Speciation without gene-flow in hybridising deer

Camille Kessler[1], Eric Wootton[2], Aaron B.A. Shafer[1]

[1]Environmental and Life Sciences Graduate Program, Trent University, 7K9J 7B8 Peterborough, Ontario, Canada

[2]Biochemistry & Molecular Biology, Trent University, K9J 7B8 Peterborough, Ontario, Canada

Corresponding authors:

Camille Kessler: camillekessler@trentu.ca

Aaron Shafer: aaronshafer@trentu.ca

Abstract

Under the ecological speciation model, divergent selection acts on ecological differences between populations, gradually creating barriers to gene flow and ultimately leading to reproductive isolation. Hybridisation is part of this continuum and can both promote and inhibit the speciation process. Here, we used white-tailed (*Odocoileus virginianus*) and mule deer (*O. hemionus*) to investigate patterns of speciation in hybridising sister species. We quantified genome-wide historical introgression and performed genome scans to look for signatures of four different selection scenarios. Despite ample modern evidence of hybridisation, we found negligible patterns of ancestral introgression and no signatures of divergence with gene flow, rather localised patterns of allopatric and balancing selection were detected across the genome. Genes under balancing selection were related to immunity, MHC and sensory perception of smell, the latter of which is consistent with deer biology. The deficiency of historical gene-flow suggests that white-tailed and mule deer were spatially separated during the glaciation cycles of the Pleistocene and genome wide differentiation accrued via genetic drift. Dobzhansky-Muller incompatibilities and selection against hybrids are hypothesised to be acting, and diversity correlations to recombination rates suggests these sister species are far along the speciation continuum.

Keywords: Speciation genomics, hybridisation, allopatric speciation, secondary contact, *Odocoileus*.

Introduction

Hybridisation is a widespread phenomenon that occurs at variable rates (Adavoudi & Pilot, 2021; Iacolina et al., 2019; Mallet, 2005; Ragavan et al., 2017; Taylor & Larson, 2019). The prevalence of hybridisation suggests speciation follows a continuum as opposed to the more discrete and allopatric view introduced by Mayr (1942) (Mallet, 2005; Stankowski & Ravinet, 2021). However, both ideas are not incompatible as the speciation continuum ultimately leads to reproductive isolation with hybridisation being a natural outcome (Stankowski & Ravinet, 2021). Under the ecological speciation model (sensu Darwin, 1859), hybrids are common and can even facilitate speciation (Nosil, 2012). Ecological speciation works when ecological differences lead to divergent selection between populations; the underlying genetic mechanisms (i.e. causative loci) gradually create barriers to gene flow that accumulate which leads to reproductive isolation (Nosil, 2012; Rundle & Nosil, 2005; Shafer & Wolf, 2013).

The genetic mechanisms involved in ecological speciation can take the form of barrier loci which act as a restraint to gene flow in different ways: for example, such loci can be under divergent selection, involved in mate choice, contribute to assortative mating, or reduce hybrid fitness (Abbott et al., 2013; Ravinet et al., 2017). Divergent selection should produce genomic islands in this context, also referred to as genomic islands of speciation, that are regions where the differentiation (e.g. F_{ST}) is higher than the neutral genomic background (Campbell et al., 2018; Ravinet et al., 2017). Such loci and islands have been observed in several species underpinning a wide array of speciation and divergence processes (Lavretsky et al., 2019; Marques et al., 2019; Momigliano et al., 2017; Poelstra et al., 2014). Scanning the genome for just F_{ST} peaks when speciation with gene flow is ongoing, however, is problematic as F_{ST} is dependent on genetic diversity (π), and other mechanisms can create similar F_{ST} profiles such

as genetic drift, global adaptation, or simply reduced genetic diversity due to background selection (Booker et al., 2021; Burri, 2017; Cruickshank & Hahn, 2014; Ravinet et al., 2017). Joint genome scans including F_{ST}, d_{xy} and π facilitate a better depiction of the processes of selection and ecological speciation, particularly the use of d_{xy} which is not influenced by current levels of diversity (Campbell et al., 2018; Cruickshank & Hahn, 2014; Ravinet et al., 2017; Shang et al., 2021). Shang et al. (2021) recently conceptualised the expected patterns F_{ST}, d_{xy} and π under four main modes of selection: divergence with gene flow, allopatric selection (i.e., positive selection post species split), balancing selection and background selection. Using pairs of *Populus* species along the speciation continuum, they investigated the genomic landscape to test clear predictions on the behaviour of F_{ST}, d_{xy} and π. Including correlations of π and F_{ST} to recombination rate (ρ) can help place a species on the continuum, with for example the correlation between F_{ST} and ρ is expected to become stronger with increasing divergence (Burri, 2017; Shang et al., 2021). This approach ultimately allows for disentangling selection patterns from a neutral background based on those statistics, and helps decipher the evolutionary history of species.

White-tailed (*Odocoileus virginianus*; WTD) and mule deer (*O. hemionus*; MD) are abundant in North America with similar morphology, activity patterns and life-history traits, but they differ in several ecological aspects (Douzery & Randi, 1997; Gilbert et al., 2006; Pitra et al., 2004; Brunjes et al., 2006; Berry et al., 2019). Mule deer favour food availability whereas WTD prioritise security and thermoregulation (Whittaker & Lindzey, 2004); consequently, WTD prefer habitats at lower altitude and with denser visual cover than MD that favour open areas at higher elevation (Anthony & Smith, 1977; Brunjes et al., 2006). White-tailed deer and MD also differ in their sociality and associated predator response. Mule deer live in large cohesive groups

78 including both sexes, and, as a group, adopt an aggressive behaviour in presence of predators,
79 whereas WTD live in smaller female-biased groups and flee in response to predators (Lingle,
80 2001, 2003). Both species represent a high economic value in North America as hunting-related
81 activities generate billions of dollars annually (Cambronne, 2013), and both species are an
82 important cultural component of Indigenous communities (Adams & Hamilton, 2011; Peres &
83 Altman, 2018).

84 Despite a species divergence date estimated at ~3.13 mya (Wright et al., 2022), WTD and
85 MD hybridise in areas of sympatry with estimated hybridisation rates ranging from 1 to 19%
86 depending on the region (Carr & Hughes, 1993; Combe et al., 2021; Hornbeck & Mahoney, 2000;
87 Russell et al., 2021). Both species likely spent considerable time in allopatry (i.e. separate
88 refugia) during glacial events (e.g Moscarella et al. 2003; Latch et al. 2009). Recent hybrid zone
89 analyses mainly found backcrossed individuals rather than F1 hybrids (Combe et al., 2021;
90 Russell et al., 2021), and an interspecies F_{ST} up to 0.4 (Combe et al., 2021). It has also been
91 shown that gene flow appears restricted but bidirectional, and there are signs of introgression
92 at both mitochondrial and nuclear levels (Bradley et al., 2003; Carr et al., 1986; Carr & Hughes,
93 1993; Cathey et al., 1998; Cronin, 1991; Derr, 1991; Russell et al., 2019), such that some MD
94 acquired WTD mitochondrial DNA around ~1.32 Mya (Wright et al., 2022). This pattern would
95 suggest ancestral hybridisation and gene flow has taken place. Given their clear behavioural
96 differences, and their documented hybridisation, white-tailed and mule deer do not present
97 the reproductive isolation required by Mayr's biological species concept (Mayr, 1942). Here, we
98 hypothesised that WTD and MD have evolved via ecological speciation, and more specifically
99 a speciation with gene flow scenario. To test this hypothesis, we quantified genome-wide
100 introgression and past admixture events and performed genome scans to look for signatures

101 of four different divergence scenarios. We expected to find genetic signs of divergent selection,
102 such as speciation islands, consistent with patterns of divergence with gene flow (DwGF), and
103 a stronger F_{ST} - ρ correlation between species compared to within. We also predicted higher
104 rates of admixture and historical introgression in areas of sympatry.

105 ## *Materials and Methods*

106 ### Sampling & Sequencing

107 We obtained tissue samples from harvested deer collected across the range of WTD & MD
108 including areas of sympatry and allopatry (Fig. 1A, Table 1). These areas were determined from
109 NatureServe records and adjusted with IUCN range data (IUCN, 2015a, 2015b; NatureServe,
110 2021a, 2021b). Specifically, the state of Washington and the province of British Columbia, for
111 which the coasts are largely MD allopatric whereas the eastern parts are sympatric, were
112 divided in two by the 120th meridian West. We extracted DNA from tissue using the Qiagen
113 DNeasy Blood and Tissue Kit following manufacturer's instructions and checked the samples
114 concentration using Invitrogen Qubit assays. WGS libraries were generated at The Centre for
115 Applied Genomics in Toronto, Canada and sequenced to an average of 4x coverage on an
116 Illumina HiSeqX.

117 ### Data processing

118 Raw reads quality was examined using FastQC (v0.11.9; Andrews S., 2010). We trimmed the
119 reads using Trimmomatic (v0.36; Bolger et al., 2014), and aligned them to the WTD genome
120 (GCA_014726795.1) with bwa-mem (v0.7.17; Li & Durbin, 2009). We sorted the reads using
121 SAMtools sort (v1.10; Li et al., 2009), identified duplicates reads with Picard MarkDuplicates
122 (v2.23.2; Broad Institute, GitHub Repository., 2019) and removed them using Sambamba view
123 (v0.7.0; Tarasov et al., 2015). We performed a local re-alignment using GATK

124 RealignerTargetCreator and IndelRealigner (v4.1.7.0; McKenna et al., 2010). For further quality
125 checks, we used Sambamba flagstat, mosdepth (v0.3.1; Pedersen and Quinlan 2018) and
126 MultiQC (v1.10; Ewels et al., 2016). For the MSMC analysis, some samples were sequenced to a
127 higher depth on multiple lanes. For those, before the duplicate removal step, the read groups
128 were specified and both sequencing files merged using Picard AddOrReplaceReadGroups and
129 MergeSamFiles. The rest of the pipeline was performed as described above with the addition
130 of a final step: to be able to carry out other analyses with a coverage consistent across samples,
131 we reduced the final coverage to the average coverage of all other samples (4x) using Picard
132 DownsampleSam.

133 We created three datasets for our different analyses. The first contained the variants
134 among all deer individuals (hereafter called DeerSNP dataset). We produced it using ANGSD
135 (v0.918; Korneliussen et al., 2014) and estimated genotype likelihoods following the GATK
136 model (-gl 2) and called genotypes (-doGeno 4) with the following filtering: SNPs with a
137 minimum p-value of $1e^{-6}$, a minimum base quality of 20 and Minimum mapQ quality of 20 (-
138 SNP_pval 1e-6, -minMapQ 20 & -minQ 20). The second dataset additionally included invariant
139 sites (referred to as DeerALL), it was produced following the same method as DeerSNP except
140 for the filtering on the SNP p-value (-SNP_pval 1e-6) which was removed. Our third dataset was
141 designed for an ABBA-BABA analysis which requires samples from an additional species as
142 outgroup. It comprises the 28 deer individuals and one caribou sample, all mapped to the
143 caribou genome (GCA_014898785.1; afterwards called DeerCAR dataset). For this dataset, we
144 first sorted the deer bam files by read name and converted them to fastq using SAMtools sort
145 and fastq respectively. We mapped these fastq to the caribou reference genome using bwa-

146 mem. We ran ANGSD using the same procedure as for DeerSNP dataset, with the addition of a
147 caribou genome sequence in bam format, obtained for this analysis from Dedato et al. (2022).

148 **Population structure analyses and cross-species coalescent rate**

149 To infer admixture proportions and compute a PCA on allele frequencies based on genotype
150 likelihood data we used PCAngsd (Meisner & Albrechtsen, 2018) and NGSadmix (Skotte et al.,
151 2013) on the DeerSNP dataset. NGSadmix was run with K = 1 up to K =7 and the best K value
152 was determined with CLUMPAK (Kopelman et al., 2015) following the Evanno method (Evanno
153 et al., 2005) with 7 replicates for each K value. We investigated the correlation between the
154 distribution on each PC and each individual's range in R using the package "stats" with a linear
155 model including the species and either latitude or longitude, and a Pearson's correlation (e.g.
156 lm(PC ~ Species + Latitude)).

157 For two samples with high coverage (Ov_ON6 at 18x - and Oh_WA1 at 22x) we
158 implemented the multiple sequentially Markovian coalescent (MSMC) to infer the cross-
159 coalescent rate. To do this we first generated a 35-mer mappability mask file using the SNPable
160 pipeline as is required for MSMC (Schiffels & Wang, 2020). We restricted analyses to scaffolds
161 >500 Kbp (Gower et al., 2018) and phased samples using whatsHap 1.3 (Martin et al., 2016) prior
162 to generating the MSMC input files. We ran MSMC2 using the time segment pattern
163 1×2+25×1+1×2+1×3 and estimated the cross-coalescence rate (CCR) across populations by
164 comparing the first haplotype from each sample in a pairwise fashion using the -P flag and
165 skipping any sites with ambiguous phasing. To quantify the migration rate over time m(t) and
166 the cumulative migration probability M(t) – which is related to the CCR (Schiffels & Durbin,
167 2014) -– we implemented the IM model to the MSMC output (Wang et al., 2020). We assumed

168 a generation time of 2 years (Demarais et al., 2000; Deyoung et al., 2003) and mutation rate of
169 1.33×10-8 mutations/site/generation.

170 Genome-wide ancestral introgression

171 To evaluate the extent of historical introgression between sympatric MD and WTD, we used the
172 ABBA-BABA test, which allows for the detection of introgression between three populations P1,
173 P2 & P3 and an outgroup (Durand et al., 2011; Green et al., 2010). We ran the ABBA-BABA
174 analysis implemented in ANGSD on the DeerCAR dataset between individuals (-doAbbababa
175 1) and between populations (-doAbbababa2 1; Soraggi et al. 2018), using the caribou as
176 outgroup. The D-statistic, standard error and Z-score were computed with ANGSD's
177 accompanying R scripts: jackKnife.R and estAvgError.R. To further our understanding of the
178 admixture events and build a maximum likelihood tree of our system, we used treemix (v1.13;
179 Pickrell & Pritchard, 2012) on both DeerCAR and DeerSNP datasets. We constructed the
180 maximum likelihood trees with migration events ranging from 0 to 5, either WTD_allopatry or
181 Caribou as root, and accounted for linkage using 1000 SNPs per block. Tree and residuals
182 visualisations were performed with associated R script plotting_funcs.R and the variance
183 explained by each migration event was computed with the get_f() function.

184 In-windows historical introgression

185 The D statistic is sensitive to genomic variation (π), and should not be used to determine
186 introgression on a small scale (Martin et al., 2015). The f_d statistic, proposed by Martin et al.
187 (2015), is less dependent on diversity than D and therefore allows for the detection of
188 potentially introgressed regions of the genome. We used the python script
189 ABBABABAwindows.py (Martin, 2021) to estimate in D and f_d in 50Kbp windows to detect
190 potentially introgressed loci in our DeerCAR dataset on two comparisons: 1) P1 = WTD

191 allopatry, P2 = WTD sympatry, P3 = MD sympatry and 2) P1 = MD allopatry, P2 = MD sympatry,
192 P3 = WTD sympatry, both with the caribou as outgroup. We then identified potentially
193 introgressed windows as those having a f_d value higher than the 97.5% quantile. Their position
194 was used in BEDTools intersect (v2.30.0; Quinlan and Hall 2010) with the caribou annotation
195 file to identify the genes present in putatively introgressed windows.

196 Genome scans

197 To detect islands of divergence between WTD & MD, we used the python script
198 popgenWindows.py (Martin, 2021) to estimate individual heterozygosity, F_{ST}, D_{xy} and π in 50Kbp
199 windows (note we filtered out scaffolds with less than 2 windows). Since the accurate
200 computation of D_{xy} and π require a dataset including the invariant sites, we performed this
201 analysis on the DeerALL dataset. To detect the outlier loci, we based our approach on the four
202 selection scenarios developed in (Shang et al., 2021): I) Divergence with gene flow that
203 presents as high F_{ST} and D_{xy} but low π; II) allopatric selection that shows high F_{ST}, low π and
204 stable D_{xy}; III) background selection that presents as high F_{ST} but low π and D_{xy}; and IV)
205 balancing selection that displays as low F_{ST} but high π and D_{xy}. Thresholds were set as the upper
206 or lower 5% for a high or low criteria, and between 45 and 55% for a stable condition. Outliers
207 were flagged as belonging to one of the four scenarios when they met the specific criteria
208 shown in Table 2.

209 We identified genes present in outlier regions by comparing the coordinates of the
210 outlier windows and the WTD annotation file in BEDTools intersect. We extracted the gene
211 names from those regions as well as those from their associated scaffolds. The gene list from
212 outlier windows was uploaded to ShinyGO for enrichment analysis with the scaffold's gene list
213 as background information and no species specification. We downloaded the enrichment

214 analysis results for the molecular function, biological process, and cellular component GOs
215 and analysed the results in R. We also uploaded the gene list to UniProt's Retrieve/ID mapping
216 tool (Pundir et al., 2016) and downloaded the information for visual inspection.

217 To estimate the position of WTD & MD along the speciation continuum through the
218 correlation between recombination rate (ρ), F_{ST}, and π, we computed ρ in windows using
219 FastEPRR (v.2.0; Gao et al., 2016). Here, we used our four categories as a proxy for a speciation
220 continuum with six comparisons: WTDa - WTDs and MDa - MDs representing no speciation,
221 WTDs - MDs, WTDs - MDa and WTDa - MDs as intermediate stages and WTDa - MDa as full
222 differentiation. We first phased the vcf using SHAPEIT (Delaneau et al., 2012) and split the
223 resulting file into scaffolds. We then ran FastEPRR separately on our four populations to obtain
224 the ρ estimation in 50kbp windows (winLength = 50000) on the 15 longest scaffolds of the WTD
225 genome, with 100 repetitions. To be able to correlate ρ with F_{ST} and π, we re-ran the genome
226 scan analysis on the population level. FastEPRR outputs ρ in $4N_er$, to convert this value to ρ in
227 cM/Mb, we computed N_e per population using the relationship $\pi = 4N_e\mu$. We used μ= 1.33×10-
228 8 and the genome scans result to average π across windows in each population. After filtering
229 to keep only windows that were present across all comparisons (12,500 windows), we
230 generated Pearson's correlations between I) ρ in both populations (ρ_1 vs ρ_2), II) average ρ
231 between populations and F_{ST} ($\rho_{1\text{-}2}$ vs F_{ST}), and III) average ρ and average π between populations
232 ($\rho_{1\text{-}2}$ vs $\pi_{1\text{-}2}$) following Shang et al. (2021). The resulting correlation coefficients were compared
233 between all comparisons to infer trends over divergence time. All results were analysed in R
234 version 4.1.0 "Camp Pontanezen" (R Core Team, 2021) .

235 ***Results***

236 We sequenced 28 individuals to an average coverage of approximately 4x. We called
237 103,970,889 SNPs in the DeerSNP dataset, which represented 4% of the genome. In the PCA
238 based on allele frequencies, PC1 accounted for 59% of the variation and separated both
239 species (Fig. 1B). PC2 explained an additional 8% of the variation and showed a spread of MD
240 individuals that is consistent with their longitudinal distribution (Pearson correlation = 0.8; *p-*
241 *value* = 0.00057). Latitude was not important in shaping PC1 or PC2 variation (Pearson
242 correlation *p-value* > 0.05). Admixture analysis showed a well-defined genetic clustering
243 between species (Evanno's method: best K = 2, Fig. 1C, Fig. S1). The WTD individual Ov_MX2
244 was partially admixed in both analyses, with NGSAdmix assigning it 92.5% WTD and 7.44% MD
245 ancestry (Fig. 1C). This individuals' admixed pattern disappeared in the admixture analysis with
246 higher K values (K = 3 to 7, Fig. S1), though additional within species clusters are observed.
247 The MSMC2-IM model showed no support for migration between species once they split, with
248 the CCR suggesting the speciation event for *Odocoileus* took place just under 1 Mya (Fig. 2).
249 To detect historical gene flow, we performed an ABBA-BABA test based on genotype
250 likelihoods between all combinations of our 28 individuals in a first analysis and between all
251 four populations in a second, with the caribou as outgroup (DeerCAR dataset, 72,438,766
252 SNPs). We computed 9828 ABBA-BABA individual topologies in ANGSD, of which 30 presented
253 signatures of introgression. We found six potential inter-species introgression events with a D-
254 statistic ranging from 0.011 to -0.013, two of which are significant (Z-score > |3| (Barlow et al.,
255 2021; Kirch et al., 2021), Fig. 3A). D-statistic in intra-species comparisons range from 0.25 to
256 0.916 in ABBA and from -0.046 to -0.914 in BABA, all intra-species comparison show a significant
257 Z-score (Fig. 3A). In the comparisons between the four populations, three ABBA-BABA
258 topologies presented signatures of introgression on a total of 12. The single inter-species

259 comparison presents a D-statistic of -0.07 (Z-score = -7.66) for an introgression between
260 sympatric MD and WTD. The other two comparisons show a reciprocal introgression between
261 MD in allopatry and sympatry of 0.9 and 0.89 (Z-score = 1342.83 & 1376.69; Fig. S3). The treemix
262 analyses showed a topology with no migration that explained over 99% of the variation (Fig.
263 3B). We further analysed historical introgression through the measure of f_d across 50kb non-
264 overlapping windows in two different introgression scenarios: 1) WTDa, WTDs, MDs, Caribou
265 and 2) MDa, MDs, WTDs, Caribou (Fig. S4). We found 1243 outlier windows with a f_d higher than
266 the 97.5% percentile in the first scenario and 1248 in the second, each representing ~2.45% of
267 the genome with 68 windows overlapping between comparisons (Fig. S4).

268 Measures of relative genomic divergence between WTD and MD were elevated (F_{ST} =
269 0.26) and absolute divergence d_{xy} was 0.011; genetic diversity (π) was higher in WTD than in MD
270 (WTD = 0.008, MD = 0.004, Mann-Whitney test: $p < 2.2e-16$). We identified a total of 1183 windows
271 presenting patterns consistent with one of the four selection scenarios (Fig. 4, Fig. S5). Of those,
272 1016 suggest a pattern of balancing selection, they were distributed across 236 scaffolds and
273 represent 1.99% of the genome. Those windows contained 121 genes, some of which showed
274 an enrichment in ontologies associated with the sense of smell and chemical stimuli, including
275 three categories presenting an enrichment above 15-fold (Fig. S6). We also detected GOs
276 related to the MHC and immunity, these include three categories with an enrichment above
277 24-fold (Fig. S6). We identified patterns of background selection in 165 windows across 58
278 scaffolds, representing 0.32% of the genome. These windows harboured 208 genes identified
279 with a UniProt ID and for which we found enrichment in GOs related to epigenetic factors such
280 as "Unmethylated CpG binding" (25-fold) or "Histone demethylation" (15-fold) (Fig. S7). These
281 windows were either isolated or clustered together into putative islands of divergence (Fig. S5).

282 We found 2 windows under allopatric selection, each window containing one gene: ACAP2, a
283 GTPase activating protein, and PCDHB4 potentially involved in cell-binding (UniProt, 2022a,
284 2022b). When we applied a more liberal cut-offs for F_{ST}, d_{xy}, and π we still observed no evidence
285 for divergence with gene flow (Table S1). The correlations between π, F_{ST} and ρ are highly
286 significant in all six comparisons (Fig. S8). As predicted correlations coefficients for F_{ST} - ρ
287 became stronger with divergent comparisons (Fig. 5). Likewise, the π - ρ correlation coefficients
288 were constant over divergence time, while the recombination landscape (ρ - ρ) differed
289 between species (Fig. 5).

290 ## ***Discussion***

291 We investigated the speciation history of the *Odocoileus* genera through introgression analyses
292 and genome scans. Our results suggest speciation between WTD & MD took place with
293 negligible historical gene flow, despite contemporary and historical evidence of interbreeding
294 (Cathey et al., 1998; Combe et al., 2021; Cronin, 1991; Cronin et al., 1988; Derr, 1991; Hornbeck
295 & Mahoney, 2000; Hughes & Carr, 1993; Russell et al., 2021; Stubblefield et al., 1986). Consistent
296 with other estimates of divergence (Combe et al., 2021; Douzery & Randi, 1997), WTD and MD
297 appear to have split ~1 Mya (Fig. 2). The main pattern of selection, however, appears to be that
298 of balancing selection, with patterns of divergence with gene flow notably absent. Moreover,
299 the scarcity of allopatric selection (i.e., positive selection post divergence) suggests isolation
300 and drift primarily underlies the species differentiation, with WTD and MD likely only recently
301 coming into secondary contact. The genome-wide scans combined with the absence of
302 introgression suggests that WTD & MD are far along the speciation continuum (i.e. Phase 3-4 in
303 Feder 2012), despite contemporary hybridisation. Some signals of selection might have been

304 lost by recombination and time, but the paucity of DwGF and AS signals even at liberal cut-offs
305 suggest speciation was driven primarily by drift.

306 Speciation with negligible historical introgression

307 Contemporary hybridisation in wild WTD & MD appears highly variable: rates between 1 and
308 19% have been observed, which can be explained by the region and degree of overlap, but
309 also by the methodologies used that vary in resolution (Cathey et al., 1998; Combe et al., 2021;
310 Cronin, 1991; Cronin et al., 1988; Derr, 1991; Hornbeck & Mahoney, 2000; Hughes & Carr, 1993;
311 Russell et al., 2021; Stubblefield et al., 1986). Moreover, a proposed historical mtDNA
312 introgression would suggest ancestral hybridisation occurred (Wright et al., 2022). We therefore
313 expected to find signs of hybridisation or introgression in our samples (e.g. Combe et al. 2021)
314 assuming that historic matings would have left some detectable signal as is seen in humans
315 (e.g. Green et al., 2010) and other species (Liu et al., 2021; Paijmans et al., 2021; Poelstra et al.,
316 2014). We found signs for some contemporary hybridisation in one sample (Fig.1C), however,
317 supplementary analysis with an additional WTD from Mexico shows that the admixture pattern
318 might reflect population structure rather than hybridisation since it is detected in both
319 individuals (Fig. S2). Our sampling surely is limited for the detection of recent hybridisation, as
320 previous studies suggest gene flow is rather restricted (Cronin, 1991; Cronin et al., 1988; Derr,
321 1991; Hughes & Carr, 1993; Russell et al., 2021). For example, Russell et al., (2021) estimated a
322 1% hybridisation rate in Alberta by sampling 987 individuals in a range overlap of
323 approximately 230,000 km^2. While signatures of fd identified potentially introgressed windows
324 in ~1% of the genome, this metric can be a sign of incomplete lineage sorting (ILS) rather than
325 introgression, particularly with no signs of genome-wide introgression as evidenced in the D-
326 statistic (Durand et al., 2011). Moreover, the independent treemix and MSMC analyses do not

support historical gene flow, thus collectively, we consider the fd signal to be reflective of ILS which would be expected given the high genome-wide diversity and a relatively recent species-split (Cutter, 2013).

The absence of signatures of ancestral gene flow suggest that WTD & MD evolved in allopatry and have recently come in secondary contact. The dynamic history of the North American continent, notably the glacial cycles of the Quaternary, shaped the evolution and distribution of many taxa (Avise et al., 1998; G. Hewitt, 2000; Shafer et al., 2010). The use of different refugia for prolonged periods of time during the glaciation events has increased the differentiation between populations in several species, including deer (Colella et al., 2021; Dussex et al., 2020; Ito et al., 2021; Kinoshita et al., 2020; Latch et al., 2014). Previous studies show that MD persisted in several refugia during the glacial cycles of the Pleistocene, increasing the intraspecific divergence (Latch et al., 2009, 2014; Wright et al., 2022). This divergence is highlighted here in the PCA which shows two distinct MD clusters, potentially segregating MD and its black-tailed deer subspecies (Fig. 1B). Environmental shifts following the LGM also seem to have impacted the divergence between subspecies of WTD & MD in Florida and Mexico for example (Alminas et al., 2021; Ellsworth et al., 1994). More recently, a form of allopatry was mediated through overharvest and habitat destruction of WTD where populations greatly decreased, even extirpated, in some regions (Budd et al., 2018; Chafin et al., 2021; Deyoung et al., 2003; McDonald et al., 2004). Here, the core of the current sympatric zone was greatly affected, notably Colorado, Montana, Idaho, Nebraska and Wyoming where WTD was almost extirpated before restocking efforts took place starting in the mid 40's (McDonald et al., 2004). Historical introgression could further be hindered by selection against hybrids, either through sexual selection, intrinsic or ecological incompatibilities (Nosil, 2012; Rundle & Nosil, 2005,

350 2005). Given the genome-wide level of differentiation, we hypothesise that Dobzhansky-Muller
351 incompatibles (DMIs) are in play, noting that they often occur early and rapidly (Schumer et al.,
352 2015)

353 Little is known about hybrid fitness in the *Odocoileus* genus. Assessment of a single F1
354 male suggested it was fertile with no overt genetic defects (Derr et al., 1991). In three hybrids
355 the spermatozoid phenotype showed a gradient of fertility with one sterile individual, one
356 subfertile or infertile, and one potentially fertile but did not reproduce (Wishart et al., 1988);
357 sperm phenotypic variation has been implicated in speciation (Albrechtová et al., 2012) and
358 DMIs (Bhattacharyya et al., 2013). Sexual selection can also be a powerful driver of selection
359 against hybrids (Servedio, 2004). In WTD, while females might favour males with larger antlers
360 (Morina et al., 2018), sexual selection is poorly understood and most studies focus on male
361 breeding success (DeYoung et al. 2006; DeYoung et al. 2009; Jones et al. 2011; Newbolt et al.
362 2017). Morphologically, however, WTD & MD are similar enough where the identification of their
363 hybrids is often unreliable with the metatarsal gland and their escape gaits being the only
364 characters consistently distinguishing hybrids (Lingle, 1992; Wishart, 1980). Thus, we propose
365 a testable hypothesis that ecological incompatibilities combined with DMI maintain the
366 species boundary. More research is also needed on WTD/MD hybrids to quantify their fitness
367 and genomic regions underlying any selected hybrid phenotypes.

368 **Mixed signatures of selection underly speciation in North American deer**

369 Given ongoing hybridisation between WTD & MD and a history of recurrent glacial cycles, we
370 expected to find deer genomes showing patterns of divergence with gene flow. The absence of
371 introgression and the nonexistence of DwGF in our genome scans, even with more liberal
372 thresholds, is surprising (Table S1). The further absence of allopatric selection suggests a

373 diminished role of selection in deer speciation, lending support to drift-induced model more
374 consistent with Mayr's (1942) view. While the number of windows in allopatric selection
375 increased with more liberal thresholds (Table S1), they are never the dominant pattern. We
376 suggest this absence could be a result of species being far along on the speciation continuum
377 such that patterns of AS (specifically high F_{ST}) are no longer detectable or meaningful because
378 they were reduced with time and recombination, which is consistent with late stages of
379 ecological speciation (Feder, 2012). Recombination rates, demography and selection can still
380 influence the patterns investigated here. We inferred ρ estimates that are consistent with those
381 in other mammals (Jensen-Seaman et al., 2004), their correlations with divergence would
382 suggest *Odocoileus* are well defined species and reflects a shift in recombination landscape
383 between WTD & MD (Fig. 5; (Burri, 2017; Shang et al., 2021).

384 Signatures of balancing selection (BLS) were widespread with windows containing
385 genes consistent with deer biology (olfactory receptors) and previous literature (i.e., MHC
386 genes). The major histocompatibility complex (MHC) is involved in pathogen recognition in
387 vertebrate species (Bernatchez & Landry, 2003; Piertney & Oliver, 2006). Polymorphism in this
388 complex is directly linked to disease susceptibility, the most diverse are the most resistant, and
389 there is evidence that MHC diversity is maintained by BLS (Bernatchez and Landry 2003;
390 Piertney and Oliver 2006; Aguilar et al. 2004; Pierini and Lenz 2018; Zhang et al. 2018). The same
391 is true for the results on sensory perception and particularly the olfactory receptors genes, both
392 of which are under BLS (Liu et al., 2021). Each allele in the olfactory receptors genes is
393 expressed by a single sensory neuron (Degl'Innocenti & D'Errico, 2017), it is therefore plausible
394 that diversity in these genes would be maintained by BLS. Olfactory receptors in WTD and MD,
395 and more broadly deer, is critical to both predator detection and rutting behaviour (Ditchkoff,

2011). Anatomically, WTD has an architecture optimised for the delivery of sensory stimuli to the receptors in the nose (Ranslow et al., 2014). Chamaillé-Jammes et al. (2014) showed that MD inspect and respond differently to different predators' olfactory cues, and both human and mammalian predators utilise wind direction and other mechanism to minimise scent detection by deer (Cherry & Barton, 2017; Zagata & Haugen, 1974). In other ungulates , olfactory reception is associated with sexual activity (Cann et al., 2019), maternal behaviour (Blank & Yang, 2017; Keller & Lévy, 2012), territory choice (Deutsch & Nefdt, 1992), predator response (Kuijper et al., 2014; Wikenros et al., 2015) and foraging (Hirata & Kusatake, 2021). Low differentiation (F_{ST}) but high diversity (π) of olfactory receptors in deer is consistent with their biology and behaviour and would be expected to be under strong selection pressures.

Conclusion

Our results suggests that WTD & MD do not conform to a speciation with gene flow scenario despite evidence of contemporary hybridisation. We propose they were spatially separated during the Quaternary glaciation cycles where genome-wide differentiation accrued via drift. This is evidenced by the majority of the genome (> 97%) not matching a selection scenario. Increased sampling and model-based demographic assessment should help clarify the role of glaciers and secondary contact in North American deer. The near absence of patterns of allopatry and ancestral gene flow suggest that white-tailed and mule deer are far along the speciation continuum, the absence of introgression signs could suggest DMIs and selection against hybrids which would contribute to the reinforcement of reproductive isolation. Future studies should focus on assaying hybrid phenotypes and vigour.

Acknowledgements

Camille Kessler was supported by an International Graduate Scholarship. Eric Wootton was supported by three Natural Sciences and Engineering Research Council (NSERC) Undergraduate Student Research Awards. This work was supported by NSERC Discovery Grant (Grant Number: RGPIN-2017- 03934); ComputeCanada Resources for Research Groups (Grant Number: RRG gme-665-ab); Canadian Foundation for Innovation: John R. Evans Leaders Fund and the Ontario Early Researcher Award (Grant Number: #36905). We thank Catherine Cullingham, Anh Dao, Orrin Duvuvuei, Brad Fulk, Steve Griffin, Levi Jaster, Lee Jeffers, Emily Latch, the NRDPFC, Charles Ruth, David Walter, Geoff Williams, Kevin White, Mark Wong, Kiana Young and Liana Zanette, for providing samples. We are also grateful to Andrew Foote, Jose Alberto Lòpez Alemàn and Emily Latch for their comments on the manuscript. Trent University is located on the traditional territory of the Michi Saagiig Anishnaabeg. Our sampling also covers most of North America, a stolen land which remains a home to many First Peoples. We are grateful to have had the opportunity to work on this land, would like to show our respect to the First Peoples, and thank them for their care, stewardship, and teachings. Miigwetch.

References

Abbott, R., Albach, D., Ansell, S., Arntzen, J. W., Baird, S. J. E., Bierne, N., Boughman, J., Brelsford, A., Buerkle, C. A., Buggs, R., Butlin, R. K., Dieckmann, U., Eroukhmanoff, F., Grill, A., Cahan, S. H., Hermansen, J. S., Hewitt, G., Hudson, A. G., Jiggins, C., … Zinner, D. (2013). Hybridization and Speciation. *Journal of Evolutionary Biology*, *26*(2), 229–246. https://doi.org/10.1111/j.1420-9101.2012.02599.x

Adams, K. P., & Hamilton, R. J. (2011). Management History. In D. G. Hewitt (Ed.), *Biology and management of white-tailed deer*. CRC Press. http://site.ebrary.com/id/11007093

Adavoudi, R., & Pilot, M. (2021). Consequences of Hybridization in Mammals: A Systematic Review. *Genes*, *13*(1), 50. https://doi.org/10.3390/genes13010050

Aguilar, A., Roemer, G., Debenham, S., Binns, M., Garcelon, D., & Wayne, R. K. (2004). High MHC Diversity Maintained by Balancing Selection in an otherwise Genetically Monomorphic Mammal. *Proceedings of the National Academy of Sciences*, *101*(10), 3490–3494. https://doi.org/10.1073/pnas.0306582101

Albrechtová, J., Albrecht, T., Baird, S. J. E., Macholán, M., Rudolfsen, G., Munclinger, P., Tucker, P. K., & Piálek, J. (2012). Sperm-related phenotypes implicated in both maintenance and breakdown of a natural species barrier in the house mouse. *Proceedings of the Royal Society B: Biological Sciences*, *279*(1748), 4803–4810. https://doi.org/10.1098/rspb.2012.1802

Alminas, O. S. V., Heffelfinger, J. R., Statham, M. J., & Latch, E. K. (2021). Phylogeography of Cedros and Tiburón Island Mule Deer in North America's Desert Southwest. *Journal of Heredity*, *112*(3), 260–275. https://doi.org/10.1093/jhered/esab013

Andrews S. (2010). *FastQC: a Quality Control Tool for High Throughput Sequence Data. Available online at: Http://www.bioinformatics.babraham.ac.uk/projects/fastqc*.

Anthony, R. G., & Smith, N. S. (1977). Ecological Relationships between Mule Deer and White-Tailed Deer in Southeastern Arizona. *Ecological Monographs*, *47*(3), 255–277. https://doi.org/10.2307/1942517

Avise, J. C., Walker, D., & Johns, G. C. (1998). Speciation Durations and Pleistocene Effects on Vertebrate Phylogeography. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *265*(1407), 1707–1712. https://doi.org/10.1098/rspb.1998.0492

Barlow, A., Paijmans, J. L. A., Alberti, F., Gasparyan, B., Bar-Oz, G., Pinhasi, R., Foronova, I., Puzachenko, A. Y., Pacher, M., Dalén, L., Baryshnikov, G., & Hofreiter, M. (2021). Middle Pleistocene genome calibrates a revised evolutionary history of extinct cave bears. *Current Biology*, *31*(8), 1771-1779.e7. https://doi.org/10.1016/j.cub.2021.01.073

Bernatchez, L., & Landry, C. (2003). MHC Studies in Nonmodel Vertebrates: What Have we Learned about Natural Selection in 15 Years? *Journal of Evolutionary Biology*, *16*(3), 363–377. https://doi.org/10.1046/j.1420-9101.2003.00531.x

Berry, S. L., Shipley, L. A., Long, R. A., & Loggers, C. (2019). Differences in Dietary Niche and Foraging Behavior of Sympatric Mule and White-Tailed Deer. *Ecosphere*, *10*(7). https://doi.org/10.1002/ecs2.2815

Bhattacharyya, T., Gregorova, S., Mihola, O., Anger, M., Sebestova, J., Denny, P., Simecek, P., & Forejt, J. (2013). Mechanistic basis of infertility of mouse intersubspecific hybrids. *Proceedings of the National Academy of Sciences*, *110*(6). https://doi.org/10.1073/pnas.1219126110

Blank, D. A., & Yang, W. (2017). Mother-young Recognition in Goitered Gazelle during Hiding Period. *Behavioural Processes*, *142*, 21–28. https://doi.org/10.1016/j.beproc.2017.05.016

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*, *30*(15), 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Booker, T. R., Yeaman, S., & Whitlock, M. C. (2021). Global Adaptation Complicates the Interpretation of Genome Scans for Local Adaptation. *Evolution Letters*, *5*(1), 4–15. https://doi.org/10.1002/evl3.208

Bradley, R. D., Bryant, F. C., Bradley, L. C., Haynie, M. L., & Baker, R. J. (2003). IMPLICATIONS OF HYBRIDIZATION BETWEEN WHITE-TAILED DEER AND MULE DEER. *The Southwestern Naturalist*, *48*(4), 7.

Broad Institute, GitHub Repository. (2019). *"Picard Toolkit."* http://broadinstitute.github.io/picard/

Brunjes, K. J., Ballard, W. B., Humphrey, M. H., Harwell, F., Mcintyre, N. E., Krausman, P. R., & Wallace, M. C. (2006). Habitat Use by Sympatric Mule and White-Tailed Deer in Texas. *Journal of Wildlife Management*, *70*(5), 1351–1359. https://doi.org/10.2193/0022-541X(2006)70[1351:HUBSMA]2.0.CO;2

Budd, K., Berkman, L. K., Anderson, M., Koppelman, J., & Eggert, L. S. (2018). Genetic Structure and Recovery of White-tailed Deer in Missouri. *The Journal of Wildlife Management*, *82*(8), 1598–1607. https://doi.org/10.1002/jwmg.21546

Burri, R. (2017). Interpreting Differentiation Landscapes in the Light of Long-term Linked Selection. *Evolution Letters*, *1*(3), 118–131. https://doi.org/10.1002/evl3.14

Cambronne, A. (2013). *Deerland: America's Hunt for Ecological Balance and the Essence of Wildness*. http://www.freading.com/ebooks/details/r:download/MDAwMDA5LTEwOTA1NDYz

Campbell, C. R., Poelstra, J. W., & Yoder, A. D. (2018). What is Speciation Genomics? The Roles of Ecology, Gene Flow, and Genomic Architecture in the Formation of Species. *Biological Journal of the Linnean Society*, *124*(4), 561–583. https://doi.org/10.1093/biolinnean/bly063

Cann, P., Chabi, M., Delsart, A., Le Danvic, C., Saliou, J.-M., Chasles, M., Keller, M., & Nagnan-Le Meillour, P. (2019). The Olfactory Secretome Varies According to Season in Female Sheep and Goat. *BMC Genomics*, *20*(1), 794. https://doi.org/10.1186/s12864-019-6194-z

Carr, S. M., Ballinger, S. W., Derr, J. N., Blankenship, L. H., & Bickham, J. W. (1986). Mitochondrial DNA Analysis of Hybridization between Sympatric White-tailed Deer and Mule Deer in West Texas. *Proceedings of the National Academy of Sciences*, *83*(24), 9576–9580. https://doi.org/10.1073/pnas.83.24.9576

Carr, S. M., & Hughes, G. A. (1993). Direction of Introgressive Hybridization between Species of North American Deer (Odocoileus) as Inferred from Mitochondrial-Cytochrome-b Sequences. *Journal of Mammalogy*, *74*(2), 331–342. https://doi.org/10.2307/1382388

Cathey, J. C., Bickham, J. W., & Patton, J. C. (1998). Introgressive Hybridization and Nonconcordant Evolutionary History of Maternal and Paternal Lineages in North American Deer. *Evolution*, *52*(4), 1224–1229. https://doi.org/10.1111/j.1558-5646.1998.tb01850.x

Chafin, T. K., Zbinden, Z. D., Douglas, M. R., Martin, B. T., Middaugh, C. R., Gray, M. C., Ballard, J. R., & Douglas, M. E. (2021). Spatial Population Genetics in Heavily Managed Species:

Separating Patterns of Historical Translocation from Contemporary Gene Flow in White-tailed Deer. *Evolutionary Applications*, *14*(6), 1673–1689. https://doi.org/10.1111/eva.13233

Chamaillé-Jammes, S., Malcuit, H., Le Saout, S., & Martin, J.-L. (2014). Innate Threat-sensitive Foraging: Black-tailed Deer Remain More Fearful of Wolf than of the less Dangerous Black Bear even after 100 years of Wolf Absence. *Oecologia*, *174*(4), 1151–1158. https://doi.org/10.1007/s00442-013-2843-0

Cherry, M. J., & Barton, B. T. (2017). Effects of Wind on Predator-Prey Interactions. *Food Webs*, *13*, 92–97. https://doi.org/10.1016/j.fooweb.2017.02.005

Colella, J. P., Lan, T., Talbot, S. L., Lindqvist, C., & Cook, J. A. (2021). Whole-genome Resequencing Reveals Persistence of Forest-Associated Mammals in Late Pleistocene Refugia along North America's North Pacific Coast. *Journal of Biogeography*, *48*(5), 1153–1169. https://doi.org/10.1111/jbi.14068

Combe, F. J., Jaster, L., Ricketts, A., Haukos, D., & Hope, A. G. (2021). Population Genomics of Free-ranging Great Plains White-tailed and Mule deer Reflects a Long History of Inter-specific Hybridization. *Evolutionary Applications*, eva.13330. https://doi.org/10.1111/eva.13330

Cronin, M. A. (1991). Mitochondrial and Nuclear Genetic Relationships of Deer (Odocoileus spp.) in Western North America. *Canadian Journal of Zoology*, *69*(5), 1270–1279. https://doi.org/10.1139/z91-179

Cronin, M. A., Vyse, E. R., & Cameron, D. G. (1988). Genetic Relationships between Mule Deer and White-Tailed Deer in Montana. *The Journal of Wildlife Management*, *52*(2), 320. https://doi.org/10.2307/3801243

Cruickshank, T. E., & Hahn, M. W. (2014). Reanalysis Suggests that Genomic Islands of Speciation are Due to Reduced Diversity, not Reduced Gene Flow. *Molecular Ecology*, *23*(13), 3133–3157. https://doi.org/10.1111/mec.12796

Cutter, A. D. (2013). Integrating Phylogenetics, Phylogeography and Population Genetics through Genomes and Evolutionary Theory. *Molecular Phylogenetics and Evolution*, *69*(3), 1172–1185. https://doi.org/10.1016/j.ympev.2013.06.006

Darwin, C., 1809-1882. (1859). *On the Origin of Species by Means of Natural Selection, or Preservation of Favoured Races in the Struggle for Life*. London: John Murray, 1859. https://search.library.wisc.edu/catalog/9934839413602122

Dedato, M., Robert, C., Taillon, J., Shafer, A., & Cote, S. (2022). *Demographic History and Conservation Genomics of Caribou (Rangifer tarandus) in Québec* [Preprint]. Preprints. https://doi.org/10.22541/au.164165840.02956352/v1

Degl'Innocenti, A., & D'Errico, A. (2017). Regulatory Features for Odorant Receptor Genes in the Mouse Genome. *Frontiers in Genetics*, *8*, 8.

Delaneau, O., Marchini, J., & Zagury, J.-F. (2012). A linear complexity phasing method for thousands of genomes. *Nature Methods*, *9*(2), 179–181. https://doi.org/10.1038/nmeth.1785

Demarais, S., Miller, K. V., & Jacobson, H. A. (2000). White-tailed deer. In S. Demarais & P. R. Krausman (Eds.), *Ecology and management of large mammals in North America*. Prentice Hall.

Derr, J. N. (1991). Genetic Interactions between White-Tailed and Mule Deer in the Southwestern United States. *The Journal of Wildlife Management*, *55*(2), 228. https://doi.org/10.2307/3809144

Derr, J. N., Hale, D. W., Ellsworth, D. L., & Bickham, J. W. (1991). *Fertility in an F1 Male Hybrid of White-tailed Deer (Odocoileus virginianus) x Mule deer (O. hemionus)*. 7.

Deutsch, J. C., & Nefdt, R. J. C. (1992). Olfactory Cues Influence Female Choice in two Lek-breeding Antelopes. *Nature*, *356*(6370), 596–598. https://doi.org/10.1038/356596a0

Deyoung, R. W., Demarais, S., Honeycutt, R. L., Rooney, A. P., Gonzales, R. A., & Gee, K. L. (2003). Genetic Consequences of White-tailed Deer (Odocoileus virginianus) Restoration in Mississippi. *Molecular Ecology*, *12*(12), 3237–3252. https://doi.org/10.1046/j.1365-294X.2003.01996.x

Ditchkoff, S. S. (2011). Anatomy and Physiology. In D. G. Hewitt (Ed.), *Biology and management of white-tailed deer*. CRC Press. http://site.ebrary.com/id/11007093

Douzery, E., & Randi, E. (1997). The Mitochondrial Control Region of Cervidae: Evolutionary Patterns and Phylogenetic Content. *Molecular Biology and Evolution*, *14*(11), 1154–1166. https://doi.org/10.1093/oxfordjournals.molbev.a025725

Durand, E. Y., Patterson, N., Reich, D., & Slatkin, M. (2011). Testing for Ancient Admixture between Closely Related Populations. *Molecular Biology and Evolution*, *28*(8), 2239–2252. https://doi.org/10.1093/molbev/msr048

Dussex, N., Alberti, F., Heino, M. T., Olsen, R.-A., van der Valk, T., Ryman, N., Laikre, L., Ahlgren, H., Askeyev, I. V., Askeyev, O. V., Shaymuratova, D. N., Askeyev, A. O., Döppes, D., Friedrich, R., Lindauer, S., Rosendahl, W., Aspi, J., Hofreiter, M., Lidén, K., … Díez-del-Molino, D. (2020). Moose Genomes Reveal Past Glacial Demography and the Origin of Modern Lineages. *BMC Genomics*, *21*(1), 854. https://doi.org/10.1186/s12864-020-07208-3

591 Ellsworth, D. L., Honeycutt, R. L., Silvy, N. J., Bickham, J. W., & Klimstra, W. D. (1994). Historical
592 Biogeography and Contemporary Patterns Of Mitochondrial DNA Variation In White-
593 Tailed Deer From The Southeastern United States. *Evolution*, *48*(1), 122–136.
594 https://doi.org/10.1111/j.1558-5646.1994.tb01299.x
595 Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the Number of Clusters of Individuals
596 Using the Software Structure: A Simulation Study. *Molecular Ecology*, *14*(8), 2611–2620.
597 https://doi.org/10.1111/j.1365-294X.2005.02553.x
598 Ewels, P., Magnusson, M., Lundin, S., & Käller, M. (2016). MultiQC: Summarize Analysis Results
599 for Multiple Tools and Samples in a Single Report. *Bioinformatics*, *32*(19), 3047–3048.
600 https://doi.org/10.1093/bioinformatics/btw354
601 Feder, J. L. (2012). *The Genomics of Speciation-With-Gene-Flow*. *28*(7), 9.
602 Gao, F., Ming, C., Hu, W., & Li, H. (2016). New Software for the Fast Estimation of Population
603 Recombination Rates (FastEPRR) in the Genomic Era. *G3 Genes|Genomes|Genetics*, *6*(6),
604 1563–1571. https://doi.org/10.1534/g3.116.028233
605 Gilbert, C., Ropiquet, A., & Hassanin, A. (2006). Mitochondrial and Nuclear Phylogenies of
606 Cervidae (Mammalia, Ruminantia): Systematics, Morphology, and Biogeography.
607 *Molecular Phylogenetics and Evolution*, *40*(1), 101–117.
608 https://doi.org/10.1016/j.ympev.2006.02.017
609 Gower, G., Tuke, J., Rohrlach, A., Soubrier, J., Llamas, B., Bean, N., & Cooper, A. (2018).
610 *Population Size History from Short Genomic Scaffolds: How Short is Too Short?* [Preprint].
611 Genetics. https://doi.org/10.1101/382036
612 Green, R. E., Krause, J., Briggs, A. W., Maricic, T., Stenzel, U., Kircher, M., Patterson, N., Li, H., Zhai,
613 W., Fritz, M. H. Y., Hansen, N. F., Durand, E. Y., Malaspinas, A. S., Jensen, J. D., Marques-

Bonet, T., Alkan, C., Prufer, K., Meyer, M., Burbano, H. A., … Paabo, S. (2010). A Draft Sequence of the Neandertal Genome. *Science*, *328*(5979), 710–722. https://doi.org/10.1126/science.1188021

Hewitt, G. (2000). The Genetic Legacy of the Quaternary Ice Ages. *Nature*, *405*(6789), 907–913. https://doi.org/10.1038/35016000

Hirata, M., & Kusatake, N. (2021). Relative Importance of Senses in Forage Discrimination by Cattle Depends on the Sensory Contrast between the Discrimination Targets: A Preliminary Study. *Animal Cognition*, *24*(1), 99–106. https://doi.org/10.1007/s10071-020-01422-y

Hornbeck, G. E., & Mahoney, J. M. (2000). *Introgressive Hybridization of Mule Deer and White-Tailed Deer in Southwestern Alberta*. 5.

Hughes, G. A., & Carr, S. M. (1993). Reciprocal Hybridization between White-tailed Deer (Odocoileus virginianus) and Mule Deer (O.hemionus) in Western Canada: Evidence from Serum Albumin and mtDNA Sequences. *Canadian Journal of Zoology*, *71*(3), 524–530. https://doi.org/10.1139/z93-073

Iacolina, L., Corlatti, L., Buzan, E., Safner, T., & Šprem, N. (2019). Hybridisation in European Ungulates: An Overview of the Current Status, Causes, and Consequences. *Mammal Review*, *49*(1), 45–59. https://doi.org/10.1111/mam.12140

Ito, T., Hayakawa, T., Suzuki–Hashido, N., Hamada, Y., Kurihara, Y., Hanya, G., Kaneko, A., Natsume, T., Aisu, S., Honda, T., Yachimori, S., Anezaki, T., Omi, T., Hayama, S., Tanaka, M., Wakamori, H., Imai, H., & Kawamoto, Y. (2021). Phylogeographic History of Japanese Macaques. *Journal of Biogeography*, jbi.14087. https://doi.org/10.1111/jbi.14087

IUCN. (2015a). *Odocoileus virginianus: Gallina, S. and Lopez Arevalo, H.: The IUCN Red List of Threatened Species 2016: e.T42394A22162580* [Data set]. International Union for Conservation of Nature. https://doi.org/10.2305/IUCN.UK.2016-2.RLTS.T42394A22162580.en

IUCN. (2015b). *Odocoileus hemionus: Sanchez Rojas, G. and Gallina Tessaro, S.: The IUCN Red List of Threatened Species 2016: e.T42393A22162113* [Data set]. International Union for Conservation of Nature. https://doi.org/10.2305/IUCN.UK.2016-1.RLTS.T42393A22162113.en

Jensen-Seaman, M. I., Furey, T. S., Payseur, B. A., Lu, Y., Roskin, K. M., Chen, C.-F., Thomas, M. A., Haussler, D., & Jacob, H. J. (2004). Comparative Recombination Rates in the Rat, Mouse, and Human Genomes. *Genome Research*, *14*(4), 528–538. https://doi.org/10.1101/gr.1970304

Keller, M., & Lévy, F. (2012). The Main but not the Accessory Olfactory System is Involved in the Processing of Socially Relevant Chemosignals in Ungulates. *Frontiers in Neuroanatomy*, *6*. https://doi.org/10.3389/fnana.2012.00039

Kinoshita, E., Kosintsev, P. A., Abramov, A. V., Solovyev, V. A., Saveljev, A. P., Nishita, Y., & Masuda, R. (2020). Holocene Changes in the Distributions of Asian and European Badgers (Carnivora: Mustelidae: Meles) Inferred from Ancient DNA Analysis. *Biological Journal of the Linnean Society*, *129*(3), 594–602. https://doi.org/10.1093/biolinnean/blaa007

Kirch, M., Romundset, A., Gilbert, M. T. P., Jones, F. C., & Foote, A. D. (2021). Ancient and modern stickleback genomes reveal the demographic constraints on adaptation. *Current Biology*, S0960982221002682. https://doi.org/10.1016/j.cub.2021.02.027

Kopelman, N. M., Mayzel, J., Jakobsson, M., Rosenberg, N. A., & Mayrose, I. (2015). Clumpak: A Program for Identifying Clustering Modes and Packaging Population Structure Inferences Across K. *Molecular Ecology Resources*, *15*(5), 1179–1191. https://doi.org/10.1111/1755-0998.12387

Korneliussen, T. S., Albrechtsen, A., & Nielsen, R. (2014). ANGSD: Analysis of Next Generation Sequencing Data. *BMC Bioinformatics*, *15*(1), 356. https://doi.org/10.1186/s12859-014-0356-4

Kuijper, D. P. J., Verwijmeren, M., Churski, M., Zbyryt, A., Schmidt, K., Jędrzejewska, B., & Smit, C. (2014). What Cues Do Ungulates Use to Assess Predation Risk in Dense Temperate Forests? *PLoS ONE*, *9*(1), e84607. https://doi.org/10.1371/journal.pone.0084607

Latch, E. K., Heffelfinger, J. R., Fike, J. A., & RHODES Jr, O. E. (2009). Species-wide Phylogeography of North American Mule Deer (Odocoileus hemionus): Cryptic Glacial Refugia and Postglacial Recolonization. *Molecular Ecology*, *18*(8), 1730–1745. https://doi.org/10.1111/j.1365-294X.2009.04153.x

Latch, E. K., Reding, D. M., Heffelfinger, J. R., Alcalá-Galván, C. H., & Rhodes, O. E. (2014). Range-wide Analysis of Genetic Structure in a Widespread, Highly Mobile Species (Odocoileus hemionus) Reveals the Importance of Historical Biogeography. *Molecular Ecology*, *23*(13), 3171–3190. https://doi.org/10.1111/mec.12803

Lavretsky, P., DaCosta, J. M., Sorenson, M. D., McCracken, K. G., & Peters, J. L. (2019). DdRAD-seq Data Reveal Significant Genome-Wide Population Structure and divergent Genomic Regions that Distinguish the Mallard and Close Relatives in North America. *Molecular Ecology*, *28*(10), 2594–2609. https://doi.org/10.1111/mec.15091

Li, H., & Durbin, R. (2009). Fast and Accurate Short Read Alignment with Burrows-Wheeler Transform. *Bioinformatics*, *25*(14), 1754–1760. https://doi.org/10.1093/bioinformatics/btp324

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., & 1000 Genome Project Data Processing Subgroup. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, *25*(16), 2078–2079. https://doi.org/10.1093/bioinformatics/btp352

Lingle, S. (1992). Escape Gaits of White-Tailed Deer, Mule Deer and Their Hybrids: Gaits Observed and Patterns of Limb Coordination. *Behaviour*, *122*(3–4), 153–181. https://doi.org/10.1163/156853992X00499

Lingle, S. (2001). Anti-Predator Strategies and Grouping Patterns in White-Tailed Deer and Mule Deer. *Ethology*, *107*(4), 295–314. https://doi.org/10.1046/j.1439-0310.2001.00664.x

Lingle, S. (2003). Group Composition and Cohesion in Sympatric White-tailed Deer and Mule deer. *Canadian Journal of Zoology*, *81*(7), 1119–1130. https://doi.org/10.1139/z03-097

Liu, Q., Mishra, M., Saxena, A. S., Wu, H., Qiu, Y., Zhang, X., You, X., Ding, S., & Miyamoto, M. M. (2021). Balancing Selection Maintains Ancient Polymorphisms at Conserved Enhancers for the Olfactory Receptor Genes of a Chinese Marine Fish. *Molecular Ecology*, *30*(16), 4023–4038. https://doi.org/10.1111/mec.16016

Liu, S., Westbury, M. V., Dussex, N., Mitchell, K. J., Sinding, M.-H. S., Heintzman, P. D., Duchêne, D. A., Kapp, J. D., von Seth, J., Heiniger, H., Sánchez-Barreiro, F., Margaryan, A., André-Olsen, R., De Cahsan, B., Meng, G., Yang, C., Chen, L., van der Valk, T., Moodley, Y., … Gilbert, M. T. P. (2021). Ancient and modern genomes unravel the evolutionary history

of the rhinoceros family. *Cell*, *184*(19), 4874-4885.e16. https://doi.org/10.1016/j.cell.2021.07.032

Mallet, J. (2005). Hybridization as an Invasion of the Genome. *Trends in Ecology & Evolution*, *20*(5), 229–237. https://doi.org/10.1016/j.tree.2005.02.010

Marques, D. A., Lucek, K., Sousa, V. C., Excoffier, L., & Seehausen, O. (2019). Admixture Between Old Lineages Facilitated Contemporary Ecological Speciation in Lake Constance Stickleback. *Nature Communications*, *10*(1), 4240. https://doi.org/10.1038/s41467-019-12182-w

Martin, M., Patterson, M., Garg, S., O Fischer, S., Pisanti, N., Klau, G. W., Schöenhuth, A., & Marschall, T. (2016). *WhatsHap: Fast and Accurate Read-based Phasing* [Preprint]. Bioinformatics. https://doi.org/10.1101/085050

Martin, S. (2021). *Genomics_general*. GitHub. https://github.com/simonhmartin/genomics_general

Martin, S., Davey, J. W., & Jiggins, C. D. (2015). Evaluating the Use of ABBA–BABA Statistics to Locate Introgressed Loci. *Molecular Biology and Evolution*, *32*(1), 244–257. https://doi.org/10.1093/molbev/msu269

Mayr, E. (1942). *Systematics and the Origin of Species*.

McDonald, J. S., Miller, K. V., Association, Q. D. M., & Resources, D. B. W. S. of F. (2004). *A History of White-tailed Deer Restocking in the United States, 1878 to 2004*. Quality Deer Management Assoc. https://books.google.ca/books?id=fv_oGwAACAAJ

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., & DePristo, M. A. (2010). The Genome Analysis Toolkit:

A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research*, *20*(9), 1297–1303. https://doi.org/10.1101/gr.107524.110

Meisner, J., & Albrechtsen, A. (2018). Inferring Population Structure and Admixture Proportions in Low-Depth NGS Data. *Genetics*, *210*(2), 719–731. https://doi.org/10.1534/genetics.118.301336

Momigliano, P., Jokinen, H., Fraimout, A., Florin, A.-B., Norkko, A., & Merilä, J. (2017). *Extraordinarily Rapid Speciation in a Marine Fish*. 6.

Morina, D. L., Demarais, S., Strickland, B. K., & Larson, J. E. (2018). While Males Fight, Females Choose: Male Phenotypic Quality Informs Female Mate Choice in Mammals. *Animal Behaviour*, *138*, 69–74. https://doi.org/10.1016/j.anbehav.2018.02.004

Moscarella, R. A., Aguilera, M., & Escalante, A. A. (2003). Phylogeography, Population Structure, and Implications for Conservation of White-tailed Deer (Odocoileus Virginianus) in Venezuela. *Journal of Mammalogy*, *84*(4), 1300–1315. https://doi.org/10.1644/BRB-028

NatureServe. (2021a). *NatureServe Explorer* [Web application]. NatureServe Mule Deer. https://explorer.natureserve.org/Taxon/ELEMENT_GLOBAL.2.101365/Odocoileus_hemionus

NatureServe. (2021b). *NatureServe Explorer [web application] WTD*. NatureServe, Arlington, Virginia. https://explorer.natureserve.org/Taxon/ELEMENT_GLOBAL.2.104777/Odocoileus_virginianus

Nosil, P. (2012). *Ecological Speciation*. Oxford University Press.

Paijmans, J. L. A., Barlow, A., Becker, M. S., Cahill, J. A., Fickel, J., Förster, D. W. G., Gries, K., Hartmann, S., Havmøller, R. W., Henneberger, K., Kern, C., Kitchener, A. C., Lorenzen, E.

D., Mayer, F., OBrien, S. J., von Seth, J., Sinding, M.-H. S., Spong, G., Uphyrkina, O., … Hofreiter, M. (2021). African and Asian leopards are highly differentiated at the genomic level. *Current Biology*, S0960982221004577. https://doi.org/10.1016/j.cub.2021.03.084

Pedersen, B. S., & Quinlan, A. R. (2018). Mosdepth: Quick Coverage Calculation for Genomes and Exomes. *Bioinformatics*, *34*(5), 867–868. https://doi.org/10.1093/bioinformatics/btx699

Peres, T. M., & Altman, H. (2018). The Magic of Improbable Appendages: Deer Antler Objects in the Archaeological Record of the American South. *Journal of Archaeological Science: Reports*, *20*, 888–895. https://doi.org/10.1016/j.jasrep.2017.10.028

Pickrell, J. K., & Pritchard, J. K. (2012). Inference of Population Splits and Mixtures from Genome-Wide Allele Frequency Data. *PLoS Genetics*, *8*(11), e1002967. https://doi.org/10.1371/journal.pgen.1002967

Pierini, F., & Lenz, T. L. (2018). Divergent Allele Advantage at Human MHC Genes: Signatures of Past and Ongoing Selection. *Molecular Biology and Evolution*, *35*(9), 2145–2158. https://doi.org/10.1093/molbev/msy116

Piertney, S. B., & Oliver, M. K. (2006). The Evolutionary Ecology of the Major Histocompatibility Complex. *Heredity*, *96*(1), 7–21. https://doi.org/10.1038/sj.hdy.6800724

Pitra, C., Fickel, J., Meijaard, E., & Groves, C. (2004). Evolution and Phylogeny of Old World Deer. *Molecular Phylogenetics and Evolution*, *33*(3), 880–895. https://doi.org/10.1016/j.ympev.2004.07.013

Poelstra, J. W., Vijay, N., Bossu, C. M., Lantz, H., Ryll, B., Muller, I., Baglione, V., Unneberg, P., Wikelski, M., Grabherr, M. G., & Wolf, J. B. W. (2014). The Genomic Landscape Underlying

Phenotypic Integrity in the Face of Gene Flow in Crows. *Science*, *344*(6190), 1410–1414. https://doi.org/10.1126/science.1253226

Pundir, S., Martin, M. J., O'Donovan, C., & The UniProt Consortium. (2016). UniProt Tools. *Current Protocols in Bioinformatics*, *53*(1). https://doi.org/10.1002/0471250953.bi0129s53

Quinlan, A. R., & Hall, I. M. (2010). BEDTools: A Flexible Suite of Utilities for Comparing Genomic Features. *Bioinformatics*, *26*(6), 841–842. https://doi.org/10.1093/bioinformatics/btq033

R Core Team. (2021). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Ragavan, P., Zhou, R., Ng, W. L., Rana, T. S., Mageswaran, T., Mohan, P. M., & Saxena, A. (2017). Natural Hybridization in Mangroves – an Overview. *Botanical Journal of the Linnean Society*, *185*(2), 208–224. https://doi.org/10.1093/botlinnean/box053

Ranslow, A. N., Richter, J. P., Neuberger, T., Van Valkenburgh, B., Rumple, C. R., Quigley, A. P., Pang, B., Krane, M. H., & Craven, B. A. (2014). Reconstruction and Morphometric Analysis of the Nasal Airway of the White-Tailed Deer (Odocoileus virginianus) and Implications Regarding Respiratory and Olfactory Airflow: White-Tailed Deer Nasal Anatomy and Function. *The Anatomical Record*, *297*(11), 2138–2147. https://doi.org/10.1002/ar.23037

Ravinet, M., Faria, R., Butlin, R. K., Galindo, J., Bierne, N., Rafajlović, M., Noor, M. A. F., Mehlig, B., & Westram, A. M. (2017). Interpreting the Genomic Landscape of Speciation: A Road Map for Finding Barriers to Gene Flow. *Journal of Evolutionary Biology*, *30*(8), 1450–1477. https://doi.org/10.1111/jeb.13047

Rundle, H. D., & Nosil, P. (2005). Ecological Speciation. *Ecology Letters*, *8*(3), 336–352. https://doi.org/10.1111/j.1461-0248.2004.00715.x

Russell, T., Cullingham, C., Ball, M., Pybus, M., & Coltman, D. (2021). Extent and Direction of Introgressive Hybridization of Mule and White-tailed Deer in Western Canada. *Evolutionary Applications*, eva.13250. https://doi.org/10.1111/eva.13250

Russell, T., Cullingham, C., Kommadath, A., Stothard, P., Herbst, A., & Coltman, D. (2019). Development of a Novel Mule Deer Genomic Assembly and Species-Diagnostic SNP Panel for Assessing Introgression in Mule Deer, White-Tailed Deer, and Their Interspecific Hybrids. *G3: Genes|Genomes|Genetics*, g3.200838.2018. https://doi.org/10.1534/g3.118.200838

Schiffels, S., & Durbin, R. (2014). Inferring Human Population Size and Separation History from Multiple Genome Sequences. *Nature Genetics*, *46*(8), 919–925. https://doi.org/10.1038/ng.3015

Schiffels, S., & Wang, K. (2020). MSMC and MSMC2: The Multiple Sequentially Markovian Coalescent. In J. Y. Dutheil (Ed.), *Statistical Population Genomics* (Vol. 2090, pp. 147–166). Springer US. https://doi.org/10.1007/978-1-0716-0199-0_7

Schumer, M., Cui, R., Rosenthal, G. G., & Andolfatto, P. (2015). Reproductive Isolation of Hybrid Populations Driven by Genetic Incompatibilities. *PLOS Genetics*, *11*(3), e1005041. https://doi.org/10.1371/journal.pgen.1005041

Servedio, M. R. (2004). The Evolution of Premating Isolation: Local Adaptation and Natural and Sexual Selection Against Hybrids. *Evolution*, *58*(5), 913–924. https://doi.org/10.1111/j.0014-3820.2004.tb00425.x

Shafer, A. B. A., Cullingham, C. I., Côté, S. D., & Coltman, D. W. (2010). Of Glaciers and Refugia: A Decade of Study Sheds New Light on the Phylogeography of Northwestern North America. *Molecular Ecology*, *19*(21), 4589–4621. https://doi.org/10.1111/j.1365-294X.2010.04828.x

Shafer, A. B. A., & Wolf, J. B. W. (2013). Widespread Evidence for Incipient Ecological Speciation: A Meta-analysis of Isolation-by-ecology. *Ecology Letters*, *16*(7), 940–950. https://doi.org/10.1111/ele.12120

Shang, H., Rendón-Anaya, M., Paun, O., Field, D. L., Hess, J., Vogl, C., Liu, J., Ingvarsson, P. K., Lexer, C., & Leroy, T. (2021). *Conserved Genomic Landscapes of Differentiation Across Populus Speciation Continuum* [Preprint]. Evolutionary Biology. https://doi.org/10.1101/2021.08.26.457771

Skotte, L., Korneliussen, T. S., & Albrechtsen, A. (2013). Estimating Individual Admixture Proportions from Next Generation Sequencing Data. *Genetics*, *195*(3), 693–702. https://doi.org/10.1534/genetics.113.154138

Soraggi, S., Wiuf, C., & Albrechtsen, A. (2018). Powerful Inference with the D-statistic on Low-coverage Whole-genome Data. *G3: Genes, Genomes, Genetics*, *8*(2). https://doi.org/10.1534/g3.117.300192

Stankowski, S., & Ravinet, M. (2021). Defining the Speciation Continuum. *Evolution*, *75*(6), 1256–1273. https://doi.org/10.1111/evo.14215

Stubblefield, S. S., Warren, R. J., & Murphy, B. R. (1986). Hybridization of Free-Ranging White-Tailed and Mule Deer in Texas. *The Journal of Wildlife Management*, *50*(4), 688. https://doi.org/10.2307/3800982

Tarasov, A., Vilella, A. J., Cuppen, E., Nijman, I. J., & Prins, P. (2015). Sambamba: Fast processing of NGS alignment formats. *Bioinformatics*, *31*(12), 2032–2034. https://doi.org/10.1093/bioinformatics/btv098

Taylor, S. A., & Larson, E. L. (2019). Insights from Genomes into the Evolutionary Importance and Prevalence of Hybridization in Nature. *Nature Ecology & Evolution*, *3*(2), 170–177. https://doi.org/10.1038/s41559-018-0777-y

UniProt. (2022a). *Q9Y5E5 · PCDB4_HUMAN*. UniProt Beta. https://beta.uniprot.org/uniprotkb/Q9Y5E5/entry

UniProt. (2022b). *Q15057 · ACAP2_HUMAN*. UniProt Beta. https://beta.uniprot.org/uniprotkb/Q15057/entry

Wang, K., Mathieson, I., O'Connell, J., & Schiffels, S. (2020). Tracking Human Population Structure through Time from Whole Genome Sequences. *PLOS Genetics*, *16*(3), e1008552. https://doi.org/10.1371/journal.pgen.1008552

Whittaker, D. G., & Lindzey, F. G. (2004). Habitat Use Patterns of Sympatric Deer Species on Rocky Mountain Arsenal, Colorado. *Wildlife Society Bulletin*, *32*(4), 1114–1123. https://doi.org/10.2193/0091-7648(2004)032[1114:HUPOSD]2.0.CO;2

Wikenros, C., Kuijper, D. P. J., Behnke, R., & Schmidt, K. (2015). Behavioural Responses of Ungulates to Indirect Cues of an Ambush Predator. *Behaviour*, *152*(7–8), 1019–1040. https://doi.org/10.1163/1568539X-00003266

Wishart, W. D. (1980). Hybrids of White-Tailed and Mule Deer in Alberta. *Journal of Mammalogy*, *61*(4), 716–719. https://doi.org/10.2307/1380318

Wishart, W. D., Hrudka, F., Schmutz, S. M., & Flood, P. F. (1988). Observations on Spermatogenesis, Sperm Phenotype, and Fertility in White-tailed × Mule Deer Hybrids

and a Yak × Cow Hybrid. *Canadian Journal of Zoology*, *66*(7), 1664–1671. https://doi.org/10.1139/z88-240

Wright, E. A., Roberts, E. K., Platt, R. N., Bayouth, J. V., Conway, W. C., & Bradley, R. D. (2022). Mitochondrial capture and subsequent genetic divergence generates a novel haplogroup: Evidence from ancient and ongoing hybridization in mule and white-tailed deer. *Journal of Mammalogy*. https://doi.org/10.1093/jmammal/gyab156

Zagata, M. D., & Haugen, A. O. (1974). Influence of Light and Weather on Observability of Iowa Deer. *The Journal of Wildlife Management*, *38*(2), 220. https://doi.org/10.2307/3800727

Zhang, P., Huang, K., Zhang, B., Dunn, D. W., Chen, D., Li, F., Qi, X., Guo, S., & Li, B. (2018). High Polymorphism in MHC-DRB Genes in Golden Snub-nosed Monkeys Reveals Balancing Selection in Small, Isolated Populations. *BMC Evolutionary Biology*, *18*(1), 29. https://doi.org/10.1186/s12862-018-1148-7

Data Accessibility and Benefit-Sharing

Data Accessibility

Raw reads for the 28 deer individuals were deposited on the NCBI Sequence Read Archive (Accession number PRJNA830519). All scripts are available on GitLab (https://gitlab.com/WiDGeT_TrentU/graduate_theses/-/tree/master/Kessler/CH_01).

Benefit-Sharing

The contributions of tissue samples to the research are described in the ACKNOWLEDGEMENTS, benefits from this research results in data accessible on public databases as described above.

Author contributions

882 CK and ABAS conceived the study, CK performed bioinformatic analyses with contribution
883 from ABAS, CK and EW performed the molecular laboratory work, CK and ABAS wrote the
884 manuscript. All authors commented on the text.
885

Speciation without gene-flow in hybridising deer

Camille Kessler[1], Eric Wootton[2], Aaron B.A. Shafer[1]

[1]Environmental and Life Sciences Graduate Program, Trent University, 7K9J 7B8 Peterborough, Ontario, Canada

[2]Biochemistry & Molecular Biology, Trent University, K9J 7B8 Peterborough, Ontario, Canada

Corresponding authors:

Camille Kessler: camillekessler@trentu.ca

Aaron Shafer: aaronshafer@trentu.ca

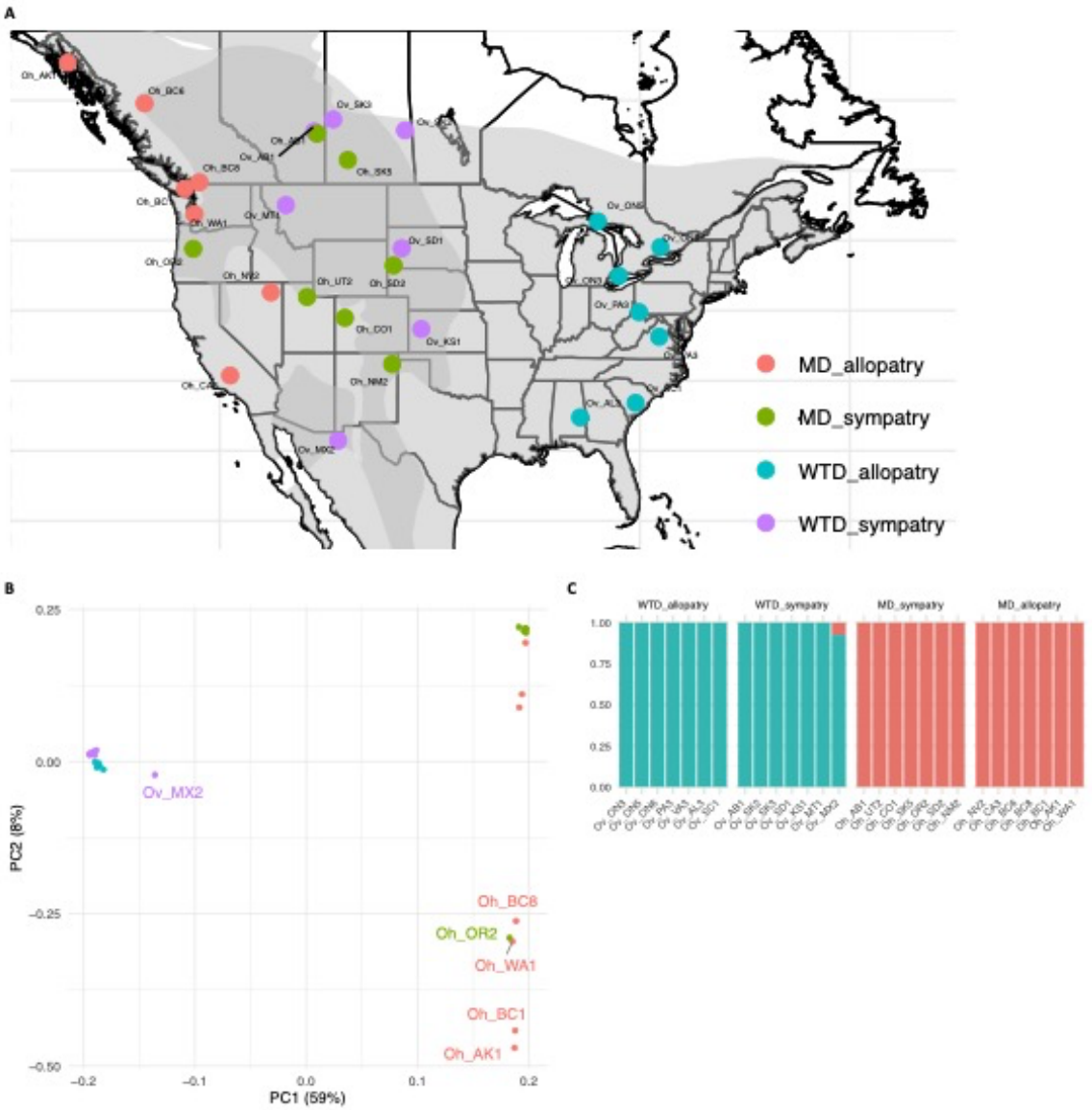

Figure 1: Deer individuals in the study. (A) Sampling locations, colours represent the different groups: WTD allopatry (WTDa, blue), WTD sympatry (WTDs, violet), MD sympatry (MDs, green), MD allopatry (MDa, orange). Areas of sympatry are coloured in dark grey and areas of allopatry in light grey, from the IUCN range data for both species (IUCN, 2015a, 2015b). (B) Principal component analysis performed in PCAngsd. (C) Admixture proportions for K = 2 calculated in NGSAdmix.

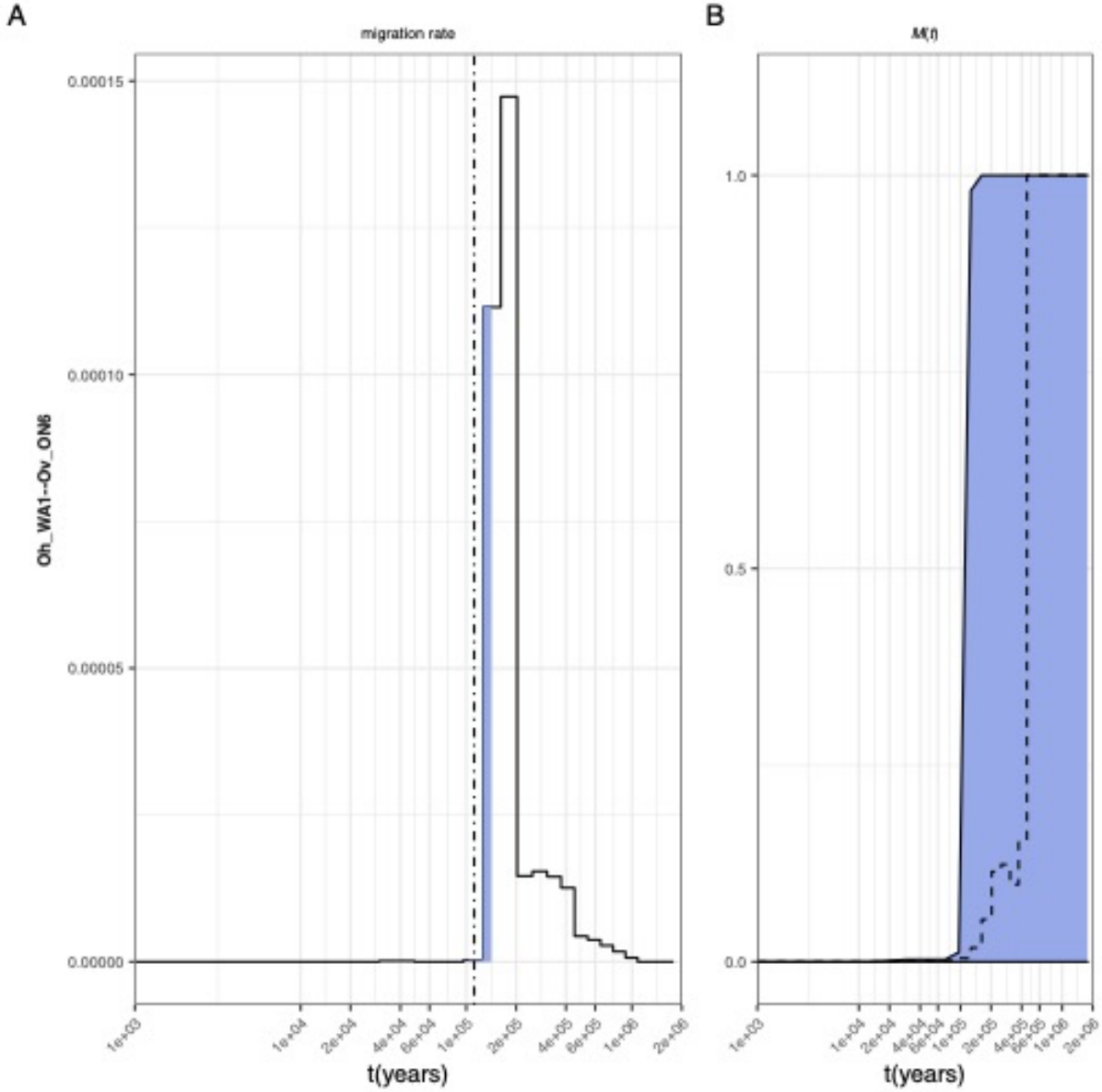

Figure 2. MSMC-IM analysis of white-tailed deer (Ov) and mule deer (Oh) from Ontario (ON) and Washington (WA). (A) Migration rates over time where dashed line indicate the time point where 50% of ancestry has merged. Blue shading indicates 99% percentile of the cumulative migration probability. (B) Cumulative migration probabilities *M*(*t*). Dashed lines indicate the relative cross coalescence rate obtained from MSMC2.

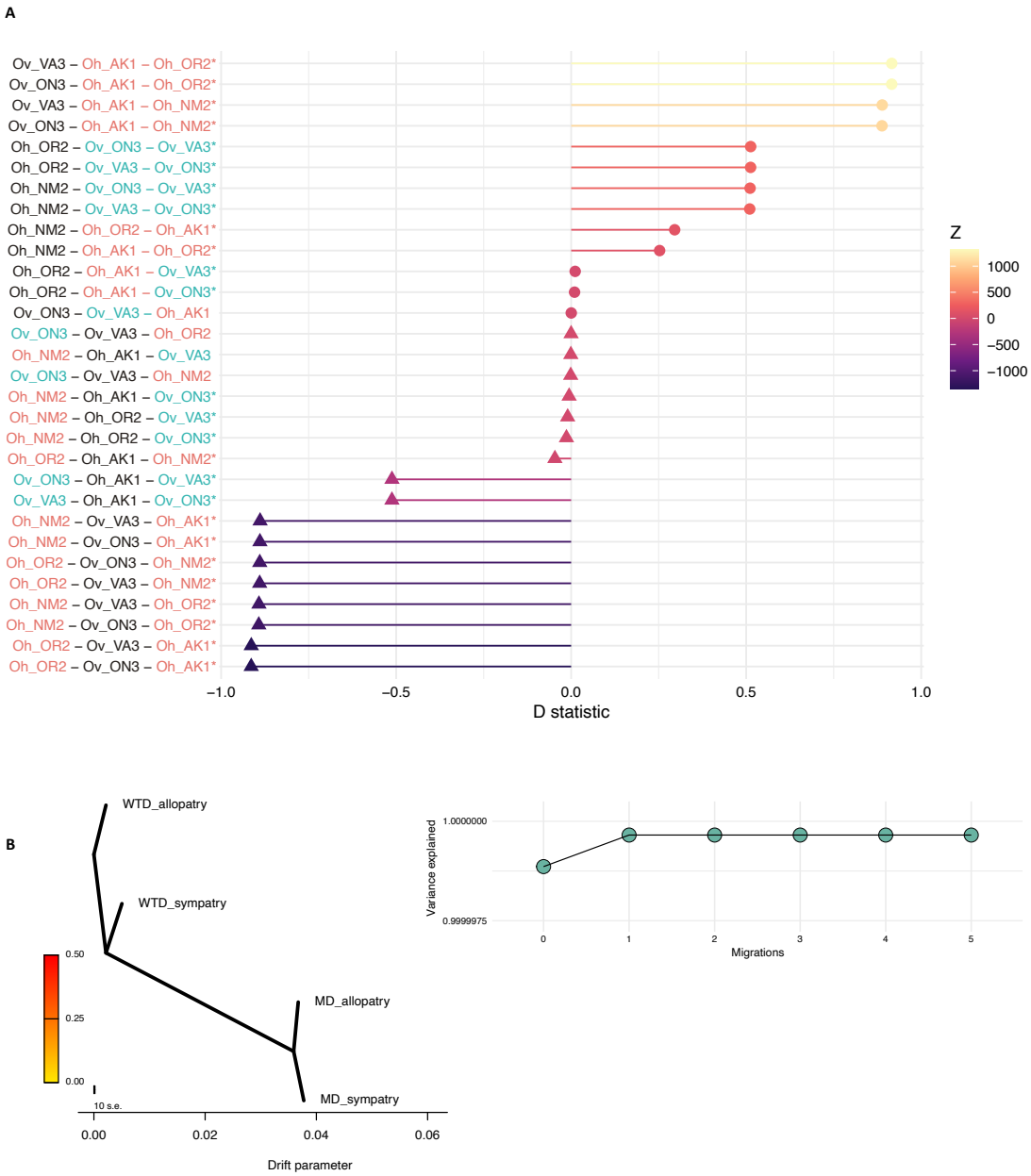

Figure 3: Introgression analyses. (A) ABBA-BABA analysis between individuals, individuals involved in gene flow are coloured depending on the species (WTD in blue, MD in orange), the colour gradient represents Z score, excess of ABBA depicted as points, excess of BABA shown as triangles, significant tests highlighted with an *(Z-score > |3|). (B) Maximum likelihood tree inferred by Treemix for 0 migration (right) and the variance explained by each migration event (left).

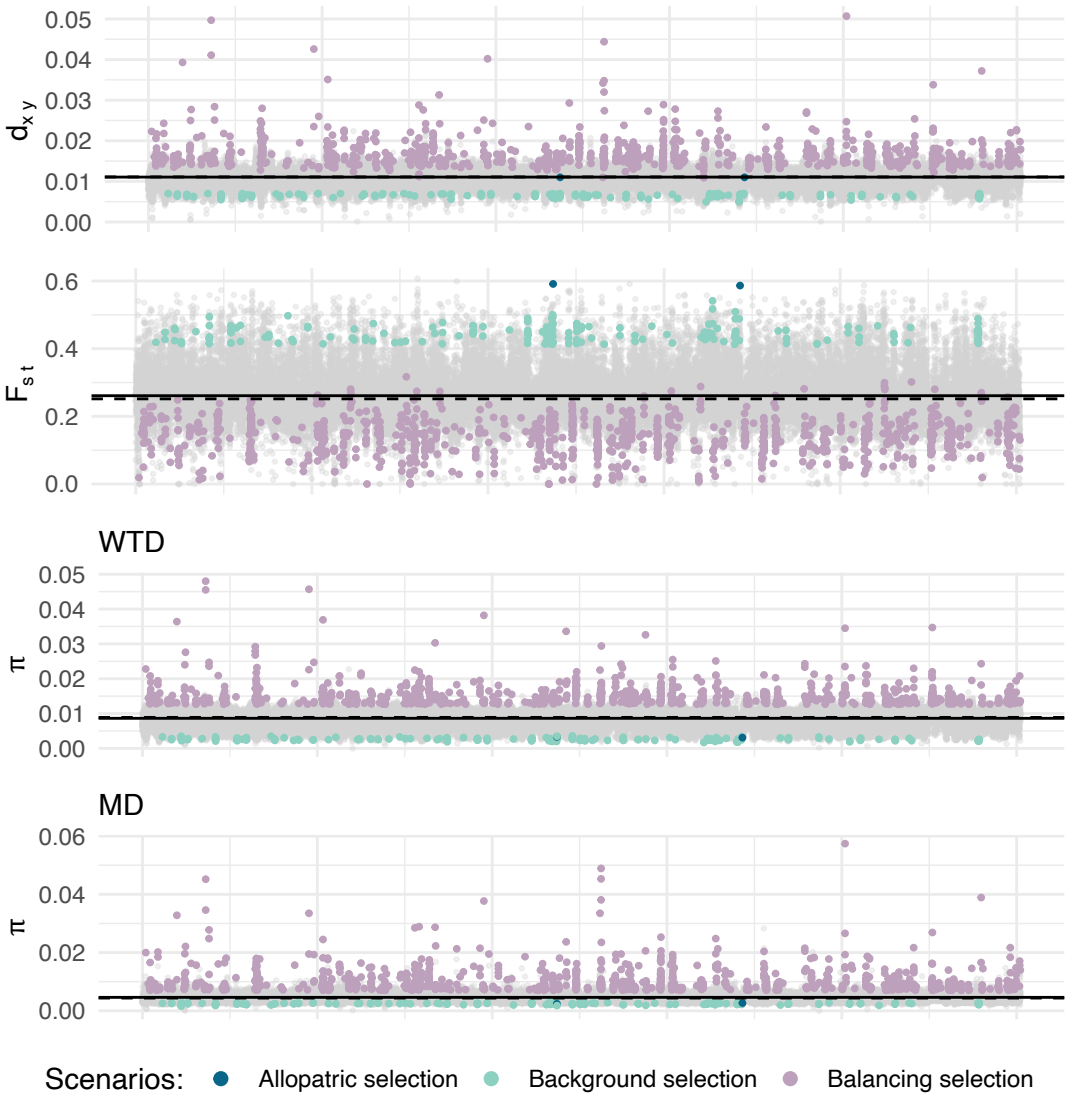

Figure 4: Distribution of windows presenting pattern of selection across the genome for d_{xy} (top), F_{ST} (middle) and π (bottom). Windows following a pattern of selection are coloured according to its corresponding evolutionary scenario, grey windows do not exhibit pattern of selection. The continuous and dashed line represent the mean and median respectively.

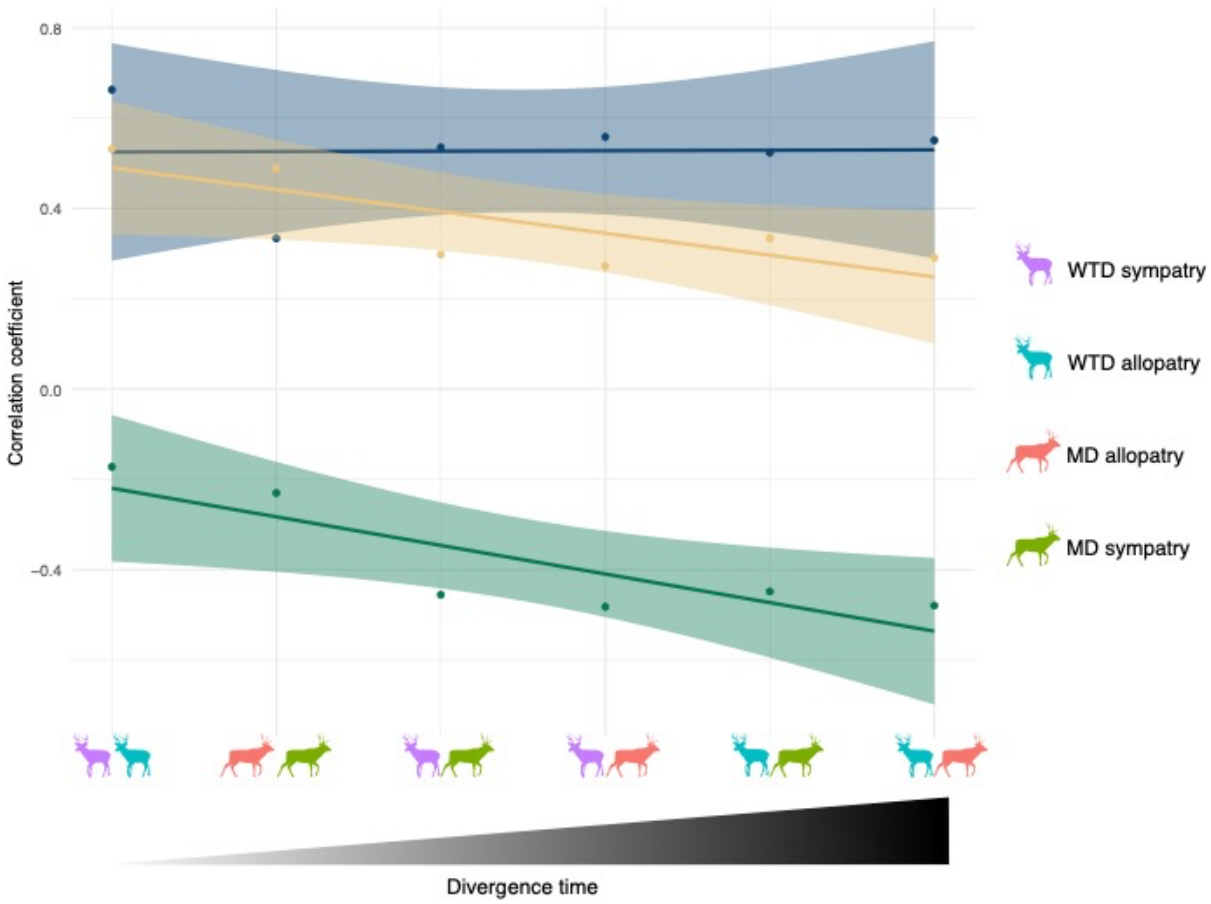

Figure 5: Correlation coefficients over divergence time for π - ρ (dark blue) F_{ST} - ρ (dark green) and ρ - ρ (yellow), trend lines represent a linear model. The speciation continuum is represented by six comparisons of our four groups).

934

935 **Table 1:** Deer sample information

Genome_ID	Species	Location	Sex	Group
Ov_ON3	*O_virginianus*	Ontario	F	WTD_allopatry
Ov_ON5	*O_virginianus*	Ontario	M	WTD_allopatry
Ov_ON6	*O_virginianus*	Ontario	M	WTD_allopatry
Ov_PA3	*O_virginianus*	Pennsylvania	F	WTD_allopatry
Ov_VA3	*O_virginianus*	Virginia	F	WTD_allopatry
Ov_AL3	*O_virginianus*	Alabama	Unk	WTD_allopatry
Ov_SC1	*O_virginianus*	South Carolina	F	WTD_allopatry
Ov_AB1	*O_virginianus*	Alberta	F	WTD_sympatry
Ov_SK2	*O_virginianus*	Saskatchewan	M	WTD_sympatry
Ov_SK3	*O_virginianus*	Saskatchewan	M	WTD_sympatry
Ov_SD1	*O_virginianus*	South Dakota	M	WTD_sympatry
Ov_KS1	*O_virginianus*	Kansas	F	WTD_sympatry
Ov_MT1	*O_virginianus*	Montana	M	WTD_sympatry
Ov_MX2	*O_virginianus*	Mexico	M	WTD_sympatry
Oh_NV2	*O_hemionus*	Nevada	M	MD_allopatry
Oh_CA3	*O_hemionus*	California	Unk	MD_allopatry
Oh_BC6	*O_hemionus*	British Columbia	M	MD_allopatry
Oh_BC8	*O_hemionus*	British Columbia	M	MD_allopatry
Oh_BC1	*O_hemionus*	British Columbia	F	MD_allopatry
Oh_AK1	*O_hemionus*	Alaska	Unk	MD_allopatry
Oh_WA1	*O_hemionus*	Washington	M	MD_allopatry
Oh_AB1	*O_hemionus*	Alberta	F	MD_sympatry
Oh_UT2	*O_hemionus*	Utah	M	MD_sympatry
Oh_CO1	*O_hemionus*	Colorado	F	MD_sympatry
Oh_SK5	*O_hemionus*	Saskatchewan	Unk	MD_sympatry
Oh_OR2	*O_hemionus*	Oregon	Unk	MD_sympatry
Oh_SD2	*O_hemionus*	South Dakota	F	MD_sympatry
Oh_NM2	*O_hemionus*	New Mexico	M	MD_sympatry

936

937

938 **Table 2:** Four different evolutionary scenario and their thresholds for F_{st}, π and D_{xy} percentiles
939 applied in the genome scans analysis.

Scenario	F_{st}	π	D_{xy}
Divergence with gene flow (DwGF)	> 0.95	< 0.05	> 0.95
Allopatric selection (AS)	> 0.95	< 0.05	Between 0.45 & 0.55
Background selection (BGS)	> 0.95	< 0.05	< 0.05
Balancing selection (BLS)	< 0.05	> 0.95	> 0.95

940

941