bioRxiv preprint doi: https://doi.org/10.1101/210229; this version posted October 28, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Genomic characterisation and conservation genetics of the indigenous Irish Kerry cattle breed

1 **Sam Browett[1], Gillian McHugo[2], Ian W. Richardson[3], David A. Magee[2], Stephen D. E. Park[3],**
2 **Alan G. Fahey[2], John F. Kearney[4], Carolina N. Correia[2], Imtiaz A. S. Randhawa[5], David E.**
3 **MacHugh[2, 6]***

4 [1] Ecosystems and Environment Research Centre, School of Environment and Life Sciences,
5 University of Salford, Salford, M5 4WT, United Kingdom.

6 [2] UCD School of Agriculture and Food Science, University College Dublin, Belfield, Dublin, D04
7 V1W8, Ireland.

8 [3] IdentiGEN, Ltd., Block 2, Blackrock Office Park, Carysfort Avenue, Blackrock, Dublin, A94 H2X4,
9 Ireland.

10 [4] Irish Cattle Breeding Federation, Highfield House, Bandon, Co. Cork, P72 X050, Ireland.

11 [5] Sydney School of Veterinary Science, The University of Sydney, Camden, 2570, NSW, Australia.

12 [6] UCD Conway Institute of Biomolecular and Biomedical Research, University College Dublin,
13 Dublin, D04 V1W8, Ireland.

14

15 *** Correspondence:** David E. MacHugh, Animal Genomics Laboratory, UCD School of

16 Agriculture and Food Science, University College Dublin, Belfield, Dublin, D04 V1W8, Ireland.

17 Email: david.machugh@ucd.ie

18 **Keywords:** cattle, conservation genomics, endangered breed, inbreeding, genetic diversity,
19 population genomics, selection signature, SNP array.

20

21

22 **Note 1:** This is an Associate Editor (D.E.M) Inaugural Article submission to *Frontiers in Genetics:*
23 *Livestock Genomics*

24 **Note 2:** British English language style preferred for publication of this article.

bioRxiv preprint doi: https://doi.org/10.1101/210229; this version posted October 28, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

25 ## Abstract

26 Kerry cattle are an endangered landrace heritage breed of cultural importance to Ireland. In the
27 present study we have used genome-wide SNP data (Illumina® BovineSNP50 array) to evaluate
28 genomic diversity within the Kerry cattle population and between Kerry cattle and other European
29 cattle breeds. Visualisation of patterns of genetic differentiation and gene flow among cattle breeds
30 using phylogenetic trees with ancestry graphs highlighted, in particular, historical gene flow from the
31 British Shorthorn breed into the ancestral population of modern Kerry cattle. Principal component
32 analysis (PCA) and genetic clustering emphasised the genetic distinctiveness of Kerry cattle relative
33 to comparator British and European cattle breeds. Modelling of genetic effective population size (N_e)
34 revealed a demographic trend of diminishing N_e over time and that recent estimated N_e values for the
35 Kerry breed may be less than the threshold for sustainable genetic conservation. In addition, analysis
36 of genome-wide autozygosity (F_{ROH}) showed that genomic inbreeding has increased significantly
37 during the 20 years between 1992 and 2012. Finally, signatures of selection revealed genomic regions
38 subject to natural and artificial selection as Kerry cattle adapted to the climate, physical geography
39 and agro-ecology of southwest Ireland.

40

41 ## 1. Introduction

42 Approximately 10,000 years ago, humans first domesticated wild aurochs (*Bos primigenius*)—the
43 progenitor of modern cattle—in the Fertile Crescent region of Southwest Asia (Edwards et al., 2010;
44 Larson and Fuller, 2014; Larson et al., 2014; Park et al., 2015; MacHugh et al., 2017). Extant domestic
45 cattle, which encompass humpless taurine (*B. taurus*), humped zebu (*B. indicus*) and myriad *B.*
46 *taurus/indicus* hybrid populations, have, through genetic drift and natural and artificial selection,
47 diversified into more than 1,100 recognised breeds. However, starting in the middle of the 20th
48 century, socioeconomic preferences for large highly productive dairy, beef and dual-purpose breeds

bioRxiv preprint doi: https://doi.org/10.1101/210229; this version posted October 28, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

49 have led to extinction and increased vulnerability of more than 200 locally-adapted landrace or native
50 cattle breeds (Gandini et al., 2004; Food and Agriculture Organization, 2007; 2015).

51 With the advent of accelerating climate change, particularly in the Arctic and circumarctic regions
52 (Vihma, 2014; Gao et al., 2015), agro-ecological environments in north-western Europe will
53 inevitably undergo significant change during the coming century (Smith and Gregory, 2013; Wheeler
54 and von Braun, 2013). It is, therefore, increasingly recognised that long-term sustainability of animal
55 production systems and food security will necessitate conservation and management of livestock
56 genetic resources in this region (Hoffmann, 2010; Boettcher et al., 2015; Kantanen et al., 2015).
57 Locally-adapted native livestock breeds with distinct microevolutionary histories and minimal
58 external gene flow will have accumulated novel genomic variation and haplotype combinations for
59 quantitative health, fertility and production traits (Hill, 2014; Felius et al., 2015; Kristensen et al.,
60 2015). These populations may therefore be key to future breeding programmes directed towards
61 adaptation of European livestock to new agro-ecological and production environments (Biscarini et
62 al., 2015; Boettcher et al., 2015; Phocas et al., 2016a; b).

63 The availability of powerful and cheap tools for genotyping large numbers of single nucleotide
64 polymorphisms (SNPs) has provided conservation biologists and animal geneticists with the
65 opportunity to characterise genomic variation and estimate population genetic parameters at very high
66 resolution in threatened or endangered livestock breeds (Pertoldi et al., 2014; Ben Jemaa et al., 2015;
67 Beynon et al., 2015; Meszaros et al., 2015; Burren et al., 2016; Decker et al., 2016; Iso-Touru et al.,
68 2016; Manunza et al., 2016; Mastrangelo et al., 2016; Visser et al., 2016; Williams et al., 2016;
69 Francois et al., 2017). These studies are already providing important baseline data for genetic
70 conservation and will underpin programmes for managed breeding and biobanking of these
71 populations (Groeneveld et al., 2016).

72 As a native breed with a claimed ancient heritage, Kerry cattle are considered culturally important
73 to Ireland (Curran, 1990). It is a landrace cattle population that remains productive in harsh upland

bioRxiv preprint doi: https://doi.org/10.1101/210229; this version posted October 28, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

74 regions with poor quality feed, which are typical of southwest Ireland where the Kerry breed evolved
75 (Food and Agriculture Organization, 2017). These cattle were often referred to anecdotally in Ireland
76 as the "poor man's cow" due to their ability to produce relatively large quantities of milk on very
77 sparse fodder; the Kerry breed is also considered to be a remnant of what was once a substantially
78 larger and more widespread historical population. Levels of inbreeding have been estimated using
79 pedigree data and the accumulated figure since the foundation of the herd book in 1887 reached 15%
80 in 1985 (O'hUigín and Cunningham, 1990).

81 In recent decades the Kerry cattle breed has experienced significant population fluctuations due to
82 changing socioeconomic and agricultural circumstances. During the 1980s, the number of breeding
83 females decreased to less than 200, prompting the Irish agricultural authorities to introduce a Kerry
84 cattle conservation scheme (McParland, 2013), which has continued to the present day in the form of
85 the Department of Agriculture, Food and the Marine (DAFM) Kerry Cattle Premium Scheme
86 (Department of Agriculture Food and the Marine, 2017).

87 The formal conservation policy and supports initiated during the early 1990s led to a significant
88 increase in the Kerry cattle population, such that by 2007 the number of breeding females had
89 increased to more than a thousand animals (Food and Agriculture Organization, 2017). In recent
90 years, however, due to deteriorating economic circumstances in Ireland post-2008, the Kerry cattle
91 population has substantially declined once again and is classified as endangered and under significant
92 threat of extinction or absorption through crossbreeding with other breeds (McParland, 2013;
93 Department of Agriculture Food and the Marine, 2014).

94 The Kerry cattle breed was one of the first European heritage cattle breeds to be surveyed using
95 molecular population genetics techniques. We have previously used autosomal microsatellite genetic
96 markers and mitochondrial DNA (mtDNA) control region sequence variation for comparative
97 evolutionary studies of genetic diversity in Kerry cattle and other British, European, African and
98 Asian breeds (MacHugh et al., 1997; MacHugh et al., 1998; MacHugh et al., 1999; Troy et al., 2001).

bioRxiv preprint doi: https://doi.org/10.1101/210229; this version posted October 28, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

99 In addition, Bray and colleagues have used microsatellites to examine admixture and ancestry in
100 Kerry cattle and the Dexter and Devon breeds (Bray et al., 2009). Results from these studies
101 demonstrated that Kerry cattle exhibit markedly low mtDNA sequence diversity, but autosomal
102 microsatellite diversity comparable to other cattle breeds native to Britain and Ireland. More recently,
103 analyses of medium- and high-density SNP genotypes generated using genome sequence data from
104 an extinct British *B. primigenius* subfossil have shown that Kerry cattle retain a significant genomic
105 signature of admixture from wild aurochs (Park et al., 2015; Upadhyay et al., 2017). This observation
106 highlights the genetic distinctiveness of the Kerry population and has major implications for
107 conservation and management of the breed.

108 For the present study, and within a genetic conservation framework, we performed high-resolution
109 comparative population genomics analyses of Kerry cattle and a range of British and European cattle
110 breeds. These analyses encompassed phylogenetic network reconstruction, evaluation of genetic
111 structure and inbreeding, modelling of historical effective population sizes and functional analyses of
112 artificial and natural selection across the Kerry genome.

113

114 2. Materials and Methods

115 2.1. Kerry Cattle Population DNA Sampling in 1991/92 and 2011/12

116 Two different population samples from the Irish Kerry cattle breed were used for this study
117 (**Figure 1**). The first population sample consisted of peripheral blood and semen straw genomic DNA
118 collected and purified from 36 male and female Kerry cattle in 1991/92, which are a subset of the
119 Kerry cattle population sample (*n* = 40) we have previously described and used for microsatellite-
120 based population genetics analyses (MacHugh et al., 1997; MacHugh et al., 1998). Pedigree records
121 and owners were consulted to ensure there was the minimum degree of genetic relatedness among
122 the animals sampled. This Kerry population sample group is coded as KY92.

bioRxiv preprint doi: https://doi.org/10.1101/210229; this version posted October 28, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

123 The second Kerry cattle population sample was collected in 2011/12 from 19 different herds
124 located across southern and western Ireland. Performagene (PG-100) nasal swab DNA collection kits
125 were used for biological sample collection (DNA Genotek Inc., Ottawa, Canada). Nasal swab DNA
126 samples were collected from a total of 75 male and female Kerry cattle and owners were consulted to
127 ensure there was the minimum degree of genetic relatedness among the animals sampled. This Kerry
128 population sample group is coded as KY12. Genomic DNA was purified from nasal swabs using the
129 Promega Maxwell® 16 DNA automated DNA extraction platform (Promega UK, Southampton,
130 United Kingdom) and established methods developed at the UCD Animal Genomics Laboratory
131 (Magee et al., 2010; Magee et al., 2011).

132 **2.2. SNP Genotyping and Assembly of Comparative SNP Data Sets**

133 Illumina® Bovine SNP50 BeadChip (Matukumalli et al., 2009) genotyping on all 111 Kerry
134 genomic DNA samples (KY92 and KY12 sample panels plus nine blinded sample duplicates for
135 quality control purposes) was performed by Weatherbys Scientific (Co. Kildare, Ireland).

136 For comparative population genomics analyses, equivalent SNP data for a range of other breeds
137 were obtained from previously published work (Decker et al., 2009; Flori et al., 2009; Gibbs et al.,
138 2009; Matukumalli et al., 2009; Gautier et al., 2010; Park et al., 2015). The breed SNP data were split
139 into two discrete composite data sets: a European breed SNP data set (EU) and a SNP data set for a
140 subset of European breeds originating from Britain and Ireland (BI). A population sample of West
141 African N'Dama *B. taurus* cattle from Guinea (NDAM) was also used as an outgroup for the
142 phylogenetic analyses. **Table 1** provides detailed biogeographical information on the cattle breed
143 samples used for the present study.

144 **2.3. Sample Removal and Quality Control and Filtering of SNPs**

145 Pedigree records and genomic relationship were used to identify animals from the KY92 and KY12
146 population samples that were either parent-offspring pairs or full-sibs. One of the two animals in each
147 pair was then randomly removed to generate the working SNP data set. Following this procedure,

bioRxiv preprint doi: https://doi.org/10.1101/210229; this version posted October 28, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

148 quality control and filtering based on recorded SNP genotypes was performed as detailed below for
149 the EU and BI data sets.

150 Prior to quality control and filtering there were 54,057 SNPs in the EU data set (608 animals,
151 including KY92 and KY12) and in the BI data set (354 animals, including KY92 and KY12). SNP
152 quality filtering was performed using PLINK version 1.07 (Purcell et al., 2007), such that individual
153 SNPs with more than 10% missing data and a minor allele frequency (MAF) of ≤ 0.01 were removed
154 from both data sets. X and Y chromosome SNPs were removed and individual animal samples with
155 a SNP call rate less than 90% were also removed from each of the two data sets.

156 SNP quality control and filtering were performed across breeds/populations (by data set) for
157 construction of phylogenies and ancestor graphs, multivariate analysis, investigation of population
158 structure and detection of signatures of selection. For intrapopulation analyses of effective population
159 size (N_e) and genomic inbreeding, all SNPs genotyped (54,057) were filtered within
160 breeds/populations as detailed above. However, an additional filtering procedure was used to remove
161 SNPs deviating from Hardy-Weinberg equilibrium (HWE) with a P value threshold of < 0.0001.
162 Also, for the N_e analysis, a more stringent MAF threshold of 0.05 was used.

163 Using the filtered genome-wide SNP data, PLINK v1.07 was also used to generate identity-by-
164 state (IBS) values between all pairs of Kerry cattle (KY92 and KY12), including the nine blinded
165 sample duplicates for quality control purposes.

166 **2.4. Construction of Phylogenetic Trees and Ancestry Graphs**

167 Maximum likelihood (ML) phylogenetic trees with ancestry graphs were generated for the EU and
168 BI data sets using the TreeMix (version 1.12) software package (Pickrell and Pritchard, 2012). The
169 West African *B. taurus* NDAM breed sample (n = 22) was used as an outgroup. TreeMix was run
170 without using SNP blocks (as described in the TreeMix software documentation) and ML

bioRxiv preprint doi: https://doi.org/10.1101/210229; this version posted October 28, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

171 phylogenetic trees were generated with no migration edges (m = 0) up to ten migration edges (m =
172 10).

173 2.5. Population Differentiation and Genetic Structure

174 To visualise the main axes of genomic variation among cattle breeds and individual animals,
175 multivariate principal component analysis (PCA) was performed for the composite EU and BI SNP
176 data sets using SMARTPCA from the EIGENSOFT package (version 4.2) with default settings
177 (Patterson et al., 2006).

178 To further investigate genetic structure and admixture history for Kerry cattle and other breeds the
179 fastSTRUCTURE software package (Raj et al., 2014) was used to analyse the EU and BI data sets
180 for a range of K possible ancestral populations (K = 2–15). For the present study, the simple prior
181 approach described by Raj and colleagues (2014) was used, which is sufficient for modelling
182 population/breed divergence. To identify the 'true' K value for the number of ancestral populations,
183 a series of fastSTRUCTURE runs with pre-defined K values were examined using the *chooseK.py*
184 script (Raj et al., 2014). Outputs from the fastSTRUCTURE analyses were visualised using the
185 DISTRUCT software program (Rosenberg et al., 2002) using standard parameters.

186 2.6. Modelling Current and Historical Effective Population Size (N_e)

187 Current and historical N_e trends were modelled with genome-wide SNP linkage disequilibrium
188 (LD) data for the KY92 and KY12 populations plus a selection of BI and EU breeds using the SNeP
189 software tool as described by Barbato and colleagues (Barbato et al., 2015). This method facilitates
190 estimation of historical N_e values from SNP linkage disequilibrium (LD) data using the following
191 equation (Corbin et al., 2012):

192
$$N_{T(t)} = \left(4f(c_t)\right)^{-1}\left(E\left[r_{adj}^2|c_t\right]^{-1} - \alpha\right)$$

193 where N_T is the effective population size t generations ago calculated as $t = \left(2f(c_t)\right)^{-1}$ (Hayes et
194 al., 2003), c_t is the recombination rate defined for a specific physical distance between SNP markers,

bioRxiv preprint doi: https://doi.org/10.1101/210229; this version posted October 28, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

195 r_{adj}^2 is the LD value adjusted for sample size and $\alpha := \{1, 2, 2.2\}$ is a correction for the occurrence of
196 mutation (Barbato et al., 2015). In addition, the SNeP program option for unphased SNP data was
197 used for the analyses described here.

198 **2.7. Evaluation of Genomic Inbreeding and Runs of Homozygosity (ROH)**

199 Individual animal genomic inbreeding was evaluated as genome-wide autozygosity estimated from
200 SNP data using runs of homozygosity (ROH) and the F_{ROH} statistic introduced by McQuillan and
201 colleagues (2008). The F_{ROH} statistic was calculated as the ratio of the total length of defined runs of
202 homozygosity (L_{ROH}) to the total length of the autosomal genome covered by SNPs:

203
$$F_{ROH} = \frac{\sum L_{ROH}}{L_{AUTO}}$$

204 PLINK v1.07 was used to define runs of homozygosity (ROH) using a sliding window approach
205 and procedures modified from previous recommendations for Illumina® Bovine SNP50 BeadChip
206 and similar SNP data sets (Purfield et al., 2012; Purfield et al., 2017). The criteria for defining
207 individual ROH were set such that the ROH was required to be at least 500 kb in length, with a
208 minimum density of one SNP per 120 kb and that there was a gap of at least 1,000 kb between each
209 ROH. A sliding window of 50 SNPs was incrementally advanced one SNP at a time along the
210 genome; each discrete window could contain a maximum of one heterozygous SNP and no more than
211 two SNPs with missing genotypes. Following Purfield and colleagues all filtered genomic SNPs,
212 including those located in centromeric regions, were used to estimate F_{ROH} values for individual
213 animals.

214 **2.8. Genome-wide Detection of Signatures of Selection and Functional Enrichment Analysis**

215 In the absence of hard selective sweeps, single selection tests using high-density SNP data do not
216 perform well in detecting signatures of selection from individual livestock breeds (Kemper et al.,
217 2014). Therefore, for the present study, genomic signatures of selection were identified using the

bioRxiv preprint doi: https://doi.org/10.1101/210229; this version posted October 28, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

218 composite selection signal (CSS) method introduced by Randhawa and colleagues (Randhawa et al.,
219 2014). The CSS method has been shown to be a robust and sensitive approach for detecting genomic
220 signatures of selection underlying microevolution of complex traits in livestock (Randhawa et al.,
221 2015). The CSS is a weighted index of signatures of selection from multiple estimates; it is a
222 nonparametric procedure that uses fractional ranks of constituent tests and does not depend on
223 assumptions about the distributions of individual test results.

224 As described in detail by Randhawa *et al.* (2014), the CSS method can be used to combine the
225 fixation index (F_{ST}), the change in selected allele frequency (ΔSAF) and the cross-population
226 extended haplotype homozygosity (*XP-EHH*) tests into one composite statistic for each SNP in a
227 population genomics data set. For the present study, we used 36,621 genome-wide SNPs genotyped
228 in 98 individual Kerry cattle samples (from both the KY92 and KY12 populations) and a sample of
229 102 randomly selected cattle (six random cattle from each breed of the EU data set). To mitigate
230 against false positives, genomic selection signatures were only considered significant if at least one
231 SNP from the set of the top 0.1% genome-wide CSS scores was flanked by at least five SNPs from
232 the set of the top 1% CSS scores.

233 The Ensembl BioMart data mining resource (Smedley et al., 2015) was used to identify genes
234 within ± 1.0 Mb of each selection peak (Ensembl release 90, August 2017). Following this,
235 Ingenuity® Pathway Analysis (IPA®: Qiagen, Redwood City, CA, USA; release date June 2017) was
236 used to perform an overrepresentation enrichment analysis (ORA) with this gene set to identify
237 canonical pathways and functional processes of biological importance. The total gene content of
238 Ensembl release 90 version of the UMD3.1 bovine genome assembly (Zimin et al., 2009) was used
239 as the most appropriate reference gene set for these analyses (Timmons et al., 2015).

240

241

bioRxiv preprint doi: https://doi.org/10.1101/210229; this version posted October 28, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

242 # 3. Results and Discussion

243 ## 3.1. Sample Removal and SNP Filtering and Quality Control

244 Inspection of the pedigree records and the genomic relationship matrix (based on identity-by-state
245 [IBS] of SNP genotypes) identified 10 animals from the KY12 population that were members of a
246 parent-offspring or full-sib pair (sample codes: KY12_01, KY12_05, KY12_13, KY12_14,
247 KY12_17, KY12_18, KY12_19, KY12_46, KY12_55, KY12_67). One sample from each of these
248 pairs was randomly removed. Thereafter, general SNP quality control and filtering led to additional
249 samples being excluded (KY12_26, KY12_28 and KY12_54), giving a total filtered KY12 population
250 sample of 62 animals for downstream population genomics analyses.

251 After SNP quality control and filtering across the two composite data sets (EU and BI), there were
252 36,621 autosomal SNPs from 605 individual animals in the EU data set and there were 37,395
253 autosomal SNPs from 351 animals in the BI data set. When the West African NDAM breed sample
254 (n = 22) was included for the ML phylogenetic tree and ancestry graph analyses, the number of SNPs
255 used was 36,000 from 627 animals for the EU data set and 37,490 from 373 animals for the BI data
256 set. The final numbers of SNPs used for individual breed/population analyses of N_e and genomic
257 inbreeding after all quality control and filtering (including additional filtering for deviations from
258 HWE) are shown in **Table 2**.

259 All data sets, including EU and BI composite data sets and individual breed/population data sets
260 had total SNP call rates of > 99%. The IBS values estimated for Kerry cattle (KY92 and KY12) from
261 filtered genome-wide SNP data are reported in Supplementary Table 1 and described further in
262 **Section 3.7**.

263 ## 3.2. Observed heterozygosity (H_o) estimated from genome-wide SNP data

264 **Table 2** provides genome-wide H_o values for each of the breeds/populations used for the present
265 study. The lowest genome-wide H_o value was observed for the West Africa NDAM *B. taurus* breed,
266 which is likely a consequence of ascertainment bias introduced by a focus on polymorphic SNPs in

bioRxiv preprint doi: https://doi.org/10.1101/210229; this version posted October 28, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

267 European *B. taurus* during design of the Illumina® Bovine SNP50 BeadChip (Matukumalli et al.,
268 2009).

269 Generally, as shown in **Table 2** for the EU and BI breeds and populations, local landrace or
270 heritage breeds display lower H_o values compared to more widespread production breeds such as the
271 Simmental (SIMM), Holstein (HOLS) or Charolais (CHAR) breeds. In addition, as might be
272 expected, production breeds originally derived from minor island populations (Jersey [JRSY] and
273 Guernsey [GNSY]) also exhibit relatively low H_o values. In the context of genetic conservation it is
274 therefore encouraging that the KY92 and KY12 population samples display intermediate H_o values
275 that are at the upper end of the range observed for the heritage breeds.

276 **3.3. Maximum Likelihood Phylogenetic Ancestry Graphs using Genome-wide SNP Data**

277 To examine microevolutionary patterns of genetic differentiation and gene flow among cattle
278 breeds and populations, ML phylogenetic ancestry graphs were generated using TreeMix. For the EU
279 data set, the ML tree topology was consistent for all values of m, with the exception of $m = 2$ migration
280 edges, where the Hereford breed (HRFD) was observed to group with the HOLS breed. The ML tree
281 generated with $m = 5$ is shown in **Figure 2**, which highlights the genetic similarity of the Northern
282 European breeds (British, Irish and Scandinavian). As expected the two Kerry population samples
283 (KY92 and KY12) are genetically very similar and emerge on the same branch as the HRFD breed.
284 It is also noteworthy that there is a high-weight migration edge between the British Shorthorn breed
285 (BSHN) and the root of the two Kerry population samples, supporting the hypothesis of historical
286 gene flow from the British Shorthorn breed into the ancestral population of modern Kerry cattle
287 (Curran, 1990).

288 For the ML trees generated using the BI data set, breed/population differentiation was less
289 apparent, possibly due to similar biogeographical origins for these breeds and/or smaller sample sizes
290 for some of the populations sampled. **Figure 3** shows the ML tree generated with $m = 5$ for the BI
291 data set. For $m = 5$, all migration edges stem from the BSHN/Lincoln Red (LNCR) branch, including

bioRxiv preprint doi: https://doi.org/10.1101/210229; this version posted October 28, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

292 a medium-weight migration edge connecting to the Kerry cattle branch. These results support the
293 hypothesis that there was significant gene flow during the 18th and 19th centuries from British
294 Shorthorn cattle into the ancestral populations for a range of modern British and Irish cattle breeds
295 (Grobet et al., 1998; Felius et al., 2011; Felius et al., 2015).

296 **3.4. Multivariate Principal Component Analysis of Genome-wide SNP Data**

297 To investigate inter- and intra-population genomic diversity and genetic relationship among
298 individual animals from multiple cattle breeds and populations, PCA was performed using genome-
299 wide SNP data. Principal component plots of the first (PC1) and second (PC2) principal components
300 are shown in **Figures 4** and **5** for the EU and BI data sets, respectively.

301 In **Figure 4**, for the EU data set, PC1 and PC2 account for 18.2% and 16.8% of the total variation
302 for PC1–10, respectively. The PC1 plot axis differentiates the British Angus (ANGU), Red Angus
303 (RANG) and BSHN and Irish KY92 and KY12 populations from the rest of the European breeds,
304 including the British HRFD and GNSY and JRSY Channel Islands breeds. In addition, the ANGU
305 and RANG and the Kerry (KY92 and KY12) emerge at the opposite extremes of the PC2 plot axis,
306 highlighting the genetic distinctiveness of the Kerry breed and supporting their status as an important
307 cattle genetic resource that should be prioritised for conservation.

308 In **Figure 5**, for the BI data set, PC1 and PC2 account for 23.7% and 22.7% of the total variation
309 for PC1–10, respectively. The PC1 plot axis recapitulates PC2 in **Figure 4** and differentiates the
310 Kerry (KY92 and KY12) from the ANGU and RANG breeds with the other British breeds emerging
311 between these two extremes. This result reiterates the genetic distinctiveness of the Kerry cattle breed
312 in comparison to a wide range of British production and heritage landrace cattle breeds, again
313 emphasising the need for genetic conservation.

314 The PC2 plot axis in **Figure 5** differentiates the HRFD breed from the other British and Irish
315 breeds and reveals substantial genetic diversity among individual HRFD animals. However, in this

bioRxiv preprint doi: https://doi.org/10.1101/210229; this version posted October 28, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

316 context, it is important to note that the pattern of genetic diversity revealed here for the HRFD
317 population sample may be due to ascertainment bias as a consequence of the strategy used to design
318 the Illumina® Bovine SNP50 BeadChip. In this regard, many of the SNPs that constitute this first-
319 generation SNP array were identified from heterozygous positions in the inbred Hereford female (L1
320 Dominette 01449) bovine genome assembly or through comparisons of random shotgun reads from
321 six diverse cattle breeds that were aligned directly to the same Hereford genome assembly
322 (Matukumalli et al., 2009). This approach to SNP array design will inevitably lead to elevated
323 intrabreed genomic variation using the Illumina® Bovine SNP50 BeadChip with Hereford cattle
324 (Meuwissen, 2009) and accounts for the dispersed pattern of individual HRFD samples in **Figure 5**.

325 Examination of **Figures 4** and **5** indicates that two of the KY12 animals sampled may exhibit a
326 genetic signature of ancestral crossbreeding with another cattle population, which, anecdotally, is
327 likely to have been due to crossbreeding with Angus cattle. Therefore, another PCA plot was
328 generated (Supplementary Figure 1) that shows PC1 and PC2 for individual animals from the KY92,
329 KY12, ANGU and RANG population samples. The two animals exhibiting a genetic signature of
330 possible ancestral crossbreeding (KY12_06 and KY12_58) are indicated on Supplementary Figure 1.
331 Under the assumption of crossbreeding with ANGU cattle, the positions of KY12_06 and KY12_58
332 on this PC plot would suggest that these animals are third generation (F_3) or greater backcrosses.
333 Notwithstanding the KY12_06 and KY12_58 data points, the genetic similarity among all Kerry
334 cattle sampled is evident by comparison of the tight KY92 and KY12 sample cluster to the dispersion
335 of the ANGU and RANG samples on the PCA plot in Supplementary Figure 1.

336 **3.5. Analysis of Genetic Structure using Genome-wide SNP data**

337 The results of the fastSTRUCTURE analyses using the EU and BI data sets are shown in **Figure**
338 **6** and **Figure 7**, respectively. For both analyses, the Kerry cattle (KY92 and KY12) cluster as a single
339 group at $K = 2$ and are differentiated from all other European or British and Irish cattle breeds. The
340 other breed group that is clearly differentiated at $K = 2$ in **Figure 7** is the cluster composed of the

bioRxiv preprint doi: https://doi.org/10.1101/210229; this version posted October 28, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

341 ANGU and RANG breeds. These results mirror the pattern shown for PC1 in **Figure 5**, and again
342 emphasise the genetic distinctiveness of Kerry cattle compared to other European production and
343 landrace heritage breeds. Using the *chooseK.py* script the 'true' number of clusters corresponding to
344 the likely number of ancestral populations was estimated to be between 12 and 14 for the EU data set
345 and either 7 or 8 for the BI data set.

346 For both data sets, animals from the KY12 population sample appear to be more genetically
347 homogenous compared to the KY92 population sample. This observation may be a consequence of
348 increasing use, since the early 1990s, of small numbers of artificial insemination (AI) Kerry sires. It
349 is also noteworthy that the two individual animals detected with a substantial signature of putative
350 historical crossbreeding (KY12_06 and KY12_58) show marked patterns of population admixture in
351 the fastSTRUCTURE results, which are indicated by red arrows in **Figure 6** and **Figure 7**.

352 **3.6. Modelling Historical Effective Population Size (N_e) using Genome-wide SNP Data**

353 The results from modelling historical N_e in a selection of production and heritage cattle breeds and
354 populations (KY92, KY12, DXTR, BSHN, BGAL, LNCR, ANGU, JRSY and HOLS) are provided
355 in Supplementary Table 2 and visualised in **Figure 8**. The 'demographic fingerprints' (Barbato et al.,
356 2015) of the two Kerry populations shown in **Figure 8** and tabulated in Supplementary Table 2 are
357 more similar to those of the production breeds with large census populations (BSHN, ANGU, JRSY,
358 HOLS) than the other heritage breeds with relatively small census population sizes (DXTR, BGAL,
359 LNCR). The KY92, KY12, BSHN, ANGU, JRSY and HOLS populations show a declining trend
360 from historical N_e peaks between 1,500 and 2,000 more than 900 generations ago to N_e values
361 estimated to be less than 200 within the last 20 generations. One the other hand, the DXTR, BSHN,
362 BGAL and LNCR populations display a more severe decline from historical N_e peaks between 2,500
363 and 4,000 more than 900 generations ago to N_e values estimated to be less than 150 within the last 20
364 generations.

bioRxiv preprint doi: https://doi.org/10.1101/210229; this version posted October 28, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

365 It is important to keep in mind these N_e trends may be partly a consequence of the relatively small
366 sample sizes for the DXTR, BGAL and LNCR breeds (see **Table 2**), coupled with different histories
367 of migration, gene flow and, in particular, strong artificial selection in the production cattle
368 populations. Notwithstanding these caveats, the most recent modelled N_e values for the KY92 and
369 KY12 population samples are 89 and 88, respectively. These values are N_e estimates for 12
370 generations in the past and assuming a generation interval of between 4 to 6 years, which is based on
371 a pedigree estimate from a similar heritage cattle population of 5.66 (Meszaros et al., 2015), this
372 corresponds to between 48 and 72 years before 2012 (for the KY12 population). This is approximately
373 the period between 1940 and 1965, which is during the time that the Kerry breed started to decline
374 precipitously in census population size and also N_e estimated from herd book data (O'hUigín and
375 Cunningham, 1990; Food and Agriculture Organization, 2017).

376 From a conservation perspective, livestock populations generally exhibit N_e values relative to total
377 census population sizes (N_c) that are substantially lower than seen in comparable wild mammal
378 populations (Hall, 2016). Also, estimation of N_e using methods such as SNeP that leverage genome-
379 wide SNP linkage disequilibrium (LD) data will tend to underestimate N_e because of physical linkage
380 between many of the SNPs in the data set (Waples et al., 2016). Nevertheless, taking this into account,
381 there is still cause for concern that the most recent N_e values modelled for the KY92 and KY12
382 population samples are below the critical N_e threshold of 100 recommended by Meuwissen (2009)
383 for long-term viability of discrete livestock breeds and populations.

384 ## 3.7. Genomic Relationship and Analysis of Inbreeding

385 Supplementary Table 1 shows a genomic relationship matrix in terms of genotype IBS for the
386 genome-wide SNP data generated for individual animals in the KY92 and KY12 population samples.
387 Close genomic relationship between individual animals sampled from the same herd is evident in the
388 SNP genotype IBS values between samples. In addition, the relatively low genomic relationship
389 between the KY12_06 and KY12_58 outlier samples (**Figures 4–7**) and the rest of the Kerry cattle

bioRxiv preprint doi: https://doi.org/10.1101/210229; this version posted October 28, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

390 sampled is also evident in Supplementary Table 1. These data emphasise the value of intrapopulation
391 genomic relationship values for identifying animals that, from a conservation genetics standpoint,
392 should be of lower priority. They also highlight the potential of genome-wide SNP data for providing
393 a systematic approach to prioritising males and females with minimum genomic relationship for
394 breeding to minimize loss of genetic diversity and maintain or increase N_e (Gandini et al., 2004;
395 Meuwissen, 2009; de Cara et al., 2011; de Cara et al., 2013).

396 Genome-wide autozygosity estimated from SNP data using runs of homozygosity (ROH) and the
397 F_{ROH} statistic are visualised in **Figure 9** for individual animals from the KY92 and KY12 population
398 samples and a range of European comparator breeds. Additional summary ROH data is provided in
399 Supplementary Table 3 and also Supplementary Figure 2, which reveals marked inter-population
400 differences in ROH length and demonstrates that the SNP density of the Illumina® Bovine SNP50
401 BeadChip is too low to reliably capture ROH below 5 Mb in length, an observation previously
402 reported by Purfield and colleagues (2012).

403 There is significant variation in F_{ROH} values among individual animals and between breeds and
404 populations. The non-parametric Wilcoxon rank sum test was performed on F_{ROH} distributions for all
405 pairwise population/breed comparisons with application of the Bonferroni correction P value
406 adjustment for multiple statistical tests (Supplementary Table 4). This analysis demonstrated that the
407 KY12 population sample exhibited a significantly higher mean F_{ROH} value than the KY92 population
408 sample ($P_{adjust} = 0.0340$). This is important from a conservation genetics perspective, indicating that
409 genome-wide autozygosity, which is highly correlated with conventional pedigree-based estimates of
410 inbreeding (F_{PED}) for cattle (Purfield et al., 2012; Ferenčaković et al., 2013; Martikainen et al., 2017),
411 has increased for the Kerry cattle population in the 20 years between sampling of the KY92 and KY12
412 populations.

413 The importance of understanding and quantifying genome-wide autozygosity for genetic
414 conservation purposes has recently been highlighted through correlation of F_{ROH} with inbreeding

bioRxiv preprint doi: https://doi.org/10.1101/210229; this version posted October 28, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

415 depression for a range of production traits in domestic cattle (Bjelland et al., 2013; Pryce et al., 2014;
416 Kim et al., 2015). Importantly, F_{ROH} has also been shown to correlate with inbreeding depression for
417 bovine fertility traits in both males (Ferencakovic et al., 2017) and females (Kim et al., 2015;
418 Martikainen et al., 2017). Finally, according to basic population genetic principles, recent inbreeding
419 captured by F_{ROH} will lead to recessive deleterious genomic variants emerging at a population level—
420 a phenomenon that has been studied in both humans and cattle (Szpiech et al., 2013; Zhang et al.,
421 2015).

422 **3.8. Genome-wide Signatures of Selection in the Kerry Cattle Breed**

423 The results of the genome-wide scan for signatures of selection using the CSS method in the Kerry
424 cattle breed are shown in **Figure 10**. Six distinct selection signatures were detected on BTA9, BTA12,
425 BTA16, BTA17, BTA19 and BTA28. A total of 178 genes were located within the genomic ranges
426 ± 1.0 Mb of selection peaks and 32 of these genes were located within the boundaries of a selection
427 peak. Supplementary Table 5 provides detailed information for these 178 genes.

428 A single gene was located within the BTA9 selection peak—the phosphodiesterase 7B gene
429 (*PDE7B*), which has been associated with neurobiological processes (de Gortari and Mengod, 2010)
430 and has been previously linked to genetic changes associated with dog (*Canis lupus familiaris*)
431 domestication and behaviour (Freedman et al., 2016). A single gene was also located within the
432 BTA16 selection peak—the dorsal inhibitory axon guidance protein gene (*DRAXIN*), which encodes
433 a protein that regulates axon guidance, neural circuit formation and vertebrate brain development
434 (Islam et al., 2009; Shinmyo et al., 2015). Twenty-four genes were located within the BTA17
435 selection peak, including *BICDL1*, *RAB35* and *RNF10*, which have been associated with
436 neurobiology and brain development (Hoshikawa et al., 2008; Schlager et al., 2010; Villarroel-
437 Campos et al., 2016) and *SIRT4* and *COQ5* that function in cellular metabolism (Kawamukai, 2015;
438 Elkhwanky and Hakkola, 2017). Six genes were located within the BTA28 selection peak, including,
439 most notably, the Rho GTPase activating protein 22 gene (*ARHGAP22*), which has recently been

bioRxiv preprint doi: https://doi.org/10.1101/210229; this version posted October 28, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

440 associated with bovine fertility as an mRNA expression biomarker for oocyte competence in cumulus
441 cells (Melo et al., 2017).

442 To obtain a broader perspective on natural and artificial selection acting at a population level on
443 the Kerry cattle genome, a functional gene set enrichment approach (GSEA) was taken using IPA
444 with the 178 genes located within ± 1.0 Mb of each selection peak (Supplementary Table 5). Of these
445 178 genes, 141 could be mapped to the IPA knowledgebase and the summary results for the IPA
446 *Physiological System Development and Function* category are shown in Supplementary Table 6,
447 revealing an enrichment of biological processes associated with nervous system development and
448 behaviour.

449 **3.9. Genomics, Genetic Distinctiveness and Microevolution of Kerry Cattle: Implications for**
450 **Breed Management and Genetic Conservation**

451 The genome-wide phylogenetic and population genetic analyses detailed here demonstrate that
452 Kerry cattle represent an important farm animal genetic resource, befitting the breed's status as a
453 livestock population with a unique history of adaptation to the climate and physical geography of
454 southwest Ireland at the edge of Western Europe. Notably, from a genetic conservation and breed
455 management perspective, high-resolution comparative PCA (**Figures 4** and **5**) and genetic clustering
456 results (**Figures 6** and **7**) demonstrate that Kerry cattle are markedly distinct from other British and
457 European cattle populations. This observation may also be placed in the context of recent
458 paleogenomic studies that have detected ancient gene flow from wild British aurochs (*B. primigenius*)
459 into the ancestors of present-day Kerry cattle (Orlando, 2015; Park et al., 2015; Upadhyay et al.,
460 2017).

461 The current genetic status of the Kerry cattle population is underlined by analyses of genetic
462 effective population size (N_e) and inbreeding using genome-wide SNP data. As shown in **Table 2**,
463 genome-wide observed heterozygosity (H_o) is relatively high in the KY92 and K12 population
464 samples, particularly for endangered heritage cattle breeds. However, it has been long recognised that

bioRxiv preprint doi: https://doi.org/10.1101/210229; this version posted October 28, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

465 monitoring N_e is a more important tool for rational breed management and long-term conservation of
466 endangered livestock populations (Notter, 1999; Gandini et al., 2004; Biscarini et al., 2015). As
467 shown in **Figure 8** and Supplementary Table 2, the Kerry cattle population has a recent demographic
468 trend of N_e decline, to the point where the most recent modelled N_e values are below the recommended
469 threshold for sustainable breed management and conservation (Meuwissen, 2009). There is also cause
470 for concern that genomic inbreeding estimated using genome-wide autozygosity (F_{ROH}) and
471 visualised in **Figure 9** has increased significantly in the 20-year period between the sampling of the
472 KY92 and KY12 Kerry cattle populations.

473 In a more positive light, as shown in the present study, detection of discrete signatures of selection
474 using the relatively low-density Illumina® Bovine SNP50 BeadChip is encouraging for wider studies
475 of genome-wide microevolution in endangered heritage livestock populations. Based on our results,
476 for example, future surveys of Kerry cattle that use higher-density SNP array platforms and ultimately
477 whole-genome sequence data will provide exquisitely detailed information on the genomic regions
478 and associated polygenic production, health, fertility and behavioural traits shaped, over many
479 centuries, by the agro-ecology and pre-industrial farming systems of southwest Ireland.

480

481 4. Conflict of Interest

482 The authors declare that the research was conducted in the absence of any commercial or financial
483 relationships that could be construed as a potential conflict of interest.

484

485 5. Ethics Statement

486 With the exception of Kerry cattle sampled during 2011-12, all samples and data was obtained
487 from previously published scientific studies. The re-use of these samples and data is consistent with
488 the 3Rs principles on replacement, refinement and reduction of animals in research

bioRxiv preprint doi: https://doi.org/10.1101/210229; this version posted October 28, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

489 (www.nc3rs.org.uk/the-3rs). For the 2011-12 Kerry cattle, population owners' consent to sample
490 DNA for research was obtained and individual owners conducted sampling of animals using non-
491 invasive nasal swabs. In this regard, scientific animal protection in Ireland is subject to European
492 Union Directive 2010/63/EU, which states that the Directive does not apply to "practices not likely
493 to cause pain, suffering, distress or lasting harm equivalent to, or higher than, that caused by the
494 introduction of a needle in accordance with good veterinary practice".

495

496 6. Author Contributions

497 DEM, DAM, AGF and JFK conceived and designed the project; DEM, IWR, DAM, AGF and JFK
498 organised sample collection and genotyping; SB, GM, IWR, DAM, SDEP, CNC, IASR and DEM
499 performed the analyses; SB and DEM wrote the manuscript and all authors reviewed and approved
500 the final manuscript.

501

502 7. Funding

503 This work was supported by Department of Agriculture, Food and the Marine (DAFM) funding
504 under the Genetic Resources for Food and Agriculture scheme (grant no: 10/GR/06); an Investigator
505 Programme Grant from Science Foundation Ireland (SFI/08/IN.1/B2038); a Research Stimulus Grant
506 from DAFM (RSF 06 406); a European Union Framework 7 Project Grant (KBBE-211602-
507 MACROSYS); the Brazilian Science Without Borders Programme (CAPES grant no. BEX-13070-
508 13-4) and the UCD MSc Programme in Evolutionary Biology.

509

510 8. Acknowledgements

511 The authors wish to express their gratitude to the Kerry Cattle Society (www.kerrycattle.ie) for
512 facilitating and encouraging this project. In particular, we thank Rosemary and Jeremy Hill, Matthew

bioRxiv preprint doi: https://doi.org/10.1101/210229; this version posted October 28, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

513 English Hayden and the Irish Cattle Breeding Federation (ICBF – www.icbf.com) for expert guidance
514 and assistance with animal sourcing and DNA sampling. We are also grateful to Professor Dan
515 Bradley (Smurfit Institute of Genetics, Trinity College Dublin) for access to DNA sample archives.
516 In addition, we thank all of the Kerry cattle owners who provided access to animals, samples and
517 pedigree information. Finally, we thank Weatherbys Scientific for provision of SNP array genotyping
518 services (www.weatherbysscientific.com).

bioRxiv preprint doi: https://doi.org/10.1101/210229; this version posted October 28, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

519 # 9. References

520 Barbato, M, Orozco-terWengel, P, Tapio, M, and Bruford, MW (2015). SNeP: a tool to estimate
521 trends in recent effective population size trajectories using genome-wide SNP data. *Front*
522 *Genet* 6**,** 109. doi: 10.3389/fgene.2015.00109.

523 Ben Jemaa, S, Boussaha, M, Ben Mehdi, M, Lee, JH, and Lee, SH (2015). Genome-wide insights
524 into population structure and genetic history of Tunisian local cattle using the Illumina
525 BovineSNP50 Beadchip. *BMC Genomics* 16**,** 677. doi: 10.1186/s12864-015-1638-6.

526 Beynon, SE, Slavov, GT, Farre, M, Sunduimijid, B, Waddams, K, Davies, B, Haresign, W, Kijas, J,
527 MacLeod, IM, Newbold, CJ, Davies, L, and Larkin, DM (2015). Population structure and
528 history of the Welsh sheep breeds determined by whole genome genotyping. *BMC Genet* 16**,**
529 65. doi: 10.1186/s12863-015-0216-x.

530 Biscarini, F, Nicolazzi, EL, Stella, A, Boettcher, PJ, and Gandini, G (2015). Challenges and
531 opportunities in genetic improvement of local livestock breeds. *Front Genet* 6**,** 33. doi:
532 10.3389/fgene.2015.00033.

533 Bjelland, DW, Weigel, KA, Vukasinovic, N, and Nkrumah, JD (2013). Evaluation of inbreeding
534 depression in Holstein cattle using whole-genome SNP markers and alternative measures of
535 genomic inbreeding. *J Dairy Sci* 96**,** 4697-4706. doi: 10.3168/jds.2012-6435.

536 Boettcher, PJ, Hoffmann, I, Baumung, R, Drucker, AG, McManus, C, Berg, P, Stella, A, Nilsen,
537 LB, Moran, D, Naves, M, and Thompson, MC (2015). Genetic resources and genomics for
538 adaptation of livestock to climate change. *Front Genet* 5**,** 461. doi:
539 10.3389/fgene.2014.00461.

540 Bray, TC, Chikhi, L, Sheppy, AJ, and Bruford, MW (2009). The population genetic effects of
541 ancestry and admixture in a subdivided cattle breed. *Anim Genet* 40**,** 393-400. doi:
542 10.1111/j.1365-2052.2009.01850.x.

543 Burren, A, Neuditschko, M, Signer-Hasler, H, Frischknecht, M, Reber, I, Menzi, F, Drogemuller, C,
544 and Flury, C (2016). Genetic diversity analyses reveal first insights into breed-specific
545 selection signatures within Swiss goat breeds. *Anim Genet* 47**,** 727-739. doi:
546 10.1111/age.12476.

547 Corbin, LJ, Liu, AYH, Bishop, SC, and Woolliams, JA (2012). Estimation of historical effective
548 population size using linkage disequilibria with marker data. *J Anim Breed Genet* 129**,** 257-
549 270. doi: 10.1111/j.1439-0388.2012.01003.x.

550 Curran, PL (1990). *Kerry and Dexter Cattle and other Ancient Irish Breeds, a History.* Dublin,
551 Ireland: Royal Dublin Society.

bioRxiv preprint doi: https://doi.org/10.1101/210229; this version posted October 28, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

552 de Cara, MA, Fernandez, J, Toro, MA, and Villanueva, B (2011). Using genome-wide information
553 to minimize the loss of diversity in conservation programmes. *J Anim Breed Genet* 128**,**
554 456-464. doi: 10.1111/j.1439-0388.2011.00971.x.
555 de Cara, MA, Villanueva, B, Toro, MA, and Fernandez, J (2013). Using genomic tools to maintain
556 diversity and fitness in conservation programmes. *Mol Ecol* 22**,** 6091-6099. doi:
557 10.1111/mec.12560.
558 de Gortari, P, and Mengod, G (2010). Dopamine D1, D2 and mu-opioid receptors are co-expressed
559 with adenylyl cyclase 5 and phosphodiesterase 7B mRNAs in striatal rat cells. *Brain Res*
560 1310**,** 37-45. doi: 10.1016/j.brainres.2009.11.009.
561 Decker, JE, Pires, JC, Conant, GC, McKay, SD, Heaton, MP, Chen, K, Cooper, A, Vilkki, J,
562 Seabury, CM, Caetano, AR, Johnson, GS, Brenneman, RA, Hanotte, O, Eggert, LS, Wiener,
563 P, Kim, JJ, Kim, KS, Sonstegard, TS, Van Tassell, CP, Neibergs, HL, McEwan, JC,
564 Brauning, R, Coutinho, LL, Babar, ME, Wilson, GA, McClure, MC, Rolf, MM, Kim, J,
565 Schnabel, RD, and Taylor, JF (2009). Resolving the evolution of extant and extinct
566 ruminants with high-throughput phylogenomics. *Proc Natl Acad Sci U S A* 106**,** 18644-
567 18649. doi: 10.1073/pnas.0904691106.
568 Decker, JE, Taylor, JF, Kantanen, J, Millbrooke, A, Schnabel, RD, Alexander, LJ, and MacNeil,
569 MD (2016). Origins of cattle on Chirikof Island, Alaska, elucidated from genome-wide SNP
570 genotypes. *Heredity* 116**,** 502-505. doi: 10.1038/hdy.2016.7.
571 Department of Agriculture Food and the Marine (2014). *State of Biodiversity for Food and*
572 *Agriculture in Ireland.* Dublin, Ireland: Department of Agriculture, Food and the Marine.
573 Department of Agriculture Food and the Marine. 2017. Kerry Cattle Premium Scheme: Scheme
574 Objectives and Terms and Conditions.
575 Edwards, CJ, Magee, DA, Park, SD, McGettigan, PA, Lohan, AJ, Murphy, A, Finlay, EK, Shapiro,
576 B, Chamberlain, AT, Richards, MB, Bradley, DG, Loftus, BJ, and MacHugh, DE (2010). A
577 complete mitochondrial genome sequence from a mesolithic wild aurochs (*Bos*
578 *primigenius*). *PLoS ONE* 5**,** e9255. doi: 10.1371/journal.pone.0009255.
579 Elkhwanky, MS, and Hakkola, J (2017). Extranuclear sirtuins and metabolic stress. *Antioxid Redox*
580 *Signal*. doi: 10.1089/ars.2017.7270.
581 Felius, M, Koolmees, PA, Theunissen, B, Consortium, ECGD, and Lenstra, JA (2011). On the
582 breeds of cattle—historic and current classifications. *Diversity* 3**,** 660.
583 Felius, M, Theunissen, B, and Lenstra, JA (2015). Conservation of cattle genetic resources: the role
584 of breeds. *J Agric Sci* 153**,** 152-162. doi: 10.1017/s0021859614000124.

bioRxiv preprint doi: https://doi.org/10.1101/210229; this version posted October 28, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

585 Ferenčaković, M, Hamzić, E, Gredler, B, Solberg, TR, Klemetsdal, G, Curik, I, and Sölkner, J
586 (2013). Estimates of autozygosity derived from runs of homozygosity: empirical evidence
587 from selected cattle populations. *J Anim Breed Genet* 130**,** 286-293. doi: 10.1111/jbg.12012.
588 Ferencakovic, M, Solkner, J, Kaps, M, and Curik, I (2017). Genome-wide mapping and estimation
589 of inbreeding depression of semen quality traits in a cattle population. *J Dairy Sci* 100**,**
590 4721-4730. doi: 10.3168/jds.2016-12164.
591 Flori, L, Fritz, S, Jaffrezic, F, Boussaha, M, Gut, I, Heath, S, Foulley, JL, and Gautier, M (2009).
592 The genome response to artificial selection: a case study in dairy cattle. *PLoS ONE* 4**,** 13.
593 doi: 10.1371/journal.pone.0006595.
594 Food and Agriculture Organization (2007). *The State of the World's Animal Genetic Resources for*
595 *Food and Agriculture.* Rome, Italy: edited by D. Pilling & B. Rischkowsky. FAO
596 Commission on Genetic Resources for Food and Agriculture Assessments.
597 Food and Agriculture Organization (2015). *The Second Report on the State of the World's Animal*
598 *Genetic Resources for Food and Agriculture.* Rome, Italy: edited by B.D. Scherf & D.
599 Pilling. FAO Commission on Genetic Resources for Food and Agriculture Assessments.
600 Food and Agriculture Organization. 2017. Domestic Animal Diversity Information System (DAD-
601 IS); available at: www.fao.org/dadis. [Accessed 01/05/2017].
602 Francois, L, Wijnrocx, K, Colinet, FG, Gengler, N, Hulsegge, B, Windig, JJ, Buys, N, and Janssens,
603 S (2017). Genomics of a revived breed: case study of the Belgian Campine cattle. *PLoS*
604 *ONE* 12**,** e0175916. doi: 10.1371/journal.pone.0175916.
605 Freedman, AH, Schweizer, RM, Ortega-Del Vecchyo, D, Han, E, Davis, BW, Gronau, I, Silva, PM,
606 Galaverni, M, Fan, Z, Marx, P, Lorente-Galdos, B, Ramirez, O, Hormozdiari, F, Alkan, C,
607 Vila, C, Squire, K, Geffen, E, Kusak, J, Boyko, AR, Parker, HG, Lee, C, Tadigotla, V,
608 Siepel, A, Bustamante, CD, Harkins, TT, Nelson, SF, Marques-Bonet, T, Ostrander, EA,
609 Wayne, RK, and Novembre, J (2016). Demographically-based evaluation of genomic
610 regions under selection in domestic dogs. *PLoS Genet* 12**,** e1005851. doi:
611 10.1371/journal.pgen.1005851.
612 Gandini, GC, Ollivier, L, Danell, B, Distl, O, Georgoudis, A, Groeneveld, E, Martyniuk, E, van
613 Arendonk, JAM, and Woolliams, JA (2004). Criteria to assess the degree of endangerment
614 of livestock breeds in Europe. *Livestock Prod Sci* 91**,** 173-182. doi:
615 http://dx.doi.org/10.1016/j.livprodsci.2004.08.001.
616 Gao, Y, Sun, J, Li, F, He, S, Sandven, S, Yan, Q, Zhang, Z, Lohmann, K, Keenlyside, N, Furevik,
617 T, and Suo, L (2015). Arctic sea ice and Eurasian climate: A review. *Adv Atmos Sci* 32**,** 92-
618 114. doi: 10.1007/s00376-014-0009-6.

bioRxiv preprint doi: https://doi.org/10.1101/210229; this version posted October 28, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

619 Gautier, M, Laloe, D, and Moazami-Goudarzi, K (2010). Insights into the genetic history of French
620 cattle from dense SNP data on 47 worldwide breeds. *PLoS ONE* 5. doi:
621 10.1371/journal.pone.0013038.
622 Gibbs, RA, Taylor, JF, Van Tassell, CP, Barendse, W, Eversoie, KA, Gill, CA, Green, RD,
623 Hamernik, DL, Kappes, SM, Lien, S, Matukumalli, LK, McEwan, JC, Nazareth, LV,
624 Schnabel, RD, Taylor, JF, Weinstock, GM, Wheeler, DA, Ajmone-Marsan, P, Barendse, W,
625 Boettcher, PJ, Caetano, AR, Garcia, JF, Hanotte, O, Mariani, P, Skow, LC, Williams, JL,
626 Caetano, AR, Diallo, B, Green, RD, Hailemariam, L, Hanotte, O, Martinez, ML, Morris,
627 CA, Silva, LOC, Spelman, RJ, Taylor, JF, Mulatu, W, Zhao, KY, Abbey, CA, Agaba, M,
628 Araujo, FR, Bunch, RJ, Burton, J, Gill, CA, Gorni, C, Olivier, H, Harrison, BE, Luff, B,
629 Machado, MA, Mariani, P, Morris, CA, Mwakaya, J, Plastow, G, Sim, W, Skow, LC, Smith,
630 T, Sonstegard, TS, Spelman, RJ, Taylor, JF, Thomas, MB, Valentini, A, Williams, P,
631 Womack, J, Wooliams, JA, Liu, Y, Qin, X, Worley, KC, Gao, C, Gill, CA, Jiang, HY, Liu,
632 Y, Moore, SS, Nazareth, LV, Ren, YR, Song, XZ, Bustamante, CD, Hernandez, RD,
633 Muzny, DM, Nazareth, LV, Patil, S, Lucas, AS, Fu, Q, Kent, MP, Moore, SS, Vega, R,
634 Abbey, CA, Gao, C, Gill, CA, Green, RD, Matukumalli, LK, Matukumalli, A, McWilliam,
635 S, Schnabel, RD, Sclep, G, Ajmone-Marsan, P, Bryc, K, Bustamante, CD, Choi, J, Gao, H,
636 Grefenstette, JJ, et al. (2009). Genome-wide survey of SNP variation uncovers the genetic
637 structure of cattle breeds. *Science* 324**,** 528-532. doi: 10.1126/science.1167936.
638 Grobet, L, Poncelet, D, Royo, LJ, Brouwers, B, Pirottin, D, Michaux, C, Menissier, F, Zanotti, M,
639 Dunner, S, and Georges, M (1998). Molecular definition of an allelic series of mutations
640 disrupting the myostatin function and causing double-muscling in cattle. *Mamm Genome* 9**,**
641 210-213.
642 Groeneveld, LF, Gregusson, S, Guldbrandtsen, B, Hiemstra, SJ, Hveem, K, Kantanen, J, Lohi, H,
643 Stroemstedt, L, and Berg, P (2016). Domesticated animal biobanking: land of opportunity.
644 *PLoS Biol* 14**,** e1002523. doi: 10.1371/journal.pbio.1002523.
645 Hall, SJ (2016). Effective population sizes in cattle, sheep, horses, pigs and goats estimated from
646 census and herdbook data. *Animal* 10**,** 1778-1785. doi: 10.1017/s1751731116000914.
647 Hayes, BJ, Visscher, PM, McPartlan, HC, and Goddard, ME (2003). Novel multilocus measure of
648 linkage disequilibrium to estimate past effective population size. *Genome Res* 13**,** 635-643.
649 doi: 10.1101/gr.387103.
650 Hill, WG (2014). Applications of population genetics to animal breeding, from Wright, Fisher and
651 Lush to genomic prediction. *Genetics* 196**,** 1-16. doi: 10.1534/genetics.112.147850.
652 Hoffmann, I (2010). Climate change and the characterization, breeding and conservation of animal
653 genetic resources. *Anim Genet* 41 Suppl 1**,** 32-46. doi: 10.1111/j.1365-2052.2010.02043.x.

bioRxiv preprint doi: https://doi.org/10.1101/210229; this version posted October 28, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

654 Hoshikawa, S, Ogata, T, Fujiwara, S, Nakamura, K, and Tanaka, S (2008). A novel function of
655 RING finger protein 10 in transcriptional regulation of the myelin-associated glycoprotein
656 gene and myelin formation in Schwann cells. *PLoS ONE* 3**,** e3464. doi:
657 10.1371/journal.pone.0003464.
658 Islam, SM, Shinmyo, Y, Okafuji, T, Su, Y, Naser, IB, Ahmed, G, Zhang, S, Chen, S, Ohta, K,
659 Kiyonari, H, Abe, T, Tanaka, S, Nishinakamura, R, Terashima, T, Kitamura, T, and Tanaka,
660 H (2009). Draxin, a repulsive guidance protein for spinal cord and forebrain commissures.
661 *Science* 323**,** 388-393. doi: 10.1126/science.1165187.
662 Iso-Touru, T, Tapio, M, Vilkki, J, Kiseleva, T, Ammosov, I, Ivanova, Z, Popov, R, Ozerov, M, and
663 Kantanen, J (2016). Genetic diversity and genomic signatures of selection among cattle
664 breeds from Siberia, eastern and northern Europe. *Anim Genet* 47**,** 647-657. doi:
665 10.1111/age.12473.
666 Kantanen, J, Lovendahl, P, Strandberg, E, Eythorsdottir, E, Li, MH, Kettunen-Praebel, A, Berg, P,
667 and Meuwissen, T (2015). Utilization of farm animal genetic resources in a changing agro-
668 ecological environment in the Nordic countries. *Front Genet* 6**,** 52. doi:
669 10.3389/fgene.2015.00052.
670 Kawamukai, M (2015). Biosynthesis of coenzyme Q in eukaryotes. *Biosci Biotechnol Biochem* 80**,**
671 23-33. doi: 10.1080/09168451.2015.1065172.
672 Kemper, KE, Saxton, SJ, Bolormaa, S, Hayes, BJ, and Goddard, ME (2014). Selection for complex
673 traits leaves little or no classic signatures of selection. *BMC Genomics* 15**,** 246. doi:
674 10.1186/1471-2164-15-246.
675 Kim, ES, Sonstegard, TS, Van Tassell, CP, Wiggans, G, and Rothschild, MF (2015). The
676 relationship between runs of homozygosity and inbreeding in Jersey cattle under selection.
677 *PLoS ONE* 10**,** e0129967. doi: 10.1371/journal.pone.0129967.
678 Kristensen, TN, Hoffmann, AA, Pertoldi, C, and Stronen, AV (2015). What can livestock breeders
679 learn from conservation genetics and vice versa? *Front Genet* 6**,** 38. doi:
680 10.3389/fgene.2015.00038.
681 Larson, G, and Fuller, DQ (2014). The evolution of animal domestication. *Annu Rev Ecol Evol Syst*
682 45**,** 115-136. doi: 10.1146/annurev-ecolsys-110512-135813.
683 Larson, G, Piperno, DR, Allaby, RG, Purugganan, MD, Andersson, L, Arroyo-Kalin, M, Barton, L,
684 Climer Vigueira, C, Denham, T, Dobney, K, Doust, AN, Gepts, P, Gilbert, MT, Gremillion,
685 KJ, Lucas, L, Lukens, L, Marshall, FB, Olsen, KM, Pires, JC, Richerson, PJ, Rubio de
686 Casas, R, Sanjur, OI, Thomas, MG, and Fuller, DQ (2014). Current perspectives and the
687 future of domestication studies. *Proc Natl Acad Sci U S A* 111**,** 6139-6146. doi:
688 10.1073/pnas.1323964111.

bioRxiv preprint doi: https://doi.org/10.1101/210229; this version posted October 28, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

689 MacHugh, DE, Larson, G, and Orlando, L (2017). Taming the past: ancient DNA and the study of
690 animal domestication. *Annu Rev Anim Biosci* 5**,** 329-351. doi: 10.1146/annurev-animal-
691 022516-022747.
692 MacHugh, DE, Loftus, RT, Cunningham, P, and Bradley, DG (1998). Genetic structure of seven
693 European cattle breeds assessed using 20 microsatellite markers. *Anim Genet* 29**,** 333-340.
694 doi: 10.1046/j.1365-2052.1998.295330.x.
695 MacHugh, DE, Shriver, MD, Loftus, RT, Cunningham, P, and Bradley, DG (1997). Microsatellite
696 DNA variation and the evolution, domestication and phylogeography of taurine and zebu
697 cattle (*Bos taurus* and *Bos indicus*). *Genetics* 146**,** 1071-1086.
698 MacHugh, DE, Troy, CS, McCormick, F, Olsaker, I, Eythorsdottir, E, and Bradley, DG (1999).
699 Early medieval cattle remains from a Scandinavian settlement in Dublin: genetic analysis
700 and comparison with extant breeds. *Philos Trans R Soc Lond B Biol Sci* 354**,** 99-108. doi:
701 10.1098/rstb.1999.0363.
702 Magee, DA, Berkowicz, EW, Sikora, KM, Berry, DP, Park, SDE, Kelly, AK, Sweeney, T, Kenny,
703 DA, Evans, RD, Wickham, BW, Spillane, C, and MacHugh, DE (2010). A catalogue of
704 validated single nucleotide polymorphisms in bovine orthologs of mammalian imprinted
705 genes and associations with beef production traits. *Animal* 4**,** 1958-1970. doi:
706 doi:10.1017/S1751731110001163.
707 Magee, DA, Berry, DP, Berkowicz, EW, Sikora, KM, Howard, DJ, Mullen, MP, Evans, RD,
708 Spillane, C, and MacHugh, DE (2011). Single nucleotide polymorphisms within the bovine
709 *DLK1-DIO3* imprinted domain are associated with economically important production traits
710 in cattle. *J Hered* 102**,** 94-101. doi: 10.1093/jhered/esq097.
711 Manunza, A, Cardoso, TF, Noce, A, Martinez, A, Pons, A, Bermejo, LA, Landi, V, Sanchez, A,
712 Jordana, J, Delgado, JV, Adan, S, Capote, J, Vidal, O, Ugarte, E, Arranz, JJ, Calvo, JH,
713 Casellas, J, and Amills, M (2016). Population structure of eleven Spanish ovine breeds and
714 detection of selective sweeps with BayeScan and hapFLK. *Sci Rep* 6**,** 27296. doi:
715 10.1038/srep27296.
716 Martikainen, K, Tyriseva, AM, Matilainen, K, Poso, J, and Uimari, P (2017). Estimation of
717 inbreeding depression on female fertility in the Finnish Ayrshire population. *J Anim Breed*
718 *Genet*. doi: 10.1111/jbg.12285.
719 Mastrangelo, S, Tolone, M, Di Gerlando, R, Fontanesi, L, Sardina, MT, and Portolano, B (2016).
720 Genomic inbreeding estimation in small populations: evaluation of runs of homozygosity in
721 three local dairy cattle breeds. *Animal* 10**,** 746-754. doi: 10.1017/s1751731115002943.
722 Matukumalli, LK, Lawley, CT, Schnabel, RD, Taylor, JF, Allan, MF, Heaton, MP, O'Connell, J,
723 Moore, SS, Smith, TPL, Sonstegard, TS, and Van Tassell, CP (2009). Development and

bioRxiv preprint doi: https://doi.org/10.1101/210229; this version posted October 28, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

724 characterization of a high density SNP genotyping assay for cattle. *PLoS ONE* 4**,** 13. doi: 10.1371/journal.pone.0005350.
726 McParland, S. 2013. National Genetic Conservation Strategy Document. Available: https://www.agriculture.gov.ie.
728 McQuillan, R, Leutenegger, AL, Abdel-Rahman, R, Franklin, CS, Pericic, M, Barac-Lauc, L, Smolej-Narancic, N, Janicijevic, B, Polasek, O, Tenesa, A, Macleod, AK, Farrington, SM, Rudan, P, Hayward, C, Vitart, V, Rudan, I, Wild, SH, Dunlop, MG, Wright, AF, Campbell, H, and Wilson, JF (2008). Runs of homozygosity in European populations. *Am J Hum Genet* 83**,** 359-372. doi: 10.1016/j.ajhg.2008.08.007.
733 Melo, EO, Cordeiro, DM, Pellegrino, R, Wei, Z, Daye, ZJ, Nishimura, RC, and Dode, MA (2017). Identification of molecular markers for oocyte competence in bovine cumulus cells. *Anim Genet* 48**,** 19-29. doi: 10.1111/age.12496.
736 Meszaros, G, Boison, SA, Perez O'Brien, AM, Ferencakovic, M, Curik, I, Da Silva, MV, Utsunomiya, YT, Garcia, JF, and Solkner, J (2015). Genomic analysis for managing small and endangered populations: a case study in Tyrol Grey cattle. *Front Genet* 6**,** 173. doi: 10.3389/fgene.2015.00173.
740 Meuwissen, T (2009). Genetic management of small populations: A review. *Acta Agriculturae Scandinavica, Section A — Animal Science* 59**,** 71-79. doi: 10.1080/09064700903118148.
742 Notter, DR (1999). The importance of genetic diversity in livestock populations of the future. *J Anim Sci* 77**,** 61-69.
744 O'hUigín, C, and Cunningham, EP (1990). Analysis of breeding structure of the Kerry breed. *J Anim Breed Genet* 107**,** 452-457. doi: 10.1111/j.1439-0388.1990.tb00057.x.
746 Orlando, L (2015). The first aurochs genome reveals the breeding history of British and European cattle. *Genome Biol* 16**,** 225. doi: 10.1186/s13059-015-0793-z.
748 Park, SD, Magee, DA, McGettigan, PA, Teasdale, MD, Edwards, CJ, Lohan, AJ, Murphy, A, Braud, M, Donoghue, MT, Liu, Y, Chamberlain, AT, Rue-Albrecht, K, Schroeder, S, Spillane, C, Tai, S, Bradley, DG, Sonstegard, TS, Loftus, BJ, and MacHugh, DE (2015). Genome sequencing of the extinct Eurasian wild aurochs, *Bos primigenius*, illuminates the phylogeography and evolution of cattle. *Genome Biol* 16**,** 234. doi: 10.1186/s13059-015-0790-2.
754 Patterson, N, Price, AL, and Reich, D (2006). Population structure and eigenanalysis. *PLoS Genet* 2**,** e190. doi: 10.1371/journal.pgen.0020190.
756 Pertoldi, C, Purfield, DC, Berg, P, Jensen, TH, Bach, OS, Vingborg, R, and Kristensen, TN (2014). Genetic characterization of a herd of the endangered Danish Jutland cattle. *J Anim Sci* 92**,** 2372-2376. doi: 10.2527/jas.2013-7206.

bioRxiv preprint doi: https://doi.org/10.1101/210229; this version posted October 28, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

759 Phocas, F, Belloc, C, Bidanel, J, Delaby, L, Dourmad, JY, Dumont, B, Ezanno, P, Fortun-Lamothe,
760 L, Foucras, G, Frappat, B, Gonzalez-Garcia, E, Hazard, D, Larzul, C, Lubac, S, Mignon-
761 Grasteau, S, Moreno, CR, Tixier-Boichard, M, and Brochard, M (2016a). Review: Towards
762 the agroecological management of ruminants, pigs and poultry through the development of
763 sustainable breeding programmes. II. Breeding strategies. *Animal* 10**,** 1760-1769. doi:
764 10.1017/s1751731116001051.
765 Phocas, F, Belloc, C, Bidanel, J, Delaby, L, Dourmad, JY, Dumont, B, Ezanno, P, Fortun-Lamothe,
766 L, Foucras, G, Frappat, B, Gonzalez-Garcia, E, Hazard, D, Larzul, C, Lubac, S, Mignon-
767 Grasteau, S, Moreno, CR, Tixier-Boichard, M, and Brochard, M (2016b). Review: Towards
768 the agroecological management of ruminants, pigs and poultry through the development of
769 sustainable breeding programmes: I-selection goals and criteria. *Animal* 10**,** 1749-1759. doi:
770 10.1017/s1751731116000926.
771 Pickrell, JK, and Pritchard, JK (2012). Inference of population splits and mixtures from genome-
772 wide allele frequency data. *PLoS Genet* 8**,** e1002967. doi: 10.1371/journal.pgen.1002967.
773 Pryce, JE, Haile-Mariam, M, Goddard, ME, and Hayes, BJ (2014). Identification of genomic
774 regions associated with inbreeding depression in Holstein and Jersey dairy cattle. *Genet Sel
775 Evol* 46**,** 71. doi: 10.1186/s12711-014-0071-7.
776 Purcell, S, Neale, B, Todd-Brown, K, Thomas, L, Ferreira, MAR, Bender, D, Maller, J, Sklar, P, de
777 Bakker, PIW, Daly, MJ, and Sham, PC (2007). PLINK: A tool set for whole-genome
778 association and population-based linkage analyses. *Am J Hum Genet* 81**,** 559-575. doi:
779 10.1086/519795.
780 Purfield, DC, Berry, DP, McParland, S, and Bradley, DG (2012). Runs of homozygosity and
781 population history in cattle. *BMC Genet* 13**,** 70. doi: 10.1186/1471-2156-13-70.
782 Purfield, DC, McParland, S, Wall, E, and Berry, DP (2017). The distribution of runs of
783 homozygosity and selection signatures in six commercial meat sheep breeds. *PLoS ONE* 12**,**
784 e0176780. doi: 10.1371/journal.pone.0176780.
785 Raj, A, Stephens, M, and Pritchard, JK (2014). fastSTRUCTURE: variational inference of
786 population structure in large SNP data sets. *Genetics* 197**,** 573-589. doi:
787 10.1534/genetics.114.164350.
788 Randhawa, IA, Khatkar, MS, Thomson, PC, and Raadsma, HW (2014). Composite selection signals
789 can localize the trait specific genomic regions in multi-breed populations of cattle and sheep.
790 *BMC Genet* 15**,** 34. doi: 10.1186/1471-2156-15-34.
791 Randhawa, IA, Khatkar, MS, Thomson, PC, and Raadsma, HW (2015). Composite selection signals
792 for complex traits exemplified through bovine stature using multibreed cohorts of European
793 and African *Bos taurus*. *G3 (Bethesda)* 5**,** 1391-1401. doi: 10.1534/g3.115.017772.

bioRxiv preprint doi: https://doi.org/10.1101/210229; this version posted October 28, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

794 Rosenberg, NA, Pritchard, JK, Weber, JL, Cann, HM, Kidd, KK, Zhivotovsky, LA, and Feldman,
795 MW (2002). Genetic structure of human populations. *Science* 298**,** 2381-2385. doi:
796 10.1126/science.1078311.
797 Schlager, MA, Kapitein, LC, Grigoriev, I, Burzynski, GM, Wulf, PS, Keijzer, N, de Graaff, E,
798 Fukuda, M, Shepherd, IT, Akhmanova, A, and Hoogenraad, CC (2010). Pericentrosomal
799 targeting of Rab6 secretory vesicles by Bicaudal-D-related protein 1 (BICDR-1) regulates
800 neuritogenesis. *EMBO J* 29**,** 1637-1651. doi: 10.1038/emboj.2010.51.
801 Shinmyo, Y, Asrafuzzaman Riyadh, M, Ahmed, G, Bin Naser, I, Hossain, M, Takebayashi, H,
802 Kawasaki, H, Ohta, K, and Tanaka, H (2015). Draxin from neocortical neurons controls the
803 guidance of thalamocortical projections into the neocortex. *Nat Commun* 6**,** 10232. doi:
804 10.1038/ncomms10232.
805 Smedley, D, Haider, S, Durinck, S, Pandini, L, Provero, P, Allen, J, Arnaiz, O, Awedh, MH,
806 Baldock, R, Barbiera, G, Bardou, P, Beck, T, Blake, A, Bonierbale, M, Brookes, AJ, Bucci,
807 G, Buetti, I, Burge, S, Cabau, C, Carlson, JW, Chelala, C, Chrysostomou, C, Cittaro, D,
808 Collin, O, Cordova, R, Cutts, RJ, Dassi, E, Di Genova, A, Djari, A, Esposito, A, Estrella, H,
809 Eyras, E, Fernandez-Banet, J, Forbes, S, Free, RC, Fujisawa, T, Gadaleta, E, Garcia-
810 Manteiga, JM, Goodstein, D, Gray, K, Guerra-Assuncao, JA, Haggarty, B, Han, DJ, Han,
811 BW, Harris, T, Harshbarger, J, Hastings, RK, Hayes, RD, Hoede, C, Hu, S, Hu, ZL,
812 Hutchins, L, Kan, Z, Kawaji, H, Keliet, A, Kerhornou, A, Kim, S, Kinsella, R, Klopp, C,
813 Kong, L, Lawson, D, Lazarevic, D, Lee, JH, Letellier, T, Li, CY, Lio, P, Liu, CJ, Luo, J,
814 Maass, A, Mariette, J, Maurel, T, Merella, S, Mohamed, AM, Moreews, F, Nabihoudine, I,
815 Ndegwa, N, Noirot, C, Perez-Llamas, C, Primig, M, Quattrone, A, Quesneville, H,
816 Rambaldi, D, Reecy, J, Riba, M, Rosanoff, S, Saddiq, AA, Salas, E, Sallou, O, Shepherd, R,
817 Simon, R, Sperling, L, Spooner, W, Staines, DM, Steinbach, D, Stone, K, Stupka, E,
818 Teague, JW, Dayem Ullah, AZ, Wang, J, Ware, D, et al. (2015). The BioMart community
819 portal: an innovative alternative to large, centralized data repositories. *Nucleic Acids Res* 43**,**
820 W589-598. doi: 10.1093/nar/gkv350.
821 Smith, P, and Gregory, PJ (2013). Climate change and sustainable food production. *Proc Nutr Soc*
822 72**,** 21-28. doi: 10.1017/s0029665112002832.
823 Szpiech, ZA, Xu, J, Pemberton, TJ, Peng, W, Zollner, S, Rosenberg, NA, and Li, JZ (2013). Long
824 runs of homozygosity are enriched for deleterious variation. *Am J Hum Genet* 93**,** 90-102.
825 doi: 10.1016/j.ajhg.2013.05.003.
826 Timmons, JA, Szkop, KJ, and Gallagher, IJ (2015). Multiple sources of bias confound functional
827 enrichment analysis of global -omics data. *Genome Biol* 16**,** 186. doi: 10.1186/s13059-015-
828 0761-7.

bioRxiv preprint doi: https://doi.org/10.1101/210229; this version posted October 28, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

829 Troy, CS, MacHugh, DE, Bailey, JF, Magee, DA, Loftus, RT, Cunningham, P, Chamberlain, AT,
830 Sykes, BC, and Bradley, DG (2001). Genetic evidence for Near-Eastern origins of European
831 cattle. *Nature* 410**,** 1088-1091. doi: 10.1038/35074088.
832 Upadhyay, MR, Chen, W, Lenstra, JA, Goderie, CR, MacHugh, DE, Park, SD, Magee, DA,
833 Matassino, D, Ciani, F, Megens, HJ, van Arendonk, JA, Groenen, MA, European Cattle
834 Genetic Diversity, C, and Crooijmans, R (2017). Genetic origin, admixture and population
835 history of aurochs (*Bos primigenius*) and primitive European cattle. *Heredity (Edinb)* 118**,**
836 169-176. doi: 10.1038/hdy.2016.79.
837 Vihma, T (2014). Effects of Arctic sea ice decline on weather and climate: a review. *Surveys in*
838 *Geophysics* 35**,** 1175-1214. doi: 10.1007/s10712-014-9284-0.
839 Villarroel-Campos, D, Henriquez, DR, Bodaleo, FJ, Oguchi, ME, Bronfman, FC, Fukuda, M, and
840 Gonzalez-Billault, C (2016). Rab35 functions in axon elongation are regulated by P53-
841 related protein kinase in a mechanism that involves Rab35 protein degradation and the
842 microtubule-associated protein 1B. *J Neurosci* 36**,** 7298-7313. doi: 10.1523/jneurosci.4064-
843 15.2016.
844 Visser, C, Lashmar, SF, Van Marle-Koster, E, Poli, MA, and Allain, D (2016). Genetic diversity
845 and population structure in South African, French and Argentinian Angora goats from
846 genome-wide SNP data. *PLoS ONE* 11**,** e0154353. doi: 10.1371/journal.pone.0154353.
847 Waples, RK, Larson, WA, and Waples, RS (2016). Estimating contemporary effective population
848 size in non-model species using linkage disequilibrium across thousands of loci. *Heredity*
849 *(Edinb)* 117**,** 233-240. doi: 10.1038/hdy.2016.60.
850 Wheeler, T, and von Braun, J (2013). Climate change impacts on global food security. *Science* 341**,**
851 508-513. doi: 10.1126/science.1239402.
852 Williams, JL, Hall, SJ, Del Corvo, M, Ballingall, KT, Colli, L, Ajmone Marsan, P, and Biscarini, F
853 (2016). Inbreeding and purging at the genomic level: the Chillingham cattle reveal
854 extensive, non-random SNP heterozygosity. *Anim Genet* 47**,** 19-27. doi: 10.1111/age.12376.
855 Zhang, Q, Guldbrandtsen, B, Bosse, M, Lund, MS, and Sahana, G (2015). Runs of homozygosity
856 and distribution of functional variants in the cattle genome. *BMC Genomics* 16**,** 542. doi:
857 10.1186/s12864-015-1715-x.
858 Zimin, AV, Delcher, AL, Florea, L, Kelley, DR, Schatz, MC, Puiu, D, Hanrahan, F, Pertea, G, Van
859 Tassell, CP, Sonstegard, TS, Marcais, G, Roberts, M, Subramanian, P, Yorke, JA, and
860 Salzberg, SL (2009). A whole-genome assembly of the domestic cow, *Bos taurus*. *Genome*
861 *Biol* 10**,** R42. doi: gb-2009-10-4-r42 [pii]

862 10.1186/gb-2009-10-4-r42.

bioRxiv preprint doi: https://doi.org/10.1101/210229; this version posted October 28, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

863 # 10. Figures

864

865 **FIGURE 1 |** Photograph of a Kerry cow and locations of Kerry cattle herd DNA sampling in southern
866 Ireland. The area of each circle corresponds to the size of each population sample. Dark green =
867 animals sampled during 1991/92 (KY92); light green = animals sampled during 2011/12 (KY12).
868 Kerry cow image is copyright of the Kerry Cattle Society Ltd.

869 **FIGURE 2 |** Maximum likelihood (ML) phylogenetic tree network graph with five migration edges
870 ($m = 5$) generated for genome-wide SNP data (36,000 autosomal SNPs) from European cattle breeds
871 (EU data set). The West African taurine N'Dama breed sampled in Guinea is included as a population
872 outgroup. Coloured lines and arrows show migration edges that model gene flow between lineages
873 with different migration weights represented by the colour gradient.

874 **FIGURE 3 |** Maximum likelihood (ML) phylogenetic tree network graph with five migration edges
875 ($m = 5$) generated for genome-wide SNP data (37,490 autosomal SNPs) from cattle breeds of British
876 and Irish origin (BI data set). The West African taurine N'Dama breed sampled in Guinea is included
877 as a population outgroup. Coloured lines and arrows show migration edges that model gene flow
878 between lineages with different migration weights represented by the colour gradient.

879 **FIGURE 4 |** Principal component analysis plot constructed for PC1 and PC2 from genome-wide SNP
880 data (36,621 autosomal SNPs) for the EU data set of 605 individual animals. The smaller histogram
881 plot shows the relative variance contributions for the first 10 PCs.

882 **FIGURE 5 |** Principal component analysis plot constructed for PC1 and PC2 from genome-wide SNP
883 data (37,395 autosomal SNPs) for the BI data set of 351 individual animals. The smaller histogram
884 plot shows the relative variance contributions for the first 10 PCs.

bioRxiv preprint doi: https://doi.org/10.1101/210229; this version posted October 28, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

885 **FIGURE 6 |** Hierarchical clustering of individual animals using genome-wide SNP data (36,621
886 autosomal SNPs) for the EU data set of 605 individual animals. Results are shown for modelled
887 ancestral populations $K = 2$ to 14. The cluster numbers corresponding to the likely number of ancestral
888 populations are highlighted with a light red overlay and the two outlier Kerry samples (KY12_06 and
889 KY12_58) are indicated with red arrows.

890 **FIGURE 7 |** Hierarchical clustering of individual animals using genome-wide SNP data (37,395
891 autosomal SNPs) for the BI data set of 351 individual animals. Results are shown for modelled
892 ancestral populations $K = 2$ to 9. The cluster numbers corresponding to the likely number of ancestral
893 populations are highlighted with a light red overlay and the two outlier Kerry samples (KY12_06 and
894 KY12_58) are indicated with red arrows.

895 **FIGURE 8 |** Genetic effective population size (N_e) trends modelled using genome-wide SNP data.
896 Results for the KY92 and KY12 populations are shown with seven comparator heritage and
897 production cattle breeds.

898 **FIGURE 9 |** Tukey box plots showing distribution of F_{ROH} values estimated with genome-wide SNP
899 data for the KY92 and KY12 populations and nine comparator heritage and production cattle breeds.

900 **FIGURE 10 |** Manhattan plots of composite selection signal (CSS) results for Kerry cattle ($n = 98$)
901 contrasted with EU cattle ($n = 102$). A) Unsmoothed results. B) Smoothed results obtained by
902 averaging CSS of SNPs within each 1 Mb window. Red dotted line on each plot denotes the genome-
903 wide 0.1% threshold for the empirical CSS scores. Red vertical arrows indicate selection peaks
904 detected on BTA09, BTA12, BTA16, BTA17, BTA19 and BTA28.

bioRxiv preprint doi: https://doi.org/10.1101/210229; this version posted October 28, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

905 **Table 1:** Cattle breed/population samples used for the present study

Breed/population	Code	Data set	Breed purpose	Country of origin	Source
Angus	ANGU	BI/EU	Beef	Scotland	1, 2
Belted Galloway	BGAL	BI	Beef	Scotland	2
British Shorthorn	BSHN	BI/EU	Dual purpose	England	2
Brown Swiss	BRSW	EU	Dairy	Switzerland	1, 2, 3
Charolais	CHAR	EU	Beef	France	1, 2, 3
Devon	DEVN	BI	Beef	England	2
Dexter	DXTR	BI	Dual purpose	Ireland	2
English Longhorn	ELHN	BI	Beef	England	2
Finnish Ayrshire	FAYR	BI/EU	Dairy	Scotland/Finland	2
Galloway	GALL	BI	Beef	Scotland	2
Gelbvieh	GELB	EU	Dual purpose	Germany	2, 4
Guernsey	GNSY	BI/EU	Dairy	Channel Islands	1, 2
Hereford	HRFD	BI/EU	Beef	England	1, 2
Holstein	HOLS	EU	Dairy	The Netherlands	1, 2, 5
Jersey	JRSY	BI/EU	Dairy	Channel Islands	1, 2, 3
Kerry sampled 1991/92	KY92	BI/EU	Dairy	Ireland	Current
Kerry sampled 2011/12	KY12	BI/EU	Dairy	Ireland	Current
Limousin	LIMS	EU	Beef/draft	France	1, 2
Lincoln Red	LNCR	BI	Beef	England	2
Montbeliarde	MONT	EU	Dairy	France	2, 5
N'Dama	NDAM	---	Dual purpose	Guinea (West Africa)	1
Norwegian Red	NRED	EU	Dairy	Norway	1
Piedmontese	PDMT	EU	Dual purpose	Italy	1, 2
Red Angus	RANG	BI/EU	Beef	Scotland	1
Red Poll	REDP	BI	Beef	England	2
Romagnola	ROMG	EU	Beef/draft	Italy	1
Scottish Highland	SCHL	BI	Beef	Scotland	2
Simmental	SIMM	EU	Dual purpose/draft	Switzerland	2, 4
South Devon	SDEV	BI	Beef	England	2
Sussex	SUSX	BI	Beef/draft	England	2
Welsh Black	WBLK	BI	Dual purpose	Wales	2
White Park	WHPK	BI	Dual purpose/draft	England	2

906 [1] Gibbs et al. (2009); [2] Decker et al. (2009); [3] Gautier et al. (2010); [4] Matukumalli et al. (2009); [5] Flori et al. (2009)
907

bioRxiv preprint doi: https://doi.org/10.1101/210229; this version posted October 28, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

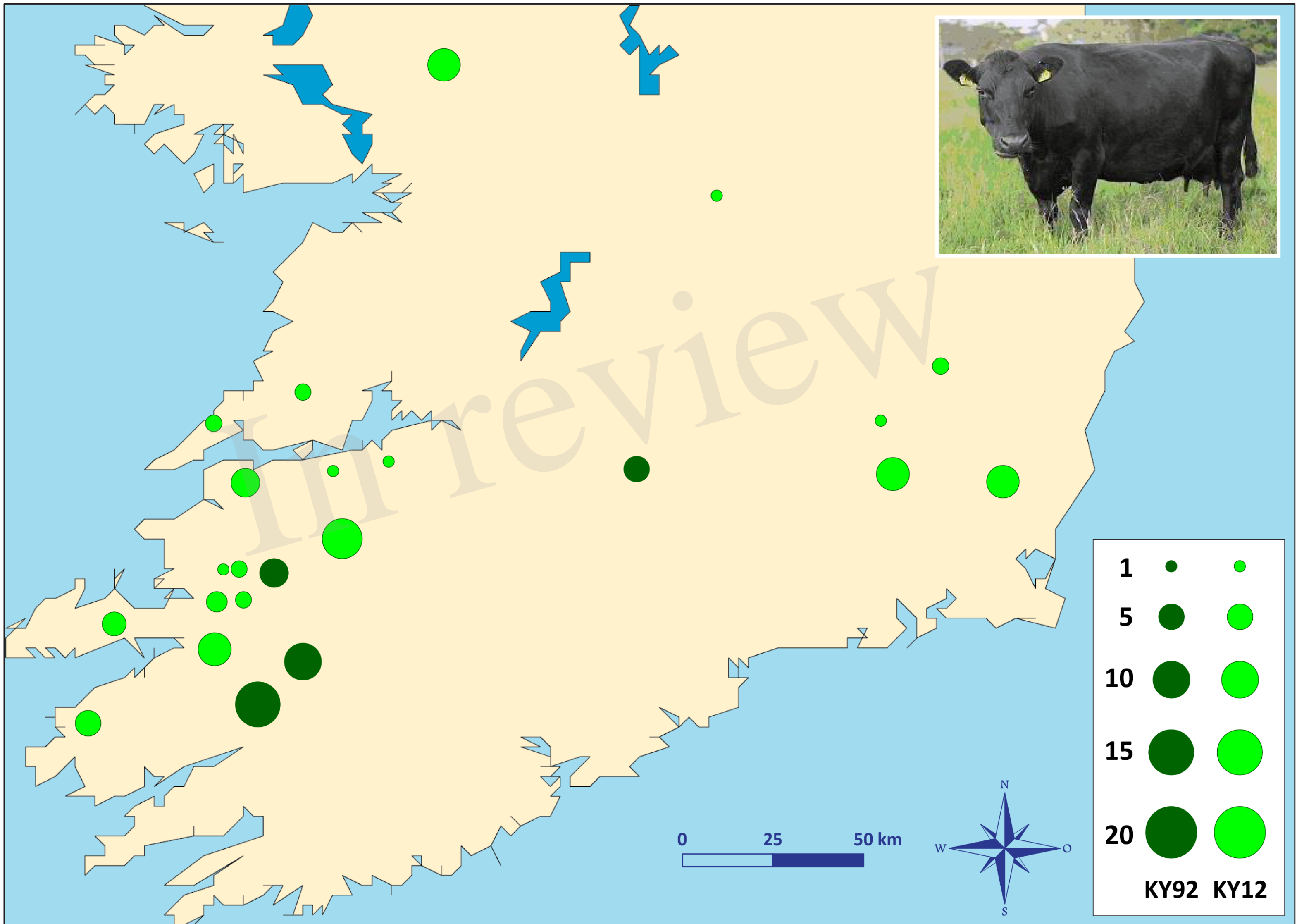
908 **Table 2:** Breed/population sample size, observed heterozygosity and SNP filtering information

Breed/population	Code	Data set	Sample size (n) post-filtering	*Observed heterozygosity* H_o	No. SNPs N_e modelling	No. SNPs genomic inbreeding
Angus	ANGU	BI/EU	72	0.3048	31,413	34,531
Belted Galloway	BGAL	BI	4	0.2902	25,997	25,997
British Shorthorn	BSHN	BI/EU	10	0.2549	29,038	29,038
Brown Swiss	BRSW	EU	31	0.28943	---	---
Charolais	CHAR	EU	48	0.3209	---	---
Devon	DEVN	BI	4	0.2859	---	---
Dexter	DXTR	BI	4	0.2458	24,753	24,753
English Longhorn	ELHN	BI	3	0.2232	---	---
Finnish Ayrshire	FAYR	BI/EU	7	0.3064	---	---
Galloway	GALL	BI	4	0.2942	---	---
Gelbvieh	GELB	EU	8	0.3125	---	---
Guernsey	GNSY	BI/EU	19	0.2764	---	38,437
Hereford	HRFD	BI/EU	35	0.2964	---	35,156
Holstein	HOLS	EU	70	0.3192	36,152	38,717
Jersey	JRSY	BI/EU	44	0.2718	31,358	35,990
Kerry sampled 1991/92	KY92	BI/EU	36	0.2965	37,556	40,801
Kerry sampled 2011/12	KY12	BI/EU	62	0.3042	36,428	40,730
Limousin	LIMS	EU	45	0.3122	---	---
Lincoln Red	LNCR	BI	7	0.2789	26 350	26,350
Montbeliarde	MONT	EU	31	0.3019	---	---
N'Dama	NDAM	---	22	0.2158	---	---
Norwegian Red	NRED	EU	20	0.3190	---	---
Piedmontese	PDMT	EU	23	0.3240	---	---
Red Angus	RANG	BI/EU	14	0.3092	---	---
Red Poll	REDP	BI	5	0.2905	---	---
Romagnola	ROMG	EU	21	0.2943	---	---
Scottish Highland	SCHL	BI	8	0.2823	---	---
Simmental	SIMM	EU	9	0.3136	---	---
South Devon	SDEV	BI	3	0.3070	---	---
Sussex	SUSX	BI	4	0.2792	---	---
Welsh Black	WBLK	BI	2	0.3203	---	---
White Park	WHPK	BI	4	0.2270	---	---

909

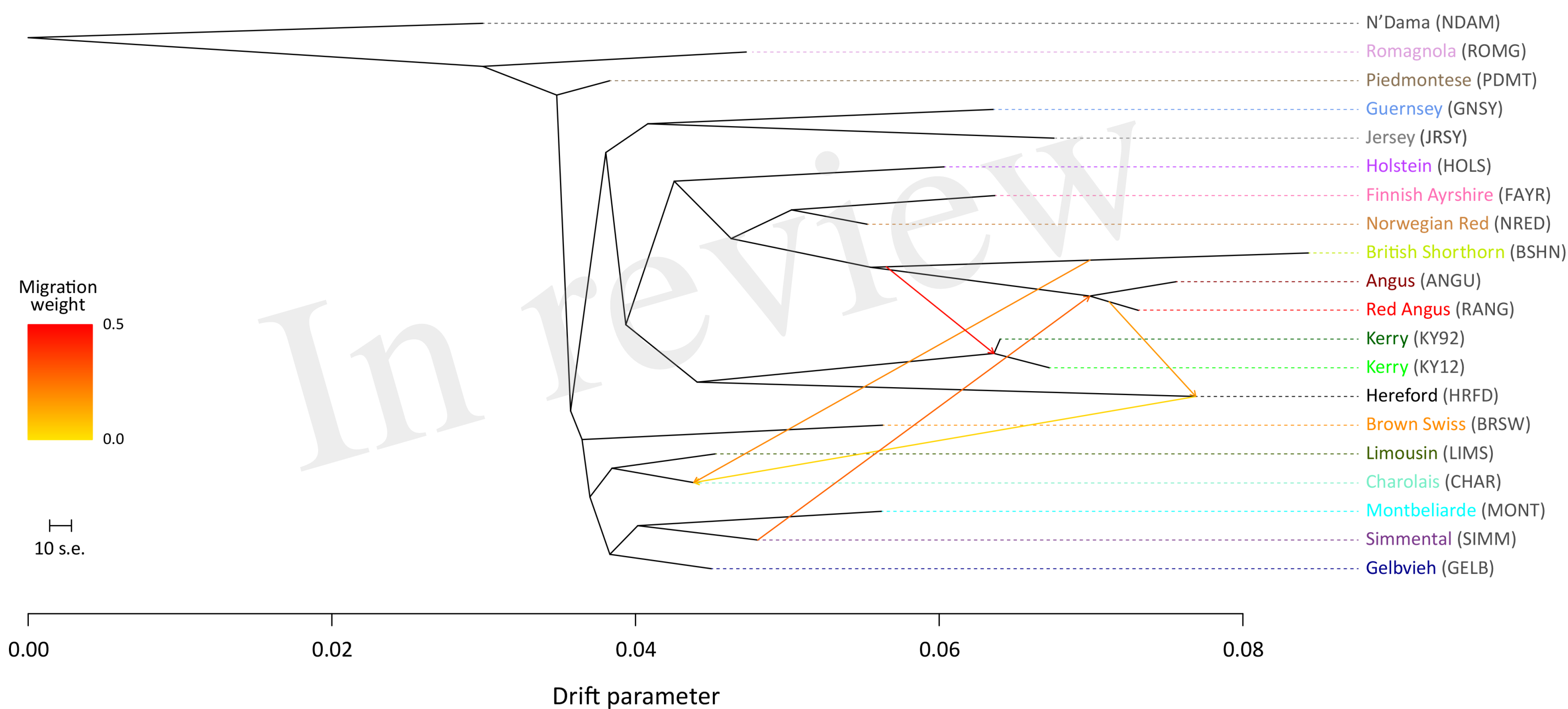
bioRxiv preprint doi: https://doi.org/10.1101/210229; this version posted October 28, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Figure 1.TIF



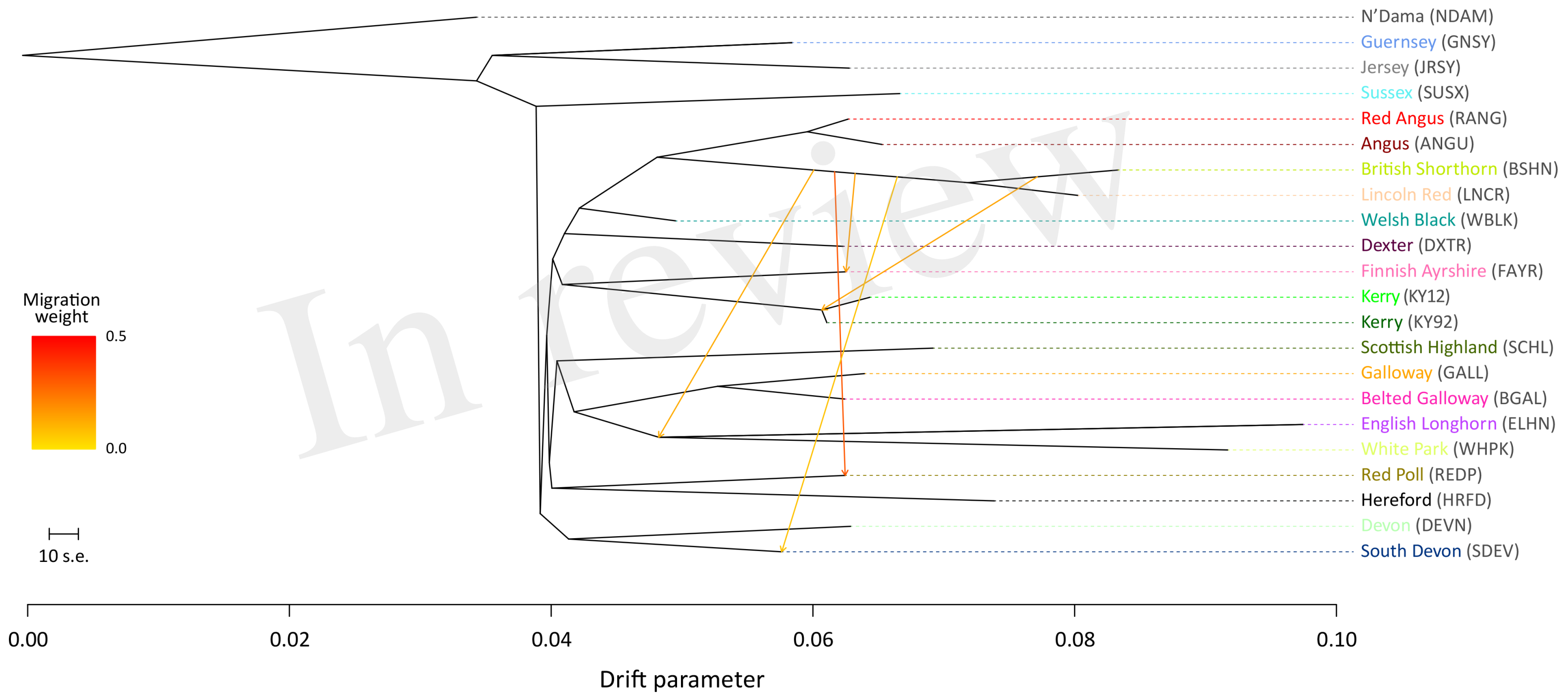
bioRxiv preprint doi: https://doi.org/10.1101/210229; this version posted October 28, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Figure 2.TIF



bioRxiv preprint doi: https://doi.org/10.1101/210229; this version posted October 28, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

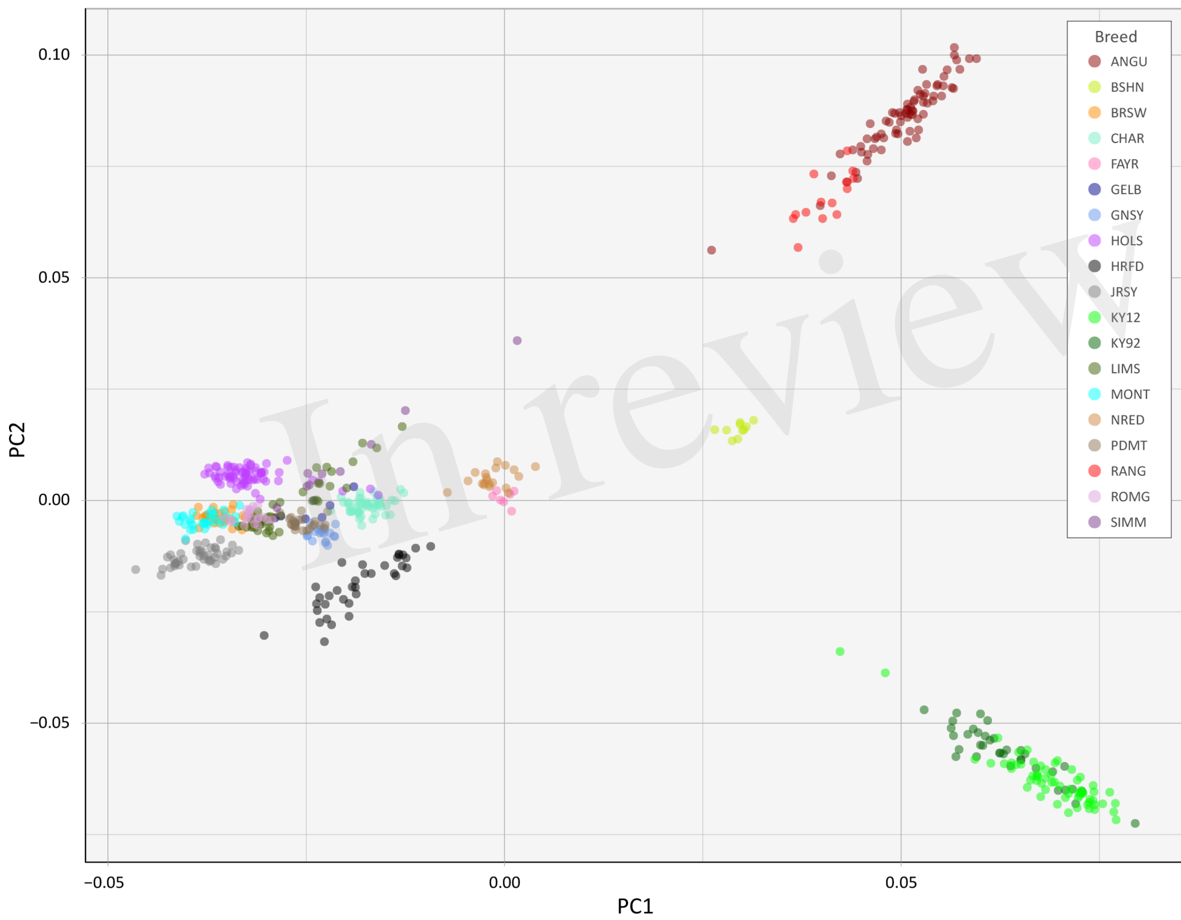
Figure 3.TIF

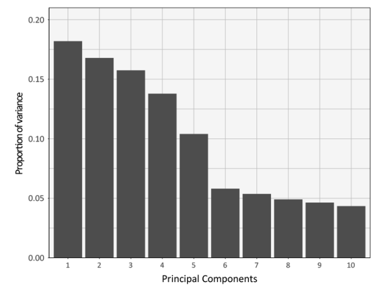


In review

bioRxiv preprint doi: https://doi.org/10.1101/210229; this version posted October 28, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

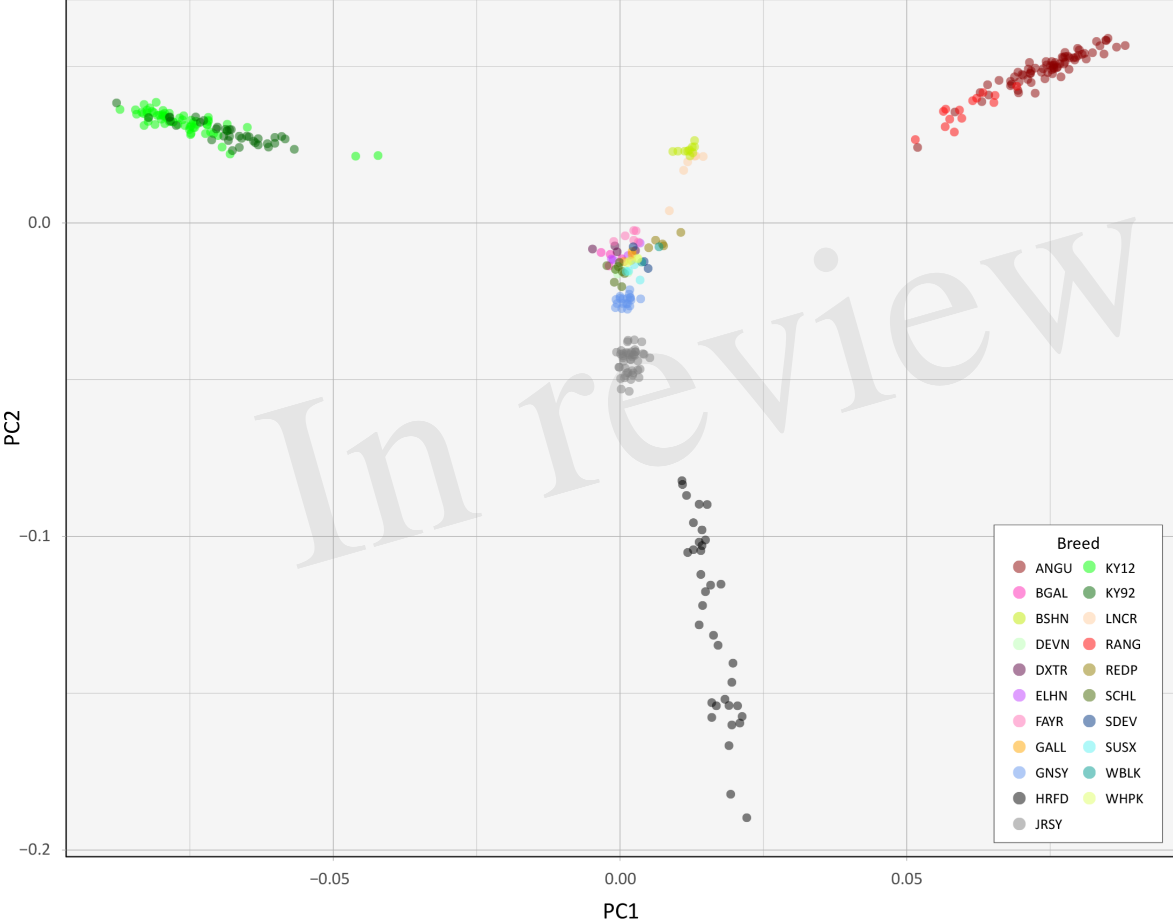
Figure 4.TIF

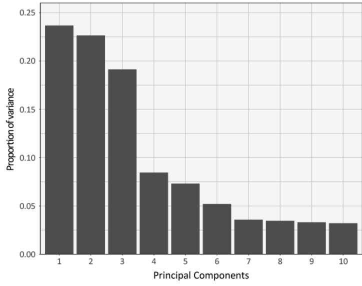




bioRxiv preprint doi: https://doi.org/10.1101/210229; this version posted October 28, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

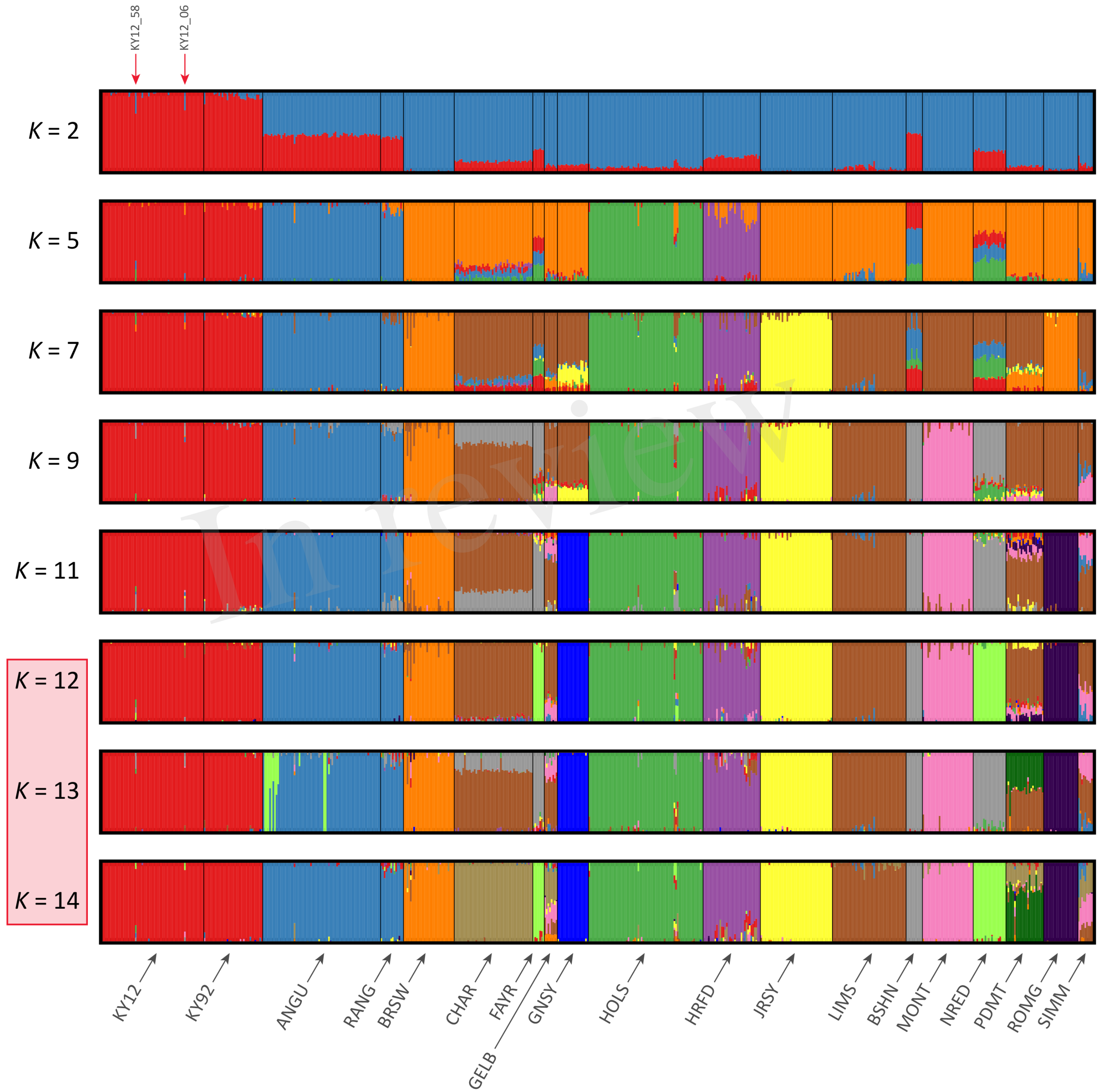
Figure 5.TIF





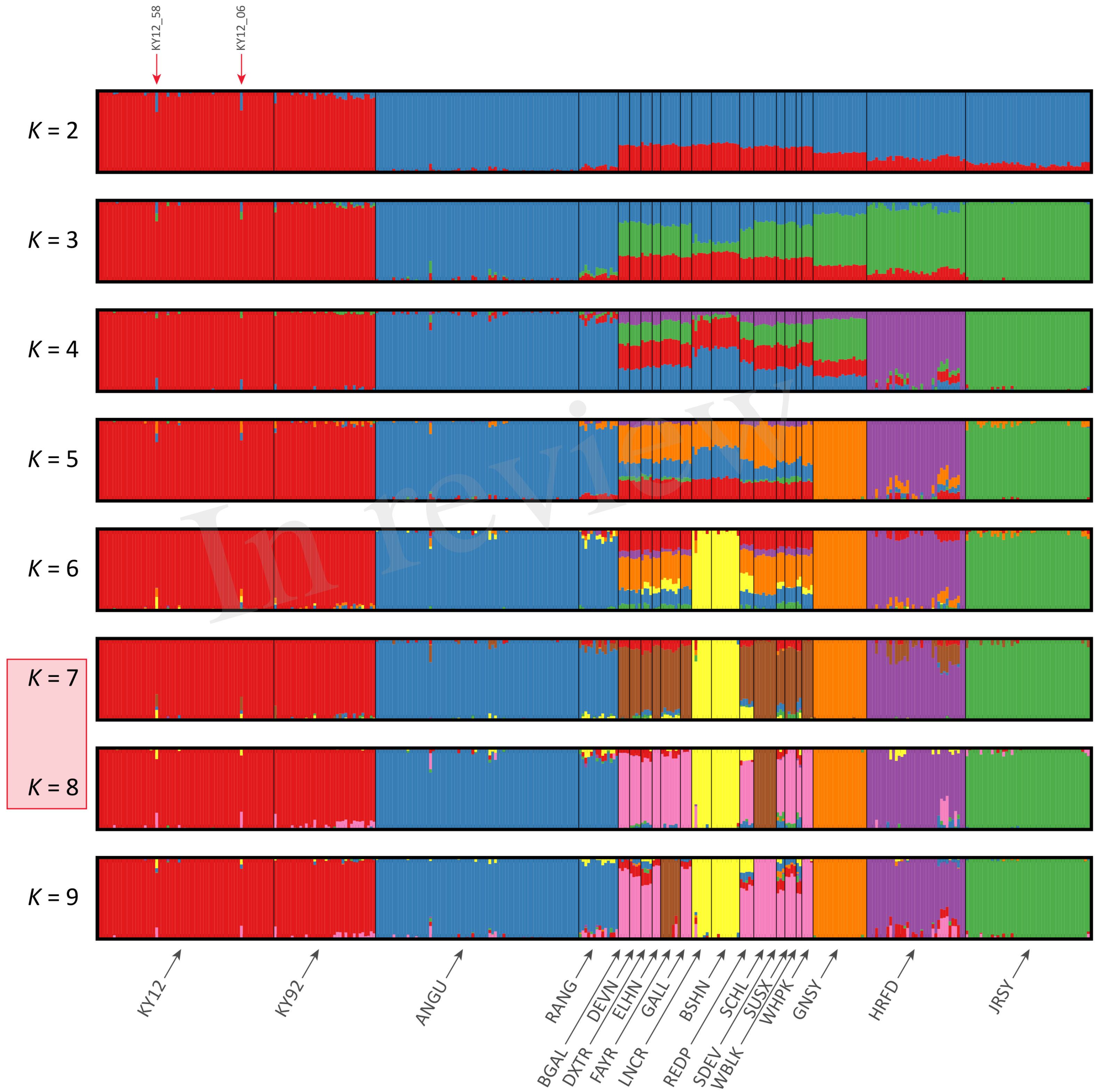
bioRxiv preprint doi: https://doi.org/10.1101/210229; this version posted October 28, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Figure 6.TIF



bioRxiv preprint doi: https://doi.org/10.1101/210229; this version posted October 28, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Figure 7.tif



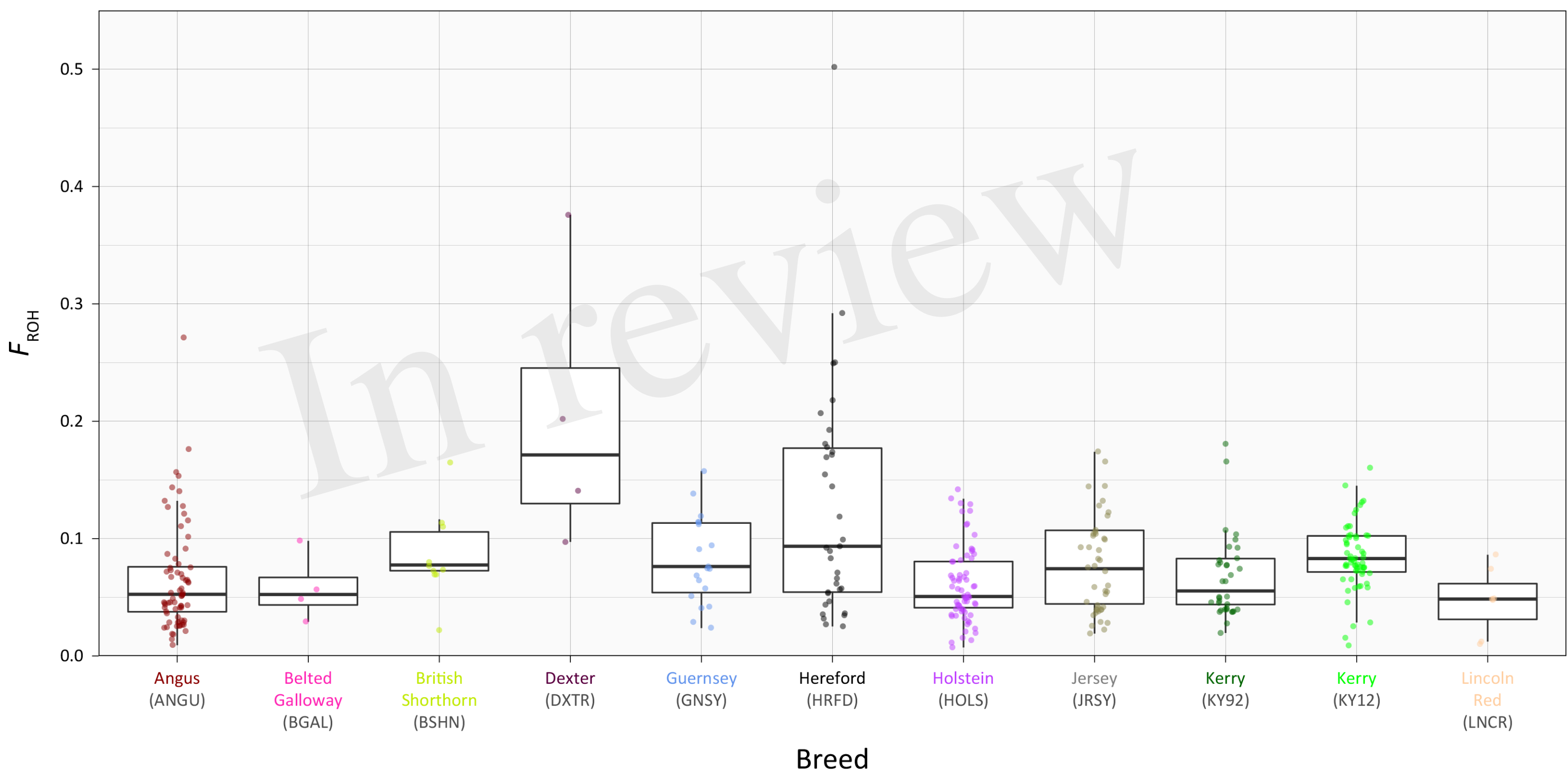
bioRxiv preprint doi: https://doi.org/10.1101/210229; this version posted October 28, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Figure 8.TIF

In review

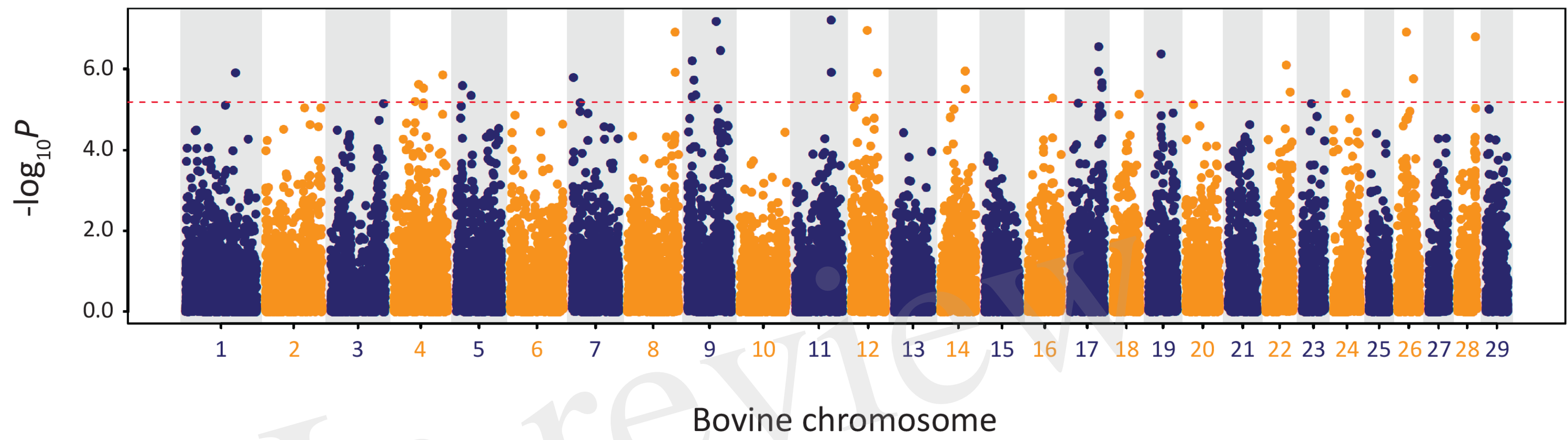
bioRxiv preprint doi: https://doi.org/10.1101/210229; this version posted October 28, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

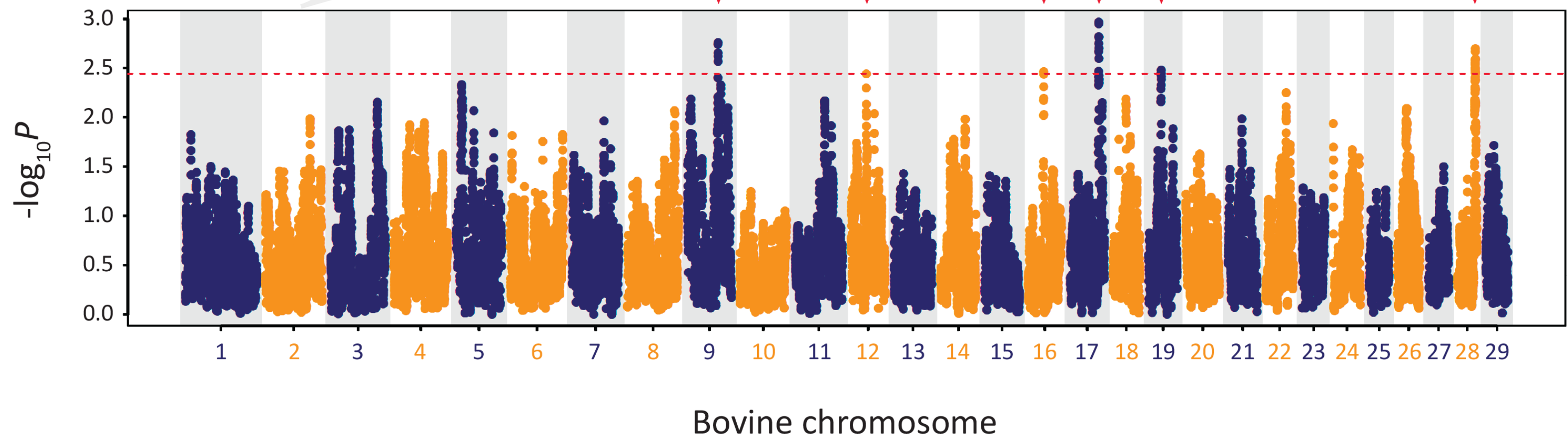
Figure 9.TIF



bioRxiv preprint doi: https://doi.org/10.1101/210229; this version posted October 28, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Figure 10.TIF





In review