

1 **Complex admixture preceded and followed the extinction of wisent in the wild**

2

3 Karolina Węcsek^{1,2}, Stefanie Hartmann², Johanna L. A. Paijmans², Ulrike Taron², Georgios

4 Xenikoudakis², James A. Cahill³, Peter D. Heintzman³, Beth Shapiro^{3,4}, Gennady

5 Baryshnikov⁵, Aleksei N. Bunevich⁶, Jennifer J. Crees⁷, Roland Dobosz^{8,9}, Ninna

6 Manaserian¹⁰, Henryk Okarma¹¹, Małgorzata Tokarska¹², Samuel T. Turvey⁷, Jan M.

7 Wójcik¹², Waldemar Żyła⁸, Jacek M. Szymura¹, Michael Hofreiter², Axel Barlow²

8

9 ¹ Department of Comparative Anatomy, Institute of Zoology, Jagiellonian University,

10 Kraków, Poland

11 ² Institute for Biochemistry and Biology, University of Potsdam, Potsdam, Germany

12 ³ Department of Ecology and Evolutionary Biology, University of California, Santa Cruz,

13 CA, 95064, USA

14 ⁴ University of California Santa Cruz Genomics Institute, University of California, Santa

15 Cruz, CA, 95064, USA

16 ⁵ Gennady Baryshnikov, Laboratory of Theriology, Zoological Institute of the Russian

17 Academy of Sciences, Russia

18 ⁶ State National Park, Belovezhskaya Pushcha, Kamieniuki, Belarus

19 ⁷ Institute of Zoology, Zoological Society of London, Regent's Park, London NW1 4RY, UK

20 ⁸ Upper Silesian Museum, Bytom, Poland

21 ⁹ Department of Zoology, Faculty of Biology and Environmental Protection, University of

22 Silesia, Katowice, Poland

23 ¹⁰ Institute of Zoology, Armenian National Academy of Sciences, Yerevan, Armenia

24 ¹¹ Institute of Nature Conservation, Polish Academy of Sciences, Kraków, Poland

25 ¹² Mammal Research Institute, Polish Academy of Sciences, Białowieża, Poland

26 Keywords: Admixture, ancient DNA, European bison, *Bison bonasus*, domestic cattle,
27 aurochs, hybridisation

28

29 Corresponding authors:

30 Karolina Węcek

31 Department of Comparative Anatomy, Institute of Zoology, Jagiellonian University,

32 Gronostajowa 9, 30-387 Kraków, Poland

33 Fax: (+48) 12 664 51, email: k.wecek@gmail.com

34 Axel Barlow

35 Institute for Biochemistry and Biology. University of Potsdam, Karl-Liebknecht-Str. 24-25,

36 14476 Potsdam OT Golm, Germany

37 Fax: (+49) 33 197 72 921, email: axel.barlow.ab@gmail.com

38

39 Running title: Genomic evidence of wisent admixture.

40

41 **Abstract**

42 Retracing complex population processes that precede extreme bottlenecks may be impossible
43 using data from living individuals. The wisent (*Bison bonasus*), Europe's largest terrestrial
44 mammal, exemplifies such a population history, having gone extinct in the wild but
45 subsequently restored by captive breeding efforts. Using low coverage genomic data from
46 modern and historical individuals, we investigate population processes occurring before and
47 after this extinction. Analysis of aligned genomes supports the division of wisent into two
48 previously recognised subspecies, but almost half of the genomic alignment contradicts this
49 population history as a result of incomplete lineage sorting and admixture. Admixture
50 between subspecies populations occurred prior to extinction and subsequently during the
51 captive breeding program. Admixture with the *Bos* cattle lineage is also widespread but
52 results from ancient events rather than recent hybridisation with domestics. Our study
53 demonstrates the huge potential of historical genomes for both studying evolutionary histories
54 and for guiding conservation strategies.

55

56

57 **Introduction**

58 The last known wild wisent, or European bison (*Bison bonasus*), was shot and killed in 1927,
59 marking the extinction of this species in the wild (Pucek, 1991). As a result of an intensive
60 captive breeding program and a series of re-establishments, today the species again occupies
61 part of its former range in Central and Eastern Europe. The total population of free-ranging
62 wisent now stands at 5,553 (Raczyński, 2014), and the International Union for Conservation
63 of Nature no longer considers the wisent as an endangered species (IUCN 2008).

64

65 The decline of the original wisent population was protracted, and their subsequent restitution
66 has been complex. In historical times, wisent ranged extensively across semi-open habitats
67 and broadleaved, mixed and coniferous forests in Western Europe, from what is today France
68 in the west to the Volga River and the Caucasus in the east, with the northernmost range
69 limits around 60° north (Fig. 1; Kuemmerle et al., 2011; Kerley et al., 2012; Bocherens et al.,
70 2015). However, ongoing habitat fragmentation and overhunting eradicated most populations.
71 By the end of the 19th century, there were only two populations of wisent left in the wild that
72 were assigned to separate subspecies: in Białowieża Forest (Lowland wisent, *B. b. bonasus*)
73 and in the western Caucasus Mountains (Caucasian wisent, *B. b. caucasicus*). Finally, even
74 these populations collapsed; the last wild Lowland wisent was shot in Poland in 1919
75 followed by the last Caucasian animal in 1927 (Pucek, 1991).

76

77 In 1924, the captive population consisted of only 54 individuals (29 males and 25 females).
78 However, detailed analysis of pedigrees (Olech 2009, Slatis 1960) has shown that some
79 individuals in this population were the direct descendents of others, and that some individuals
80 left no modern descendents. Thus, the true founding population of wisent was considerably
81 smaller, and is thought to comprise of just 12 individuals (Slatis 1960). All but one of these
82 12 founders were Lowland wisent, almost half of which came from a population established
83 in 1865 in Pless (now Pszczyna, Poland). The remaining founder was a Caucasian wisent bull
84 named M100 KAUKASUS, that represented the last surviving pure Caucasian wisent in
85 captivity. The modern herds that are derived from this founding population are managed as
86 two separate genetic lines. The Lowland line (L) derives from 7 Lowland founders (4 males
87 and 3 females), and is thus considered to represent a pure Lowland wisent lineage. The
88 Lowland-Caucasian (LC) line originates from all 12 founders (5 males, 7 females), which

89 included the last remaining Caucasian wisent bull (Slatis, 1960). Descendants of the LC line
90 thus represent a mixture of Lowland and Caucasian wisent ancestry.

91

92 Although the wisent restitution undoubtedly represents a tremendous conservation success,
93 several factors may limit the long-term viability of the species, many of which are applicable
94 to ex-situ conservation strategies in general. A factor that has received particular attention is
95 that of reduced genetic variability, which may be correlated with a lowered resistance to
96 disease and parasites in wisent (e.g.: Krasieńska & Krasieński, 2007 and references therein;
97 Adaszek et al., 2014; Karbowski et al., 2014a; 2014b and the references therein; Krzysiak et
98 al., 2014; Majewska et al., 2014; Moskwa et al., 2014; Oleński et al., 2015; Panasiewicz et al.,
99 2015), and also seriously impacts conservation programs for other threatened species (Altizer
100 et al., 2007). Although genetic variability and geneflow among living wisent herds has been
101 investigated using limited sets of genetic markers (e.g. Gralak et al., 2004; Luenser et al.,
102 2005; Radwan et al., 2007; Wójcik et al., 2009; Babik et al., 2012; Tokarska et al., 2009a,
103 2009b, 2015), to date no published study has investigated the genetic composition of extant
104 wisent at the level of the complete genome. Similarly, while preserved specimens of some
105 founding individuals that may retain genetic information still exist within museum
106 collections, the genetic variability among these founders and their subsequent contribution to
107 living herds remains unquantified.

108

109 A second threat for wisent is potential hybridisation with domestic cattle (*Bos taurus*).
110 Although F1 hybrid bulls are sterile, hybrid female offspring are not (Basrur, 1968), and
111 would therefore have the potential to reintegrate back into the wild population. It is known
112 that domestic cattle were grazed in the Białowieża Forest of Poland and Belarus for many
113 years at the same time as wisent were present there, though no births of hybrids in natural

114 conditions were recorded (Kraśńska & Kraśński 2007). Hybridisation with domestics in
115 general is of wider concern for threatened species (Wolf et al., 2001). Well known examples
116 of this kind of hybridisation include Przewalski's horse and domestic horse (Sarkissian et al.,
117 2015), mouflon and domestic sheep (Lühken et al., 2009), wolf/coyote and dog (Bohling &
118 Waits, 2011), wildcat and domestic cat (Steyer et al., 2016), and game birds and domestic
119 fowl (Arrieta et al., 2013). However, in the case of wisent, the extent of cattle admixture into
120 living herds remains undetermined, as does the admixture that may have occurred prior to the
121 establishment of the captive breeding program. Such information is critical to assess the
122 magnitude of the threat that cattle admixture represents to the long term viability and integrity
123 of the species.

124

125 Here we present low-coverage whole genome sequencing data from both modern genetic
126 lines, as well as from four historical samples representing two of the original founding
127 individuals and two individuals of the now extinct Caucasian wisent subspecies. Using these
128 data, we compare the relative contributions of the founding individuals to the genomes of
129 living wisent. We also determine the magnitude and distribution of Caucasian wisent ancestry
130 in modern and founding individuals. Furthermore, we uncover evidence of admixture with the
131 cattle/aurochs lineage that has resulted in a significant component of mixed ancestry in the
132 genomes of both founding and modern wisent. Our results have important implications for
133 both understanding the evolutionary history and for the future conservation management of
134 wisent. Finally, we demonstrate the huge potential of genomic approaches, in particular
135 applied to historical samples, for conservation management of endangered species.

136

137

138

139 **RESULTS**

140

141 *Sequencing of wisent genomes*

142 We conducted shotgun sequencing of wisent genomes using Illumina technology and mapped
143 the resulting sequence reads to a reference genome assembly of the Asian water buffalo,
144 *Bubalus bubalis* (GenBank accession no. GCA_000471725.1), which represents an outgroup
145 to the wisent/cattle clade (Fig. S1; Hassanin et al., 2013). In total we sequenced seven
146 individuals that we divide into three categories. Detailed sample information, including
147 provenance, is provided in Table 1 and sample localities are shown in Figure 1.

148

149 *Modern wisent* – three modern individuals representing both genetic lines. These comprise
150 two individuals from the L line (MdL1, mean read depth 1.59x; and MdL2, mean read depth
151 1.63x; from the Polish and Belarussian parts of the Białowieża Forest, respectively) and one
152 individual from the LC line (MdLC, from Dydiowa in the Bieszczady Mountains, mean read
153 depth 1.49x).

154

155 *Founding wisent* – two individuals assignable to the Lowland wisent subspecies *B. b. bonasus*
156 from the initial breeding population originating from Pszczyna, both of which contributed to
157 the establishment of both the L and the LC genetic lines: foundress F42 PLANTA (1904-
158 1931, mean read depth 1.82x) and her male offspring, M158 PLATEN (1926-1933, mean read
159 depth 1.36x), who was fathered by another founder M45 PLEBEJER (1917-1937).

160

161 *Caucasian wisent* – two individuals from the early 1900's representing the now extinct
162 Caucasian wisent subspecies *B. b. caucasicus* (Cc1, from Kubańskaja Obłast, mean read
163 depth 1.17x; and Cc2, from North Ossetia-Alania, mean read depth 0.92x).

164

165 For each individual, we collapsed mapped reads into a single pseudo-haploid genome
166 sequence by randomly selecting a single high quality nucleotide from the read stack at each
167 position of the reference genome, following the procedure described by Cahill et al. (2013;
168 2015). This procedure disregards heterozygous positions, where only one allele will be
169 sampled, but should not introduce any biases in allele sampling. Ancient DNA fragments
170 frequently contain miscoding lesions resulting from postmortem DNA degradation, the most
171 common of which involves the deamination of cytosine to uracil, which causes C to T
172 substitutions in the resulting data (e.g. Dabney et al., 2013b). This pattern is present at varying
173 levels in sequence data from our historical samples (Fig. S2-S5), and so we restricted all
174 subsequent analyses to transversion sites only to avoid any confounding effects of DNA
175 damage.

176

177 *Genomic divergence*

178 We investigated patterns of nuclear genomic divergence among wisent by conducting
179 pairwise comparisons of the number of transversion differences occurring along a sliding
180 window of 1Mb, producing a distribution of genomic divergence for each wisent pair. The
181 resulting probability densities showed that nuclear genomic divergence is broadly similar
182 among all modern and founding wisent (Fig. 2a). The two founding individuals, PLANTA
183 and PLATEN, are somewhat less diverged from one another than either is from all modern
184 wisent (Fig. 2a), reflecting their mother-son relationship. Slightly increased divergence is
185 observed between the modern LC line individual (MdLC) and all modern and founding wisent
186 (Fig. 2a), which may reflect the increased component of Caucasian wisent ancestry in this
187 individual resulting from the captive breeding program.

188

189 Genomic divergence between Caucasian and both modern and founding wisent greatly
190 exceeds that occurring between the latter two groups (Fig. 2a-c). Substantial divergence is
191 also found between the two Caucasian wisent individuals. One of these Caucasian wisent
192 (Cc1) was found to be less diverged from modern and founding wisent than other Caucasian
193 wisent individual (Cc2), suggesting the presence of not only substantial genetic diversity but
194 also substantial population structure in the extinct Caucasian wisent subspecies.

195

196 We also investigated mitochondrial genome variability among all individuals subjected to
197 nuclear genome sequencing, in addition to eight other modern wisent (Tab. S1). Sequence
198 analysis revealed that all investigated modern wisent, both founding wisent, and a single
199 historical Caucasian wisent (Cc1), share a single haplotype. The haplotype occurring in the
200 second historical Caucasian wisent (Cc2) differed from this widely shared haplotype by a
201 single transition site. These results hint at a major loss of mitochondrial haplotype diversity
202 prior to the extinction of wisent in the wild. This inference is supported by additional
203 haplotypes that we recovered from three ancient middle Holocene wisent from Austria and
204 one ancient middle Holocene Caucasian wisent from Armenia (Tab. 1). All these ancient
205 haplotypes are substantially divergent from those found in modern and historical wisent,
206 suggesting a substantial loss of haplotype diversity, potentially within the last ~1500 years.

207

208 Neighbour-joining phylogenetic analysis of total nuclear genomic divergence
209 supports paraphyly of Caucasian wisent (Fig. 2d), as does analysis of mitochondrial
210 haplotypes (Fig. S1). We further investigated the population history of wisent by dividing
211 aligned nuclear genome sequences into non-overlapping 1MB blocks and subjecting each
212 block to maximum-likelihood phylogenetic analysis. For this analysis, we included both
213 Caucasian wisent and both modern L line wisent as representatives of the Lowland wisent

214 subspecies, with water buffalo as outgroup. Founding wisent and the modern LC line wisent
215 were not included to avoid any confounding effects of direct ancestor-descendent
216 relationships and documented Caucasian wisent introgression (Slatis, 1960) on phylogenetic
217 interpretation, respectively. We found that 57% of the investigated genomic blocks support
218 reciprocal monophyly of Caucasian and Lowland wisent (Fig. 3). We therefore conclude that
219 this most likely represents the true population history. All alternative topologies occur,
220 individually, at a much lower frequency. Nevertheless, almost half of the genome sequence
221 alignment of these individuals contradicts the true population history.

222

223 In order to interpret wisent genomic divergence in the context of total species genetic
224 diversity, we obtained data from the NCBI Short Read Archive for seven domestic cattle and
225 seven yak (*Bos grunniens*; Tab. S4) and subjected them to the same analysis pipeline.

226 Genomic divergence among modern wisent was found to be similar to that found among
227 domestic cattle, and exceeded that found among yak (Fig. 4). We conducted equivalent
228 comparisons for modern wisent and these other bovid species but with the inclusion of
229 transition as well as transversion sites. Interestingly, the distribution of genomic divergence
230 for pairs of wisent was bimodal in all cases (Fig. 4), but the relative levels of genomic
231 divergence between species were similar to that measured using only transversion sites.

232

233 **Wisent gene flow and admixture**

234 We investigated patterns of admixture among wisent using the D statistic test for admixture
235 (Green et al., 2010). This test identifies any imbalance in the number of derived alleles that
236 either of two closely related individuals share with a candidate introgressor. A significant
237 excess of derived alleles shared between one individual and the introgressor provides

238 evidence of admixture between them (Durand et al., 2011). For all D statistic tests, we used
239 water buffalo (*Bubalus bubalis*) as outgroup for allele polarisation.

240

241 We first investigated patterns of derived allele sharing among modern wisent, and found no
242 statistically significant signal of admixture between the modern LC line individual and either
243 modern L line individual (Tab. S5). Between modern and founding wisent, we found that
244 modern L line wisent share a significantly greater number of derived alleles with founding
245 wisent than the modern LC line individual does (Fig. 5a), indicating a greater contribution of
246 the two founding wisent investigated here to the L line, relative to the LC line, which is
247 consistent with pedigrees (Slatis, 1960).

248

249 We then investigated admixture involving Caucasian wisent. We found a significant excess of
250 derived allele sharing between one founding wisent, PLANTA, and one Caucasian wisent,
251 Cc2, relative to the other founding wisent, her son PLATEN (Fig. 5b). This indicates that a
252 proportion of the genome of PLANTA can be attributed to admixture with Caucasian wisent.
253 Furthermore, we can deduce that PLEBEJER, the father of PLATEN, must have possessed a
254 lower level of Caucasian wisent admixture than PLANTA, and that PLATEN himself was
255 likely admixed to some degree through inheritance from PLANTA. The detection of
256 admixture involving one Caucasian wisent (Cc2) but not the other (Cc1) further supports the
257 existence of genetic structure in Caucasian wisent inferred from estimates of genomic
258 divergence (Figs. 2b & 2c).

259

260 Next, we investigated evidence of Caucasian wisent admixture among modern wisent.
261 Consistent with expectations, we found that the modern LC line individual (MdLC) shares an
262 excess of derived alleles with one of the Caucasian wisent (Cc1) relative to modern L line

263 individuals (Fig. 5c). We did not, however, detect such an excess between the modern LC line
264 individual and the second Caucasian wisent (Cc2), relative to the modern L line individuals.
265 We can therefore infer that the last surviving Caucasian wisent, KAUKASUS, whose living
266 descendants comprise the modern LC line, was more closely related to Caucasian wisent
267 individual Cc1 than to individual Cc2.

268

269 We further investigated Caucasian ancestry in the genome of the modern LC line individual
270 by performing phylogenetic analysis of non-overlapping 1MB genomic blocks. This analysis
271 involved the modern LC line wisent (MdLC), the founding wisent (PLATEN) that was found
272 to be least admixed with Caucasian wisent, Caucasian wisent (Cc1) and domestic cattle, with
273 water buffalo as outgroup. These datasets were mapped to the reference genome of the zebu,
274 *Bos indicus* (Canavez et al., 2012), to benefit from the chromosome-level assembly of that
275 genome. Of the investigated genomic blocks, we find that 22% return the modern LC line and
276 Caucasian wisent as monophyletic (Fig. 6), and may therefore represent introgressed
277 segments of Caucasian wisent ancestry in this modern LC line individual. Around 8% of these
278 blocks are likely to result from incomplete lineage sorting, based on the frequency of
279 occurrence of the opposing topology (Fig. 6), producing an overall estimate of 14% of the
280 genome of the modern LC individual that results from Caucasian wisent admixture, most
281 likely inherited from the bull KAUKASUS. In addition to providing an estimate of admixture
282 proportions, our method is also able to accurately map admixed segments of the genome (Fig.
283 6a). Many of these segments span multiple megabase blocks. For example, a contiguous
284 22MB admixed block is observed on chromosome 4, which may span as much as 33MB
285 under the assumption that intervening blocks with missing data are linked to adjacent ones.
286 Relative admixture proportions also vary among chromosomes in this individual. For

287 example, chromosome 27 almost entirely lacks Caucasian wisent ancestry whereas around
288 50% of chromosomes 4 and 15 are likely derived from admixture.
289
290 Finally, we investigated evidence of Caucasian wisent ancestry in the modern L line. We
291 found that both modern L line wisent share a significant excess of derived alleles with
292 Caucasian wisent (Fig. 5b), relative to founding wisent. Thus, modern L line individuals
293 appear more admixed with Caucasian wisent than the two founding wisent investigated here.
294 This admixture signal could result from either variable admixture proportions among
295 founding individuals, or recent geneflow between the L and LC lines, although D statistic
296 comparison of modern individuals failed to detect the latter (see above). We further
297 investigated these alternative hypotheses by comparing the sizes of putatively admixed
298 genomic blocks. Recent geneflow results in large contiguous genomic blocks derived from
299 admixture in the genomes of the recipient population, which are broken up over time as a
300 result of recombination. Pedigree information provides an approximate date for geneflow
301 from Caucasian wisent into the modern LC line around 90 years ago (15-22 generations), the
302 result of which are many intact multi-megabase genomic blocks derived from Caucasian
303 wisent in the modern LC line individual (Fig. 6). We compared the sizes of these blocks with
304 those of putative Caucasian wisent ancestry in a modern L line individual, and found the
305 abundance of large blocks to be considerably lower in the latter (Fig. 7). This rejects recent
306 admixture and instead supports variable admixture proportions among the founding herd in
307 explaining the observed signal of Caucasian wisent admixture in this modern L line
308 individual.
309
310
311

312 **Admixture with the cattle/aurochs lineage**

313 We investigated potential admixture between wisent and the cattle/aurochs lineage using
314 pseudo-haploid sequences generated from short read data of two domestic Holstein cows and
315 an ancient aurochs (*Bos primigenius*; Park et al., 2015), the extinct species from which cattle
316 were domesticated and that lived sympatrically with wisent up until its extinction around 400
317 years ago (van Vuure, 2005). First, we looked for significant differences in derived allele
318 sharing between cattle and wisent, relative to the aurochs. We found that all investigated
319 wisent share a significant excess of derived alleles with cattle relative to aurochs (Fig. 5d).
320 This suggests either admixture between wisent and domestic cattle, or alternatively, admixture
321 with aurochs, if aurochs populations were highly structured and the admixing individuals
322 were from a population more closely related to domestic cattle than the British aurochs used
323 in this analysis.

324

325 We also compared derived allele sharing with cattle among the individual wisent. We found
326 that all modern wisent investigated here share a significant excess of derived alleles with
327 cattle relative to any founding or Caucasian wisent (Fig. 5e). Variable admixture was also
328 observed among founding wisent. Specifically, PLANTA shares more derived alleles with
329 domestic cattle than either PLATEN or the Caucasian wisent do. Finally, we estimated the
330 genomic fraction in the modern wisent that could be attributed to cattle/aurochs admixture,
331 above that occurring in the founding wisent, using the f statistic (Durand et al., 2011). This
332 analysis showed that 2.4% to 3.2% of the sampled modern wisent genomes can be attributed
333 to admixture with the cattle/aurochs lineage, respectively, that is not the result of direct
334 inheritance from the sampled founding wisent (Tab. S6).

335

336 The increased admixture signal observed in all modern wisent relative to founding wisent
337 could result from either variable cattle/aurochs admixture proportions among the founding
338 herd from which all modern wisent are derived, or alternatively, from recent admixture with
339 modern cattle. However, the fact that we do not find evidence of any complete genomic 1MB
340 blocks resulting from cattle admixture in the modern LC line individual (Fig. 6a) argues
341 strongly against recent cattle admixture, and instead supports variable cattle/aurochs
342 admixture among the founding herd in explaining the excess of derived alleles shared among
343 domestic cattle and modern wisent.

344

345 **DISCUSSION**

346 Retracing complex population histories can be challenging. In particular, admixture involving
347 populations or species that are now extinct may be impossible based solely on data from
348 living individuals (Hofreiter et al., 2015). Through the use of low-coverage genomic data
349 from modern and historical wisent, including from the now extinct Caucasian wisent
350 subspecies, we have revealed the complexity of wisent evolution. This complex history
351 involved not only admixture resulting directly from the captive breeding program, but also
352 older processes occurring prior to their extinction in the wild, which included admixture with
353 another bovid lineage (Fig. 8).

354

355 ***Wisent evolution and admixture***

356 The accepted view of wisent evolution is of two distinct subspecies, Lowland wisent and
357 Caucasian wisent, that both underwent dramatic population declines, with the last few
358 surviving individuals serving as founders of the modern L (Lowland only) and LC (Lowland
359 and Caucasian) lines (Pucek et al., 2004). Our results show that this model is an
360 oversimplification. We find evidence of at least two highly differentiated populations within

361 Caucasian wisent, with one of these showing greater pairwise similarity with Lowland wisent
362 than with the second Caucasian population, at the level of the complete genome (Fig. 2).
363 However, analysis of aligned nuclear genomic blocks from four individuals returns Caucasian
364 and Lowland wisent as reciprocally monophyletic across slightly more than half of the
365 genomic alignment (Fig. 3), providing support that this topology reflects the true history of
366 population divergence. Thus, among any two Caucasian wisent and any two modern (or
367 founding Lowland) wisent, we may expect that any single locus has only around 50%
368 probability of reflecting the true history of population divergence. Moreover, increased
369 sampling of individuals is likely to further reduce this proportion, potentially to such an extent
370 that, at a given sampling level, the true evolutionary history cannot be untangled from the
371 effects of random drift and more recent admixture. This result reinforces the notion that
372 phylogeny-based interpretation may be inappropriate at the level of the complete genome, and
373 that alternative, more flexible models will be required to keep pace with our ability to
374 generate such data (Hofreiter et al., 2015).

375

376 A further implied assumption of the traditional view of wisent evolution is that, with the
377 exception of the Caucasian bull KAUkasus, all founding wisent represented “pure”
378 Lowland wisent (Pucek et al., 2004; Tokarska et al., 2015). On this basis, the modern L line
379 that is derived only from the latter can also be considered as pure Lowland wisent, referable to
380 the subspecies *B. b. bonasus*. Our results also demonstrate that the notion of wisent subspecies
381 purity is flawed in the sense that founding Lowland individuals were in fact admixed with
382 Caucasian wisent to varying degrees. We demonstrate this both directly for the founder
383 PLANTA in comparison to another founder, her son PLATEN, and also indirectly, by the
384 elevated signal of Caucasian wisent admixture in modern L line wisent, relative to these
385 founders, most likely a result of inheritance from other founding individuals not included in

386 this analysis (Fig. 7). The notion of subspecies purity therefore disregards the fact that
387 admixture between Caucasian and Lowland wisent almost certainly occurred prior to the
388 extinction of wisent in the wild, and such admixture could therefore be regarded as part of the
389 normal population processes and dynamics of this species.

390

391 The notion of subspecies purity has driven efforts to ensure that free-living L and LC herds do
392 not come into contact (Pucek et al., 2004), and also motivated genetic investigations of living
393 populations that may have been recipients of geneflow from the opposing genetic line
394 (Tokarska et al., 2015). The latter study investigated the modern L line population that
395 currently inhabits the eastern, Belarussian part of the Białowieża Forest. Some individuals
396 were found to possess a microsatellite allele that was common among Caucasian wisent but
397 absent in all studied Lowland wisent, which these authors interpreted as evidence of recent
398 admixture with the modern LC line. Although the individual from this population (modern L
399 line, MdL2) that we sequenced does not possess this putative Caucasian wisent allele, we
400 nevertheless detected evidence of Caucasian wisent admixture above that occurring in
401 founding wisent for this individual (Fig. 5). However, the small size of admixed blocks, in
402 addition to non-significant D comparisons of modern lines, supports variable Caucasian
403 wisent admixture among founding wisent in explaining this result, which may also account
404 for the occurrence of putative Caucasian wisent alleles in other individuals from this
405 population. Future studies of such individuals using the methodology applied here would
406 provide a robust test of these alternative hypotheses.

407

408 ***Wisent conservation and de-extinction***

409 The issue of low genetic variability among living wisent is considered as cause for concern,
410 and has been demonstrated by several population-level studies using various molecular and

411 biochemical markers, such as blood-group systems and blood serum proteins (Sipko et al.,
412 1995; Gębczyński & Tomaszewska-Guszkiewicz 1987; Hartl & Pucek, 1994; Sipko et al.,
413 1997), mtDNA (this study; Tiedemann et al., 1998; Burzyńska et al., 1999; Wójcik et al.,
414 2009; Hassanin et al., 2012), nuclear gene sequences (Sipko et al., 1994; Udina et al., 1994;
415 Kamiński & Zabolewicz 1997; Udina & Shakhaev 1998; Burzyńska & Topczewski, 1999;
416 Radwan et al., 2007; Babik et al., 2012; Hassanin et al., 2013), microsatellites (Gralak et al.,
417 2004; Luenser et al., 2005; Tokarska et al., 2009a; 2009b) and SNPs (Tokarska et al., 2009b,
418 2015). In apparent contrast to these results, we find relatively high levels of genomic
419 divergence among both modern and founding wisent, being approximately equal to or
420 exceeding that found between pairs of cattle or yak, respectively. Pairwise divergence
421 between modern and Caucasian wisent was found to be even higher, indicating a major loss of
422 genetic diversity from the species as a whole following the extinction of the latter population.

423
424 The discrepancy between estimates of genetic variation obtained from population-level
425 studies and from pairwise nuclear genomic divergences, as reported here, may result from
426 several factors. First, application of microsatellites or SNP markers developed for cattle to
427 wisent will likely lead to ascertainment bias (Albrechtsen et al., 2010), particularly if diversity
428 estimates for wisent are then compared back to those obtained for cattle using the same
429 marker set. Admixture with the cattle/aurochs lineage detected in this study adds a further
430 layer of complexity in interpreting data from molecular markers developed for cattle. Second,
431 although genetic diversity and pairwise genetic divergence are both measures of genetic
432 variation in the broad sense, these measures are different and not necessarily interchangeable.
433 Thus, wisent may lack allelic diversity, but the alleles that do exist may be highly divergent
434 from one another. Indeed, the bimodal distributions observed for pairwise comparisons based
435 on transitions and transversions are compatible with high levels of relatedness -- likely

436 inbreeding -- among modern wisent, and low mitochondrial haplotype diversity among
437 modern and historical individuals in comparison to that found among middle Holocene wisent
438 (Fig. S1) provides further evidence of a loss of allelic variation. Under this scenario of low
439 allelic diversity but high nuclear allelic divergence, given continued population expansion and
440 sufficient time, nuclear genomic diversity may potentially be restored as recombination
441 dissects and shuffles divergent chromosomes over successive generations.

442

443 The Caucasian wisent is extinct in the wild, but a fraction of its genepool survives in the
444 genomes of modern wisent. Our results provide not only a direct measure of this admixture in
445 a modern LC line individual, but also allow us to map with relative accuracy chromosomal
446 segments that are inherited from Caucasian wisent. Although sometimes controversial, the
447 concept of de-extinction has generated considerable interest (Sherkow & Greely, 2013), and
448 attempts are currently underway to generate animals that, at least superficially, resemble the
449 quagga (*Equus quagga*; Heywood, 2013), an extinct subspecies of plains zebra, and also the
450 aurochs (van Vuure, 2005), by careful selective breeding of their living relatives. In both of
451 these cases, selective breeding and the ultimate success of the project are based solely on
452 morphological criteria. Our study demonstrates that, at least in principle, by generating
453 chromosomal admixture maps for multiple living representatives of the LC line, it would be
454 possible to selectively breed an animal that is, at the genomic level, highly similar to a
455 Caucasian wisent.

456

457 ***Admixture with the cattle/aurochs lineage***

458 Hybridisation of wild species with their domesticated close relatives is a subject of
459 considerable discussion and concern for conservation management (Ellstrand et al., 2010).
460 Frequently, such events are deemed to be detrimental to the recipient wild species, as is the

461 case for American bison (*Bison bison*) herds found to be admixed with cattle (Halbert & Derr,
462 2007). However, the introgression of alleles from domestics into wild populations may also
463 provide the basis for adaptation (Fuelner et al., 2013; Anderson et al., 2009), and allow
464 populations to take advantage of new ecological niches (Monzón et al., 2014). In any case, the
465 identification of admixture is the essential first step to guide conservation policy, and
466 measuring admixture proportions among individuals is likely the second essential step for
467 implementing conservation measures. Our results achieve both of these objectives.

468

469 We detect a significant proportion of admixture with cattle relative to aurochs in the genomes
470 of all wisent except founder PLATEN. However, since D statistic is a relative test, and
471 PLATEN is the offspring of admixed founder PLANTA, it is reasonable to infer the PLATEN
472 is also admixed to some extent. Although our comparisons consistently found an excess of
473 derived alleles shared between cattle and wisent, relative to aurochs and wisent, it is difficult
474 to completely exclude admixture with aurochs as an alternative explanation, given the lack of
475 knowledge of population structure in aurochs. Testing these two alternatives would require
476 data from additional aurochs populations from within the core distribution of wisent, and
477 would be a valuable direction for future research.

478

479 The timing of admixture also has implications for conservation management. Specifically, the
480 removal of individuals resulting from very recent hybridisation may be deemed appropriate
481 (Halbert & Derr, 2007). The small size of cattle admixed blocks in modern wisent (at least
482 undetectable at a 1MB scale) clearly rejects recent cattle admixture for the individuals
483 investigated here. Instead, admixture must have occurred prior to the establishment of the
484 captive breeding program, and the admixture signal detected in modern wisent results from
485 inheritance from the founders that were admixed with cattle to varying degrees. Thus, based

486 on the current evidence, cattle introgression appears of low concern for wisent conservation
487 for the following reasons: 1. admixture does not appear to have occurred since the
488 establishment of the captive breeding program, although screening of additional individuals
489 may be desirable to further support this generalisation; 2. the number of intervening
490 generations separating living wisent from the F1 hybrids is likely sufficient that all living
491 wisent are admixed to some extent (Chang, 1999); and 3. our results may in fact reflect
492 admixture with aurochs, rather than domestic cattle, although this hypothesis requires further
493 investigation.

494

495 *Conclusions*

496 The ability to detect admixture is of key importance for both evolutionary and applied
497 conservation studies. However, interpretation of a significant signal of admixture, in terms of
498 both evolutionary inference and the formulation of management strategies, may require
499 information its timing. Using seven low-coverage wisent genomes from both modern and
500 historical wisent we have revealed multiple instances of admixture, but moreover, because the
501 approximate age of introgression of Caucasian wisent into the modern LC line is known,
502 through comparisons of the sizes of likely admixed genomic segments we have inferred the
503 relative ages of other admixture events. This unique historical information, coupled with the
504 ability to recover genomic data from historical samples, establish wisent as an exemplary
505 taxon for the study of admixture in wild populations. As new analytical methods for studying
506 admixture are developed, wisent can serve as a valuable empirical test of both their
507 performance and utility.

508

509

510

511 **MATERIALS AND METHODS**

512 Complete details of all samples and specimens used in this study are shown in Table 1.

513

514 *Laboratory methods, modern samples*

515 DNA was extracted from tissue samples of three modern wisent using either a DNeasy Blood
516 & Tissue Kit (Qiagen) according to the manufacturer's instructions (sample MdL1) or by
517 phenol/chloroform extraction (Sambrook & Russell, 2001). We mechanically sheared the
518 DNA of the modern samples using a Covaris S220 sonicator to an average fragment length of
519 500 bp and prepared indexed Illumina libraries from 500 ng of each modern DNA extract
520 using a published double-stranded protocol (Meyer & Kircher, 2010) with modifications
521 (Fortes & Paijmans, 2015). Library molecules from 450 bp to 1000 bp were then selected
522 using a Pippin Prep Instrument (Sage Science).

523

524 *Laboratory methods, historical samples*

525 DNA extraction from four museum specimens as well as sequencing library preparation steps
526 preceding amplification were performed in a dedicated ancient DNA laboratory (Evolutionary
527 Adaptive Genomics Group, Potsdam University, Germany). DNA extracts were prepared
528 from horn and bone powder obtained by grinding in a mixer mill (MM 400, RETSCH). DNA
529 extraction followed the protocol of Dabney et al. (2013a), except for horn samples where we
530 used a different digestion buffer containing 10mM Tris buffer (pH 8.0), 10 mM NaCl, 5 mM
531 CaCl₂, 2.5 mM EDTA (pH 8.0), 2% SDS (Shapiro & Hofreiter, 2012). The museum samples
532 were already fragmented due to degradation, so were not sonicated. We used 25 µl of each
533 DNA extract to construct single-stranded indexed Illumina libraries according to the protocol
534 of Gansauge & Meyer (2013).

535

536 ***Sequencing***

537 Final library concentrations and the distribution of insert sizes were determined using a 2200
538 TapeStation (Agilent Technologies) and Qubit HS-assay (Thermo Fisher Scientific),
539 respectively. Each library was then sequenced using an Illumina NextSeq 500 instrument. For
540 modern libraries we used a High Output Kit (75 bp paired-end sequencing), for libraries
541 obtained from historical horn samples we used High Output Kits (75 bp single-end and 150 bp
542 paired-end) and each library built from historical bone samples was sequenced separately with
543 High Output Kits (75 bp single-end and paired-end). Full details of sequencing results are
544 provided in Supporting Tables S2 and S3.

545

546 ***Data processing, mapping and pseudo-haploidisation***

547 For paired-end data, we trimmed adapter sequences and merged overlapping read pairs using
548 SeqPrep (<https://github.com/jstjohn/SeqPrep>), requiring a minimum read length of 30 bp (-L
549 30), minimum overlap of 15 bp (-o 15), and a minimum merge quality of 13 (-q 13). Adapters
550 occurring at the 3' ends of single-end reads were trimmed using cutadapt (Martin, 2011), also
551 requiring a minimum length of reads of 30 (-m 30). We then mapped the resulting data to the
552 zebu (*Bos indicus*; GenBank accession no. GCA_000247795.2) and water buffalo (*Bubalus*
553 *bubalis*; GenBank accession no. GCA_000471725.1) nuclear genomes and wisent
554 mitochondrial genome (KW, unpublished) using BWA aln version 0.7.8 (Li & Durbin, 2009)
555 with default 0.04 mismatch value. We removed duplicate reads likely resulting from PCR
556 amplification using samtools rmdup (Li et al., 2009). Detailed descriptions of the mapping
557 results are provided in Supporting Tables S2 and S3. We then generated pseudo-haploid
558 sequences as described by Cahill et al. (2015) and used these for further analysis.

559

560

561 ***Pairwise genomic divergence***

562 Pairwise genomic divergence was calculated by dividing genomic alignments into non-
563 overlapping 1MB blocks and calculating the proportion of transversions, or transitions plus
564 transversions (comparisons of modern individuals only), for each pair of individuals,
565 accounting for the presence of missing data. Blocks with > 75% missing data were
566 disregarded. Probability densities were generated by kernel density estimation in R (R Core
567 Team, 2014) using default parameters. Full details of comparative data generated for domestic
568 cattle and yak (data from the NCBI Short Read Archive) are provided in Supporting Table S4.
569

570 ***Mitochondrial genome analysis***

571 In addition to the mitochondrial genomes generated from the seven modern and four historical
572 specimens, we generated four mitogenomes from ancient middle Holocene wisent individuals
573 (Tab. 1). These ancient samples were radiocarbon dated at either the Oxford University
574 Radiocarbon Accelerator Unit (Oxford, UK) using ultrafiltered collagen and accelerator mass
575 spectrometry (Ramsey et al. 2004a, 2004b) or the VERA-Laboratorium Institut für
576 Isotopenforschung und Kernphysik (Vienna, Austria). We calibrated radiocarbon dates using
577 the IntCal13 calibration curve (Reimer et al., 2013) in OxCal v4.2
578 (<https://c14.arch.ox.ac.uk/oxcal/OxCal.html>).

579
580 DNA was extracted from the Holocene *B. bonasus* samples (Bb1-Bb3) at the Henry
581 Wellcome Ancient Biomolecules Centre (Oxford University, UK), following Shapiro et al.
582 (2004). We extracted the Holocene *B. b. caucasicus* sample (Cc3) in the specialist
583 Paleogenomics facility at UC Santa Cruz, following Rohland et al. (2010). DNA library
584 construction, mitochondrial target enrichment, sequencing, and sequence data processing
585 protocols for the four Holocene samples followed approach four in Heintzman et al. (2016),

586 except that the whole mitochondrial genome consensus sequence was retained. The mean read
587 depth of these Holocene consensus sequences ranged from 14.1 to 165.7x. The consensus
588 sequences for the four Holocene and one historic Caucasian bison have been deposited in
589 GenBank with accession numbers XXXXXXXXXX-XXXXXXX. The mitogenomic
590 sequences from the remaining modern and historic samples were identical to either haplotypes
591 previously published or the historic Caucasian bison haplotype (Tab. S1).

592

593 We assessed phylogenetic relationships among wisent mitochondrial haplotypes, as well as
594 their placement within the wider bovin (tribe Bovini) tree. Wisent sequences were aligned
595 with those of 12 other bovin taxa, including the extinct steppe bison (*Bison priscus*) and
596 aurochs (*Bos primigenius*) (Tab. S1). We excluded two previously published wisent
597 mitochondrial genomes from the phylogenetic analysis (GenBank accessions: HQ223450,
598 HM045017/NC_014044), as these sequences were considered problematic. Specifically,
599 HQ223450 has multiple insertions totaling 9 bp in the ND4 coding region, and
600 HM045017/NC_014044 has multiple indels and point mutations concentrated in the large
601 rRNA and ND3 coding regions. Sequence alignment, partitioning, model testing, and
602 phylogenetic and associated statistical support methods followed the ordinal-level analyses of
603 Heintzman et al. (2015), except that we used the *B. bison* reference mitochondrial genome
604 (NC_012346) for partitioning. We selected the following models of molecular evolution for
605 the six partitions: GTR+I+G (CP1, 3803 bp; rRNAs, 2541 bp), GTR+G (CP3, 3803 bp),
606 HKY+I+G (CP2, 3803 bp; tRNAs, 1526 bp; control region, 927 bp). We used the saola
607 (*Pseudoryx nghetinhensis*; Tab. S1) as outgroup in both the maximum likelihood and
608 Bayesian analyses, following Bibi (2013).

609

610

611 *D statistic tests*

612 The *D* statistic involves 4 genomes: a genome from each of two sister populations (P1 and
613 P2), a genome from a third population as the potential source of introgression (P3), and an
614 outgroup genome (O) to identify the ancestral state (identified as the A allele). We identified
615 variable positions at which P3 possessed the derived allele (B) and presence of the derived
616 allele is variable among P1 and P2, leading to two possible patterns: either ABBA or BABA.
617 Under the scenario of incomplete lineage sorting without geneflow these patterns should
618 occur with equal frequency and the expected *D* value will be zero. An excess of ABBA or
619 BABA patterns is interpreted as evidence of admixture. However, it might also arise from
620 nonrandom mating in the ancestral population due to population structure (Eriksson &
621 Manica, 2012). To determine the ancestral state we used the water buffalo genome. In all tests
622 involving data mapped to the zebu genome, we took into consideration the autosomes only.
623 We performed a total of 105 comparisons considering all possible combinations of wisent, all
624 wisent with either domestic cattle or aurochs as candidate admixer, and domestic cattle and
625 aurochs with all wisent as candidate admixer. These results are reported in Supporting Table
626 S4. \hat{f} test (Green et al., 2010; Durand et al., 2011) was used to estimate the proportion of the
627 genome derived from admixture. This test requires two individuals of the candidate
628 introgressor species that are not themselves admixed. For our datasets this was possible only
629 for admixture involving the cattle/aurochs lineage. For both *D* statistic and \hat{f} test, significance
630 was assessed using a weighted block jackknife using 1Mb blocks (Green et al., 2010; Durand
631 et al., 2011). The weighted block jackknife tests if admixture signals are uniform across the
632 whole genome and therefore reflect the same population history. By removing one at a time
633 blocks of adjacent sites (larger than the extent of linkage disequilibrium) and computing the
634 variance of the *D* statistic or \hat{f} values over the entire genome *M* times leaving each block of
635 the genome in turn, and then multiplying by *M* and taking the square root we generated the

636 standard error. The number of standard errors by which D or \hat{f} differs from zero is the Z
637 score. The results with Z -scores greater than 3 in absolute value were qualified as statistically
638 significant (Green et al., 2010).

639

640 *Nuclear genome phylogenetic tests*

641 The aligned pseudohaploid sequences were divided into non-overlapping blocks of 1MB. If
642 each of the five taxa contained no more than 50% gaps within a window, the sequence data
643 were recoded into binary characters to only score transversions (Rs: 0, Ys: 1), otherwise the
644 window was recorded as having insufficient data. A Maximum Likelihood phylogeny under
645 the BINGAMMA model and with the water buffalo as outgroup was then computed for each
646 alignment with sufficient data using RaxML (Stamatakis, 2014). The topology of each
647 phylogeny was evaluated using a custom Perl script that made use of the ETE3 software
648 (Huerta-Cepas et al., 2016). The lengths of admixed genomic regions was estimated by
649 counting the number of consecutive 1MB blocks returning the respective tree topology. Due
650 to the presence of blocks with insufficient data, these measurements are likely to be
651 underestimates. Evaluation of the lengths of genomic regions was conducted using the
652 empirical cumulative distribution function in R, with default parameters.

653

654 *ACKNOWLEDGEMENTS*

655 We would like to thank Andrzej Imiołczyk and his colleagues from the Specimen Preparation
656 Laboratory of the Upper Silesian Museum in Bytom. We thank Andrew Fields and Joshua
657 Kapp for technical assistance. The study has been funded by National Science Center (NCN)
658 grant 2011/03/N/NZ8/00086 to KW.

659

660 *AUTHOR CONTRIBUTIONS*

661 Conceived the study - KW, JMS, AB, MH
662 Designed laboratory experiments - AB, KW, JLAP, PDH
663 Performed lab work - KW, UT, GX, PDH, BS
664 Coordinated data analysis - KW, SH, AB, JAC, JLAP, PDH
665 Performed data analysis - KW, SH, AB, PDH
666 Coordinated writing of the manuscript - KW, AB, MH
667 Obtained funding - KW, JMS, MH
668 Provided samples - GB, ANB, JJC, RD, NM, HO, MT, STT, JMW, WŻ
669 All authors read, gave comments and helped revise the final version of the manuscript.
670

671 ***COMPETING INTERESTS***

672 The authors declare that they have no competing financial interests.
673
674

675 ***REFERENCES***

676
677 Adaszek L, Dzięgiel B, Krzysiak M, Skrzypczak M, Adaszek M, Staniec M, Wylupek D,
678 Winiarczyk S. 2014. Detection of *Borrelia burgdorferi* sensu lato DNA in the blood of
679 wild bison from Białowieża primeval forest in eastern Poland. *Pol J Vet Sci* **17**:713-715.
680 Albrechtsen A, Nielsen FC, Nielsen R. 2010. Ascertainment biases in SNP chips affect
681 measures of population divergence. *Mol Biol Evol* **27**:2534-2547. doi:
682 10.1093/molbev/msq148.
683 Altizer S, Nunn CL, Lindenfors P. 2007. Do threatened hosts have fewer parasites? A
684 comparative study in primates. *J Anim Ecol* **76**:304-314. doi: 10.1111/j.1365-
685 2656.2007.01214.x

- 686 Anderson LA, Christianson GB, Linden JF. 2009. Stimulus-specific adaptation occurs in the
687 auditory thalamus. *J. Neurosci* **29**:7359–7363. doi: 10.1523/JNEUROSCI.0793-09.2009.
- 688 Arrieta RS, Lijtmaer DA, Tubaro PL. 2013. Evolution of postzygotic reproductive isolation in
689 galliform birds: analysis of first and second hybrid generations and backcrosses. *Biol J*
690 *Linnean Soc* **110**:528–542. doi: 10.1111/bij.12153.
- 691 Babik W, Kawalko A, Wójcik JM, Radwan J. 2012. Low major histocompatibility complex
692 class I (MHC I) variation in the European bison (*Bison bonasus*). *J Hered* **103**:349-359.
693 doi: 10.1093/jhered/ess005.
- 694 Basrur PK. 1968. Hybrid Sterility. In: Benirschke K, editor. *Comparative Mammalian*
695 *Cytogenetics*. Springer. 107-131.
- 696 Benecke, N. 2005. The Holocene distribution of European bison – the archaeozoological
697 record. *Munibe* **57**:421-28.
- 698 Bibi F. 2013. A multi-calibrated mitochondrial phylogeny of extant Bovidae (Artiodactyla,
699 Ruminantia) and the importance of the fossil record to systematics. *BMC Evol Biol* **13**:166.
700 doi: 10.1186/1471-2148-13-166.
- 701 Bocherens H, Hofman-Kamińska E, Drucker DG, Schmölcke U, Kowalczyk R. 2015.
702 European bison as a refugee species? Evidence from isotopic data on Early Holocene bison
703 and other large herbivores in northern Europe. *PLoS One* **10**:e0115090. doi:
704 10.1371/journal.pone.0115090.
- 705 Bohling JH, Waits LP. 2011. Assessing the prevalence of hybridization between sympatric
706 *Canis* species surrounding the red wolf (*Canis rufus*) recovery area in North Carolina. *Mol*
707 *Ecol* **20**:2142-2156. doi: 10.1111/j.1365-294X.2011.05084.x.
- 708 Burzyńska B, Olech W, Topczewski J. 1999. Phylogeny and genetic variation of the European
709 bison *Bison bonasus* based on mitochondrial DNA D-loop sequences. *Acta Theriol* **44**:253-
710 262.

- 711 Burzyńska B, Topczewski J. 1995. Genotyping of *Bison bonasus* kappa-casein gene following
712 DNA sequence amplification. *Anim Genet* **26**:335-336.
- 713 Cahill JA, Green RE, Fulton TL, Stiller M, Jay F, Ovsyanikov N, Salamzade R, St John J,
714 Stirling I, Slatkin M, Shapiro B. 2013. Genomic evidence for island population conversion
715 resolves conflicting theories of polar bear evolution. *PLoS Genet* **9**:e1003345. doi:
716 10.1371/journal.pgen.1003345.
- 717 Cahill JA, Stirling I, Kistler L, Salamzade R, Ersmark E, Fulton TL, Stiller M, Green RE,
718 Shapiro B. 2015. Genomic evidence of geographically widespread effect of gene flow from
719 polar bears into brown bears. *Mol Ecol* **24**:1205–1217. doi: 10.1111/mec.13038.
- 720 Canavez FC, Luche DD, Stothard P, Leite KR, Sousa-Canavez JM, Plastow G, Meidanis J,
721 Souza MA, Feijao P, Moore SS, Camara-Lopes LH. 2012. Genome sequence and assembly
722 of *Bos indicus*. *J Hered* **103**:342-8. doi: 10.1093/jhered/esr153.
- 723 Chang, JT. 1999. Recent common ancestors of all present-day individuals. *Adv. in Appl.*
724 *Probab* 1002-1026.
- 725 Dabney J, Knapp M, Glocke I, Gansauge MT, Weihmann A, Nickel B, Valdiosera C, Garcia
726 N, Paabo S, Arsuaga JL, Meyer M. 2013a. Complete mitochondrial genome sequence of a
727 Middle Pleistocene cave bear reconstructed from ultrashort DNA fragments. *Proc Natl*
728 *Acad Sci U S A* **110**:15758–15763. doi: 10.1073/pnas.1314445110.
- 729 Dabney J, Meyer M, Pääbo S. 2013b. Ancient DNA damage. *Cold Spring Harb Perspect Biol*
730 **5**: a012567. doi: 10.1101/cshperspect.a012567.
- 731 Der Sarkissian C, Ermini L, Schubert M, Yang MA, Librado P, Fumagalli M, Jónsson H, Bar-
732 Gal GK, Albrechtsen A, Vieira FG, Petersen B, Ginolhac A, Seguin-Orlando A,
733 Magnussen K, Fages A, Gamba C, Lorente-Galdos B, Polani S, Steiner C, Neuditschko M,
734 Jagannathan V, Feh C, Greenblatt CL, Ludwig A, Abramson NI, Zimmermann W,
735 Schafberg R, Tikhonov A, Slicheritz-Ponten T, Willerslev E, Marques-Bonet T, Ryder OA,

- 736 McCue M, Rieder S, Leeb T, Slatkin M, Orlando L. 2015. Evolutionary Genomics and
737 Conservation of the Endangered Przewalski's Horse. *Curr Biol* **25**:2577-2583. doi:
738 10.1016/j.cub.2015.08.032.
- 739 Durand EY, Patterson N, Reich D, Slatkin M. 2011. Testing for ancient admixture between
740 closely related populations. *Mol Biol Evol* **28**:2239–2252. doi: 10.1093/molbev/msr048.
- 741 Ellstrand NC, Heredia SM, Leak-Garcia JA, Heraty JM, Burger JC, Yao L, Nohzadeh-
742 Malakshah S, Ridley CE. 2010. Crops gone wild: evolution of weeds and invasives from
743 domesticated ancestors. *Evol Appl* **3**:494-504. doi: 10.1111/j.1752-4571.2010.00140.x.
- 744 Eriksson A, Manica A. 2012. Effect of ancient population structure on the degree of
745 polymorphism shared between modern human populations and ancient hominins. *Proc*
746 *Natl Acad Sci U S A* **109**:13956–13960. doi: 10.1073/pnas.1200567109.
- 747 Feulner PG, Chain FJ, Panchal M, Eizaguirre C, Kalbe M, Lenz TL, Mundry M, Samonte IE,
748 Stoll M, Milinski M, Reusch TB, Bornberg-Bauer E. 2013. Genome-wide patterns of
749 standing genetic variation in a marine population of three-spined sticklebacks. *Mol Ecol*
750 **22**:635-649. doi: 10.1111/j.1365-294X.2012.05680.x.
- 751 Fortes GG, Paijmans JLA. 2015. Analysis of whole mitogenomes from ancient samples. In
752 Whole Genome Amplification. Editor T. Kroneis, Humana Press, USA. Available on
753 ArXiv.
- 754 Gansauge MT, Meyer M. 2013. Single-stranded DNA library preparation for the sequencing
755 of ancient or damaged DNA. *Nat Protoc* **8**:737–748. doi: 10.1038/nprot.2013.038.
- 756 Gębczyński M, Tomaszewska-Guszkiewicz K. 1987. Genetic variability in the European
757 bison. *Biochem Syst Ecol* **15**:285-288.
- 758 Gralak B, Krasieńska M, Niemczewski C, Krasieński ZA, Żurkowski M. 2004. Polymorphism
759 of bovine microsatellite DNA sequences in the lowland European bison. *Acta Theriol*
760 **49**:449-456.

761 Green RE, Krause J, Briggs AW, Maricic T, Stenzel U, Kircher M, Patterson N, Li H, Zhai
762 W, Fritz MH, Hansen NF, Durand EY, Malaspinas AS, Jensen JD, Marques-Bonet T,
763 Alkan C, Prüfer K, Meyer M, Burbano HA, Good JM, Schultz R, Aximu-Petri A, Butthof
764 A, Höber B, Höffner B, Siegemund M, Weihmann A, Nusbaum C, Lander ES, Russ C,
765 Novod N, Affourtit J, Egholm M, Verna C, Rudan P, Brajkovic D, Kucan Z, Gusic I,
766 Doronichev VB, Golovanova LV, Lalueza-Fox C, de la Rasilla M, Fortea J, Rosas A,
767 Schmitz RW, Johnson PL, Eichler EE, Falush D, Birney E, Mullikin JC, Slatkin M,
768 Nielsen R, Kelso J, Lachmann M, Reich D, Pääbo S. 2010. A draft sequence of the
769 Neandertal genome. *Science* **328**:710–722. doi: 10.1126/science.1188021.

770 Halbert ND, Derr JN. 2007. A comprehensive evaluation of cattle introgression into US
771 federal bison herds. *J Hered* **98**:1-12.

772 Hartl GB, Pucek Z. 1994. Genetic depletion in the European bison *Bison bonasus* and the
773 significance of electrophoretic heterozygosity for conservation. *Conserv Biol* **8**:167-174.

774 Hassanin A, An J, Ropiquet A, Nguyen TT, Couloux A. 2013. Combining multiple autosomal
775 introns for studying shallow phylogeny and taxonomy of Laurasiatherian mammals:
776 Application to the tribe Bovini (Cetartiodactyla, Bovidae). *Mol Phylogenet Evol* **66**:766–
777 75. doi: 10.1016/j.ympev.2012.11.003.

778 Hassanin A, Delsuc F, Ropiquet A, Hammer C, Jansen van Vuuren B, Matthee C, Ruiz-
779 Garcia M, Catzeflis F, Areskoug V, Nguyen TT, Couloux A. 2012. Pattern and timing of
780 diversification of Cetartiodactyla (Mammalia, Laurasiatheria), as revealed by a
781 comprehensive analysis of mitochondrial genomes. *C R Biol* **335**:32-50. doi:
782 10.1016/j.crv.2011.11.002.

783 Hassanin A, Ropiquet A. 2007. Resolving a zoological mystery: the kouprey is a real species.
784 *Proc Biol Sci* **274**:2849–55. doi: 10.1098/rspb.2007.0830

- 785 Heintzman PD, Froese D, Ives JW, Soares AE, Zazula GD, Letts B, Andrews TD, Driver JC,
786 Hall E, Hare PG, Jass CN, MacKay G, Southon JR, Stiller M, Woywitka R, Suchard MA,
787 Shapiro B. 2016. Bison phylogeography constrains dispersal and viability of the Ice Free
788 Corridor in western Canada. *PNAS USA*. doi: 10.1073/pnas.1601077113.
- 789 Heintzman PD, Zazula GD, Cahill JA, Reyes AV, MacPhee RD, Shapiro B. 2015. Genomic
790 data from extinct North American Camelops revise camel evolutionary history. *Mol Biol*
791 *Evol* **32**:2433–2440. doi: 10.1093/molbev/msv128.
- 792 Heywood P. 2013. The quagga and science: what does the future hold for this extinct zebra?
793 *Perspect Biol Med* **56**:53-64. doi: 10.1353/pbm.2013.0008.
- 794 Hofreiter M, Paijmans JLA, Goodchild H, Speller CF, Barlow A, Fortes GG, Thomas JA,
795 Ludwig A, Collins MJ. 2015. The future of ancient DNA: Technical advances and
796 conceptual shifts. *Bioessays* **37**:284–293. doi: 10.1002/bies.201400160.
- 797 Huerta-Cepas J, Serra F, Bork P. 2016. ETE 3: Reconstruction, analysis and visualization of
798 phylogenomic data. *Mol Biol Evol* **33**:1635-1638. doi: 10.1093/molbev/msw046.
- 799 Kamiński S, Zabolewicz T. 1997. DDE I restriction fragment length polymorphism in the
800 *Bison bonasus* untranslated region of kappa-casein gene. *Acta Academiae AC Technicae*
801 *Olstenensis, Zootechnica* **47**:174-153.
- 802 Karbowski G, Demiaszkiewicz AW, Pyziel AM, Wita I, Moskwa B, Werszko J, Bień J,
803 Goździk K, Lachowicz J, Cabaj W. 2014a. The parasitic fauna of the European bison
804 (*Bison bonasus*) (Linnaeus, 1758) and their impact on the conservation. Part 1. The
805 summarising list of parasites noted. *Acta Parasitol* **59**:363-371. doi:10.2478/s11686-014-
806 0252-0.
- 807 Karbowski G, Demiaszkiewicz AW, Pyziel AM, Wita I, Moskwa B, Werszko J, Bień J,
808 Goździk K, Lachowicz J, Cabaj W. 2014b. The parasitic fauna of the European bison

- 809 (*Bison bonasus*) (Linnaeus, 1758) and their impact on the conservation. Part 2. The
810 structure and changes over time. *Acta Parasitol* **59**:372-9. doi:10.2478/s11686-014-0253-z.
- 811 Kerley GIH, Kowalczyk R, Cromsigt JPGM. 2012. Conservation implications of the refugee
812 species concept and the European bison: king of the forest or refugee in a marginal habitat?
813 *Ecography* **35**:519–529. doi: 10.1111/j.1600-0587.2011.07146.x
- 814 Krasińska M, Krasiński ZA. 2007. European bison. The Nature Monograph. Mammal
815 Research Institute Polish Academy of Science, Białowieża.
- 816 Krzysiak MK, Dudek K, Krajewska M, Bednarek D, Szulowski K. 2014. Serological studies
817 to determine the occurrence of Johne's disease and mycoplasma infection in the Northern-
818 East Polish population of European bison (*Bison bonasus*). *Pol J Vet Sci* **17**:721-723. doi:
819 10.2478/pjvs-2014-0105.
- 820 Kuemmerle T, Hickler T, Olofsson J, Schurgers G, Radeloff VC. 2011. Reconstructing range
821 dynamics and range fragmentation of European bison for the last 8000 years. *Diversity*
822 *Distrib* **18**:47–59. doi: 10.1111/j.1472-4642.2011.00849.x.
- 823 Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler
824 transform. *Bioinformatics* **25**:1754–1760. doi: 10.1093/bioinformatics/btp324.
- 825 Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin
826 R; 1000 Genome Project Data Processing Subgroup. 2009. The Sequence Alignment/Map
827 format and SAMtools. *Bioinformatics* **25**:2078–2079. doi: 10.1093/bioinformatics/btp352.
- 828 Luenser K, Fickel J, Lehnen A, Speck S, Ludwig A. 2005. Low level of genetic variability in
829 European bisons *Bison bonasus* from the Bialowieza National Park in Poland. *Eur. J Wildl*
830 *Res* **51**:84-87. doi: 10.1007/s10344-005-0081-4.
- 831 Lühken G, Wagner HW, Seichter D, Hecht W, Erhardt G. 2009. Genetic Characterization of a
832 Sheep-Dwarf Goat Hybrid. *Cytogenet Genome Res* **125**:158–161. doi:10.1159/000227841.

- 833 Majewska AC, Werner A, Cabaj W, Moskwa B. 2014. The first report of *Toxoplasma gondii*
834 antibodies in free-living European bison (*Bison bonasus bonasus* Linnaeus). *Folia*
835 *Parasitol* **61**:18-20.
- 836 Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing
837 reads. *EMBnet.journal* **17**:10-12. doi: 10.14806/ej.17.1.200.
- 838 Meyer M, Kircher M. 2010. Illumina sequencing library preparation for highly multiplexed
839 target capture and sequencing. *Cold Spring Harb Protoc* **6**:pdb.prot5448. doi:
840 10.1101/pdb.prot5448.
- 841 Monzón J. 2014. First regional evaluation of nuclear genetic diversity and population
842 structure in northeastern coyotes (*Canis latrans*). *F1000Res* **3**:66. doi:
843 10.12688/f1000research.3567.1.
- 844 Moskwa B, Bień J, Goździk K, Cabaj W. 2014. The usefulness of DNA derived from third
845 stage larvae in the detection of *Ashworthius sidemi* infection in European bison, by a
846 simple polymerase chain reaction. *Parasit Vectors* **7**:215. doi:10.1186/1756-3305-7-215.
- 847 Oleński K, Tokarska M, Hering DM, Puckowska P, Ruś A, Pertoldi C, Kamiński S. 2015.
848 Genome-wide association study for posthitis in the free-living population of European
849 bison (*Bison bonasus*). *Biol Direct*: **10**:2. doi: 10.1186/s13062-014-0033-6.
- 850 Panasiewicz G, Zamojska A, Bieniek M, Gizejewski Z, Szafranska B. 2015. Persistent
851 Müllerian duct syndrome (PMDS) in the Polish free-ranged bull populations of the
852 European bison (*Bison bonasus* L.). 2014. *Anim Reprod Sci* **152**:123-136.
853 doi:10.1016/j.anireprosci.2014.11.012.
- 854 Park SD, Magee DA, McGettigan PA, Teasdale MD, Edwards CJ, Lohan AJ, Murphy A,
855 Braud M, Donoghue MT, Liu Y, Chamberlain AT, Rue-Albrecht K, Schroeder S, Spillane
856 C, Tai S, Bradley DG, Sonstegard TS, Loftus BJ, MacHugh DE. 2015. Genome
857 sequencing of the extinct Eurasian wild aurochs, *Bos primigenius*, illuminates the

- 858 phylogeography and evolution of cattle. *Genome Biol* **16**:234. doi: 10.1186/s13059-015-
859 0790-2.
- 860 Pucek Z, Belousova IP, Krasieńska M, Krasieński ZA, Olech W. 2004. European bison. Status
861 Survey and Conservation Action Plan. IUCN/SSC Bison Specialist Group, IUCN. Gland,
862 Switzerland and Cambridge, ix+54.
- 863 Pucek Z. 1991. History of the European bison and problems of its protection and
864 management. In: Bobek B, Perzanowski K, Regelin WL, editors. *Global trends in wildlife*
865 *management*. Kraków-Warszawa (Poland). Świat Press.
- 866 Raczyński J. 2014. Księga Rodowodowa Żubrów. Białowiecki Park Narodowy, Białowieża.
- 867 Radwan J, Kawałko A, Wójcik JM, Babik W. 2007. MHC-DRB3 variation in a free-living
868 population of the European bison, *Bison bonasus*. *Mol Ecol* **16**:531-540.
- 869 Ramsey CB, Gigham T, Leach P. 2004a. Towards high-precision AMS: progress and
870 limitations. *Radiocarbon* **46**:17-24.
- 871 Ramsey CB, Higham T, Bowles A, Hedgeset R. 2004b. Improvements to the pretreatment of
872 bone at Oxford. *Radiocarbon* **46**: 155-63.
- 873 Rohland N, Siedel H, Hofreiter M. 2010. A rapid column-based ancient DNA extraction
874 method for increased sample throughput. *Mol Ecol Resour* **10**:677–683. doi:
875 10.1111/j.1755-0998.2009.02824.x.
- 876 R Core Team (2014). R: A language and environment for statistical computing. R Foundation
877 for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>.
- 878 Reimer PJ, Bard E, Bayliss A, Beck JW, Blackwell PG, Ramsey CB, Buck CE, Cheng H,
879 Edwards RL, Friedrich M, Grootes PM, Guilderson TP, Haflidason H, Hajdas I, Hatté C,
880 Heaton TJ, Hoffmann DL, Hogg AG, Hughen KA, Kaiser KF, Kromer B, Manning SW,
881 Niu M, Reimer RW, Richards DA, Scott EM, Southon JR, Staff RA, Turney CSM, van der

- 882 Plicht J. 2013. IntCal13 and Marine13 radiocarbon age calibration curves 0-50,000 years
883 cal BP. *Radiocarbon* **55**:1869–1887.
- 884 Sambrook J, Russell DW. 2001. Molecular cloning. A laboratory manual. Cold Spring Harbor
885 Laboratory Press.
- 886 Sawyer S, Krause J, Guschanski K, Savolainen V, Pääbo S. 2012. Temporal patterns of
887 nucleotide misincorporations and DNA fragmentation in ancient DNA. *PLoS One*
888 **7**:e34131. doi: 10.1371/journal.pone.0034131.
- 889 Shapiro B, Drummond AJ, Rambaut A, Wilson MC, Matheus PE, Sher AV, Pybus OG,
890 Gilbert MT, Barnes I, Binladen J, Willerslev E, Hansen AJ, Baryshnikov GF, Burns JA,
891 Davydov S, Driver JC, Froese DG, Harington CR, Keddie G, Kosintsev P, Kunz ML,
892 Martin LD, Stephenson RO, Storer J, Tedford R, Zimov S, Cooper A. 2004. Rise and fall
893 of the Beringian steppe bison. *Science* **306**:1561–1565.
- 894 Shapiro B, Hofreiter M. 2012. Ancient DNA. Methods and protocols. New York, Humana
895 Press.
- 896 Sherkow JS, Greely HT. 2013. Genomics. What if extinction is not forever? *Science* **340**:32-
897 33. doi: 10.1126/science.1236965.
- 898 Sipko TP, Rautian GS, Udina IG, Rakitskaya, TA. 1995. Investigation of blood group
899 polymorphism in European bison (*Bison bonasus*). *Russ J Genet* **32**:356-351.
- 900 Sipko TP, Rautian GS, Udina IG, Strelchenko, NS. 1997. Conservation of genetic material
901 from endangered and economically important species in the establishment of cryobanks.
902 In: Physiology and general biology reviews. TM Tupaev and NK Kolcov, eds, 35-99.
- 903 Sipko TP, Udina IG, Badagueva YN, Sulimowa GE. 1994. Polymorphism of kappa-casein
904 gene in the family Bovidae: comparative analysis. *Russ J Genet* **30**:203-207.
- 905 Slatis HM. 1960. An analysis of inbreeding in the European bison. *Genetics* **45**:275–287.
906 PMID: 17247924.

- 907 Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of
908 large phylogenies. *Bioinformatics* **30**:1312-3. doi: 10.1093/bioinformatics/btu033.
- 909 Steyer K, Kraus RHS, Mölich T, Anders O, Cocchiararo B, Frosch C, Geib A, Götz M,
910 Herrmann M, Hupe K, Kohnen A, Krüger M, Müller F, Pir JB, Reiners TE, Roch S,
911 Schade U, Schiefenhövel P, Siemund M, Simon O, Steeb S, Streif S, Streit B, Thein J,
912 Tiesmeyer A, Trinzen M, Vogel B, Nowak C. 2016. Large-scale genetic census of an
913 elusive carnivore, the European wildcat (*Felis s. silvestris*). *Conserv Genet* 1-17.
914 doi:10.1007/s10592-016-0853-2.
- 915 Tiedemann R, Nadlinger K, Pucek Z. 1998. Mitochondrial DNA-RFLP analysis reveals low
916 levels of genetic variation in European bison *Bison bonasus*. *Acta Theriol Suppl* **5**:83-87.
- 917 Tokarska M, Bunevich AN, Demontis D, Sipko T, Perzanowski K, Baryshnikov G,
918 Kowalczyk R, Voitukhovskaya Y, Wójcik JM, Marczuk B, Ruczyńska I, Pertoldi C. 2015.
919 Genes of the extinct Caucasian bison still roam the Białowieża Forest and are the source of
920 genetic discrepancies between Polish and Belarusian populations of the European bison,
921 *Bison bonasus*. *Biol J Linnean Soc* **114**:752–763. doi:10.1111/bij.12470.
- 922 Tokarska M, Kawalko A, Wójcik JM, Pertoldi C. 2009a. Genetic variability in the European
923 bison (*Bison bonasus*) population from Białowieża forest over 50 years. *Biol J Linn Soc*
924 **97**:801-809. doi: 10.1111/j.1095-8312.2009.01203.x
- 925 Tokarska M, Marshall T, Kowalczyk R, Kawałko A, Wójcik JM, Pertoldi C, Kristensen TN,
926 Loeschcke V, Gregersen VR, Bendixen C. 2009b. Effectiveness of microsatellite and SNP
927 loci (BovineSNP50 BeadChip, Illumina®) in parentage assignment and identity analysis in
928 species with low genetic diversity: the case of European bison. *Heredity* **103**:326-332.
929 doi:10.1038/hdy.2009.73.
- 930 Udina IG, Shakhav GO. 1998. Restriction fragment length polymorphism (RFLP) of exon 2
931 of the MHC Bibo-DRB3 gene in European bison *Bison bonasus*. *Acta Theriol* **43**:75-82.

- 932 Udina IG, Sokolova SS, Sipko TP, Sulimova GE. 1994. Comparative characteristics of DNA
933 polymorphism for major histocompatibility complex loci DQB and DRB in representatives
934 of the Bovidae family. *Genetika* **30**:356-360.
- 935 van Vuure C. 2005. Retracing the Aurochs: History, Morphology and Ecology of an extinct
936 wild Ox. Pensoft Publishers. Sofia-Moscow. ISBN 954-642-235-5.
- 937 Wolf DE, Takebayashi N, Riesberg LH, 2001. Predicting the risk of extinction through
938 hybridization. *Conserv Biol* **15**:1039–1053.
- 939 Wójcik JM, Kawalko A, Tokarska M, Jaarola M, Vallenback P, Pertoldi C. 2009. Post-
940 bottleneck mtDNA diversity in a free-living population of European bison: implications for
941 conservation. *J Zool* **277**:81-87. doi: 10.1111/j.1469-7998.2008.00515.x

942

943

944 ***FIGURE AND TABLE LEGENDS***

945

946 Figure 1. Map of Western Europe showing the putative historical range of Lowland wisent
947 (shaded green) and Caucasian wisent (shaded grey) based on bone remains and written
948 records (according to Benecke et al., 2005; Kuemmerle et al., 2011; Tokarska et al., 2011;
949 Bocherens et al., 2015) and sample locations. Black circles indicate contemporary free-
950 ranging modern L line herds and white circles indicate modern LC line herds. Purple and
951 peach circles denote the locations of investigated modern L (MdL1, MdL2) and LC (MdLC)
952 line wisent, respectively, orange squares show the location of the Holocene wisent (Bb1-3)
953 and blue and yellow triangles indicate historical founding wisent from the Pszczyna
954 population (PLANTA and PLATEN) and the extinct Caucasian wisent (Cc1-3), respectively.

955

956 Figure 2. Pairwise genomic divergence among wisent. A, B, C, show probability densities for
957 pairwise transversion divergence (x axes) along a sliding window of 1Mb. Individual plots
958 show all pairwise comparisons among modern and founding individuals (A); comparisons of
959 modern and founding individuals and Caucasian wisent Cc1 (B); and comparisons of
960 Caucasian wisent Cc2 and all other individuals (C). Specific comparisons discussed in the text
961 are identified by colours, according to the key at the top right of each plot. Schematic
962 neighbour-joining phylogeny based on whole genome distances (D).

963

964 Figure 3. Population history of Lowland and Caucasian wisent, estimated using two
965 representatives of each. The pie chart shows the percentage of 1Mb genomic blocks
966 supporting each alternatively rooted tree topology. The indicates fraction of genome blocks
967 returning both wisent subspecies as reciprocally monophyletic. Dark and light blue colours
968 show the next most frequently encountered topologies in which Caucasian wisent are
969 paraphyletic (dark blue: Cc2 most divergent, and light blue: Cc1 most divergent, the first of
970 which is compatible with estimates pairwise genomic divergence (see Fig 2d).

971

972 Figure 4. Comparison of pairwise genomic divergence within three bovid species: wisent
973 (*Bison bonasus*), domestic cattle (*Bos taurus*), and yak (*Bos grunniens*). Probability densities
974 were calculated along a sliding window of 1Mb from transversions only (A), and from
975 transitions and transversions (B). For wisent, only modern samples included.

976

977 Figure 5. Results of *D* statistic analysis. Red and grey points show significant and
978 nonsignificant *D* values (x axis), respectively, and show: the genetic contribution of the
979 founders (F) to the modern individuals (A); Caucasian wisent (Cc) admixture with modern L
980 and LC herds (MdL and MdLC) and one founder, PLANTA, relative to the least Caucasian

981 admixed wisent, founder PLATEN (B); Caucasian wisent admixture among modern wisent
982 (C); apparent cattle (DC) admixture with all investigated wisent (W) relative to aurochs (Aur)
983 (D); variance in cattle/aurochs admixture among wisent (W) compared to PLATEN (E).
984 Detailed *D* statistic results are provided in Supporting Table S5.

985

986 Figure 6. Genomic admixture map (A) of Caucasian bison ancestry in the modern LC line
987 individual (MdLC). Colored blocks indicate 1Mb genomic blocks returning alternative tree
988 topologies, blue blocks are compatible with the species tree; yellow blocks return the
989 monophyly of the modern LC line and Caucasian wisent, and likely result from admixture and
990 to a lesser extent incomplete lineage sorting; red blocks return the monophyly of PLATEN
991 and Caucasian wisent and likely result from incomplete lineage sorting. “X” shows blocks
992 with missing data. The pie chart (B) shows the percentage of 1Mb genomic supporting each
993 tree topology identified by colours, according to the key presented above.

994

995 Figure 7. Variation in the sizes of genomic blocks in modern L (blue) and LC (red) likely
996 resulting from Caucasian wisent admixture. Plots show cumulative probability densities
997 calculated at a scale of 1MB. Genomic blocks in the LC line wisent (red) result from
998 admixture occurring around 90 years ago; the lower abundance of larger admixed blocks in
999 the modern L line wisent support that this admixture event preceded the former. The plots
1000 have been truncated to aid visualisation, and single blocks of 18MB and 22MB in the LC line
1001 individual are not shown. The largest block size detected in the modern L line individual was
1002 8MB.

1003

1004 Figure 8. Schematic diagram showing inferred admixture among wisent, and among wisent
1005 and the cattle/aurochs lineage. Arrows indicate the direction of the geneflow. Black lines

1006 indicate admixture between wisent and cattle/aurochs, yellow lines/arrow - between
1007 Caucasian and founding or modern wisent respectively, and the blue arrow - from founders to
1008 modern wisent.

1009

1010 Table 1. Basic information about sampled individuals. The last column is the average
1011 coverage for each sample after mapping it to either the water buffalo (*Bubalus bubalis*)
1012 nuclear reference genome (GenBank accession no. GCA_000471725.1) or the American
1013 bison (*Bison bison*) mitochondrial reference genome (GenBank accession no. NC_012346.1).

1014

1015 **SUPPORTING INFORMATION**

1016

1017 Figure S1. A phylogeny of wisent (A) and the Bovini (B), inferred from a partitioned
1018 maximum likelihood (ML) analysis of whole mitochondrial genomes. (A) is an expansion of
1019 the region in the blue box in (B). The green box in (A) highlights the two haplotypes found in
1020 sampled historic and modern individuals: one in Caucasian wisent (Cc2) and the other one
1021 identical to haplotype published already (JN632602; Hassanin et al., 2012) in all remaining
1022 historical and modern wisent (see Table S1). Tips are coloured based on whether the
1023 haplotype occurs in modern individuals (black), or only in historical or ancient individuals,
1024 and therefore likely extinct (red). The outgroup, *Pseudoryx nghetinhensis*, is not shown.

1025 Branch support is indicated by bootstrap percentages based on 500 ML bootstrap replicates
1026 (above branches) and Bayesian posterior probabilities (below branches).

1027

1028 Figure S2. DNA fragmentation (upper four plots) and deamination (lower two plots) patterns
1029 of founding wisent, PLANTA, sequencing reads. For deamination plots, red lines show the
1030 frequency of C to T substitutions (Y axes) in the sequenced historical DNA fragments relative

1031 to the reference genome at the 5' (left plot) and 3' (right plot) fragment ends. X axes show
1032 sequenced positions moving internally from the 5' (positive values) and 3' (negative values)
1033 fragment ends. Elevated rate of C to T substitutions at fragment ends are indicative of DNA
1034 damage.

1035

1036 Figure S3. DNA fragmentation and deamination patterns of founding wisent, PLATEN,
1037 sequencing reads.

1038

1039 Figure S4. DNA fragmentation and deamination patterns of Caucasian wisent, Cc1,
1040 sequencing reads.

1041

1042 Figure S5. DNA fragmentation and deamination patterns of Caucasian wisent, Cc2,
1043 sequencing reads.

1044

1045 Table S1. Species and accession data for sequences included in the mitochondrial genome
1046 phylogenetic analysis.

1047

1048 Table S2. Full details of sequencing and mapping results for modern wisent samples. SR –
1049 merged reads, PE – unmerged reads.

1050

1051 Table S3. Full details of sequencing and mapping results for archival wisent samples. SR –
1052 merged reads, PE – unmerged reads.

1053

1054 Table S4. Full details of mapping results for data downloaded from the NCBI Short Read
1055 Archive. SR – merged reads, PE – unmerged reads.

1056 Table S5. Detailed D statistic results. In red significant results are shown. Results highlighted
1057 in grey are presented also in Figure 5.

1058

1059 Table S6. \hat{f} results. In the last column percentage of the genome resulted from hybridisation is
1060 given.

1061

1062

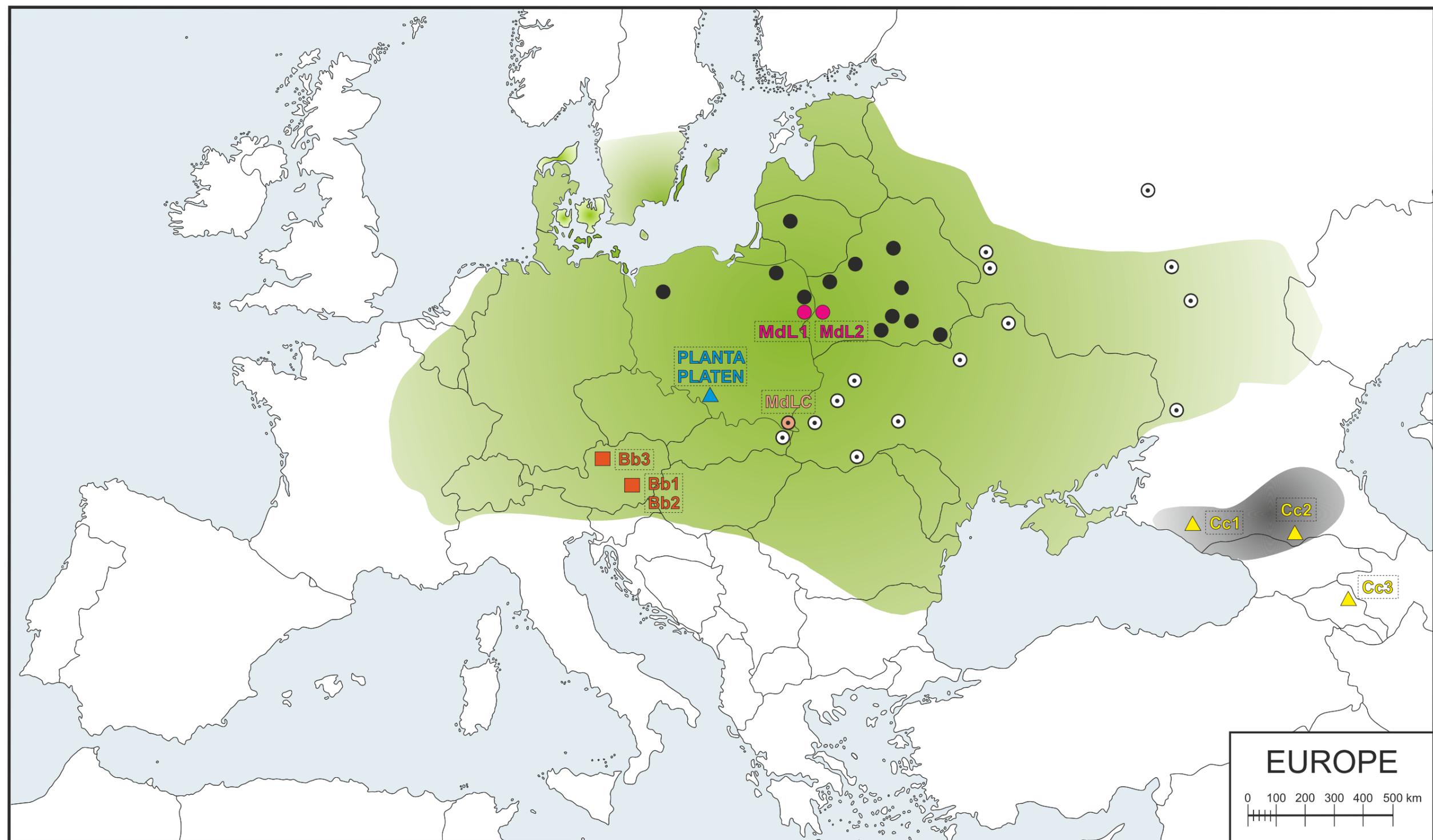
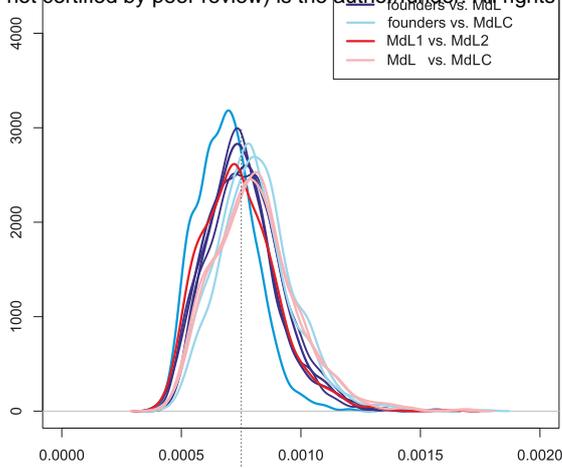
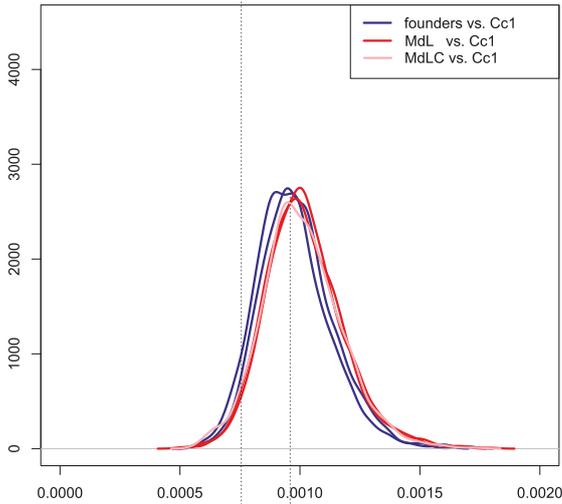


Figure 1. Map of Western Europe showing the putative historical range of Lowland wisent (shaded green) and Caucasian wisent (shaded grey) based on bone remains and written records (according to Benecke et al., 2005; Kuemmerle et al., 2011; Tokarska et al., 2011; Bocherens et al., 2015) and sample locations. Black circles indicate contemporary free-ranging modern L line herds and white circles indicate modern LC line herds. Purple and peach circles denote the locations of investigated modern L (MdL1, MdL2) and LC (MdLC) line wisent, respectively, orange squares show the location of the Holocene wisent (Bb1-3) and blue and yellow triangles indicate historical founding wisent from the Pszczyna population (PLANTA and PLATEN) and the extinct Caucasian wisent (Cc1-3), respectively.

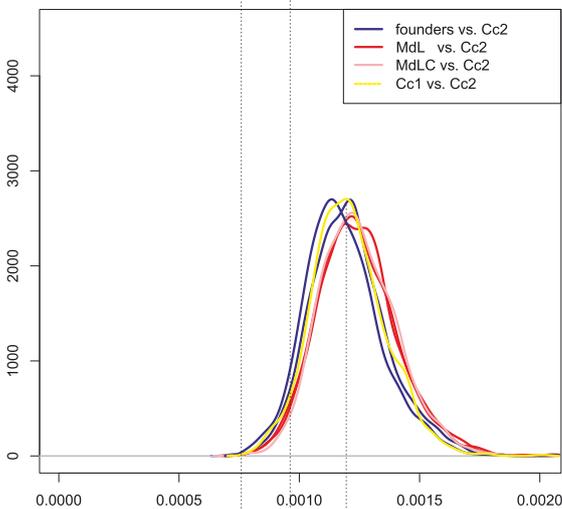
A



B



C



D

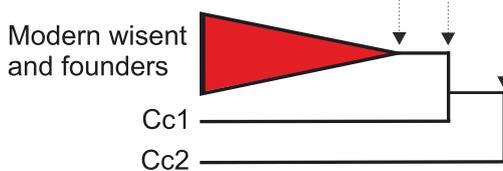


Figure 2. Pairwise genomic divergence among wisent. A, B, C, show probability densities for pairwise transversion divergence (x axes) along a sliding window of 1Mb. Individual plots show all pairwise comparisons among modern and founding individuals (A); comparisons of modern and founding individuals and Caucasian wisent Cc1 (B); and comparisons of Caucasian wisent Cc2 and all other individuals (C). Specific comparisons discussed in the text are identified by colours, according to the key at the top right of each plot. Schematic neighbour-joining phylogeny based on whole genome distances (D).

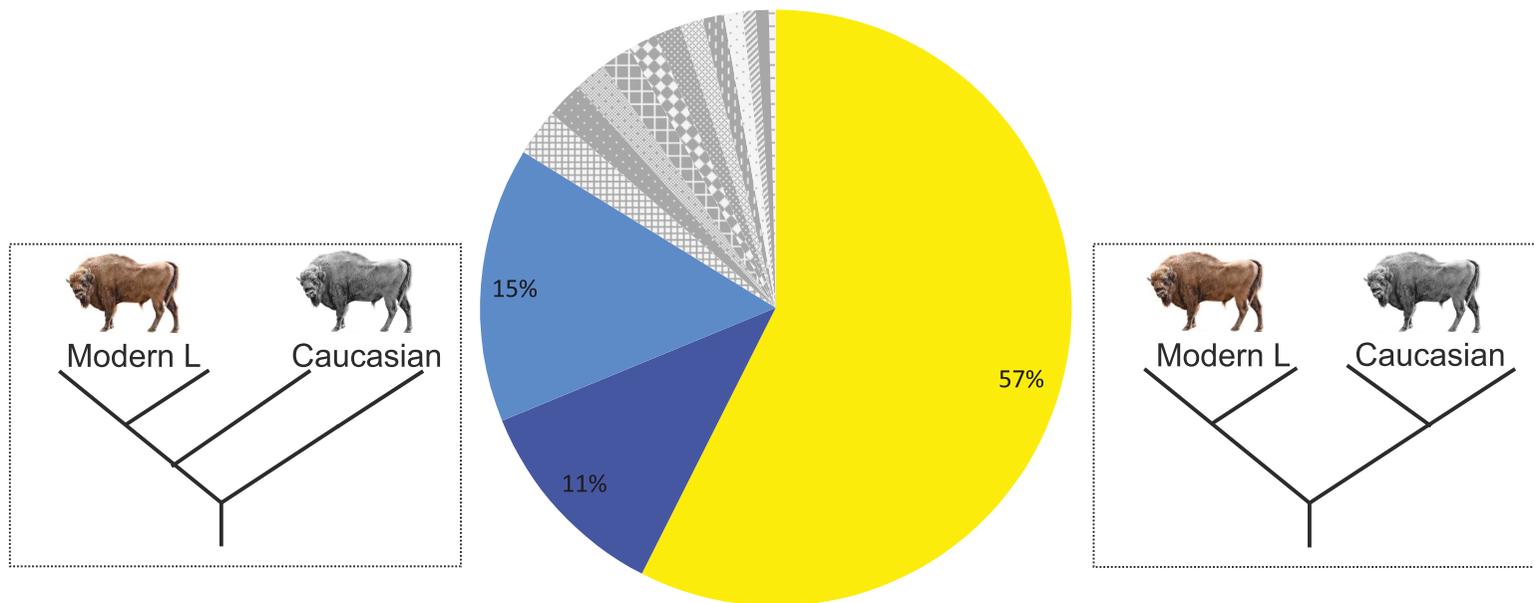


Figure 3. Population history of Lowland and Caucasian wisent, estimated using two representatives of each. The pie chart shows the percentage of 1Mb genomic blocks supporting each alternatively rooted tree topology. The indicates fraction of genome blocks returning both wisent subspecies as reciprocally monophyletic. Dark and light blue colours show the next most frequently encountered topologies in which Caucasian wisent are paraphyletic (dark blue: Cc2 most divergent, and light blue: Cc1 most divergent, the first of which is compatible with estimates pairwise genomic divergence (see Fig 2d).

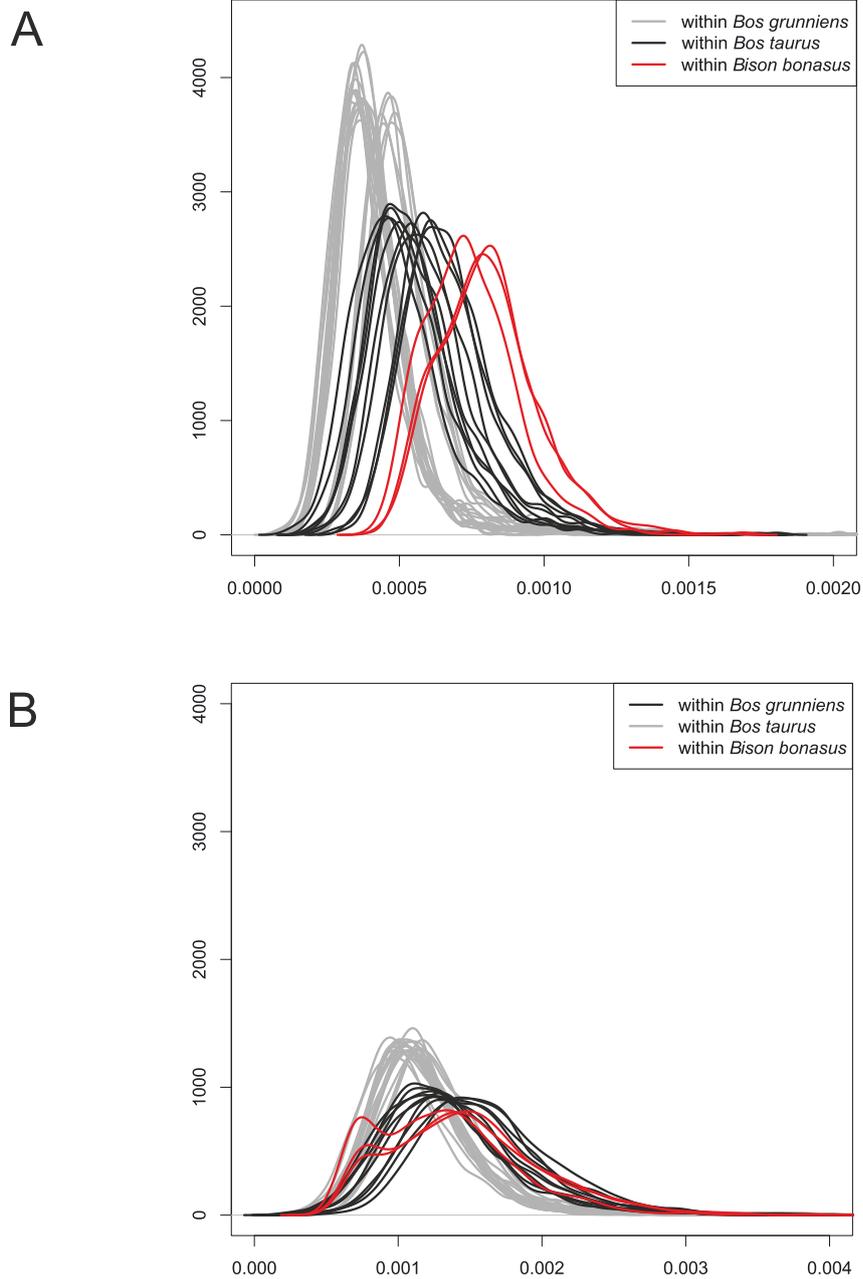


Figure 4. Comparison of pairwise genomic divergence within three bovid species: wisent (*Bison bonasus*), domestic cattle (*Bos taurus*), and yak (*Bos grunniens*). Probability densities were calculated along a sliding window of 1Mb from transversions only (A), and from transitions and transversions (B). For wisent, only modern samples included.

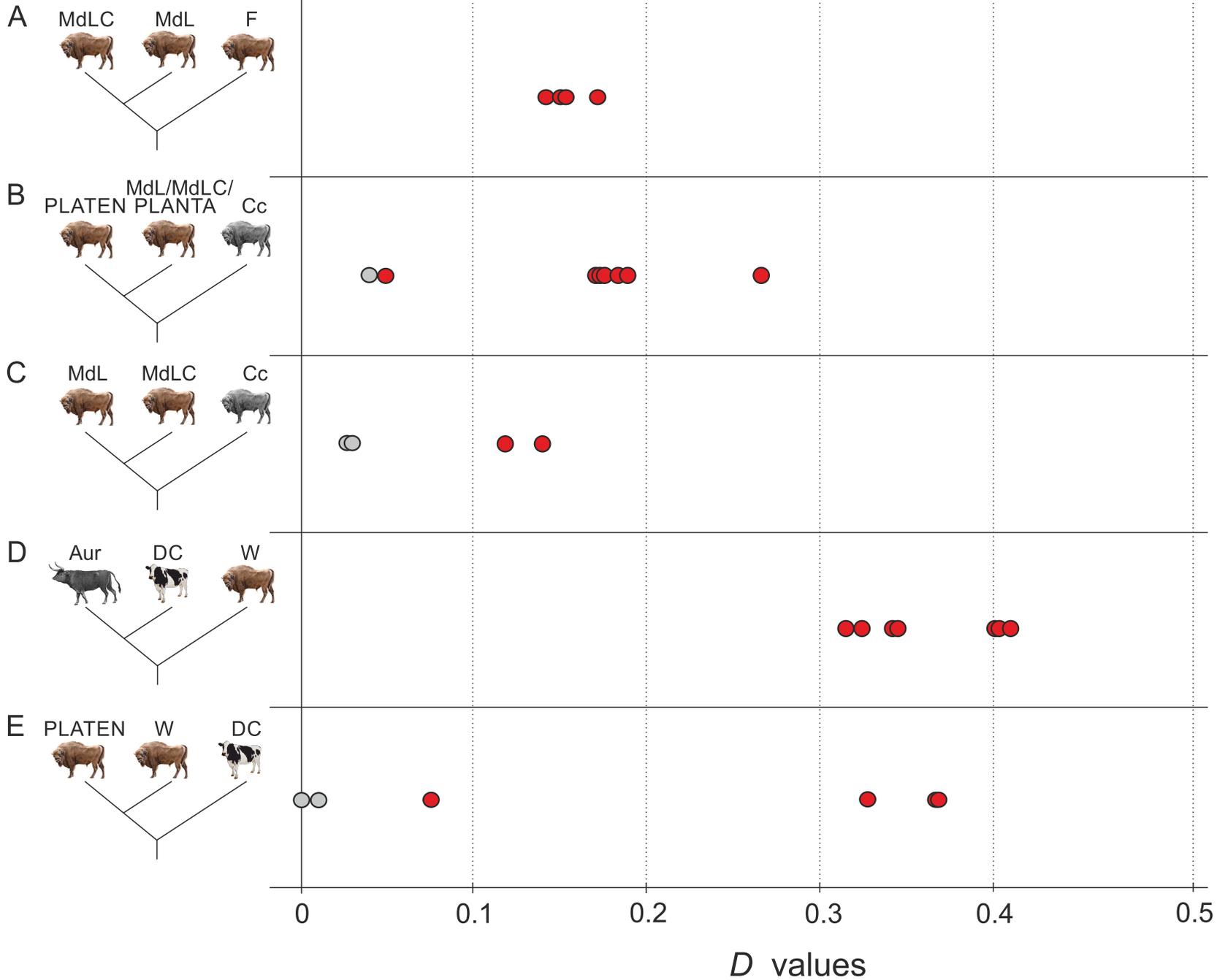


Figure 5. Results of D statistic analysis. Red and grey points show significant and nonsignificant D values (x axis), respectively, and show: the genetic contribution of the founders (F) to the modern individuals (A); Caucasian wisent (Cc) admixture with modern L and LC herds (MdL and MdLC) and one founder, PLANTA, relative to the least Caucasian admixed wisent, founder PLATEN (B); Caucasian wisent admixture among modern wisent (C); apparent cattle (DC) admixture with all investigated wisent (W) relative to aurochs (Aur) (D); variance in cattle/aurochs admixture among wisent (W) compared to PLATEN (E). Detailed D statistic results are provided in Supporting Table S5.

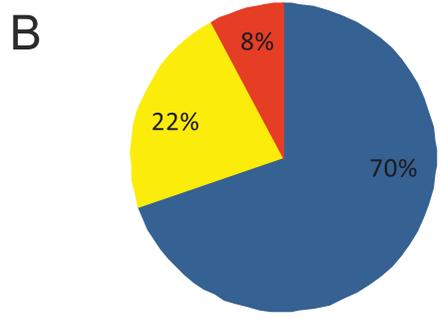
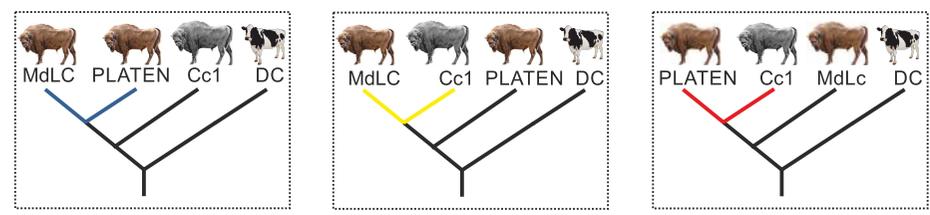


Figure 6. Genomic admixture map (A) of Caucasian bison ancestry in the modern LC line individual (MdLC). Colored blocks indicate 1Mb genomic blocks returning alternative tree topologies, blue blocks are compatible with the species tree; yellow blocks return the monophyly of the modern LC line and Caucasian wisent, and likely result from admixture and to a lesser extent incomplete lineage sorting; red blocks return the monophyly of PLATEN and Caucasian wisent and likely result from incomplete lineage sorting. “X” shows blocks with missing data. The pie chart (B) shows the percentage of 1Mb genomic supporting each tree topology identified by colours, according to the key presented above.

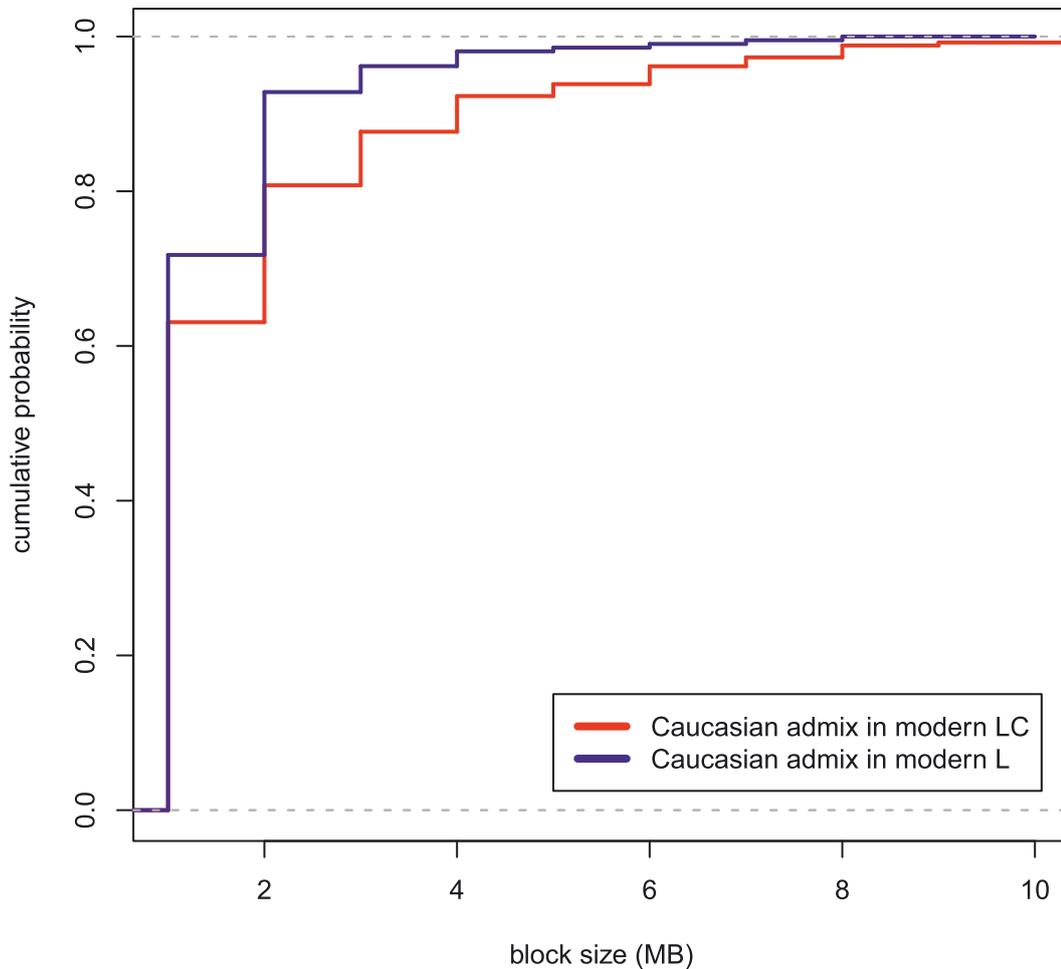


Figure 7. Variation in the sizes of genomic blocks in modern L (blue) and LC (red) likely resulting from Caucasian wisent admixture. Plots show cumulative probability densities calculated at a scale of 1MB. Genomic blocks in the LC line wisent (red) result from admixture occurring around 90 years ago; the lower abundance of larger admixed blocks in the modern L line wisent support that this admixture event preceded the former. The plots have been truncated to aid visualisation, and single blocks of 18MB and 22MB in the LC line individual are not shown. The largest block size detected in the modern L line individual was 8MB.

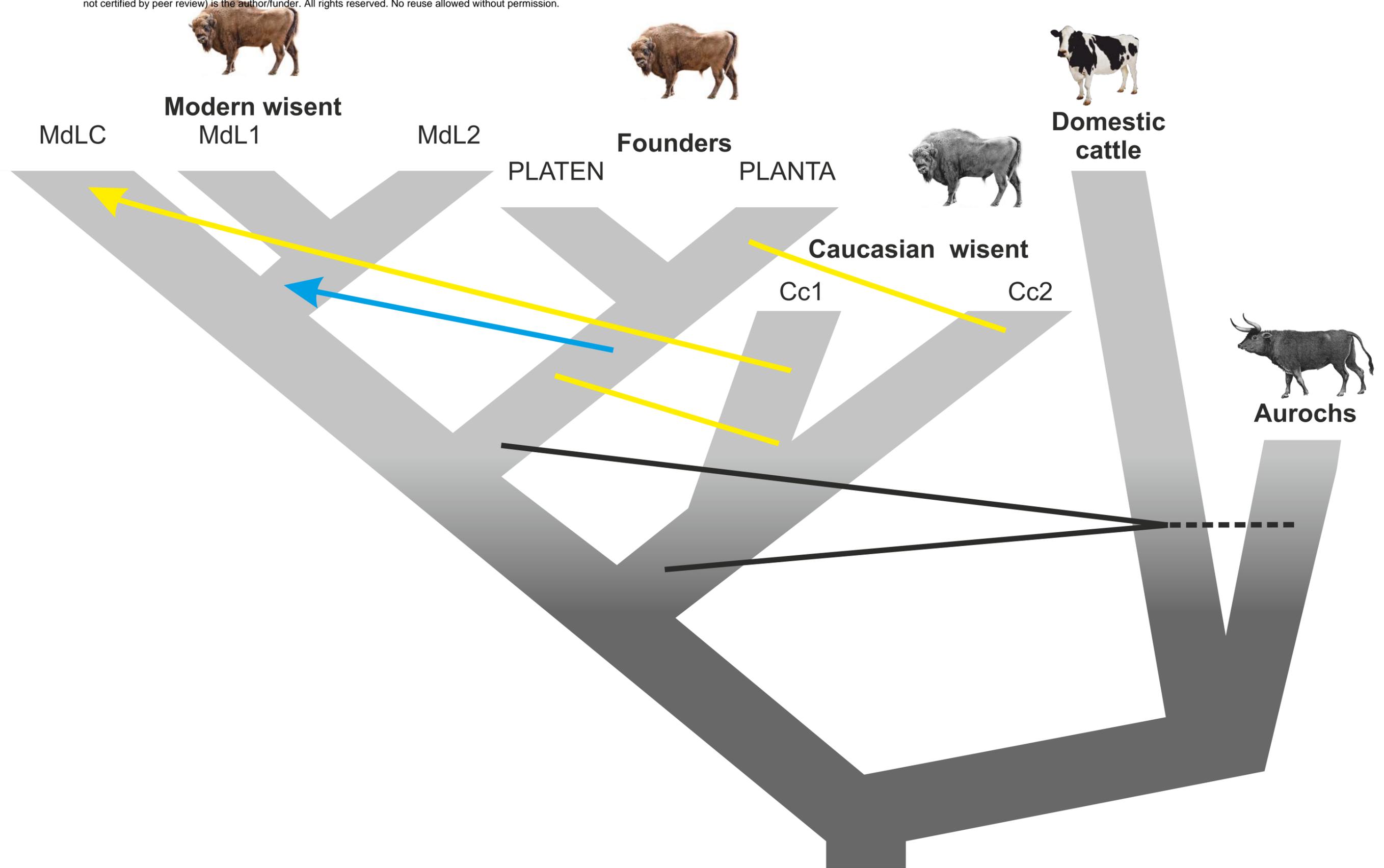


Figure 8. Schematic diagram showing inferred admixture among wisent, and among wisent and the cattle/aurochs lineage. Arrows indicate the direction of the gene flow. Black lines indicate admixture between wisent and cattle/aurochs, yellow lines/arrow - between Caucasian and founding or modern wisent respectively, and the blue arrow - from founders to modern wisent.

Table 1. Basic information about sampled individuals. The last column is the average coverage for each sample after mapping it to either the water buffalo (*Bubalus bubalis*) nuclear reference genome (GenBank accession no. GCA_000471725.1) or the American bison (*Bison bison*) mitochondrial reference genome (GenBank accession no. NC_012346.1).

Sample code	Sample name	Taxon	Group	Sample location	Sample origin	Tissue type	Year of death*/ radiocarbon age	Water buffalo nDNA mapping coverage/ mtDNA coverage
MdL1	Z3	<i>Bison bonasus</i>	Modern Lowland line (L)	Białowieża Forest, Polish part, Poland		kidney	2006	1.59 (n)
MdL2	868			Białowieża Forest, Belarussian part, Belaruss		muscles	2009	1.63 (n)
MdLC	Dy		Modern Lowland-Caucasian line (LC)	Dydiowa, Bieszczady Mountains, Poland		muscles	2012	1.49 (n)
PLANTA	F42 PLANTA	<i>Bison b. bonasus</i>	Founders	Pszczyna, Poland	Upper Silesian Museum, Bytom, Poland	horns	1931	1.50 (n)
PLATEN	M158 PLATEN						1933	1.39 (n)
Cc1	8853	<i>Bison b. caucasicus</i>	Caucasian wisent	Kubańskaja Oblast	Zoological Institute RAS, Sankt Petersburg, Russia	skull bones	1911*	1.07 (n)

Cc2	22533			North Ossetia-Alania	State Darwin Museum, Moscow, Russia		1949*	0.92 (n)
Cc3	AF005		Holocene Caucasian wisent	Sevan Lake region, Armenia		horn core	¹⁴ C: 4972±29 OxA-31935 5724-5657	58.10 (mt)
Bb1	BS587	<i>Bison b. bonasus</i>	Holocene wisent	Styria, Austria	VMNH H-1981-28-6	tibia	¹⁴ C: 1480±70 VERA 0145 1511-1302	115.56 (mt)
Bb2	BS589				VMNH H-1977-49-1	femur	¹⁴ C: 1980±45 VERA 0142 1987-1886	165.75 (mt)
Bb3	BS607			Upper Austria, Austria	VMNH H-1979-48-1	femur	¹⁴ C: 1370±50 VERA 0143 1338-1265	14.10 (mt)

* for Caucasian individuals (Cc1, Cc2) year of collecting the sample is given. VNHM, Vienna Museum of Natural History. n; nuclear, mt; mitochondrial. Radiocarbon dates include the ¹⁴C age (years before present; yr BP), ¹⁴C accession (where known), and calibrated age BP (1σ). Calibrated dates follow the IntCal13 calibration curve.