Genomic characterisation of the indigenous Irish Kerry cattle breed

Browett, S, McHugo, G, Richardson, IW, Magee, DA, Park, SDE, Fahey, AG, Kearney, JF, Correia, CN, Randhawa, IAS and MacHugh, DE

http://dx.doi.org/10.3389/fgene.2018.00051

<table>
<thead>
<tr>
<th>Title</th>
<th>Genomic characterisation of the indigenous Irish Kerry cattle breed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Authors</td>
<td>Browett, S, McHugo, G, Richardson, IW, Magee, DA, Park, SDE, Fahey, AG, Kearney, JF, Correia, CN, Randhawa, IAS and MacHugh, DE</td>
</tr>
<tr>
<td>Type</td>
<td>Article</td>
</tr>
<tr>
<td>URL</td>
<td>This version is available at: <a href="http://usir.salford.ac.uk/id/eprint/46300/">http://usir.salford.ac.uk/id/eprint/46300/</a></td>
</tr>
<tr>
<td>Published Date</td>
<td>2018</td>
</tr>
</tbody>
</table>

USIR is a digital collection of the research output of the University of Salford. Where copyright permits, full text material held in the repository is made freely available online and can be read, downloaded and copied for non-commercial private study or research purposes. Please check the manuscript for any further copyright restrictions.

For more information, including our policy and submission procedure, please contact the Repository Team at: usir@salford.ac.uk.
Genomic Characterisation of the Indigenous Irish Kerry Cattle Breed

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

1 Ecosystems and Environment Research Centre, School of Environment and Life Sciences, University of Salford, Salford, United Kingdom, 2 Animal Genomics Laboratory, UCD School of Agriculture and Food Science, University College Dublin, Dublin, Ireland, 3 IdentiGEN Ltd., Dublin, Ireland, 4 Irish Cattle Breeding Federation, Bandon, Ireland, 5 Sydney School of Veterinary Science, The University of Sydney, Camden, NSW, Australia, 6 UCD Conway Institute of Biomolecular and Biomedical Research, University College Dublin, Dublin, Ireland

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

Keywords: cattle, conservation genomics, endangered breed, inbreeding, genetic diversity, population genomics, selection signature, single nucleotide polymorphism

INTRODUCTION

Approximately 10,000 years ago, humans first domesticated wild aurochs (Bos primigenius)—the progenitor of modern cattle—in the Fertile Crescent region of Southwest Asia (Larson and Fuller, 2014; Larson et al., 2014; MacHugh et al., 2017). Extant domestic cattle, which encompass humpless taurox (B. taurus), humped zebu (B. indicus) and myriad B. taurus/indicus hybrid populations, have, through genetic drift and natural and artificial selection, diversified into more than 1,100 recognised breeds. However, beginning in the middle of the twentieth century, socioeconomic preferences for large highly productive dairy, beef and dual-purpose breeds have led to extinction and increased vulnerability of more than 200 locally-adapted landrace or native cattle breeds (Gandini et al., 2004; Food and Agriculture Organization, 2007, 2015).

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

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

As a native breed with a claimed ancient heritage, Kerry cattle are considered culturally important to Ireland (Curran, 1990). It is a landrace cattle population that remains productive in harsh upland regions with poor quality feed, which are typical of southwest Ireland where the Kerry breed evolved (Food and Agriculture Organization, 2017). These cattle were often referred to anecdotally in Ireland as the “poor man’s cow” due to their ability to produce relatively large quantities of milk on very sparse fodder; the Kerry breed is also considered to be a remnant of what was once a substantially larger and more widespread historical population. Levels of inbreeding have been estimated using pedigree data and the accumulated figure since the foundation of the herd book in 1887 reached 15% in 1985 (O’Huigin and Cunningham, 1990).

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

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

The Kerry cattle breed was one of the first European heritage cattle breeds to be surveyed using molecular population genetics techniques. We have previously used autosomal microsatellite genetic markers and mitochondrial DNA (mtDNA) control region sequence variation for comparative evolutionary studies of genetic diversity in Kerry cattle and other British, European, African and Asian breeds (MacHugh et al., 1997, 1998, 1999; Troy et al., 2001). In addition, Bray et al. have used microsatellites to examine admixture and ancestry in Kerry cattle and the Dexter and Devon breeds (Bray et al., 2009). Results from these studies demonstrated that Kerry cattle exhibit markedly low mtDNA sequence diversity, but autosomal microsatellite diversity comparable to other cattle breeds native to Britain and Ireland. More recently, analyses of medium- and high-density SNP genotypes generated using genome sequence data from an extinct British B. primigenius subfossil have shown that Kerry cattle retain a significant genomic signature of admixture from wild aurochs (Park et al., 2015; Upadhyay et al., 2017). This observation highlights the genetic distinctiveness of the Kerry population and has major implications for conservation and management of the breed.

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

MATERIALS AND METHODS
Kerry Cattle Population DNA Sampling in 1991/92 and 2011/12

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

The second Kerry cattle population sample was collected in 2011/12 from 19 different herds located across southern and western Ireland. Performagene (PG-100) nasal swab DNA collection kits were used for biological sample collection (DNA
Genetek Inc., Ottawa, Canada). Nasal swab DNA samples were collected from a total of 75 male and female Kerry cattle and owners were consulted to ensure that a representative sample of animals was obtained. This Kerry population sample group is coded as KY12. Genomic DNA was purified from 0.5 ml of each PG-100 nasal swab sample using the laboratory protocol recommended by the manufacturer (DNA Genotek Inc.).

SNP Genotyping and Assembly of Comparative SNP Data Sets

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

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

Sample Removal and Quality Control and Filtering of SNPs

Genomic non-exclusion—in other words, genome-wide SNP profiles completely compatible with parent-offspring relationship (with allowance for very low-level genotyping error)—were used to identify animals from the KY92 and KY12 population samples that were parent-offspring pairs. One of the two animals in each pair was then randomly removed to generate the working SNP data set. Following this procedure, quality control and filtering based on recorded SNP genotypes was performed as detailed below for the EU and BI data sets.

Prior to quality control and filtering there were 54,057 SNPs in the EU data set (608 animals, including KY92 and KY12) and in the BI data set (354 animals, including KY92 and KY12). SNP quality filtering was performed using PLINK version 1.07 (Purcell et al., 2007), such that individual SNPs with more than 10% missing data and a minor allele frequency (MAF) of ≤0.01 (1%) were removed from both data sets; however, for analyses of genomic inbreeding and runs of homozygosity (ROH) the MAF filtering threshold was not imposed. Only autosomal SNPs were retained and individual animal samples with a SNP call rate less than 90% were also removed from each of the two data sets.

SNP quality control and filtering were performed across breeds/populations (by data set) for construction of phylogenies and ancestor graphs, multivariate analysis, investigation of population structure and detection of signatures of selection. For intrapopulation analyses of effective population size ($N_e$)}
TABLE 1 | Cattle breed/population samples used for the present study.

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

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

and genomic inbreeding, all SNPs genotyped (54,057) were filtered within breeds/populations as detailed above. However, an additional filtering procedure was used to remove SNPs deviating from Hardy-Weinberg equilibrium (HWE) with a P-value threshold of < 0.0001. Also, for the N_e analysis, a more stringent MAF threshold of 0.05 was used.

Generation of Identity-by-State (IBS) Matrix
Using the filtered genome-wide SNP data, PLINK v1.07 was also used to generate identity-by-state (IBS) values for all pairs of Kerry cattle (KY92 and KY12), including the nine blinded sample duplicates for sample verification and tracking purposes.

Construction of Phylogenetic Trees and Ancestry Graphs
Maximum likelihood (ML) phylogenetic trees with ancestry graphs were generated for the EU and BI data sets using the TreeMix (version 1.12) software package (Pickrell and Pritchard, 2012). The West African B. taurus NDAM breed sample (n = 22) was used as an outgroup. TreeMix was run without using SNP blocks (as described in the TreeMix software documentation) and ML phylogenetic trees were generated with no migration edges (m = 0) up to ten migration edges (m = 10).

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

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

**Modelling Current and Historical Effective Population Size \( (N_e) \)**

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

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

where \( N_{T} \) is the effective population size \( t \) generations ago calculated as \( t = \left(2f(c_t)\right)^{-1} \) (Hayes et al., 2003), \( c_t \) is the recombination rate defined for a specific physical distance between SNP markers, \( r_{adj}^2 \) is the LD value adjusted for sample size and the recommended default \( \alpha \) value = 1 was used to correct for the occurrence of mutation (Barbato et al., 2015). In addition, the SNeP program option for unphased SNP data was used for the analyses described here.

**Evaluation of Genomic Inbreeding and Runs of Homozygosity (ROH)**

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

\[
F_{ROH} = \frac{\sum L_{ROH}}{L_{AUTO}}
\]

PLINK v1.07 was used to define ROH using a sliding window approach and procedures modified from previous recommendations for Illumina® Bovine SNP50 BeadChip and similar SNP data sets (Purfield et al., 2012, 2017). The criteria for defining individual ROH were set such that the ROH was required to be at least 500 kb in length, with a minimum density of one SNP per 120 kb and that there was a gap of at least 1,000 kb between each ROH. A sliding window of 50 SNPs was incrementally advanced one SNP at a time along the genome; each discrete window could contain a maximum of one heterozygous SNP and no more than two SNPs with missing genotypes. Following Purfield et al. (2012) all filtered genomic SNPs (without a MAF threshold), including those located in centromeric regions, were used to estimate \( F_{ROH} \) values for individual animals.

**Genome-Wide Detection of Signatures of Selection and Functional Enrichment Analysis**

In the absence of hard selective sweeps, single selection tests using high-density SNP data do not perform well in detecting signatures of selection from individual livestock breeds (Kemper et al., 2014). Therefore, for the present study, genomic signatures of selection were identified using the composite selection signal (CSS) method introduced by Randhawa et al. (2014). The CSS method has been shown to be a robust and sensitive approach for detecting genomic signatures of selection underlying microevolution of complex traits in livestock (Randhawa et al., 2015). The CSS is a weighted index of signatures of selection from multiple estimates; it is a nonparametric procedure that uses fractional ranks of constituent tests and does not depend on assumptions about the distributions of individual test results.

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

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

**RESULTS AND DISCUSSION**

**Sample Removal and SNP Filtering and Quality Control**

Genomic non-exclusion identified 20 parent-offspring pairs from the KY12 population and one sample from each pair was randomly removed (sample codes: KY12_01, KY12_05, KY12_13, KY12_14, KY12_17, KY12_18, KY12_19, KY12_46, KY12_55, KY12_67). Thereafter, general SNP quality control and filtering led to additional samples being excluded (KY12_26, KY12_28 and KY12_54), giving a total filtered KY12 population sample of 62 animals for downstream population genomics analyses.
After SNP quality control and filtering across the two composite data sets (EU and BI), there were 36,621 autosomal SNPs from 605 individual animals in the EU data set and there were 37,395 autosomal SNPs from 351 animals in the BI data set. When the West African NDAM breed sample \((n = 22)\) was included for the ML phylogenetic tree and ancestry graph analyses, the number of SNPs used was 36,000 from 627 animals for the EU data set and 37,490 from 373 animals for the BI data set. The final numbers of SNPs used for individual breed/population analyses of \(N_e\) and genomic inbreeding after all quality control and filtering (including additional filtering for deviations from HWE) are shown in Table 2.

All data sets, including EU and BI composite data sets and individual breed/population data sets had total SNP call rates of \(> 99\%\). The IBS values estimated for Kerry cattle (KY92 and KY12) from filtered genome-wide SNP data are reported in Supplementary Table 1 and described further in section Genomic Relationship and Analysis of Inbreeding.

### Observed Heterozygosity \((H_o)\) Estimated from Genome-Wide SNP Data

Table 2 provides genome-wide \(H_o\) values for each of the breeds/populations used for the present study. The lowest genome-wide \(H_o\) value was observed for the West Africa NDAM \(B. taurus\) breed, which is likely a consequence of ascertainment bias introduced by a focus on polymorphic SNPs in European \(B. taurus\) during design of the Illumina\textsuperscript{®} Bovine SNP50 BeadChip (Matukumalli et al., 2009).

Generally, as shown in Table 2 for the EU and BI breeds and populations, local landrace or heritage breeds display lower \(H_o\) values compared to more widespread production breeds such as the Simmental (SIMM), Holstein (HOLS) or Charolais (CHAR) breeds. In addition, as might be expected, production breeds

<table>
<thead>
<tr>
<th>Breed/population</th>
<th>Code</th>
<th>Data set</th>
<th>Sample size (n)</th>
<th>Sample size (n) post-filtering</th>
<th>Observed heterozygosity (H_o)</th>
<th>No. SNPs (N_e) modelling</th>
<th>No. SNPs genomic inbreeding</th>
</tr>
</thead>
<tbody>
<tr>
<td>Angus</td>
<td>ANGU</td>
<td>BI/EU</td>
<td>72</td>
<td>31,413</td>
<td>0.3048</td>
<td>31,413</td>
<td>39,576</td>
</tr>
<tr>
<td>Belted Galloway</td>
<td>BGAL</td>
<td>BI</td>
<td>4</td>
<td>25,997</td>
<td>0.2902</td>
<td>25,997</td>
<td>39,582</td>
</tr>
<tr>
<td>British Shorthorn</td>
<td>BSHN</td>
<td>BI/EU</td>
<td>10</td>
<td>29,038</td>
<td>0.2549</td>
<td>29,038</td>
<td>39,576</td>
</tr>
<tr>
<td>Brown Swiss</td>
<td>BRSW</td>
<td>EU</td>
<td>31</td>
<td>29,576</td>
<td>0.2804</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Charolais</td>
<td>CHAR</td>
<td>EU</td>
<td>48</td>
<td>29,576</td>
<td>0.3209</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Devon</td>
<td>DEVN</td>
<td>BI</td>
<td>4</td>
<td>27,576</td>
<td>0.2859</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Dexter</td>
<td>DXTR</td>
<td>BI</td>
<td>4</td>
<td>24,753</td>
<td>0.2458</td>
<td>24,753</td>
<td>37,903</td>
</tr>
<tr>
<td>English Longhorn</td>
<td>ELHN</td>
<td>BI</td>
<td>3</td>
<td></td>
<td>0.2232</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Finnish Ayrshire</td>
<td>FAYR</td>
<td>BI/EU</td>
<td>7</td>
<td>29,576</td>
<td>0.3064</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Galloway</td>
<td>GALL</td>
<td>BI</td>
<td>4</td>
<td>24,753</td>
<td>0.2942</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Gelbvieh</td>
<td>GELB</td>
<td>EU</td>
<td>8</td>
<td>24,753</td>
<td>0.3125</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Guernsey</td>
<td>GNNS</td>
<td>BI/EU</td>
<td>19</td>
<td>50,323</td>
<td>0.2764</td>
<td>–</td>
<td>50,323</td>
</tr>
<tr>
<td>Hereford</td>
<td>HRFD</td>
<td>BI/EU</td>
<td>35</td>
<td></td>
<td>0.2964</td>
<td>–</td>
<td>39,535</td>
</tr>
<tr>
<td>Holstein</td>
<td>HOLS</td>
<td>EU</td>
<td>70</td>
<td>39,152</td>
<td>0.3192</td>
<td>36,152</td>
<td>43,135</td>
</tr>
<tr>
<td>Jersey</td>
<td>JPSY</td>
<td>BI/EU</td>
<td>44</td>
<td>31,358</td>
<td>0.2718</td>
<td>31,358</td>
<td>43,136</td>
</tr>
<tr>
<td>Kerry sampled 1991/92</td>
<td>KY92</td>
<td>BI/EU</td>
<td>36</td>
<td>37,556</td>
<td>0.2965</td>
<td>37,556</td>
<td>51,731</td>
</tr>
<tr>
<td>Kerry sampled 2011/12</td>
<td>KY12</td>
<td>BI/EU</td>
<td>62</td>
<td>36,428</td>
<td>0.3042</td>
<td>36,428</td>
<td>50,756</td>
</tr>
<tr>
<td>Limousin</td>
<td>LIMS</td>
<td>EU</td>
<td>45</td>
<td></td>
<td>0.3122</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Lincoln Red</td>
<td>LNCR</td>
<td>BI</td>
<td>7</td>
<td>26,350</td>
<td>0.2789</td>
<td>26,350</td>
<td>34,173</td>
</tr>
<tr>
<td>Montbeliard</td>
<td>MONT</td>
<td>EU</td>
<td>31</td>
<td></td>
<td>0.3019</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>N’Dama</td>
<td>NDAM</td>
<td>–</td>
<td>22</td>
<td>21,516</td>
<td>0.2158</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Norwegian Red</td>
<td>NRED</td>
<td>EU</td>
<td>20</td>
<td></td>
<td>0.3190</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Piedmontese</td>
<td>PDMT</td>
<td>EU</td>
<td>23</td>
<td></td>
<td>0.3240</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Red Angus</td>
<td>RANG</td>
<td>BI/EU</td>
<td>14</td>
<td></td>
<td>0.3092</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Red Poll</td>
<td>REDP</td>
<td>BI</td>
<td>5</td>
<td></td>
<td>0.2905</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Romagnola</td>
<td>ROMG</td>
<td>EU</td>
<td>21</td>
<td></td>
<td>0.2943</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Scottish Highland</td>
<td>SCHL</td>
<td>BI</td>
<td>8</td>
<td></td>
<td>0.2823</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Simmental</td>
<td>SIMM</td>
<td>EU</td>
<td>9</td>
<td></td>
<td>0.3136</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>South Devon</td>
<td>SDEV</td>
<td>BI</td>
<td>3</td>
<td></td>
<td>0.3070</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Sussex</td>
<td>SUSX</td>
<td>BI</td>
<td>4</td>
<td></td>
<td>0.2792</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Welsh Black</td>
<td>WBLK</td>
<td>BI</td>
<td>2</td>
<td></td>
<td>0.3203</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>White Park</td>
<td>WHPK</td>
<td>BI</td>
<td>4</td>
<td></td>
<td>0.2270</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>
originally derived from minor island populations (Jersey [JRSY] and Guernsey [GNSY]) also exhibit relatively low $H_o$ values. In the context of genetic conservation it is therefore encouraging that the KY92 and KY12 population samples display intermediate $H_o$ values that are at the upper end of the range observed for the heritage breeds.

**Maximum Likelihood Phylogenetic Ancestry Graphs Using Genome-Wide SNP Data**

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

For the ML trees generated using the BI data set, breed/population differentiation was less apparent, possibly due to similar biogeographical origins for these breeds and/or smaller sample sizes for some of the populations sampled. Figure 3 shows the ML tree generated with $m = 5$ for the BI data set. For $m = 5$, all migration edges stem from the BSHN/Lincoln Red (LNCR)
branch, including a medium-weight migration edge connecting to the Kerry cattle branch. These results support the hypothesis that there was significant gene flow during the eighteenth and nineteenth centuries from British Shorthorn cattle into the ancestral populations for a range of modern British and Irish cattle breeds (Grobet et al., 1998; Felius et al., 2011, 2015).

**Multivariate Principal Component Analysis of Genome-Wide SNP Data**

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

In Figure 4, for the EU data set, PC1 and PC2 account for 18.2% and 16.8% of the total variation for PC1–10, respectively. The PC1 plot axis differentiates the British Angus (ANGU), Red Angus (RANG) and BSHN and Irish KY92 and KY12 populations from the rest of the European breeds, including the British HRFD and GNSY and JRSY Channel Islands breeds. In addition, the ANGU and RANG and the Kerry (KY92 and KY12) emerge at the opposite extremes of the PC2 plot axis. In Figure 5, for the BI data set, PC1 and PC2 account for 23.7% and 22.7% of the total variation for PC1–10, respectively. The PC1 plot axis recapitulates PC2 in Figure 4 and differentiates the Kerry (KY92 and KY12) from the ANGU and RANG breeds with the other British breeds emerging between these two extremes. These results highlight the genetic distinctiveness of the Kerry cattle breed in comparison to a wide range of British production and heritage landrace cattle breeds and support their status as an important cattle genetic resource that should be prioritised for conservation.

The PC2 plot axis in Figure 5 differentiates the HRFD breed from the other British and Irish breeds and reveals substantial genetic diversity among individual HRFD animals. However, in this context, it is important to note that the pattern of genetic diversity revealed here for the HRFD population sample may be due to ascertainment bias as a consequence of the strategy used to design the Illumina® Bovine SNP50 BeadChip. In this regard, many of the SNPs that constitute this first-generation SNP array were identified from heterozygous positions in the inbred Hereford female (L1 Dominette 01449) bovine genome assembly or through comparisons of random shotgun reads from six diverse cattle breeds that were aligned directly to the same Hereford genome assembly (Matukumalli et al., 2009). This approach to SNP array design will inevitably lead to elevated intrabreed genomic variation using the Illumina® Bovine SNP50 BeadChip with Hereford cattle (Meuwissen, 2009) and accounts for the dispersed pattern of individual HRFD samples in Figure 5.

Examination of Figures 4, 5 indicates that two of the KY12 animals sampled may exhibit a genetic signature of ancestral crossbreeding with another cattle population, which, anecdotally, is likely to have been due to crossbreeding with Angus cattle. Therefore, another PCA plot was generated (Supplementary Figure 1) that shows PC1 and PC2 for individual animals from the KY92, KY12, ANGU and RANG population samples. The two

---

**FIGURE 4** | Principal component analysis plot constructed for PC1 and PC2 from genome-wide SNP data (36,621 autosomal SNPs) for the EU data set of 605 individual animals. The smaller histogram plot shows the relative variance contributions for the first 10 PCs and PC1 and PC2 account for 18.2% and 16.8% of the total variation for PC1–10, respectively.
animals exhibiting a genetic signature of possible crossbreeding (KY12_06 and KY12_58) are indicated on Supplementary Figure 1. Notwithstanding the KY12_06 and KY12_58 data points, the genetic similarity among all Kerry cattle sampled is evident by comparison of the tight KY92 and KY12 sample cluster to the dispersion of the ANGU and RANG samples on the PCA plot in Supplementary Figure 1.

The values for the variation accounted for by PC3, PC4 and PC5 in Figure 4 (EU data set) are relatively high (15.7, 13.8, and 10.4%, respectively). For Figure 5 (BI data set), the variation accounted for by PC3 is also relatively high (19.1%). Therefore, we generated additional PCA plots of PC1 for each of the two data sets vs. these additional principal components (see Supplementary Figures 2–5).

### Analysis of Genetic Structure Using Genome-Wide SNP Data

The results of the fastSTRUCTURE analyses using the EU and BI data sets are shown in Figures 6, 7, respectively. For both analyses, the Kerry cattle (KY92 and KY12) cluster as a single group at $K = 2$ and are differentiated from all other European or British and Irish cattle breeds. The other breed group that is clearly differentiated at $K = 2$ in Figure 7 is the cluster composed of the ANGU and RANG breeds. These results mirror the pattern shown for PC1 in Figure 5, and again emphasise the genetic distinctiveness of Kerry cattle compared to other European production and landrace heritage breeds. Using the `chooseK.py` script the “true” number of clusters corresponding to the likely number of ancestral populations was estimated to be between 12 and 14 for the EU data set and either 7 or 8 for the BI data set.

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

### Modelling Historical Effective Population Size ($N_e$) Using Genome-Wide SNP Data

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

It is important to keep in mind these $N_e$ trends may be partly a consequence of the relatively small sample sizes for the DXTR, BGAL, and LNCR breeds (see Table 2), coupled with different histories of migration, gene flow and, in particular, strong artificial selection in the production cattle populations. Notwithstanding these caveats, the most recent modelled $N_e$ values for the KY92 and KY12 population samples are 89 and 88, respectively. These values are $N_e$ estimates for 12 generations in the past and assuming a generation interval of between 4 and 6 years, which is based on a pedigree estimate from a similar heritage cattle population of 5.66 (Mészáros et al., 2015), this corresponds to between 48 and 72 years before 2012 (for the KY12 population). This is approximately the period between 1940 and 1965, which is during the time that the Kerry breed started to decline precipitously in census population size and also $N_e$ estimated from herd book data (O’Huigin and Cunningham, 1990; Food and Agriculture Organization, 2007).

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

**Genomic Relationship and Analysis of Inbreeding**

Supplementary Table 1 shows a genomic relationship matrix in terms of genotype IBS for the genome-wide SNP data generated for individual animals in the KY92 and KY12 population
samples. Close genomic relationship between individual animals sampled from the same herd is evident in the SNP genotype IBS values between samples. In addition, the relatively low genomic relationship between the KY12_06 and KY12_58 outlier samples (Figures 4–7) and the rest of the Kerry cattle sampled is also evident in Supplementary Table 1. These data emphasise the value of intrapopulation genomic relationship values for identifying animals (in this case, KY12_06 and KY12_58) that should not be used in breeding programmes. They also highlight the potential of genome-wide SNP data for providing a systematic approach to prioritising males and females with minimum genomic relationship for breeding to minimize loss of genetic diversity and maintain or increase $N_e$ (Gandini et al., 2004; Meuwissen, 2009; de Cara et al., 2011, 2013).

Genome-wide autozygosity estimated from SNP data using $F_{ROH}$ and the $F_{ROH}$ statistic are visualised in Figure 9 for individual animals from the KY92 and KY12 population samples and a range of European comparator breeds. Additional summary ROH data is provided in Supplementary Table 3 and also Supplementary Figure 6, which reveals marked inter-population differences in ROH length and demonstrates that the SNP density of the Illumina® Bovine SNP50 BeadChip is too low to reliably capture ROH below 5 Mb in length, an observation previously reported by Purfield et al. (2012).

There is significant variation in $F_{ROH}$ values among individual animals and between breeds and populations. The non-parametric Wilcoxon rank sum test was performed on $F_{ROH}$ distributions for all pairwise population/breed comparisons with application of the Bonferroni correction $P$-value adjustment for multiple statistical tests (Supplementary Table 4). This analysis demonstrated that the KY12 population sample exhibited a significantly higher mean $F_{ROH}$ value than the KY92 population sample (0.098 vs. 0.079; $P_{\text{adjust}} = 0.0081$). This is important from a conservation genetics perspective, indicating that genome-wide autozygosity, which is highly correlated with conventional pedigree-based estimates of inbreeding ($F_{PED}$) for cattle (Purfield et al., 2012; Ferencaković et al., 2013; Martikainen et al., 2017), has increased for the Kerry cattle population in the 20 years between sampling of the KY92 and KY12 populations.
The importance of understanding and quantifying genome-wide autozygosity for genetic conservation purposes has recently been highlighted through correlation of $F_{ROH}$ with inbreeding depression for a range of production traits in domestic cattle (Bjelland et al., 2013; Pryce et al., 2014; Kim et al., 2015). Importantly, $F_{ROH}$ has also been shown to correlate with inbreeding depression for bovine fertility traits in both males (Ferencaković et al., 2017) and females (Kim et al., 2015; Martikainen et al., 2017). Finally, according to basic population genetic principles, recent inbreeding captured by $F_{ROH}$ will lead to recessive deleterious genomic variants emerging at a population level—a phenomenon that has been studied in both humans and cattle (Szpiech et al., 2013; Zhang et al., 2015).

**Genome-Wide Signatures of Selection in the Kerry Cattle Breed**

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

A single gene was located within the BTA9 selection peak—the phosphodiesterase 7B gene ($PDE7B$), which has been associated with neurobiological processes (de Gortari and Mengod, 2010) and has been previously linked to genetic changes associated with dog ($Canis lupus familiaris$) domestication and behaviour (Freedman et al., 2016). A single gene was also located within the BTA16 selection peak—the dorsal inhibitory axon guidance protein gene ($DRAXIN$), which encodes a protein that regulates axon guidance, neural circuit formation and vertebrate brain development (Islam et al., 2009; Shinmyyo et al., 2015). Twenty-four genes were located within the BTA17 selection peak, including $BICDL1$, $RAB35$, and $RNF10$, which have been associated with neurobiology and brain development (Hoshikawa et al., 2008; Schlager et al., 2010; Villarroel-Campos et al., 2016) and $SIRT4$ and $COQ5$ that function in cellular metabolism (Kawamukai, 2015; Elkhwanky and Hakkola, 2017). Six genes were located within the BTA28 selection peak, including, most notably, the Rho GTPase activating protein 22 gene ($ARHGAP22$), which has recently been associated with bovine fertility as an mRNA expression biomarker for oocyte competence in cumulus cells (Melo et al., 2017).

To obtain a broader perspective on natural and artificial selection acting at a population level on the Kerry cattle genome, a functional gene set enrichment approach (GSEA) was taken using IPA with the 178 genes located within $\pm$ 1.0 Mb of each selection peak (Supplementary Table 5). Of these 178 genes, 141 could be mapped to the IPA knowledgebase and the summary results for the IPA **Physiological System Development and Function** category are shown in Supplementary Table 6, revealing an enrichment of biological processes associated with nervous system development and behaviour. This functional enrichment coupled with the neurobiologically relevant single-gene selection peaks on BTA9 ($PDE7B$) and BTA16 ($DRAXIN$) suggests that natural and/or artificial selection related to brain development and behaviour has been important in the microevolution of the Kerry cattle breed. In this regard, it is therefore noteworthy.
that Kerry cattle, including bulls, are recognised as being comparatively docile and easy to manage (Curran, 1990).

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

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

The current genetic status of the Kerry cattle population is underlined by analyses of genetic effective population size (*N_e*) and inbreeding using genome-wide SNP data. As shown in Table 2, genome-wide observed heterozygosity (*H_o*) is relatively high in the KY92 and K12 population samples, particularly for endangered heritage cattle breeds. However, it has been long recognised that monitoring *N_e* is a more important tool for rational breed management and long-term conservation of endangered livestock populations (Notter, 1999; Gandini et al., 2004; Biscarini et al., 2015). As shown in Figure 8 and Supplementary Table 2, the Kerry cattle population has a recent demographic trend of *N_e* decline, to the point where the most recent modelled *N_e* values are below the recommended threshold for sustainable breed management and conservation (Meuwissen, 2009). There is also cause for concern that genomic inbreeding estimated using genome-wide autozygosity (*F_{ROH}* and visualised in Figure 9) has increased significantly in the 20-year period between the sampling of the KY92 and K12 Kerry cattle populations.

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

In conclusion, the results presented here for the Kerry cattle population demonstrate that population genomics analyses of large SNP data sets can provide useful information concerning...
the microevolution and recent genetic history of heritage livestock breeds. In particular, we would recommend that comparable surveys in other populations consider the use of genome-wide scans for signatures of selection, which can provide a functional genomics perspective on evolutionary adaptations to particular agricultural environments and production systems.

DATA ACCESSIBILITY

The Illumina® Bovine SNP50 BeadChip SNP genotype data for the Kerry cattle KY92 and KY12 population samples generated for this study are available from the Dryad Digital Repository: https://doi.org/10.5061/dryad.8fk81.

ETHICS STATEMENT

With the exception of Kerry cattle sampled during 2011–12, all samples and data was obtained from previously published scientific studies. The re-use of these samples and data is consistent with the 3Rs principles on replacement, refinement and reduction of animals in research (www.nc3rs.org.uk/the-3rs). For the 2011–12 Kerry cattle, population owners’ consent to sample DNA for research was obtained and individual owners conducted sampling of animals using non-invasive nasal swabs. In this regard, scientific animal protection in Ireland is subject to European Union Directive 2010/63/EU, which states that the Directive does not apply to “practices not likely to cause pain, suffering, distress or lasting harm equivalent to, or higher than, that caused by the introduction of a needle in accordance with good veterinary practice.”

AUTHOR CONTRIBUTIONS

DEM, DAM, AF, and JK conceived and designed the project; DEM, IWR, DAM, AF, and JK organised sample collection and genotyping; SB, GM, IWR, DAM, SP, CC, IASR, and DEM performed the analyses; SB and DEM wrote the manuscript and all authors reviewed and approved the final manuscript.

FUNDING

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

The authors wish to express their gratitude to the Kerry Cattle Society (www.kerrycastle.ie) for facilitating and encouraging this project. In particular, we thank Rosemary and Jeremy Hill, Matthew English Hayden, Jeanne Hendrick and the Irish Cattle Breeding Federation (ICBF – www.icbf.com) for expert guidance and assistance with animal sourcing and DNA sampling. We are also grateful to Professor Dan Bradley (Smurfit Institute of Genetics, Trinity College Dublin) for access to DNA sample archives. In addition, we thank all of the Kerry cattle owners who provided access to animals, samples and pedigree information. Finally, we thank Weatherbys Scientific for provision of SNP array genotyping services (www.weatherbyscientific.com).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2018.00051/full#supplementary-material

REFERENCES


de Gortari, P., and Mengod, G. (2010). Dopamine D1, D2 and mu-opioid receptors are co-expressed with adenyl cyclase 5 and phosphodiesterase 7B mRNA in striatal rat cells. Brain Res. 1310, 37–45. doi: 10.1016/j.brainres.2009.11.009


Conflict of Interest Statement: The authors IWR and SP are employed by IdentiGEN, Ltd.

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

Copyright © 2018 Browett, McHugo, Richardson, Magee, Park, Fahey, Kearney, Correia, Randhawa and MacHugh. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.