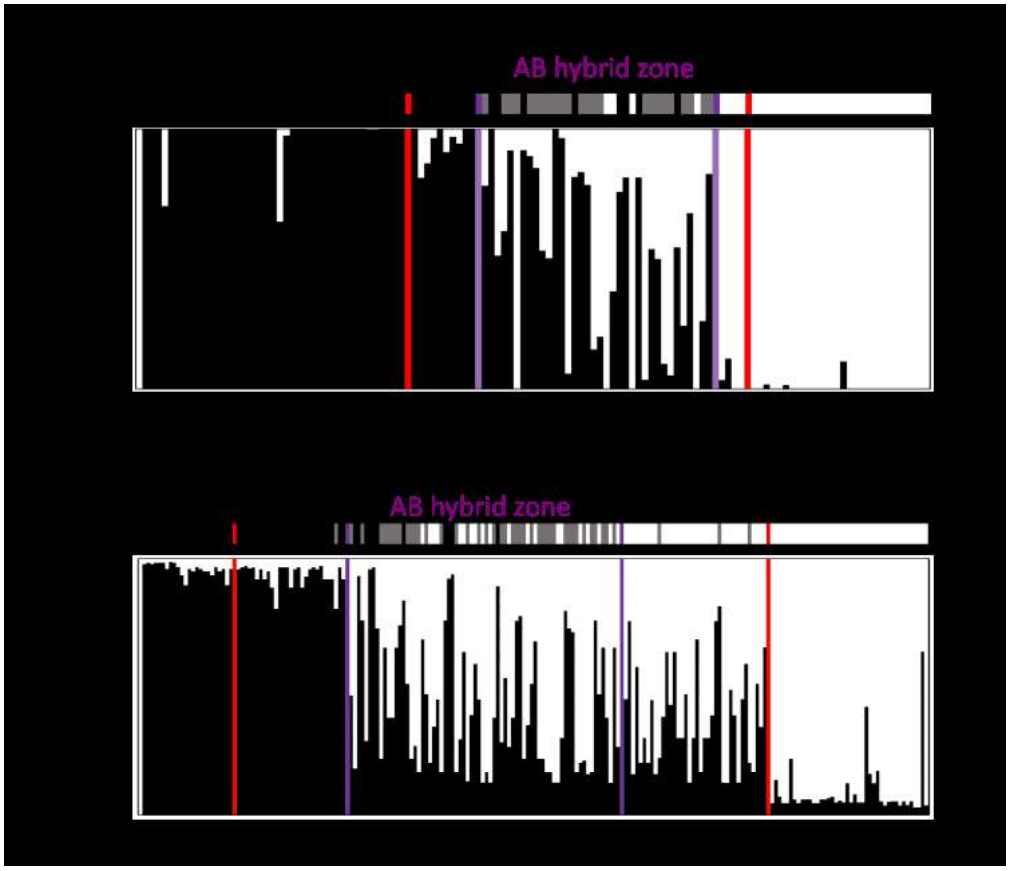


Figure 2. ADMIXTURE plot of the sapsuckers using RAD-Seq data (a) and STRUCTURE plot using traditional genetic markers (b) showing *S. nuchalis* ancestry (black) and *S. varius* ancestry (white). PID columns show phenotypic species identification for each individual in the STRUCTURE and ADMIXTURE plots, with black representing *S. nuchalis*, white for *S. varius*, and grey for hybrids. Red lines separate AB birds from birds sampled outside AB. The individuals within AB are organized south to north based on latitude. Violet separates birds trapped within the 275 km hybrid zone identified using RAD-Seq Q values. See Figure 4 and Supplemental Materials 3 and 4 for additional details on locations.

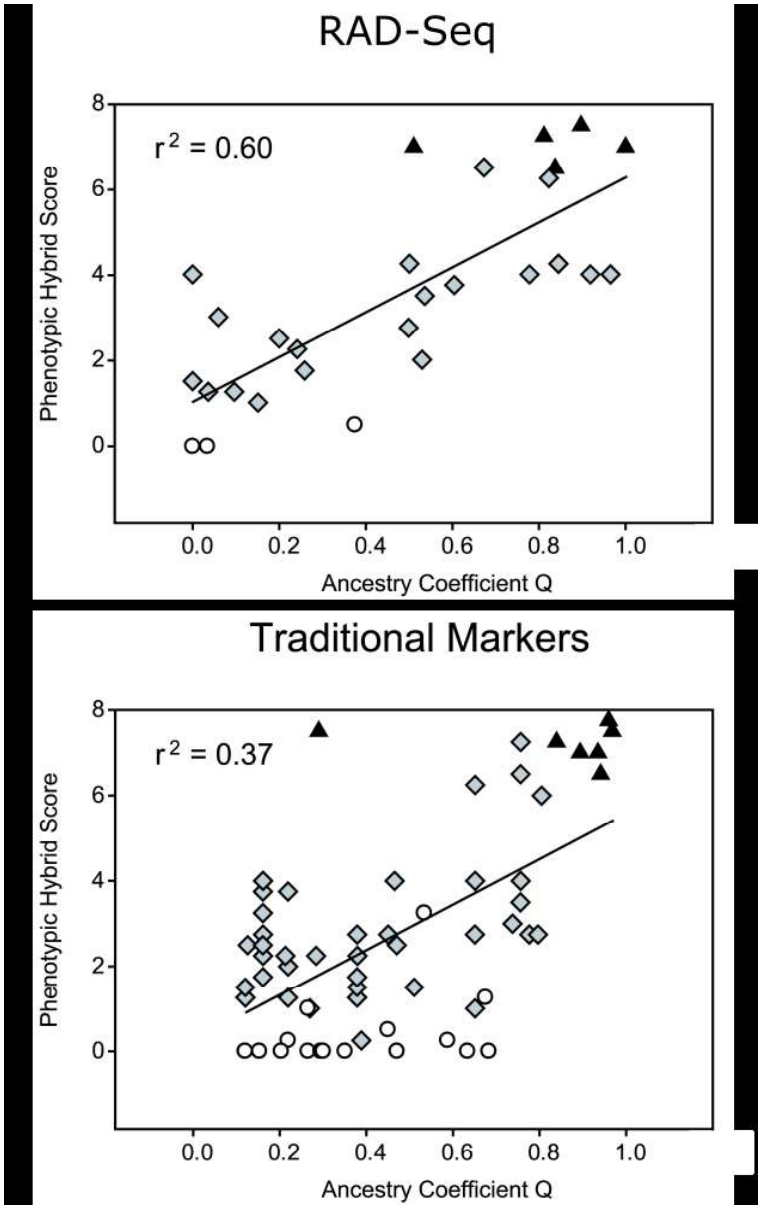


Figure 3. Relationship between the phenotypic hybrid scores of sapsuckers in the hybrid zone in west-central Alberta and the birds' ancestry coefficient (Q) as determined using RAD-Seq (a) or traditional genetic markers (b) data. Black triangles represent *S. nuchalis*, white circles represent *S. varius*, and grey diamonds are hybrids.

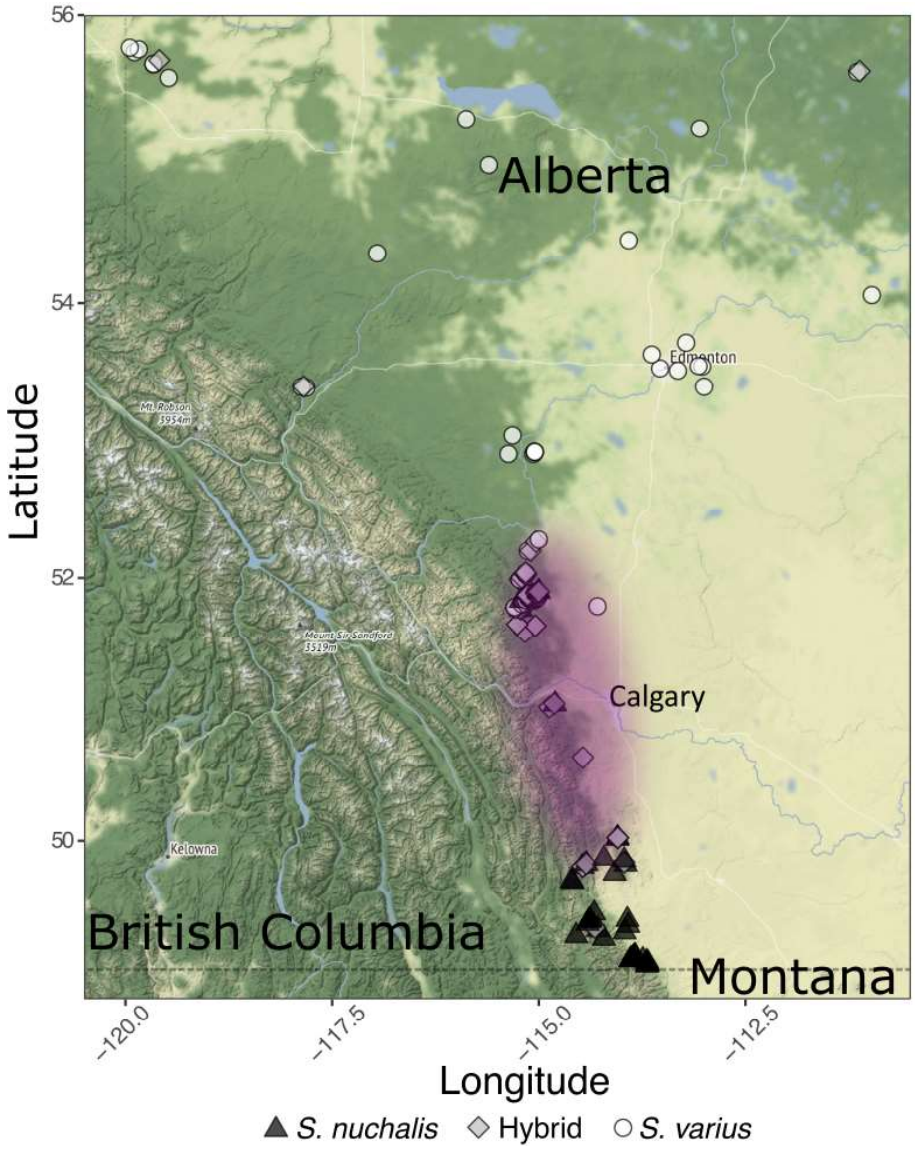


Figure 4. Sampling locations of individuals phenotypically identified as *S. nuchalis* (black triangles), *S. varius* (white circles), and hybrids (grey diamonds). Violet indicates location of hybrid zone identified using RAD-Seq loci.

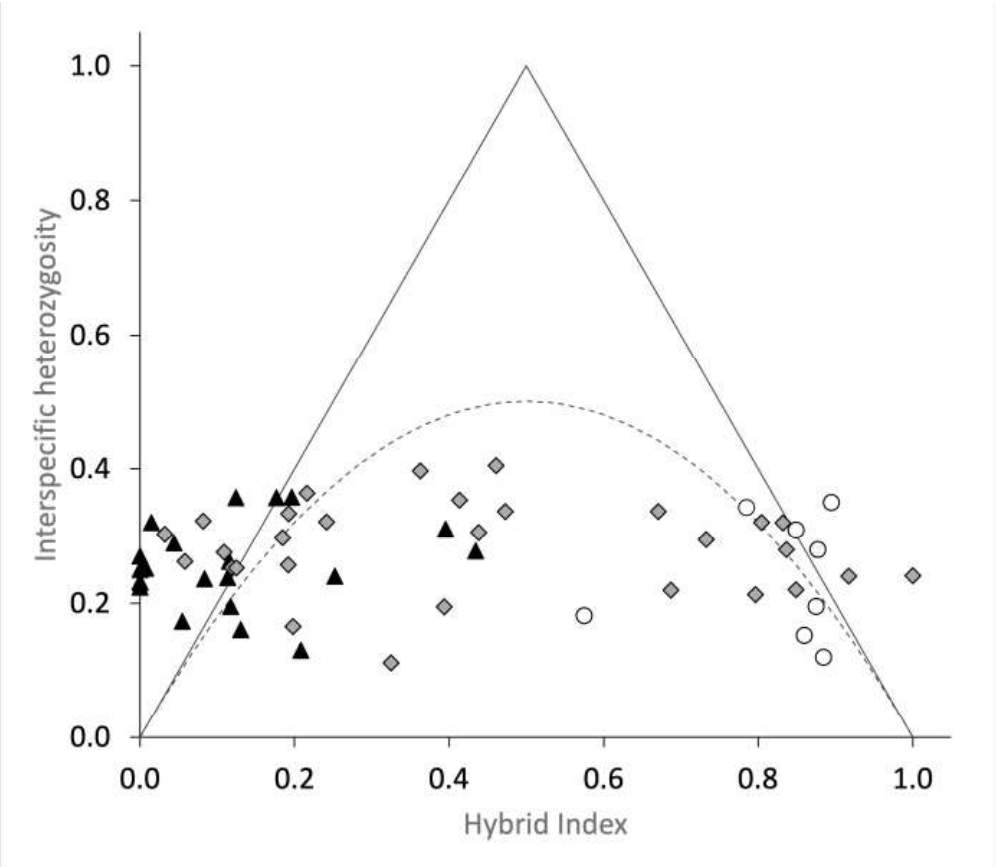


Figure 5. Interspecific heterozygosity and hybrid index (HINDEX) scores of each individual in AB for both GBS and ddRAD datasets combined. Symbols identify the forms sampled based on phenotype of the birds: pure *S. nuchalis* (black triangles), pure *S. varius* (white circles), and hybrids (grey diamonds). Individuals located on the solid line are back-crossed individuals. The dotted line indicates the expected distribution for individuals if mating was random throughout the zone.

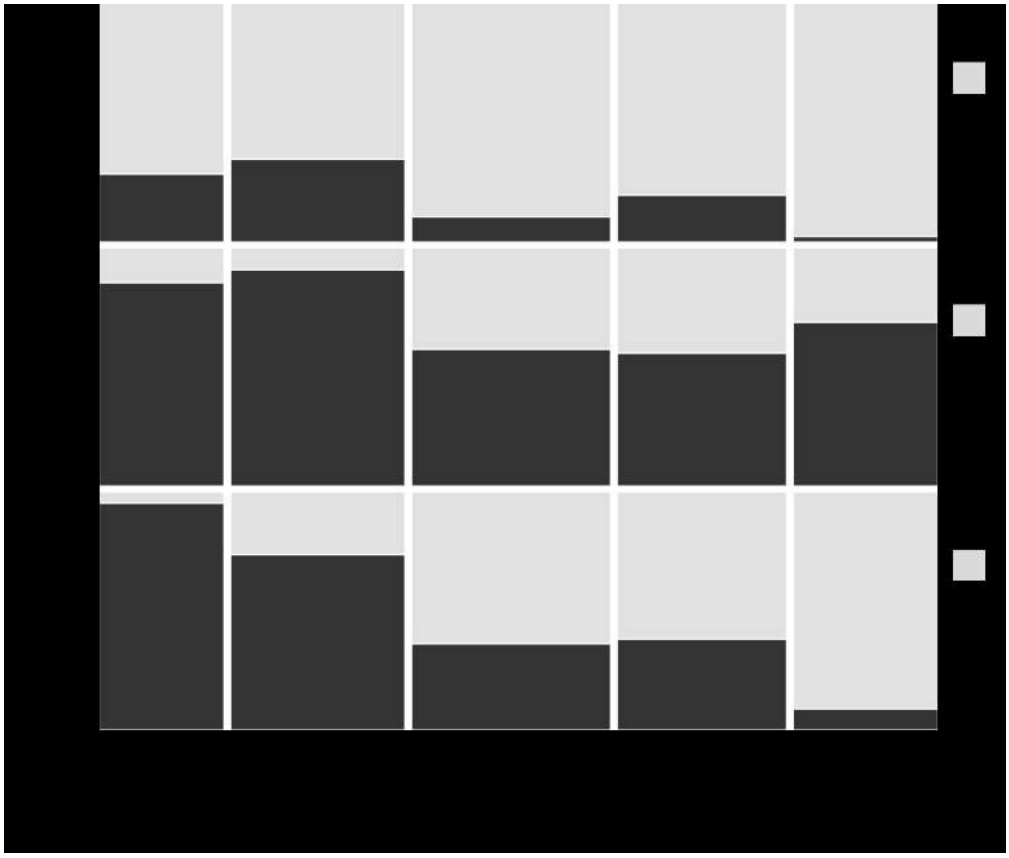


Figure 6. SNP assignments for Enol, GAPD, and anonymous nuclear marker for each population sampled. Sample sizes in Table 1.