Speciation without gene-flow in hybridizing 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

Hybridisation in nature is part of the continuum of speciation. 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. 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 introgression and performed genome scans to look for signatures of four different selection scenarios. Despite modern evidence of hybridization, we found no sign of introgression, no signature of divergence with gene flow, and localized patterns of allopatric and balancing selection in the genome. Of note, genes under balancing selection were related to immunity and MHC or sensory perception of smell, the latter of which is consistent with deer biology. The deficiency of patterns selection suggests that white-tailed and mule deer were spatially separated during the glaciation cycles of the Pleistocene where genome wide differentiation accrued via drift. Absence of historical introgression signs could suggest Dobzhansky-Muller incompatibilities and selection against hybrids, and that both species are now far along the speciation continuum. Our results suggests that WTD & MD do not conform to a speciation with gene flow scenario, but that they evolved via drift in allopatry during the Quaternary and that both species are currently advanced along the speciation continuum.

Introduction

Hybridisation is a widespread phenomenon that occurs at variable rates (Mallet 2005; Ragavan et al. 2017; Iacolina et al. 2019; Taylor and Larson 2019; Adavoudi and Pilot 2021). 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 and Ravinet 2021). However, both ideas are not incompatible as the speciation continuum ultimately leads to reproductive isolation with hybridisation being a natural outcome (Stankowski and 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; eventually, the underlying genetic mechanisms (i.e. causative loci) gradually create barriers to gene flow, ultimately leading to reproductive isolation (Rundle and Nosil 2005; Nosil 2012; Shafer and 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 or contribute to assortative mating, or reduce hybrids 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 of the genome where the differentiation (e.g. F_{ST}) is higher than the neutral genomic background (Ravinet et al. 2017; Campbell et al. 2018). Such loci and islands have been observed in several species, underpinning a wide array of speciation and divergence processes (Poelstra et al. 2014; Momigliano et al. 2017; Lavretsky et al. 2019; Marques et al. 2019). Scanning the genome for just F_{ST} peaks when speciation with gene flow is ongoing, however, is problematic as other mechanisms can create similar F_{ST} profiles such as genetic drift, global

55 adaptation, or simply reduced genetic diversity due to background selection (Cruickshank and
56 Hahn 2014; Burri 2017; Ravinet et al. 2017; Booker et al. 2021). Joint genome scans including
57 F_{ST}, d_{xy} and π facilitate a better depiction of the processes of selection and ecological
58 speciation (Cruickshank and Hahn 2014; Ravinet et al. 2017; Campbell et al. 2018; Shang et al.
59 2021). Of note, Shang et al. (2021) recently conceptualized the expected patterns F_{ST}, d_{xy} and π
60 under the main modes of selection; this approach allows for neutral patterns of evolution to
61 be disentangled from selection based on contrasting patterns of the aforementioned summary
62 statistics.

63 White-tailed deer (*Odocoileus* virginianus; WTD) and mule deer (*O.* hemionus; MD) are
64 abundant in North America with similar morphology, activity patterns and life-history traits,
65 but they differ in several ecological aspects (Douzery and Randi 1997; Gilbert et al. 2006; Pitra
66 et al. 2004; Brunjes et al. 2006; Berry et al. 2019). WTD prioritise security and thermoregulation
67 whereas MD favour food availability (Whittaker and Lindzey 2004); consequently, WTD prefer
68 habitats at lower altitude and with denser visual cover than MD that favour open areas at higher
69 elevation (Anthony and Smith 1977; Brunjes et al. 2006). White-tailed deer and MD also differ in
70 their sociality and associated predator response. Mule deer live in large cohesive groups
71 including both sexes, and, as a group, adopt an aggressive behaviour in presence of predators,
72 whereas WTD live in smaller female-biased groups and flee in response to predators (Lingle
73 2001; Lingle 2003). Both species represent a high economic value in North America as hunting-
74 related activities generate billions of dollars annually (Cambronne 2013), and both species are
75 important cultural component of Indigenous communities (Adams and Hamilton 2011; Peres
76 and Altman 2018).

77 Despite a species divergence date estimated at ~3.13 mya (Wright et al. 2021), WTD and
78 MD hybridise in areas of sympatry with estimated hybridisation rates ranging from 1 to 19%
79 depending on the region (Carr and Hughes 1993; Hornbeck and Mahoney 2000; Combe et al.
80 2021; Russell et al. 2021). Both species likely spent considerable time in allopatry (i.e. separate
81 refugia) during glacial events (e.g Moscarella et al. 2003; Latch et al. 2009). It has also been
82 shown that however restricted, gene flow appears bidirectional as there are signs of
83 introgression at both mitochondrial and nuclear levels, such that some MD acquired WTD
84 mitochondrial DNA around ~1.32 Mya (Carr et al. 1986; Cronin 1991; Derr 1991; Carr and Hughes
85 1993; Cathey et al. 1998; Bradley et al. 2003; Russell et al. 2019; Wright et al. 2021). This pattern
86 also suggests ancestral hybridisation and gene flow has taken place. Given their clear
87 phenotypic and behavioural differences, and their documented hybridisation, white-tailed
88 and mule deer do not conform to the biological species concept (Mayr 1942). Here, we
89 hypothesize that WTD and MD have evolved via ecological speciation, and more specifically a
90 speciation with gene flow scenario. To test this hypothesis, we quantified genome-wide
91 introgression and past admixture events and performed genome scans to look for signatures
92 of four different selection regimes. We expected to find genetic signs of divergent selection,
93 such as speciation islands, consistent with patterns of divergence with gene flow (DwGF). We
94 also predicted higher rates of admixture and historical introgression in areas of sympatry.

95 ***Results***

96 We sequenced 28 individuals to an average coverage of approximately 4x. We called 10,397,088
97 SNPs, which represented 0.4% of the genome. In the PCA based on allele frequencies, PC1
98 accounted for 59% of the variation and separated both species (Fig.1B). PC2 explained an
99 additional 8% of the variation and showed a spread of MD individuals that is consistent with

100 their longitudinal distribution (Pearson correlation = 0.8; *p*-value = 0.00057). Latitude was not
101 important in shaping PC1 or PC2 variation ($p > 0.05$). Admixture analysis showed a well-defined
102 genetic clustering between species (Evanno's method: best K = 2; Fig.1C). The WTD individual
103 Ov_MX2 was partially admixed in both analyses, with NGSAdmix assigning it 92.5% WTD and
104 7.44% MD ancestry (Fig.1C). This individuals' admixed pattern disappeared in the admixture
105 analysis with higher K values (K = 3 to 7, Fig.S1), though additional within species clusters are
106 observed.

107 We performed an ABBA-BABA test based on genotype likelihoods between all
108 combinations of our 28 individuals in a first analysis and between all four populations in a
109 second, with the caribou as outgroup. We computed 9828 ABBA-BABA individual topologies
110 in ANGSD; inter-species introgression estimates were all very close to 0, with a D-statistic
111 ranging from 0.011 to -0.013 and Z-score from 6.92 to -8.85. D-statistic in intra-species
112 comparisons range from 0.25 to 0.916 (Z-score from 114.54 to 13111.97) in ABBA and from -
113 0.046 to -0.914 (Z-score from -25.42 to -1348.55) in BABA (Fig.2A). In the comparisons between
114 the four populations, three ABBA-BABA topologies presented signatures of introgression on a
115 total of 12. The single inter-species comparison presents a D-statistic of -0.07 (Z-score = -7.66)
116 for an introgression between sympatric MD and WTD. The other two comparisons show a
117 reciprocal introgression between MD in allopatry and sympatry of 0.9 and 0.89 (Z-score =
118 1342.83 & 1376.69; Fig.S2). The treemix analyses showed a topology with no migration that
119 explained over 99% of the variation (Fig.2B).

120 We further analysed introgression through the measure of f_d across 50kb non-
121 overlapping windows in two different introgression scenarios: 1) WTD_allopatry,
122 WTD_sympatry, MD_sympatry, Caribou and 2) MD_allopatry, MD_sympatry, WTD_sympatry,

123 Caribou (Fig.S3). We found 1243 outlier windows with a f_d higher than the 97.5% percentile in
124 the first scenario and 1248 in the second, each representing ~2.45% of the genome with 68
125 windows overlapping between comparisons (Fig.S3).
126 Measures of relative genomic divergence between WTD and MD were elevated (F_{ST} = 0.26) but
127 absolute divergence was relatively low (d_{xy} = 0.011); genetic diversity (π) was higher in WTD than
128 in MD (WTD = 0.008, MD = 0.004, Mann-Whitney test: $p < 2.2e-16$). We identified a total of 1183
129 windows presenting patterns consistent with one of the four selection scenarios (Fig.3, Fig.S4).
130 Of those, 1016 suggest a pattern of balancing selection, they were distributed across 236
131 scaffolds and represent 1.99% of the genome. Those windows contained 121 genes, some of
132 which showed an enrichment in ontologies associated with the sense of smell and chemical
133 stimuli, including three categories presenting an enrichment above 15-fold (Fig.4A, Fig. S5). We
134 also detected GOs related to the MHC and immunity, these include three categories with an
135 enrichment above 24-fold (Fig. S5). We identified patterns of background selection in 165
136 windows across 58 scaffolds, representing 0.32% of the genome. These windows harboured
137 208 genes identified with a UniProt ID and for which we found enrichment in GOs related to
138 epigenetic factors such as “Unmethylated CpG binding” (25-fold) or “Histone demethylation”
139 (15-fold) (Fig.4B, Fig. S6). These windows were either isolated or clustered together into
140 putative islands of divergence (Fig.S4). We found 2 windows under allopatric selection, each
141 window containing one gene: ACAP2, a GTPase activating protein, and PCDHB4 potentially
142 involved in cell-binding (UniProt 2022a; UniProt 2022b). When we applied a more liberal cut-
143 offs for F_{ST}, d_{xy}, and π we still observed no evidence for divergence with gene flow (Table S3).

144 ***Discussion***

145 We investigated the speciation history of the *Odocoileus* genera through introgression analyses
146 and genome scans. Our results show the absence of historical hybridisation between WTD &
147 MD suggesting speciation between those species took place without gene flow despite
148 contemporary and historical evidence (Stubblefield et al. 1986; Cronin et al. 1988; Cronin 1991;
149 Derr 1991; Hughes and Carr 1993; Cathey et al. 1998; Hornbeck and Mahoney 2000; Combe et
150 al. 2021; Russell et al. 2021). The main pattern of selection from the genome scans is that of
151 balancing selection, with notably patterns of divergence with gene flow absent. Moreover, the
152 scarcity of allopatric selection (i.e., positive selection post divergence) suggests isolation and
153 drift primarily underlies the species differentiation, with WTD and MD likely only recently
154 coming into secondary contact. The genome-wide scans combined with the absence of
155 introgression suggests that WTD & MD are far along the speciation continuum (i.e. Phase 3-4 in
156 Feder 2012), despite contemporary hybridization.

157 Speciation with an absence of historical introgression

158 Contemporary hybridisation in wild WTD & MD is documented at rates ranging from 1 to 19%
159 depending on the region (Stubblefield et al. 1986; Cronin et al. 1988; Cronin 1991; Derr 1991;
160 Hughes and Carr 1993; Cathey et al. 1998; Hornbeck and Mahoney 2000; Combe et al. 2021;
161 Russell et al. 2021). Moreover, a proposed historical mtDNA introgression would suggest
162 ancestral hybridisation occurred (Wright et al. 2022). We therefore expected to find signs of
163 hybridisation or introgression in our samples (e.g. Combe et al. 2021). We found evidence for
164 some contemporary hybridisation in one sample (Fig.1C), but no strong signals of historical
165 interbreeding at the genome-wide level were detected. Our sampling surely is limited for the
166 detection of recent hybridisation, but previous studies suggest gene flow is rather restricted
167 (Cronin et al. 1988; Cronin 1991; Derr 1991; Hughes and Carr 1993; Russell et al. 2021). For

168 example, Russell et al. (2019) discovered a 1% hybridisation rate in Alberta by sampling 987
169 individuals in a range overlap of approximately 230,000 km^2. While signatures of fd identified
170 potentially introgressed windows in ~1% of the genome, this metric can be a sign of
171 incomplete lineage sorting (ILS) rather than introgression, particularly with no signs of
172 genome-wide introgression as evidenced in the D statistic. Moreover, the independent treemix
173 analysis does not support historical gene flow, thus collectively, we consider the fd signal to be
174 reflective of ILS which would be expected given the high genome-wide diversity and a relatively
175 recent species-split (Cutter 2013).

176 The absence of signatures of ancestral hybridisation suggest that WTD & MD evolved in
177 allopatry and have recently come in secondary contact. The dynamic history of the North
178 American continent, notably the glacial cycles of the Quaternary, shaped the evolution and
179 distribution of many species (Avise et al. 1998; Hewitt 2000; Shafer et al. 2010). The use of
180 different refugia for prolonged periods of time during the glaciation events has increased the
181 differentiation between populations in several species, including deer (Latch et al. 2014;
182 Dussex et al. 2020; Kinoshita et al. 2020; Colella et al. 2021; Ito et al. 2021). Previous studies
183 show that MD persisted in several refugia during the glacial cycles of the Pleistocene, increasing
184 the intraspecific divergence (Latch et al. 2009; Latch et al. 2014; Wright et al. 2022).
185 Environmental shifts following the LGM also seem to have impacted the divergence between
186 subspecies of WTD & MD in Florida and Mexico for example (Ellsworth et al. 1994; Alminas et al.
187 2021). More recently, a form of allopatry was mediated through overharvest and habitat
188 destruction of WTD where populations greatly decreased to extirpation in some regions
189 (Deyoung et al. 2003; McDonald et al. 2004; Budd et al. 2018; Chafin et al. 2021). Here, the core
190 of the current sympatric zone was greatly affected, notably Colorado, Montana, Idaho,

191 Nebraska and Wyoming where WTD was almost extirpated before restocking efforts took place
192 starting in the mid 40's (McDonald et al. 2004). Historical introgression could further be
193 hindered by selection against hybrids, either through sexual selection, intrinsic or ecological
194 incompatibilities (Rundle and Nosil 2005; Rundle and Nosil 2005; Nosil 2012). Given the
195 genome-wide level of differentiation, we hypothesize that Dobzhansky-Muller incompatibles
196 (DMIs) are in play, noting that they often occur early and rapidly (Schumer et al. 2015)

197 Little is known about hybrid fitness in the *Odocoileus* genera. Assessment of a single F1
198 male suggested it was fertile and it showed no overt genetic defects (Derr et al. 1991), but
199 conclusions can not be drawn from a single observation. Sexual selection can be a powerful
200 driver of selection against hybrids (Servedio 2004). In WTD, while females might favour males
201 with larger antlers (Morina et al. 2018), sexual selection is overall poorly understood in
202 *Odocoileus* and most studies focus on male breeding success rather than sexual selection by
203 females (DeYoung et al. 2006; DeYoung et al. 2009; Jones et al. 2011; Newbolt et al. 2017).
204 Morphologically, however, WTD & MD are very similar and the identification of their hybrids is
205 often unreliable with the metatarsal gland and their escape gaits being the only characters
206 consistently distinguishing hybrids (Lingle 1992; Wishart 1980). Thus, we propose a testable
207 hypothesis that ecological incompatibilities combined with DMI maintain the species
208 boundary. More research is also needed on WTD/MD hybrids to quantify their fitness and
209 genomic regions underlying any selected hybrid phenotypes.

210 Mixed signatures of selection underly speciation in North American deer

211 Given ongoing hybridisation between WTD & MD and a history of recurrent glacial cycles, we
212 expected to find deer genomes showing patterns of divergence with gene flow. The absence of
213 introgression and the nonexistence of DwGF in our genome scans, even with more liberal

214 thresholds, is surprising (Table S1). The further absence of allopatric selection suggests a
215 diminished role of selection in deer speciation, lending support to drift-induced model more
216 consistent with Mayr's (1942) view. While the number of windows in allopatric selection
217 increased with more liberal thresholds (Table S1), they are never the dominant pattern. We
218 suggest this absence could be a result of species being far along on the speciation continuum
219 such that patterns of AS (specifically high F_{ST}) are no longer detectable or meaningful, which is
220 consistent with late stages of ecological speciation (Feder 2012). Recombination rates,
221 demography and selection can influence the patterns investigated here (Shang et al. 2021),
222 and thus future studies could integrate those factors to confirm the position of WTD & MD on
223 the speciation continuum.

224 Signatures of balancing selection (BLS) were widespread with windows containing
225 genes consistent with deer biology (olfactory receptors) and previous literature (i.e., MHC
226 genes). The major histocompatibility complex (MHC) is involved in pathogen recognition in
227 vertebrate species (Bernatchez and Landry 2003; Piertney and Oliver 2006). Polymorphism in
228 this complex is directly linked to disease susceptibility, the most diverse are the most resistant,
229 and there is evidence that MHC diversity is maintained by BLS (Bernatchez and Landry 2003;
230 Piertney and Oliver 2006; Aguilar et al. 2004; Pierini and Lenz 2018; Zhang et al. 2018). The same
231 is true for the results on sensory perception and particularly the olfactory receptors genes, both
232 of which are under BLS (Liu et al. 2021). Each allele in the olfactory receptors genes is expressed
233 by a single sensory neuron (Degl'Innocenti and D'Errico 2017), it is therefore plausible that
234 diversity in these genes would be maintained by BLS. Olfactory receptors in WTD and MD, and
235 more broadly deer, is critical to both predator detection and rutting behaviour (Ditchkoff 2011).
236 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 utilize wind direction and other mechanism to minimize scent detection by deer (Zagata and Haugen 1974; Cherry and Barton 2017). In other ungulates species, olfactory reception is associated with sexual activity (Cann et al. 2019), maternal behaviour (Keller and Lévy 2012; Blank and Yang 2017), territory choice (Deutsch and Nefdt 1992), predator response (Kuijper et al. 2014; Wikenros et al. 2015) and foraging (Hirata and 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 hybridization. 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 hybridisation 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.

Materials and Methods

Sampling & Sequencing

260 We obtained tissue samples from harvested deer collected across the range of WTD & MD
261 including areas of sympatry and allopatry (Fig.1A, Table 1). These areas were determined from
262 NatureServe records and adjusted with IUCN range data (IUCN 2015a; IUCN 2015b;
263 NatureServe 2021a; NatureServe 2021b). Specifically, the state of Washington and the province
264 of British Columbia, for which the coasts are largely MD allopatric whereas the eastern parts
265 are sympatric, were divided in two by the 120th meridian West. We also removed Manitoba from
266 the sympatric list given the area is quite small and not sampled. We extracted DNA from tissue
267 using the Qiagen DNeasy Blood and Tissue Kit following manufacturer's instructions and
268 checked the samples concentration using Invitrogen Qubit assays. WGS libraries were
269 generated at The Centre for Applied Genomics in Toronto, Canada and sequenced to an
270 average of 4x coverage on an Illumina HiSeqX.

271 Data processing

272 Raw reads quality was examined using FastQC (v0.11.9; (Andrews S. 2010). We trimmed the
273 reads using Trimmomatic (v0.36; Bolger et al., 2014), and aligned them to the WTD genome
274 (GCA_014726795.1) with bwa-mem (v0.7.17; Li & Durbin, 2009). We sorted the reads using
275 SAMtools sort (v1.10; Li et al., 2009), identified duplicates reads with Picard MarkDuplicates
276 (v2.23.2; Broad Institute, GitHub Repository., 2019) and removed them using Sambamba view
277 (v0.7.0; Tarasov et al., 2015). We performed a local re-alignment using GATK
278 RealignerTargetCreator and IndelRealigner (v4.1.7.0; McKenna et al., 2010). For further quality
279 checks, we used Sambamba flagstat, mosdepth (v0.3.1; Pedersen and Quinlan 2018) and
280 MultiQC (v1.10; Ewels et al., 2016). Some samples were sequenced to a higher depth on
281 multiple lanes, for those samples, before the duplicate removal step, the read groups were
282 specified and both sequencing files merged using Picard AddOrReplaceReadGroups and

283 MergeSamFiles. The rest of the pipeline was performed as described above with the addition
284 of a final step: we reduced the final coverage to the average coverage of all other samples (4x)
285 using Picard DownsampleSam.

286 We created three datasets for our different analyses. The first contained the variants
287 among all deer individuals (hereafter called DeerSNP dataset). We produced it using ANGSD
288 (v0.918; Korneliussen et al., 2014) and estimated genotype likelihoods following the GATK
289 model (-gl 2) and called genotypes (-doGeno 4) with the following filtering: SNPs with a
290 minimum p-value of $1e^{-6}$, a minimum base quality of 20 and Minimum mapQ quality of 20 (-
291 SNP_pval 1e-6, -minMapQ 20 & -minQ 20). The second dataset additionally included invariant
292 sites (referred to as DeerALL), it was produced following the same method as DeerSNP except
293 for the filtering on the SNP p-value (-SNP_pval 1e-6) which was removed. Our third dataset
294 comprises the 28 deer individuals and one caribou sample, all mapped to the caribou genome
295 (GCA_014898785.1; afterwards called DeerCAR dataset). For this dataset, we first sorted the
296 deer data by read name and converted them to fastq using SAMtools sort and fastq
297 respectively. We mapped the deer data to the caribou reference genome using bwa-mem. We
298 ran ANGSD using the same procedure as for DeerSNP dataset and with the addition of a
299 caribou genome sequence obtained for this analysis (Dedato et al. 2022).

300 Population structure analyses

301 To infer admixture proportions and compute a PCA on allele frequencies based on genotype
302 likelihood data we used PCAngsd (Meisner and Albrechtsen 2018) and NGSadmix (Skotte et al.
303 2013) on the DeerSNP dataset. NGSadmix was run with K = 1 up to K =7 and the best K value
304 was determined with CLUMPAK (Kopelman et al. 2015) following the Evanno method (Evanno
305 et al. 2005) with 7 replicates for each K value. We investigated the correlation between the

306 distribution on each PC and each individual's range in R with a linear model including the
307 species and either latitude or longitude, and a Pearson's correlation.

308 Genome-wide introgression

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

322 In-windows introgression

323 The D statistic is sensitive to genomic variation (π), and should not be used to determine
324 introgression on a small scale (Martin et al. 2015). The f_d statistic, proposed by Martin et al.
325 (2015), is less dependent on diversity than D and therefore allows to detect potentially
326 introgressed regions of the genome. We used the python script ABBABABAwindows.py (Martin
327 2021) to estimate in D and f_d in 50Kbp windows to detect potentially introgressed loci in our
328 DeerCAR dataset on two comparisons: 1) P1 = WTD allopatry, P2 = WTD sympatry, P3 = MD

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

334 Genome scans

335 To detect islands of divergence between WTD & MD, we used the python script
336 popgenWindows.py (Martin 2021) to estimate individual heterozygosity, F_{ST}, D_{xy} and π in 50Kbp
337 windows (note we filtered out scaffolds shorter than 50kb). The accurate computation of D_{xy}
338 and π require a dataset including the invariant sites, we performed this analysis on the DeerALL
339 dataset. To detect the outlier loci, we based our approach on the four selection scenarios
340 developed in Shang et al. (2021): I) Divergence with gene flow that presents as high F_{ST} and D_{xy}
341 but low π; II) allopatric selection that shows high F_{ST}, low π and stable D_{xy}; III) background
342 selection that presents as high F_{ST} but low π and D_{xy}; and IV) balancing selection that displays
343 as low F_{ST} but high π and D_{xy}. Thresholds were set as the upper or lower 5% for a high or low
344 criteria, and between 45 and 55% for a stable condition. Outliers were flagged as belonging to
345 one of the four scenarios when they met the specific criteria shown in Table 2. For each
346 scenario, we identified genes present in outlier regions by comparing the coordinates of the
347 outlier windows and the WTD annotation file in BEDTools intersect. We extracted the gene
348 names from those regions as well as those from their associated scaffolds. The gene list from
349 outlier windows was uploaded to ShinyGO for enrichment analysis with the scaffold's gene list
350 as background information. We downloaded the enrichment analysis results for the molecular
351 function, biological process, and cellular component GOs and analysed the results in R. We

also uploaded the gene list to UniProt's Retrieve/ID mapping tool (Pundir et al. 2016) and downloaded the information for visual inspection. All results were analysed in R version 4.1.0 "Camp Pontanezen" (R Core Team 2021) .

Data availability

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

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; ComputeCanada Resources for Research Groups; Canadian Foundation for Innovation: John R. Evans Leaders Fund and the Ontario Early Researcher Award. 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 precious comments on the manuscript.

References

Abbott R, Albach D, Ansell S, Arntzen JW, Baird SJE, Bierne N, Boughman J, Brelsford A, Buerkle CA, Buggs R, et al. 2013. Hybridization and speciation. *J. Evol. Biol.* 26:229–246.

Adams KP, Hamilton RJ. 2011. Management History. In: Hewitt DG, editor. Biology and management of white-tailed deer. Boca Raton: CRC Press. Available from: http://site.ebrary.com/id/11007093

Adavoudi R, Pilot M. 2021. Consequences of Hybridization in Mammals: A Systematic Review. *Genes* 13:50.

Aguilar A, Roemer G, Debenham S, Binns M, Garcelon D, Wayne RK. 2004. High MHC diversity maintained by balancing selection in an otherwise genetically monomorphic mammal. *Proc. Natl. Acad. Sci. U.S.A.* 101:3490–3494.

Alminas OSV, Heffelfinger JR, Statham MJ, Latch EK. 2021. Phylogeography of Cedros and Tiburón Island Mule Deer in North America's Desert Southwest.Murphy WJ, editor. *Journal of Heredity* 112:260–275.

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 RG, Smith NS. 1977. Ecological Relationships between Mule Deer and White-Tailed Deer in Southeastern Arizona. *Ecological Monographs* 47:255–277.

Avise JC, Walker D, Johns GC. 1998. Speciation durations and Pleistocene effects on vertebrate phylogeography. *Proc. R. Soc. Lond. B* 265:1707–1712.

Bernatchez L, Landry C. 2003. MHC studies in nonmodel vertebrates: what have we learned about natural selection in 15 years? *Journal of Evolutionary Biology* 16:363–377.

Berry SL, Shipley LA, Long RA, Loggers C. 2019. Differences in dietary niche and foraging behavior of sympatric mule and white-tailed deer. *Ecosphere* [Internet] 10. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1002/ecs2.2815

Blank DA, Yang W. 2017. Mother-young recognition in goitered gazelle during hiding period. *Behavioural Processes* 142:21–28.

Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30:2114–2120.

Booker TR, Yeaman S, Whitlock MC. 2021. Global adaptation complicates the interpretation of genome scans for local adaptation. *Evolution Letters* 5:4–15.

Bradley RD, Bryant FC, Bradley LC, Haynie ML, Baker RJ. 2003. IMPLICATIONS OF HYBRIDIZATION BETWEEN WHITE-TAILED DEER AND MULE DEER. *The Southwestern Naturalist* 48:7.

Broad Institute, GitHub Repository. 2019. “Picard Toolkit.” Available from: http://broadinstitute.github.io/picard/

Brunjes KJ, Ballard WB, Humphrey MH, Harwell F, Mcintyre NE, Krausman PR, Wallace MC. 2006. Habitat Use by Sympatric Mule and White-Tailed Deer in Texas. *Journal of Wildlife Management* 70:1351–1359.

Budd K, Berkman LK, Anderson M, Koppelman J, Eggert LS. 2018. Genetic structure and recovery of white-tailed deer in Missouri: Genetic Structure of Deer in Missouri. *Jour. Wild. Mgmt.* 82:1598–1607.

Burri R. 2017. Interpreting differentiation landscapes in the light of long-term linked selection: DIFFERENTIATION AND LONG-TERM LINKED SELECTION. *Evolution Letters* 1:118–131.

Cambronne A. 2013. Deerland: America’s hunt for ecological balance and the essence of wildness. Available from: http://www.freading.com/ebooks/details/r:download/MDAwMDA5LTEwOTA1NDYz

Campbell CR, Poelstra JW, Yoder AD. 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* 124:561–583.

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* 20:794.

Carr SM, Ballinger SW, Derr JN, Blankenship LH, Bickham JW. 1986. Mitochondrial DNA analysis of hybridization between sympatric white-tailed deer and mule deer in west Texas. *Proceedings of the National Academy of Sciences* 83:9576–9580.

Carr SM, Hughes GA. 1993. Direction of Introgressive Hybridization between Species of North American Deer (Odocoileus) as Inferred from Mitochondrial-Cytochrome-b Sequences. *Journal of Mammalogy* 74:331–342.

Cathey JC, Bickham JW, Patton JC. 1998. INTROGRESSIVE HYBRIDIZATION AND NONCONCORDANT EVOLUTIONARY HISTORY OF MATERNAL AND PATERNAL LINEAGES IN NORTH AMERICAN DEER. *Evolution* 52:1224–1229.

Chafin TK, Zbinden ZD, Douglas MR, Martin BT, Middaugh CR, Gray MC, Ballard JR, Douglas ME. 2021. Spatial population genetics in heavily managed species: Separating patterns of historical translocation from contemporary gene flow in white-tailed deer. *Evol Appl* 14:1673–1689.

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* 174:1151–1158.

Cherry MJ, Barton BT. 2017. Effects of wind on predator-prey interactions. *Food Webs* 13:92–97.

Colella JP, Lan T, Talbot SL, Lindqvist C, Cook JA. 2021. Whole-genome resequencing reveals persistence of forest-associated mammals in Late Pleistocene refugia along North America's North Pacific Coast. *J. Biogeogr.* 48:1153–1169.

Combe FJ, Jaster L, Ricketts A, Haukos D, Hope AG. 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.

Cronin MA. 1991. Mitochondrial and nuclear genetic relationships of deer (*Odocoileus* spp.) in western North America. *Can. J. Zool.* 69:1270–1279.

Cronin MA, Vyse ER, Cameron DG. 1988. Genetic Relationships between Mule Deer and White-Tailed Deer in Montana. *The Journal of Wildlife Management* 52:320.

Cruickshank TE, Hahn MW. 2014. Reanalysis suggests that genomic islands of speciation are due to reduced diversity, not reduced gene flow. *Mol Ecol* 23:3133–3157.

Cutter AD. 2013. Integrating phylogenetics, phylogeography and population genetics through genomes and evolutionary theory. *Molecular Phylogenetics and Evolution* 69:1172–1185.

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 Available from: 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. Preprints Available from: https://www.authorea.com/users/407174/articles/552124-demographic-history-and-conservation-genomics-of-caribou-rangifer-tarandus-in-qu%C3%A9bec?commit=d3fe88936f0714b1711f9dba33bc149835a4aba3

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

Derr JN. 1991. Genetic Interactions between White-Tailed and Mule Deer in the Southwestern United States. *The Journal of Wildlife Management* 55:228.

Derr JN, Hale DW, Ellsworth DL, Bickham JW. 1991. Fertility in an F1 male hybrid of white-tailed deer (Odocoileus virginianus) \m=x\mule deer (O. hemionus). :7.

Deutsch JC, Nefdt RJC. 1992. Olfactory cues influence female choice in two lek-breeding antelopes. *Nature* 356:596–598.

DeYoung RW, Demarais S, Gee KL, Honeycutt RL, Hellickson MW, Gonzales RA. 2009. Molecular Evaluation of the White-tailed Deer (Odocoileus Virginianus) Mating System. *Journal of Mammalogy* 90:946–953.

DeYoung RW, Demarais S, Honeycutt RL, Gee KL, Gonzales RA. 2006. Social Dominance and Male Breeding Success in Captive White-Tailed Deer. *Wildlife Society Bulletin* 34:131–136.

Deyoung RW, Demarais S, Honeycutt RL, Rooney AP, Gonzales RA, Gee KL. 2003. Genetic consequences of white-tailed deer (*Odocoileus virginianus*) restoration in Mississippi. *Molecular Ecology* 12:3237–3252.

Ditchkoff SS. 2011. Anatomy and Physiology. In: Hewitt DG, editor. Biology and management of white-tailed deer. Boca Raton: CRC Press. Available from: 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* 14:1154–1166.

Durand EY, Patterson N, Reich D, Slatkin M. 2011. Testing for Ancient Admixture between Closely Related Populations. *Molecular Biology and Evolution* 28:2239–2252.

Dussex N, Alberti F, Heino MT, Olsen R-A, van der Valk T, Ryman N, Laikre L, Ahlgren H, Askeyev IV, Askeyev OV, et al. 2020. Moose genomes reveal past glacial demography and the origin of modern lineages. *BMC Genomics* 21:854.

Ellsworth DL, Honeycutt RL, Silvy NJ, Bickham JW, Klimstra WD. 1994. HISTORICAL BIOGEOGRAPHY AND CONTEMPORARY PATTERNS OF MITOCHONDRIAL DNA VARIATION IN WHITE-TAILED DEER FROM THE SOUTHEASTERN UNITED STATES. *Evolution* 48:122–136.

Evanno G, Regnaut S, Goudet J. 2005. Detecting the number of clusters of individuals using the software structure: a simulation study. *Mol Ecol* 14:2611–2620.

Ewels P, Magnusson M, Lundin S, Käller M. 2016. MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* 32:3047–3048.

Feder JL. 2012. The genomics of speciation-with-gene-flow. 28:9.

Gilbert C, Ropiquet A, Hassanin A. 2006. Mitochondrial and nuclear phylogenies of Cervidae (Mammalia, Ruminantia): Systematics, morphology, and biogeography. *Molecular Phylogenetics and Evolution* 40:101–117.

Green RE, Krause J, Briggs AW, Maricic T, Stenzel U, Kircher M, Patterson N, Li H, Zhai W, Fritz MHY, et al. 2010. A Draft Sequence of the Neandertal Genome. *Science* 328:710–722.

Hewitt G. 2000. The genetic legacy of the Quaternary ice ages. *Nature* 405:907–913.

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. *Anim Cogn* 24:99–106.

Hornbeck GE, Mahoney JM. 2000. Introgressive Hybridization of Mule Deer and White-Tailed Deer in Southwestern Alberta. :5.

Hughes GA, Carr SM. 1993. Reciprocal hybridization between white-tailed deer (*Odocoileus virginianus*) and mule deer (*O . hemionus*) in western Canada: evidence from serum albumin and mtDNA sequences. *Can. J. Zool.* 71:524–530.

Iacolina L, Corlatti L, Buzan E, Safner T, Šprem N. 2019. Hybridisation in European ungulates: an overview of the current status, causes, and consequences. *Mam Rev* 49:45–59.

Ito T, Hayakawa T, Suzuki–Hashido N, Hamada Y, Kurihara Y, Hanya G, Kaneko A, Natsume T, Aisu S, Honda T, et al. 2021. Phylogeographic history of Japanese macaques. *J Biogeogr*:jbi.14087.

IUCN. 2015a. Odocoileus virginianus: Gallina, S. and Lopez Arevalo, H.: The IUCN Red List of Threatened Species 2016: e.T42394A22162580. Available from: http://www.iucnredlist.org/details/42394/0

IUCN. 2015b. Odocoileus hemionus: Sanchez Rojas, G. and Gallina Tessaro, S.: The IUCN Red List of Threatened Species 2016: e.T42393A22162113. Available from: http://www.iucnredlist.org/details/42393/0

Jones PD, Strickland BK, Demarais S, DeYoung RW. 2011. Inconsistent association of male body mass with breeding success in captive white-tailed deer. *J Mammal* 92:527–533.

Keller M, Lévy F. 2012. The main but not the accessory olfactory system is involved in the processing of socially relevant chemosignals in ungulates. *Front. Neuroanat.* [Internet] 6. Available from: http://journal.frontiersin.org/article/10.3389/fnana.2012.00039/abstract

Kinoshita E, Kosintsev PA, Abramov AV, Solovyev VA, Saveljev AP, 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* 129:594–602.

Kopelman NM, Mayzel J, Jakobsson M, Rosenberg NA, Mayrose I. 2015. CLUMPAK : a program for identifying clustering modes and packaging population structure inferences across *K*. *Mol Ecol Resour* 15:1179–1191.

Korneliussen TS, Albrechtsen A, Nielsen R. 2014. ANGSD: Analysis of Next Generation Sequencing Data. *BMC Bioinformatics* 15:356.

Kuijper DPJ, 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?Sueur C, editor. *PLoS ONE* 9:e84607.

Latch EK, Heffelfinger JR, Fike JA, RHODES Jr OE. 2009. Species-wide phylogeography of North American mule deer (*Odocoileus hemionus*): cryptic glacial refugia and postglacial recolonization. *Molecular Ecology* 18:1730–1745.

Latch EK, Reding DM, Heffelfinger JR, Alcalá-Galván CH, Rhodes OE. 2014. Range-wide analysis of genetic structure in a widespread, highly mobile species (*Odocoileus hemionus*) reveals the importance of historical biogeography. *Mol Ecol* 23:3171–3190.

Lavretsky P, DaCosta JM, Sorenson MD, McCracken KG, Peters JL. 2019. ddRAD-seq data reveal significant genome-wide population structure and divergent genomic regions that distinguish the mallard and close relatives in North America. *Mol Ecol* 28:2594–2609.

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

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* 25:2078–2079.

Lingle S. 1992. Escape Gaits of White-Tailed Deer, Mule Deer and Their Hybrids: Gaits Observed and Patterns of Limb Coordination. *Behaviour* 122:153–181.

Lingle S. 2001. Anti-Predator Strategies and Grouping Patterns in White-Tailed Deer and Mule Deer. *Ethology* 107:295–314.

Lingle S. 2003. Group composition and cohesion in sympatric white-tailed deer and mule deer. *Can. J. Zool.* 81:1119–1130.

Liu Q, Mishra M, Saxena AS, Wu H, Qiu Y, Zhang X, You X, Ding S, Miyamoto MM. 2021. Balancing selection maintains ancient polymorphisms at conserved enhancers for the olfactory receptor genes of a Chinese marine fish. *Mol Ecol* 30:4023–4038.

Mallet J. 2005. Hybridization as an invasion of the genome. *Trends in Ecology & Evolution* 20:229–237.

Marques DA, Lucek K, Sousa VC, Excoffier L, Seehausen O. 2019. Admixture between old lineages facilitated contemporary ecological speciation in Lake Constance stickleback. *Nat Commun* 10:4240.

Martin S. 2021. genomics_general. *GitHub* [Internet]. Available from: https://github.com/simonhmartin/genomics_general

Martin SH, Davey JW, Jiggins CD. 2015. Evaluating the Use of ABBA–BABA Statistics to Locate Introgressed Loci. *Molecular Biology and Evolution* 32:244–257.

Mayr E. 1942. Systematics and the Origin of Species.

McDonald JS, Miller KV, Association QDM, Resources DBWS of F. 2004. A History of White-tailed Deer Restocking in the United States, 1878 to 2004. Quality Deer Management Assoc. Available from: 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, et al. 2010. The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research* 20:1297–1303.

Meisner J, Albrechtsen A. 2018. Inferring Population Structure and Admixture Proportions in Low-Depth NGS Data. *Genetics* 210:719–731.

Momigliano P, Jokinen H, Fraimout A, Florin A-B, Norkko A, Merilä J. 2017. Extraordinarily rapid speciation in a marine fish. :6.

Morina DL, Demarais S, Strickland BK, Larson JE. 2018. While males fight, females choose: male phenotypic quality informs female mate choice in mammals. *Animal Behaviour* 138:69–74.

Moscarella RA, Aguilera M, Escalante AA. 2003. PHYLOGEOGRAPHY, POPULATION STRUCTURE, AND IMPLICATIONS FOR CONSERVATION OF WHITE-TAILED DEER (ODOCOILEUS VIRGINIANUS) IN VENEZUELA. *Journal of Mammalogy* 84:1300–1315.

NatureServe. 2021a. NatureServe Explorer. *NatureServe Mule Deer* [Internet]. Available from: https://explorer.natureserve.org/Taxon/ELEMENT_GLOBAL.2.101365/Odocoileus_hemionus

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

Newbolt CH, Acker PK, Neuman TJ, Hoffman SI, Ditchkoff SS, Steury TD. 2017. Factors influencing reproductive success in male white-tailed deer: Male Breeding Success in Deer. *Jour. Wild. Mgmt.* 81:206–217.

Nosil P. 2012. Ecological speciation. Oxford ; New York: Oxford University Press

Pedersen BS, Quinlan AR. 2018. Mosdepth: quick coverage calculation for genomes and exomes.Hancock J, editor. *Bioinformatics* 34:867–868.

Peres TM, 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.

Pickrell JK, Pritchard JK. 2012. Inference of Population Splits and Mixtures from Genome-Wide Allele Frequency Data.Tang H, editor. *PLoS Genet* 8:e1002967.

Pierini F, Lenz TL. 2018. Divergent Allele Advantage at Human MHC Genes: Signatures of Past and Ongoing Selection.Wilke C, editor. *Molecular Biology and Evolution* 35:2145–2158.

Piertney SB, Oliver MK. 2006. The evolutionary ecology of the major histocompatibility complex. *Heredity* 96:7–21.

Pitra C, Fickel J, Meijaard E, Groves C. 2004. Evolution and phylogeny of old world deer. *Molecular Phylogenetics and Evolution* 33:880–895.

Poelstra JW, Vijay N, Bossu CM, Lantz H, Ryll B, Muller I, Baglione V, Unneberg P, Wikelski M, Grabherr MG, et al. 2014. The genomic landscape underlying phenotypic integrity in the face of gene flow in crows. *Science* 344:1410–1414.

Pundir S, Martin MJ, O'Donovan C, The UniProt Consortium. 2016. UniProt Tools. *Current Protocols in Bioinformatics* [Internet] 53. Available from: https://onlinelibrary.wiley.com/doi/10.1002/0471250953.bi0129s53

Quinlan AR, Hall IM. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 26:841–842.

R Core Team. 2021. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing Available from: https://www.R-project.org/

Ragavan P, Zhou R, Ng WL, Rana TS, Mageswaran T, Mohan PM, Saxena A. 2017. Natural hybridization in mangroves – an overview. *Botanical Journal of the Linnean Society* 185:208–224.

Ranslow AN, Richter JP, Neuberger T, Van Valkenburgh B, Rumple CR, Quigley AP, Pang B, Krane MH, Craven BA. 2014. Reconstruction and Morphometric Analysis of the Nasal Airway of the White-Tailed Deer (*O docoileus virginianus*) and Implications Regarding Respiratory and Olfactory Airflow: White-Tailed Deer Nasal Anatomy and Function. *Anat. Rec.* 297:2138–2147.

Ravinet M, Faria R, Butlin RK, Galindo J, Bierne N, Rafajlović M, Noor MAF, Mehlig B, Westram AM. 2017. Interpreting the genomic landscape of speciation: a road map for finding barriers to gene flow. *J. Evol. Biol.* 30:1450–1477.

Rundle HD, Nosil P. 2005. Ecological speciation: Ecological speciation. *Ecology Letters* 8:336–352.

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. *Evol Appl*: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*:g3.200838.2018.

Schumer M, Cui R, Rosenthal GG, Andolfatto P. 2015. Reproductive Isolation of Hybrid Populations Driven by Genetic Incompatibilities.Payseur BA, editor. *PLoS Genet* 11:e1005041.

Servedio MR. 2004. THE EVOLUTION OF PREMATING ISOLATION: LOCAL ADAPTATION AND NATURAL AND SEXUAL SELECTION AGAINST HYBRIDS. *Evolution* 58:913–924.

Shafer ABA, Cullingham CI, Côté SD, Coltman DW. 2010. Of glaciers and refugia: a decade of study sheds new light on the phylogeography of northwestern North America. *Molecular Ecology* 19:4589–4621.

Shafer ABA, Wolf JBW. 2013. Widespread evidence for incipient ecological speciation: a meta-analysis of isolation-by-ecology.Wiens J, editor. *Ecol Lett* 16:940–950.

Shang H, Rendón-Anaya M, Paun O, Field DL, Hess J, Vogl C, Liu J, Ingvarsson PK, Lexer C, Leroy T. 2021. Conserved genomic landscapes of differentiation across *Populus* speciation continuum. Evolutionary Biology Available from: http://biorxiv.org/lookup/doi/10.1101/2021.08.26.457771

Skotte L, Korneliussen TS, Albrechtsen A. 2013. Estimating Individual Admixture Proportions from Next Generation Sequencing Data. *Genetics* 195:693–702.

Soraggi S, Wiuf C, Albrechtsen A. 2018. Powerful inference with the D-statistic on low-coverage whole-genome data. *G3: Genes, Genomes, Genetics* 8.

Stankowski S, Ravinet M. 2021. Defining the speciation continuum. *Evolution* 75:1256–1273.

Stubblefield SS, Warren RJ, Murphy BR. 1986. Hybridization of Free-Ranging White-Tailed and Mule Deer in Texas. *The Journal of Wildlife Management* 50:688.

Tarasov A, Vilella AJ, Cuppen E, Nijman IJ, Prins P. 2015. Sambamba: fast processing of NGS alignment formats. *Bioinformatics* 31:2032–2034.

Taylor SA, Larson EL. 2019. Insights from genomes into the evolutionary importance and prevalence of hybridization in nature. *Nat Ecol Evol* 3:170–177.

UniProt. 2022a. Q9Y5E5 · PCDB4_HUMAN. *UniProt beta* [Internet]. Available from: https://beta.uniprot.org/uniprotkb/Q9Y5E5/entry

UniProt. 2022b. Q15057 · ACAP2_HUMAN. *UniProt beta* [Internet]. Available from: https://beta.uniprot.org/uniprotkb/Q15057/entry

Whittaker DG, Lindzey FG. 2004. Habitat use patterns of sympatric deer species on Rocky Mountain Arsenal, Colorado. *Wildlife Society Bulletin* 32:1114–1123.

Wikenros C, Kuijper DPJ, Behnke R, Schmidt K. 2015. Behavioural responses of ungulates to indirect cues of an ambush predator. *Behav* 152:1019–1040.

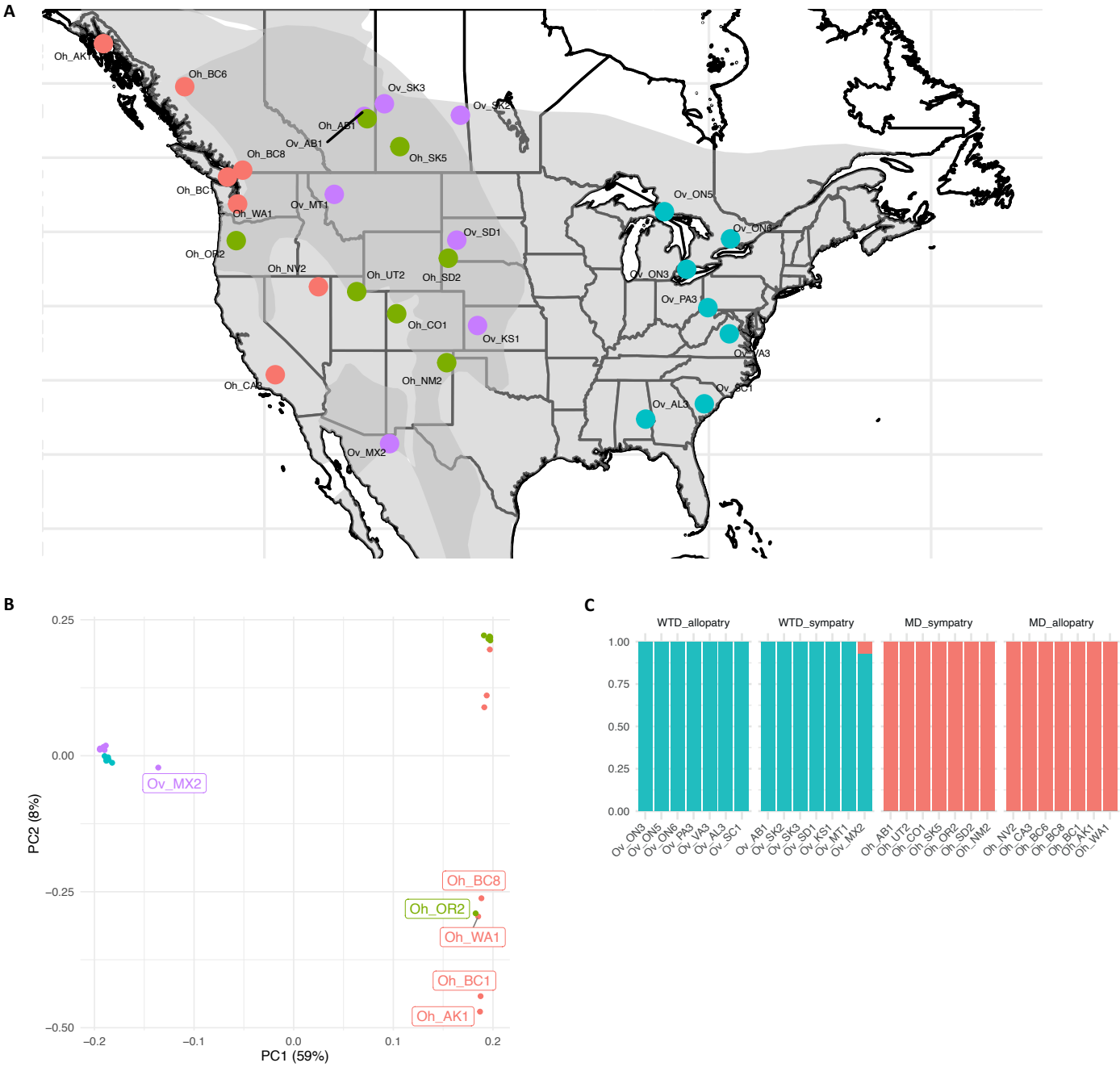
Wishart WD. 1980. Hybrids of White-Tailed and Mule Deer in Alberta. *Journal of Mammalogy* 61:716–719.

Wright CA, McRoberts JT, Rota CT, Wiskirchen KH, Keller BJ, Millspaugh JJ. 2021. Female White-tailed deer (*Odocoileus virginianus*) Behavior During Pregnancy, Parturition, and Lactation in 2 Contrasting Ecoregions. *Wildl. Soc. bull.*:wsb.1207.

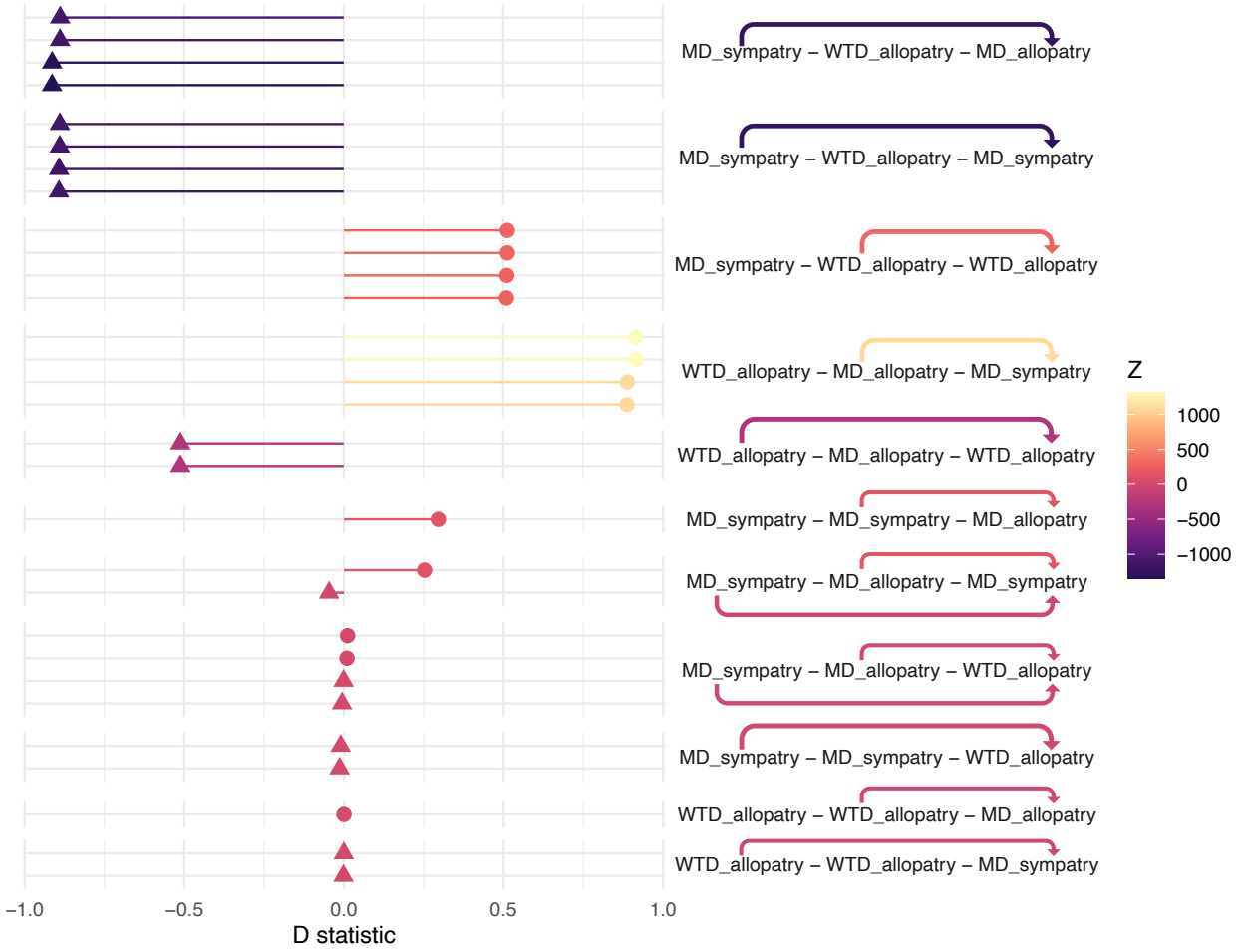
Wright EA, Roberts EK, Platt RN, Bayouth JV, Conway WC, Bradley RD. 2022. Mitochondrial capture and subsequent genetic divergence generates a novel haplogroup: evidence from ancient and ongoing hybridization in mule and white-tailed deer.Jezkova T, editor. *Journal of Mammalogy*:gyab156.

Zagata MD, Haugen AO. 1974. Influence of Light and Weather on Observability of Iowa Deer. *The Journal of Wildlife Management* 38:220.

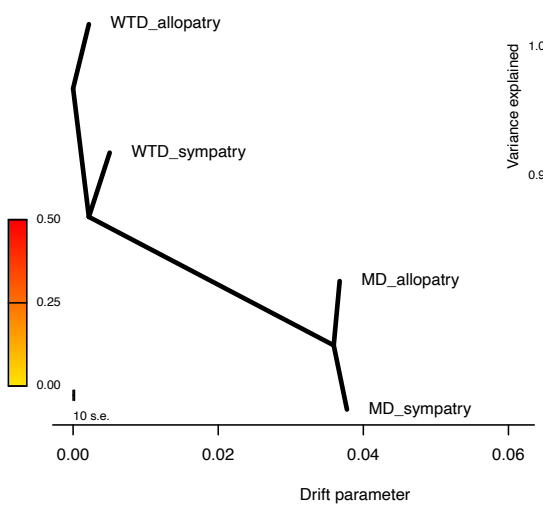
Zhang P, Huang K, Zhang B, Dunn DW, Chen D, Li F, Qi X, Guo S, Li B. 2018. High polymorphism in MHC-DRB genes in golden snub-nosed monkeys reveals balancing selection in small, isolated populations. *BMC Evol Biol* 18:29.

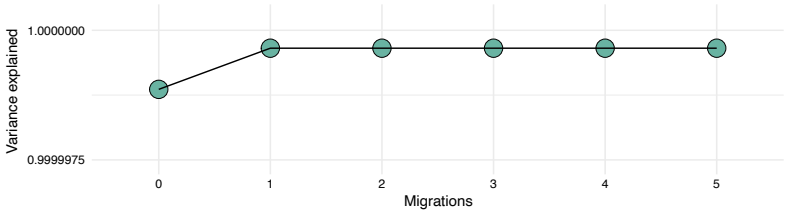

A
Oh_AK1
Oh_BC6
Ov_SK3
Ov_SK2
Oh_AB1
Ov_AB1
Oh_BC8
Oh_SK5
Oh_BC1
Oh_WA1
Ov_MT1
Ov_ON5
Ov_ON6
Ov_SD1
Oh_OR2
Oh_NV2
Oh_UT2
Oh_SD2
Ov_ON3
Ov_PA3
Oh_CO1
Ov_KS1
Ov_VA3
Oh_CA3
Oh_NM2
Ov_SC1
Ov_AL3
Ov_MX2
B
PC2 (8%)
PC1 (59%)
0.25
0.00
−0.25
−0.50
−0.2
−0.1
0.0
0.1
0.2
Ov_MX2
Oh_BC8
Oh_OR2
Oh_WA1
Oh_BC1
Oh_AK1
C
WTD_allopatry
WTD_sympatry
MD_sympatry
MD_allopatry
1.00
0.75
0.50
0.25
0.00
Ov_ON3
Ov_ON5
Ov_ON6
Ov_PA3
Ov_VA3
Ov_AL3
Ov_SC1
Ov_AB1
Ov_SK2
Ov_SK3
Ov_SD1
Ov_KS1
Ov_MT1
Ov_MX2
Oh_AB1
Oh_UT2
Oh_CO1
Oh_SK5
Oh_OR2
Oh_SD2
Oh_NM2
Oh_NV2
Oh_CA3
Oh_BC6
Oh_BC8
Oh_BC1
Oh_AK1
Oh_WA1

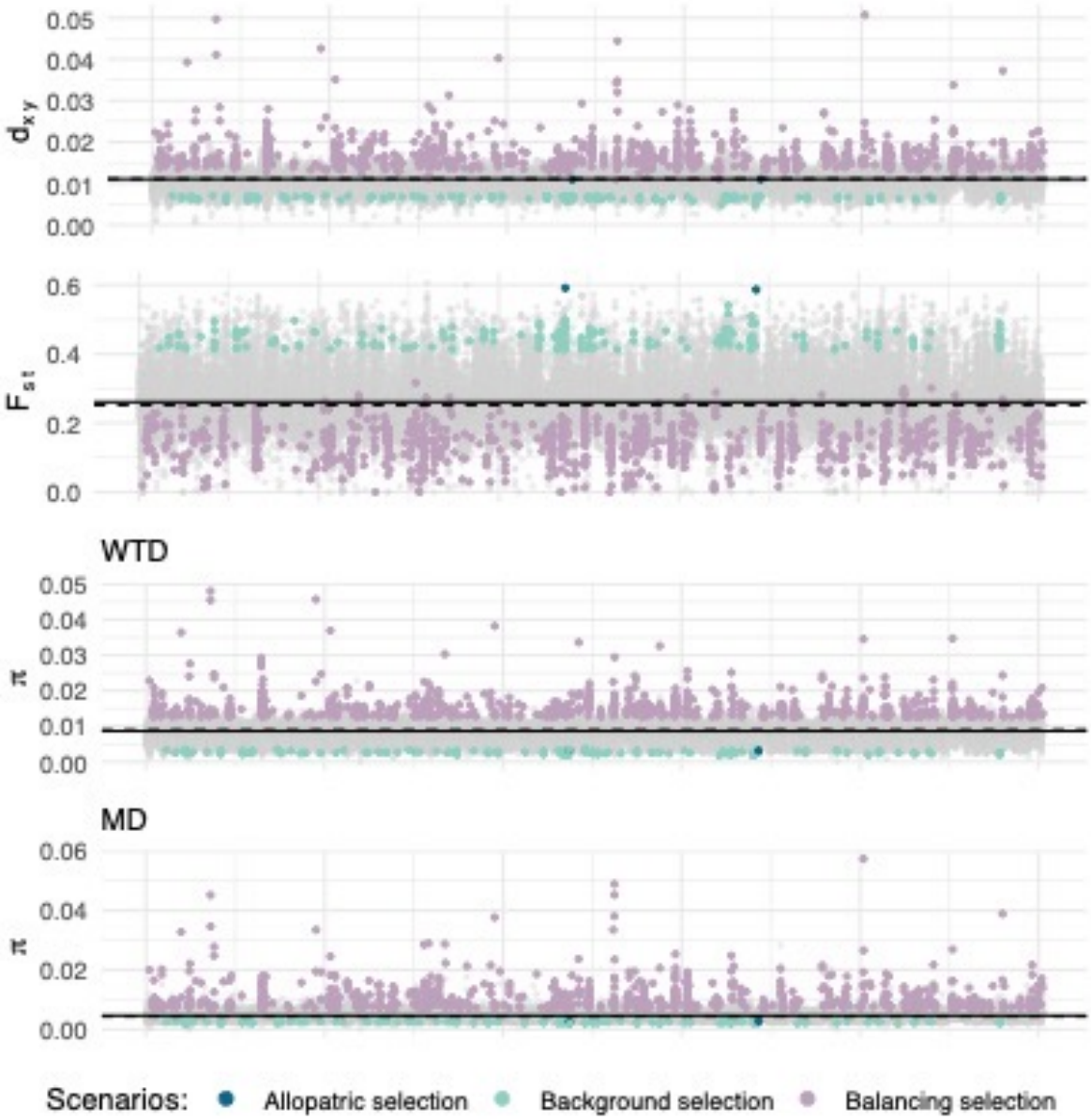
A

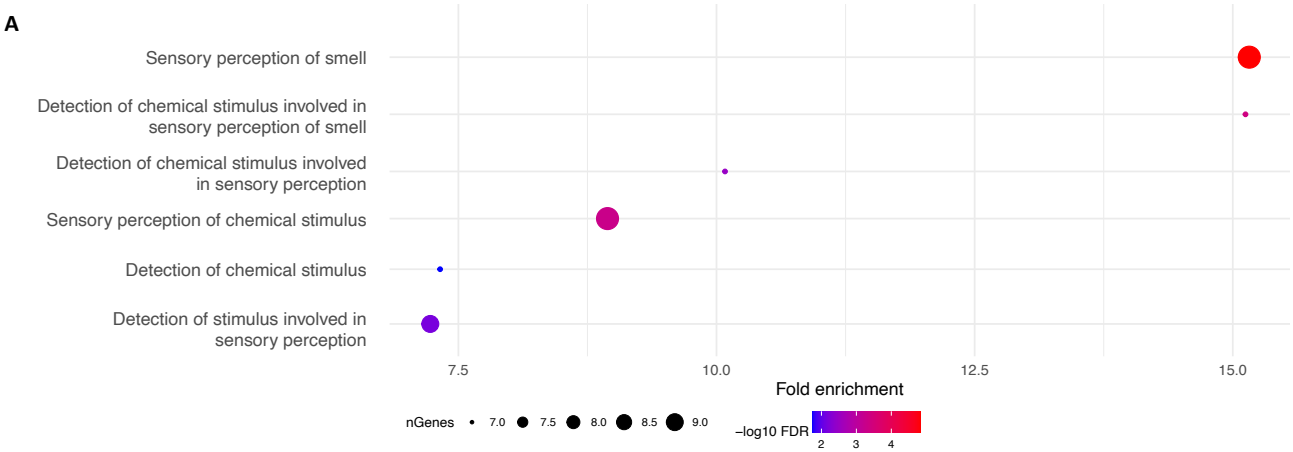

MD_sympatry – WTD_allopatry – MD_allopatry
MD_sympatry – WTD_allopatry – MD_sympatry
MD_sympatry – WTD_allopatry – WTD_allopatry
WTD_allopatry – MD_allopatry – MD_sympatry
WTD_allopatry – MD_allopatry – WTD_allopatry
MD_sympatry – MD_sympatry – MD_allopatry
MD_sympatry – MD_allopatry – MD_sympatry
MD_sympatry – MD_allopatry – WTD_allopatry
MD_sympatry – MD_sympatry – WTD_allopatry
WTD_allopatry – WTD_allopatry – MD_allopatry
WTD_allopatry – WTD_allopatry – MD_sympatry
Z
1000
500
0
−500
−1000
−1.0
−0.5
0.0
0.5
1.0
D statistic

B


WTD_allopatry
WTD_sympatry
MD_allopatry
MD_sympatry
0.50
0.25
0.00
10 s.e.
0.00
0.02
0.04
0.06
Drift parameter


Variance explained
1.0000000
0.9999975
0
1
2
3
4
5
Migrations


0.05
0.04
0.03
0.02
0.01
0.00
d_{xy}
0.6
0.4
0.2
0.0
F_{st}
WTD
0.05
0.04
0.03
0.02
0.01
0.00
π
MD
0.06
0.04
0.02
0.00
π
Scenarios:
Allopatric selection
Background selection
Balancing selection


A
Sensory perception of smell
Detection of chemical stimulus involved in sensory perception of smell
Detection of chemical stimulus involved in sensory perception
Sensory perception of chemical stimulus
Detection of chemical stimulus
Detection of stimulus involved in sensory perception
7.5
10.0
12.5
15.0
Fold enrichment
nGenes
7.0
7.5
8.0
8.5
9.0
−log10 FDR
2
3
4

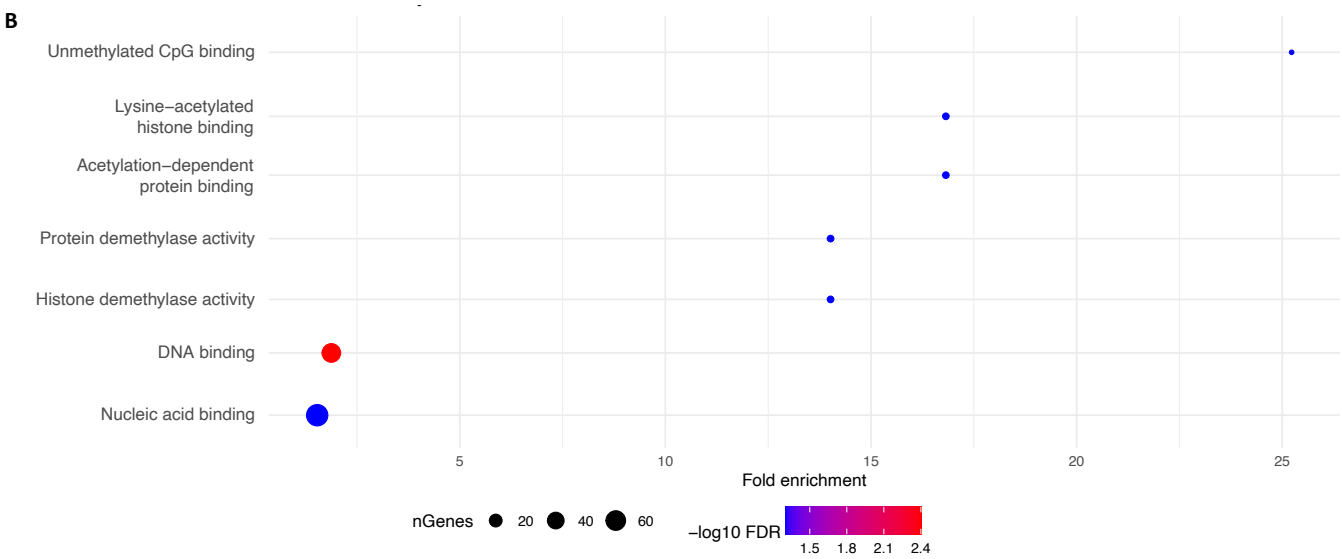

B
Unmethylated CpG binding
Lysine−acetylated histone binding
Acetylation−dependent protein binding
Protein demethylase activity
Histone demethylase activity
DNA binding
Nucleic acid binding
5
10
15
20
25
Fold enrichment
nGenes
20
40
60
−log10 FDR
1.5
1.8
2.1
2.4

Figure legends

Figure 1: Deer individuals in the study. (A) Sampling locations, colours represent the different groups: WTD allopatry (blue), WTD sympatry (violet), MD sympatry (green), MD allopatry (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: Introgression analyses. (A) ABBA-BABA analysis between individuals, we only show their assigned populations for clarity. Colour gradient represents Z score, excess of ABBA depicted as points, excess of BABA shown as triangles, arrows illustrate which populations are involved in gene flow. (B) Maximum likelihood tree inferred by Treemix for 0 migration (right) and the variance explained by each migration event (left).

Figure 3: 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 4: Gene enrichment analysis from outlier windows. (A) Biological process GOs enriched in the Balancing selection outlier windows. (B) Molecular function GOs enriched in the Background selection windows. Colour represents -log10 FDR, size indicates the number of genes in our dataset that falls under given GO category.

Tables

Table 1: Sample information

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

Table 2: Different scenario and their thresholds for F_{st}, π and D_{xy} percentiles

Scenario	F_{st}	π	D_{xy}	Panel in Fig.1A Shang et al. (2021)
Divergence with gene flow (DwGF)	> 0.95	< 0.05	> 0.95	1
Allopatric selection (AS)	> 0.95	< 0.05	Between 0.45 & 0.55	2
Background selection (BGS)	> 0.95	< 0.05	< 0.05	3
Balancing selection (BLS)	< 0.05	> 0.95	> 0.95	4