Received: 5 September 2016 | Revised: 16 February 2018 | Accepted: 23 February 2018

DOI: 10.1111/mec.14555

ORIGINAL ARTICLE

WILEY MOLECULAR ECOLOGY

Conservation of a domestic metapopulation structured into related and partly admixed strains

Jelena Ramljak[1,2] | Gjoko Bunevski[3] | Hysen Bytyqi[4] | Božidarka Marković[5] | Muhamed Brka[6] | Ante Ivanković[2] | Kristaq Kume[7] | Srđan Stojanović[8] | Vasil Nikolov[9] | Mojca Simčič[10] | Johann Sölkner[11] | Elisabeth Kunz[1] | Sophie Rothammer[1] | Doris Seichter[12] | Hans-Peter Grünenfelder[13] | Elli T. Broxham[13] | Waltraud Kugler[13] | Ivica Medugorac[1]

[1]Population Genomics Group, Department of Veterinary Sciences, LMU Munich, Munich, Germany

[2]Department of Animal Science and Technology, Faculty of Agriculture, University of Zagreb, Zagreb, Croatia

[3]Faculty of Agricultural Sciences and Food, University Ss. Cyril and Methodius, Skopje, Macedonia

[4]Department of Animal Science, Faculty of Agriculture and Veterinary, University of Prishtina, Prishtina, Kosovo-UNMIC

[5]Department of Livestock Science, Biotechnical Faculty, University of Montenegro, Podgorica, Montenegro

[6]Institute of Animal Sciences, Faculty of Agriculture, University of Sarajevo, Sarajevo, Bosnia and Herzegovina

[7]ALBAGENE Association, Tirana, Albania

[8]Ministry of Agriculture, Forestry and Water Management, Beograd, Serbia

[9]Executive Agency for Selection and Reproduction in Animal Breeding, Sofia, Bulgaria

[10]Department of Animal Science, Biotechnical Faculty, University of Ljubljana, Ljubljana, Slovenia

[11]Division of Livestock Science, Department of Sustainable Agricultural Systems, University of Natural Resources and Life Sciences Vienna, Vienna, Austria

[12]Tierzuchtforschung e.V. München, Poing, Germany

[13]SAVE Foundation, St. Gallen, Switzerland

Correspondence
Ivica Medugorac, Population Genomics Group, Department of Veterinary Sciences, LMU Munich, Munich, Germany.
Email: ivica.medjugorac@gen.vetmed.uni-muenchen.de

Funding information
Food and Agriculture Organization of the United Nations (FAO), Grant/Award Number: PO 299432

Abstract

Preservation of genetic diversity is one of the most pressing challenges in the planetary boundaries concept. Within this context, we focused on genetic diversity in a native, unselected and highly admixed domesticated metapopulation. A set of 1,828 individuals from 60 different cattle breeds was analysed using a medium density SNP chip. Among these breeds, 14 Buša strains formed a metapopulation represented by 350 individuals, while the remaining 46 breeds represented the global cattle population. Genetic analyses showed that the scarcely selected and less differentiated Buša metapopulation contributed a substantial proportion (52.6%) of the neutral allelic diversity to this global taurine population. Consequently, there is an urgent need for synchronized maintenance of this highly fragmented domestic metapopulation, which is distributed over several countries without sophisticated infrastructure and highly endangered by continuous replacement crossing as part of the global genetic homogenization process. This study collected and evaluated samples, data and genomewide information and developed genome-assisted cross-border conservation concepts. To detect and maintain genetic integrity of the metapopulation strains, we designed and applied a composite test that combines six

metrics based on additive genetic relationships, a nearest neighbour graph and the distribution of semiprivate alleles. Each metric provides distinct information components about past admixture events and offers an objective and powerful tool for the detection of admixed outliers. The here developed conservation methods and presented experiences could easily be adapted to comparable conservation programmes of domesticated or other metapopulations bred and kept in captivity or under some other sort of human control.

KEYWORDS
cattle, fine-scale structure, metapopulation conservation, SNP, sustainability

1 | INTRODUCTION

Genetic diversity is one of nine control variables within the planetary boundaries framework (Steffen et al., 2015). According to this concept, genetic diversity captures the role of a basic information bank that "provides the long-term capacity of the biosphere to persist under and adapt to abrupt and gradual abiotic change." This long-term capacity is of fundamental importance to the Earth system; however, it is currently being diminished by human activities and was assigned to the high-risk zone for Earth system functioning (Steffen et al., 2015). Two planetary boundaries at high risk (genetic diversity and biochemical flows) are clearly associated with human agricultural activities. Agriculture thus is the driving force that causes degradation of complex ecosystems (e.g., Burger et al., 2012; Sattari, Bouwman, Martinez Rodríguez, Beusen, & van Ittersum, 2016), while, at the same time, it should provide the long-term capacity to meet the needs for food and energy in the face of continued human population growth and socioeconomic development. In this study, we were interested in genetic and functional diversity as a basic genetic information bank for long-term development. In domestic species, diversity is trimmed by agricultural policies, demands for cheaper food and intensive selection focusing on a narrow range of cosmopolitan breeds with increasing production in unified artificial environments. It is general knowledge that genetic or biodiversity losses are irreversible (e.g., Mace, 2012; Mace et al., 2014). The levels or types of diversity loss that may possibly trigger irreversible changes in the Earth system (Steffen et al., 2015) or that may lead to irreversible losses of functional traits, which are necessary for innovations and persistence in the future, remain, however, unknown. From a long-term perspective, native, undifferentiated and fairly unselected domesticated crop varieties and animal breeds thus play an invaluable role in terms of future functional diversity, that is the capacity to evolve new functions or products and adapt to new environmental conditions.

The genetic diversity of cattle, which are the most important domestic animal species in terms of global output of animal-source foods, is under gradual depletion (FAO, 2015). Similar to other domesticated animal species, this process is primarily due to continuous replacement by crossing well adapted, scarcely selected and heterogeneous local populations with small numbers of highly selected and currently competitive cosmopolitan breeds (e.g., Simčič et al., 2015). This process is directly related to the genetic homogenization of domesticated species (Taberlet et al., 2008) and could be seen as a special form of biotic homogenization, which is considered one of the most prominent forms of the biotic impoverishment worldwide (Olden, LeRoy, Douglas, Douglas, & Fausch, 2004).

Conservation policies concerning genetic diversity within domestic species largely fall under the authority of local governments and focus primarily on the diversity between and within breeds or strains on a local scale. As successful long-term programmes for conservation of animal resources demand for sufficient infrastructure and gross domestic product (GDP) and most programmes so far were restricted to breeds of local interest on the verge of extinction, redundancy across and bias towards countries with relatively high GDP necessarily occurred (FAO, 2015). Consequently, most maintenance programmes focus on already genetically depleted and in part strongly selected breeds and strains. Moreover, the generation of tools, conservation methods, results and experiences is also strongly biased towards the same regions. Several whole-genome studies using SNP arrays have been carried out in African and European taurine cattle populations of which rather small parts were native and unselected (Decker et al., 2014; Flori et al., 2014; Gautier, Laloë, & Moazami-Goudarzi, 2010). However, due to the ascertainment bias introduced in the construction of such arrays (e.g., the BovineSNP50 chip; The Bovine HapMap Consortium, 2009), the multiple benefits of these valuable tools cannot be equivalently used in all populations.

Vilas, Pérez-Figueroa, Quesada, and Caballero (2015) demonstrated through simulation analyses and experimental studies that the high adaptive potential of a population is better indicated by allelic diversity of neutral markers than by expected heterozygosity. Therefore, maximizing the allelic diversity implies higher responses to selection than maximizing heterozygosity. This is in agreement with our previous observations, which were based on a smaller number of breeds and microsatellites as the marker of choice (Medugorac et al., 2009, 2011; Ramljak, Ivanković, Veit-Kensch, Förster, & Medugorac, 2011). There are considerable advantages of high-throughput genotyping assays when compared with multi-allelic microsatellite markers (for some reasons, see Decker et al., 2014;

Flori et al., 2014; Gautier et al., 2010). However, due to the ascertainment bias (The Bovine HapMap Consortium, 2009), the multiple benefits of allelic diversity cannot be used directly. Moreover, various parameters and methods developed for multi-allelic markers by the broad scientific community over the past decades have only a limited value for SNP arrays affected by ascertainment bias.

In this study, we focused on the conservation value of fairly unselected and less differentiated strains of native Buša cattle that originated from South-Eastern European countries (SE Europe) and thus from a geographic region close to the domestication centre. Compared with most Western European cattle breeds, there are remarkable differences in breeding structure as well as in body size, habitus, production traits, longevity and reproduction. There is no single generally accepted definition of the term "breed" (FAO, 2015) but due to their heterogeneous appearance and the lack of controlled mating and phenotype recording in the Buša subpopulations, we prefer to use the term "strains of a metapopulation" rather than "breeds." With regard to body size and exterior, Buša cows with their withers height varying between 95 and 115 cm and a body weight of 200–250 kg (i.e., about half of the body weight of modern European cattle breeds) are in a way comparable to the small short-horned cattle populations that have been reported in SE Europe since the Late Bronze and Iron Ages (Becker, 1986; Sachenbacher-Palavestra, 1986). The high fitness of these small Buša animals, which are well adapted to the local climatic and environmental constraints, is confirmed by their high reproductive ability and longevity in challenging environments (summarized in Text S1).

The objectives of this study were to develop and assess a concept for a genome-assisted conservation programme within a domesticated metapopulation that is characterized by multiple admixture events and a low differentiation level. To reduce ascertainment bias and take advantage of allelic diversity as an indicator of adaptive potential, we used short haplotype blocks (≤150 kb) as multi-allelic markers. To demonstrate the concept of the genome-assisted conservation programme, we focused on phylogenetic network analyses, semiprivate haplotype blocks and unified additive relationships between the members of the metapopulation, as well as on neighbouring breeds that were the donors of foreign haplotypes in the current genetic pool of the metapopulation. Phylogenetic relationships between the remaining breeds were consulted only marginally for an improved discussion of the core objectives. Last but not least, an important objective of this study was to encourage international, solutions-oriented large-scale actions for conservation programmes of endangered, not well-differentiated, local domestic populations that are more prevalent in regions with relatively low GDP and that are continuously replaced by highly selected, unified cosmopolitan breeds.

2 | MATERIALS AND METHODS

2.1 | Animals, breeds and strains

A total of 1,828 individuals from 60 cattle breeds were used in the analyses. Detailed information about the breeds and samples is presented in Table S1. The European *Bos taurus* group was represented by 57 breeds, while the West African *Bos taurus* breed N'Dama and *Bos indicus* (Zebu) cattle (represented by the Gir and Nelore breed) served as outliers for phylogenetic analyses. Furthermore, Anatolian cattle, which originate close to the domestication centre and are known as an admixture of *Bos taurus* and *Bos indicus* (Decker et al., 2014), were also included in the phylogenetic analyses. For reasons of better comparison, European *Bos taurus* animals were pre-assigned to eight groups according to their geographic origin: The (i) Buša ($N = 350$), (ii) Podolian ($N = 75$), (iii) Alpine ($N = 483$), (iv) West France ($N = 233$), (v) North Sea ($N = 264$), (vi) Iberian ($N = 28$), (vii) Channel Island ($N = 68$) and (viii) Great Britain ($N = 211$) groups (Figure 1, Table S1). Samples were collected from animals that were unrelated according to the breeders' statements and, when available, pedigree data.

The metapopulation of 14 Buša strains (i.e., the Buša group) as a core of this study will be described more precisely in the following parts. Additional characteristics of particular strains (origin, type traits and production indicators) can be found in the Supporting material (Text S1). Buša (written as Busha in Kosovo[1] and Albania) is a collective term for small *Bos taurus* cattle that originate from the SE European mountain regions, which include the Dinaric Alps, the Balkans and the Rhodopian Mountains. Comprehensive samplings in Bulgaria (Rhodope Shorthorn, RHS), Macedonia (Macedonian Buša, MKB), Albania (Prespa cattle, PRE; Middle-Albanian cattle, IMB; Lekbibaj cattle, LKB; Dibra cattle, DBB; Skodra cattle, SKB), Kosovo[1] (Dukagjini Busha, DGB; Red Metohian Busha, RMB; Sharri Busha, SHB), Montenegro (Montenegro Buša, MNB), Serbia (Serbian Buša, SRB), Bosnia and Herzegovina (Bosnian and Herzegovinian Buša, BHB) and Croatia (Croatian Buša, HRB) were performed (Text S1). All 14 Buša strains represent small, short-horned, uni-coloured cattle subpopulations with sporadic exchange of animals between geographically close countries with diffuse barriers in the past. With their low body weight and small and hard hooves, they are very well adapted to the karstic and rocky terrain.

2.2 | Genotyping, haplotyping and unified additive relationships (UAR)

All 1,828 animals were genotyped with the Bovine SNP50 BeadChip (iScan SY101-1001, Illumina) using standard procedures (http://www.illumina.com). Table S1 indicates the origin of the samples and genotypes. A total of 479 animals were genotyped in this study, while the remaining whole-genome genotypes originated from other studies. Quality control criteria were applied to remove samples with more than 10% of missing genotypes and SNPs with unknown or X-chromosomal position, call rates <95%, or MAF < 0.025 from further analysis. The 45,454 SNPs that passed the filtering criteria covered 2.509 Gb of the autosomal genome in total.

[1]This designation is without prejudice to positions on status and is in line with UNSCR 1244 and ICJ advisory opinion on the Kosovo declaration of independence.

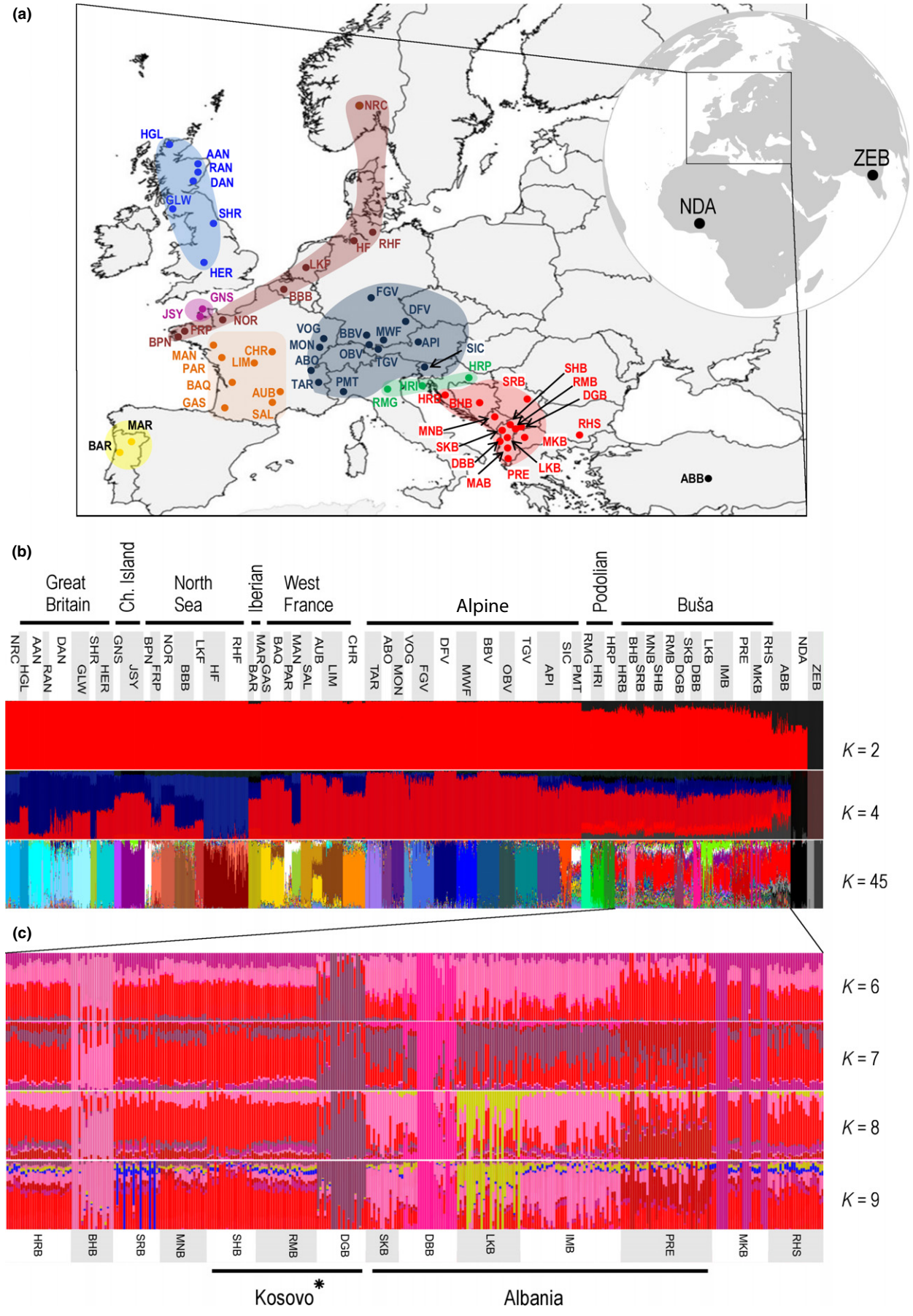
(a)
NRC
HGL
AAN
RAN
DAN
GLW
SHR
HER
HF
RHF
LKP
BBB
FGV
DFV
GNS
JSY
NOR
VOG
BBV
MWF
API
FRP
BPN
MON
OBV
TGV
SIC
CHR
MAN
LIM
ABO
HRP
PAR
TAR
PMT
HRI
SHB
SRB
RMB
BAQ
AUB
RMG
HRB
DGB
GAS
SAL
BHB
MAR
RHS
MNB
MKB
BAR
SKB
DBB
LKB
PRE
MAB
ABB
ZEB
NDA
(b)
Great Britain
Ch. Island
North Sea
Iberian
West France
Alpine
Podolian
Buša
NRC HGL AAN RAN DAN GLW SHR HER GNS JSY BPN FRP NOR BBB LKF HF RHF MAR BAR GAS BAQ PAR MAN SAL AUB LIM CHR TAR ABO MON VOG FGV DFV MWF BBV OBV TGV API SIC PMT RMG HRI HRP HRB BHB SRB MNB SHB RMB DGB SKB DBB LKB IMB PRE MKB RHS ABB NDA ZEB
K = 2
K = 4
K = 45
(c)
K = 6
K = 7
K = 8
K = 9
HRB BHB SRB MNB SHB RMB DGB SKB DBB LKB IMB PRE MKB RHS
Kosovo*
Albania

FIGURE 1 Geographic and genetic structure of the sampled breeds. (a) Geographic origin of the 60 sampled cattle breeds. Zebu (ZEB) and N'Dama (NDA) serve as outliers. Bayesian clustering performed with the ADMIXTURE software on two data sets: (b) complete data set of 60 breeds (N = 1,828) and (c) only 14 Buša strains (N = 350). Individuals (thin vertical bars) are assigned to the different clusters at different K values. The lowest cross-validation errors were cv = 45 (for b) and cv = 6 in the lower panel (c). Breed abbreviations are listed in Table S1. *This designation is without prejudice to positions on status, and is in line with UNSCR 1244 and ICJ advisory opinion on the Kosovo declaration of independence

A Hidden Markov Model implemented in the program BEAGLE version 3.0 (Browning & Browning, 2007) was used to impute missing genotypes and to estimate haplotype phase. Three cohorts were formed that consisted of trios (two parents and one offspring), pairs (one parent and one offspring), and unrelated animals, respectively. For haplotyping, all available animals including those that turned out to be irrelevant for this study were considered. Large half- and full-sib families sampled in various non-Buša breeds had the potential to improve phasing and imputation accuracy in the entire data set (Browning & Browning, 2007) but were otherwise not useful for phylogenetic studies. To infer genomewide relationships between all animal pairs, a unified additive relationship (UAR) matrix was estimated according to Yang et al. (2010). This matrix was used to reduce familial structures within the subpopulations through successive exclusion of highly related animals and also for conservation decisions (see below).

2.3 | Haplotype diversity

The Illumina Bovine SNP50 BeadChip was primarily developed by comparing whole-genome sequence reads representing five *taurine* and one *indicine* breed with the reference genome assembly obtained from a *taurine* Hereford (The Bovine HapMap Consortium, 2009). This led to the ascertainment of SNPs with increased minor allele frequencies within the sequenced breeds. Consequently, a large proportion of the included markers are less informative in breeds that were not considered in the BeadChip development (The Bovine HapMap Consortium, 2009). To reduce ascertainment bias, similar to the study of Simčič et al. (2015), haplotype blocks that spanned less than 150 kb and contained four SNPs with an intermarker distance of less than 50 kb were defined. The use of 4-SNP blocks with a maximum size of less than 150 kb presents a compromise between the maximum number of markers and the minimum recombination probability within the blocks. A total of 5,255 such SNP blocks were considered as multi-allelic markers and their haplotypes as alleles. These multi-allelic markers were used to infer unbiased allelic diversity and heterozygosity.

2.4 | Genetic diversity

Allelic diversity, heterozygosity and population differentiation parameters were estimated by programs tested and used by Simčič et al. (2015). The total number of alleles (nA), mean number of alleles per block (mA), allelic richness (AR; El Mousadik & Petit, 1996), observed (H_O) and expected heterozygosity (H_E) (Nei, 1987) were estimated. For each breed, the number of private alleles (pA; alleles observed only in one population) was counted and their frequency (fpA) was estimated. A large proportion of the included populations was genetically and geographically very close and, consequently, shared an increased proportion of alleles. To take this situation into account, alleles observed in only two populations were defined as semiprivate (spA) and their counts were used as an additional estimator of allelic diversity. Population differentiation was calculated using G_{ST} (Nei, 1973; see implemented in Jost, 2007, 2008) as well as D_{EST} (Jost, 2008), which is not dependent on heterozygosity. In addition to allelic diversity defined in terms of allele counts (nA), we estimated and partitioned total allelic diversity into within- and between-subpopulation components in a way analogous to the classical partition of gene diversity (Caballero & Rodriguez-Ramilo, 2010). The total allelic diversity (A_T) is the sum of the within-subpopulation allelic richness (A_S) and the between-subpopulations allelic differentiation (D_A), that is $A_T = A_S + D_A$. The coefficient of allelic differentiation is defined as $A_{ST} = D_A/A_T$ (see formulas 10 to 14 in Caballero & Rodriguez-Ramilo, 2010).

2.5 | Past effective population size using LD

Past effective population size (*Ne*) was estimated using an approach described in Flury et al. (2010). For this purpose, pairwise r^2 values were estimated with HAPLOVIEW (Barrett, Fry, Maller, & Daly, 2005) after exclusion of SNPs that were monomorphic and had a MAF < 0.1. Subsequently, pairwise r^2 values were grouped over all 29 autosomes in distance bins of 25 kb for each breed. To estimate *Ne* at different time points ($Ne_{(t)}$), the procedure suggested by Hayes, Visscher, McPartlan, and Goddard (2003), which considers the number of generations (t) in the past as $1/2c$, was used. Due to the absence of an accurate genetic map for cattle, the genetic distance c was replaced by distances in mega base pairs (1 Mb = 1 cM). To estimate *Ne* since the foundation of the breeds (~50 generations ago) close to the beginning of the industrial revolution (Felius, 2007) and to observe trends up to the present, 25-kb distance bins between 975 and 10,025 kb were used, thereby considering a cattle generation interval of 5 years. To eliminate downward bias in *Ne* estimated by some fraction of physically linked loci, we adjusted the native *Ne* estimate by the number of autosomes (Chr = 29) in cattle (*Ne* = (native *Ne*)/(0.098 + 0.219 * ln(Chr)); Waples, Larson, & Waples, 2016).

2.6 | Cluster analyses

To infer the population structure of species with complex breeding histories, diffuse barriers between subpopulations and recurrent admixture events, supervised and unsupervised clustering

approaches were carried out. The Nei unbiased D_A-distances (Nei, Tajima, & Tateno, 1983; abbreviated as NeiD_A to omit confusion with D_A defined above) were calculated based on the allele frequencies and presented in a phylogenetic neighbour-net with the program SPLITSTREE4 (Huson & Bryant, 2006). In addition to this supervised approach, unsupervised clustering by individual allele sharing distances, admixture analyses and phylogenetic network analyses was performed.

First, a relationship matrix as the proportion of genomewide shared alleles (PS) was estimated. This similarity matrix, which is based on haplotype blocks to reduce the ascertainment bias, was converted into an allele sharing distance matrix (D_{PS}) according to the method for microsatellites developed by Bowcock et al. (1994). Furthermore, to infer population structure, a heuristic method (Veit-Kensch, Medugorac, Jedrzejewski, Bunevich, & Foerster, 2007) was applied and the D_{PS} matrix was visualized in a two-dimensional graph (2DD). This unsupervised clustering approach (Simčič et al., 2015) aimed to allocate individuals into clusters without prior knowledge of subpopulation or sample origin. Second, to investigate population structure based on SNP genotypes, the program ADMIXTURE 1.23 (Alexander, Novembre, & Lange, 2009), which adopts the likelihood model implemented in the STRUCTURE program (Falush, Stephens, & Pritchard, 2003; Pritchard, Stephens, & Donnely, 2000) but runs considerably faster, was used. To derive the most likely number of populations (K), the cross-validation error (Alexander et al., 2009) was estimated for $K = 2$ to $K = 60$ when the whole data set was included. When only Buša strains were included in the data set, $K = 2$ to $K = 14$ were assumed. The 2DD and ADMIXTURE results were plotted using the R programming language (R Development Core Team 2008). Third, to visualize the network-based individual and population relationships, the approach of Steinig, Neuditschko, Khatkar, Raadsma, and Zenger (2015) implemented in NETVIEW (version 1.1) was used. The distance matrix (D_{PS}) was estimated from 4-SNP block genotypes as described above. The number of mutual k-nearest neighbours (mk-NN) was used to construct a Mutual Nearest Neighbour Graph where each individual (node) is connected by an edge whose weight represents the genetic distance between the nodes (X_i, X_j). K values from 1 to 100 were considered to represent network topologies at different structure scales. Network visualization was performed using the plotting function in NETVIEW and CYTOSCAPE 3.3.0 (Shannon et al., 2003).

2.7 | Conservation priorities for subpopulations

For conservation purposes, coefficients of global allelic differentiation (A_T) (Caballero & Rodriguez-Ramilo, 2010) as well as the relative contribution of each population to a theoretical pool maximizing allelic diversity ($A_{T(pool)}$) were calculated with rarefaction using the program METAPOP version 2.0 (Pérez-Figueroa, Saura, Fernández, Toro, & Caballero, 2009). This program is based on haplotype blocks to reduce the ascertainment bias of the SNP array and infer unbiased allelic diversity.

2.8 | Conservation priorities for animals within the metapopulation

According to the definition of a metapopulation (Akçakaya, Mills, & Doncaster, 2007), individuals of a strain within the metapopulation that are worth preserving should be without or with only low admixture of foreign haplotypes. Identification of admixed individuals in a metapopulation under conservation relies on different patterns of genetic variation caused by introgression of remote haplotypes. Consequently, purebred individuals will show a higher additive genetic relationship to individuals originating from the same metapopulation and a lower relationship to some individuals from remote or foreign populations. The proportion of foreign alleles will increase in admixed individuals. This is reflected in the increased proportion of semiprivate alleles in admixed individuals and the increased genetic distance of admixed to purebred animals of the own strain and the metapopulation. An unsupervised network analysis of the multidimensional genetic distance matrix aimed to detect the most reliable genetic connections between the analysed individuals independent of their origin. An increased number of network connections to individuals of foreign origin pointed to possible admixture.

All above-mentioned patterns could therefore be considered as admixture signatures. If each signature provides partly distinct information components about past admixture events, combining the signals should have greater power for detecting admixed individuals than any single test. We estimated a composite statistic by multivariate outlier analysis (*mvOutlier*; Filzmoser, Garrett, & Reimann, 2005) that used up to six parameters estimated for each animal of the metapopulation (Text S2):

1 The genetic distance to the own population ($D_{UAR(W)}$)
2 The highest UAR with a particular animal of a foreign breed (maxUAR1(*B*))
3 The second highest UAR with a particular animal of a foreign breed (maxUAR2(*B*))
4 The highest average UAR with all animals of foreign breeds (maxUAR(*P*))
5 The number of connections to foreign animals in the Nearest Neighbour Graph (k-NN(*B*))
6 The relative number of semiprivate alleles observed in the particular animal (nspAA)

The applied R code and details related to the implementation of the *mvOutlier* test are summarized in the Supporting material (Text S2). Animals detected as significant multivariate outliers could be considered as significantly admixed and thus less suitable for a conservation breeding programme.

3 | RESULTS

3.1 | Genetic diversity

The predefined Buša group in this study showed the highest average level of genetic diversity with $\overline{nA}$ = 37,157, $\overline{mA}$ = 7.61, $\overline{H_O}$ = 0.732,

$\overline{H_E}$ = 0.726, while the native Iberian, North Sea and Channel Island breeds showed the lowest genetic diversity (Table 1). Differences between breeds were the highest in the number of private and semiprivate alleles. The Buša group possessed the highest average number of private and semiprivate alleles ($\overline{pA}$ = 181, $\overline{spA}$ = 218), while almost only half of them were also found in the seven other predefined groups (proportions ranging from 26% to 50% for pA and 24% to 49% for spA; Table 1).

Within the Buša group, the IMB strain encompassed the majority of pA (290) and spA (381). Likewise, IMB showed a low average frequency of private alleles ($\overline{fpA}$ = 0.013). The lowest values of the genetic diversity parameters in the Buša group, however, were found in the BHB strain (Table 1).

As expected, the maximum allelic richness (AR) was found in the Buša group (6.47), followed by the West France group (5.82), while the other predefined groups had an average AR below 5.62. Taking all individual breeds into account, the lowest AR value was found in JSY (4.52), while the RMB strain showed the highest value (6.90). Allelic richness showed the highest correlation (0.94) with expected haplotype heterozygosity (H_E) and the lowest (0.76) with observed SNP heterozygosity ($H_{O(SNP)}$).

3.2 | Level of genetic differentiation and distances

Pairwise G_{ST} and D_{EST} values were highly correlated (0.998), but D_{EST} values showed a wider range of distribution (0.0013 to 0.2918) compared with G_{ST} (0.0050 to 0.1954) (Table S2). To enable comparison with prior studies and reduce redundancy, only G_{ST} values will be discussed.

The lowest genetic differentiation was observed among the non-selected Buša strains ($\overline{G_{ST}}$ = 0.032), while the highest value was found among the highly selected and isolated Great Britain breeds ($\overline{G_{ST}}$ = 0.114). These values were generally consistent with the NeiD_A−distances (Table S3) and with multidimensional scaling of the D_{PS}−distances (Figure S1), where animals of scarcely selected strains showed higher genetic distances between each other (D_{PS}) and dispersed broadly in two-dimensional space. The lowest NeiD_A-distances occurred between two Angus breeds (AAN and DAN; 0.032) and between two Buša strains (HRB and IMB; 0.033). Taking into account individual D_{PS}−distances, in the case of the Angus breeds, the low distance was caused by the compact and close structure of both breeds. In the case of two Buša strains, both strains showed high but overlapping diversity that resulted into a larger D_{PS}−distance between the animals of the strains but low differentiation ($\overline{G_{ST}}$ = 0.010) and distance ($\overline{\text{Nei}D_A}$ = 0.033) between the strains. The amount of genetic diversity within the Buša group could be demonstrated by the example of the Rhodope Shorthorn cattle. The RHS animals, which had been sampled in three villages over a distance of 135 km, showed only marginally lower allele sharing distances ($\overline{D_{PS}}$ = 1.027) than the eight administratively (by breeding organizations) isolated breeds from the North Sea group that had been sampled along the coast from Brittany to Norway ($\overline{D_{PS}}$ = 1.075) over a distance of more than 2,200 km (Table S3).

3.3 | Assessment of population structure based on the phylogenetic neighbour-net

We used the ZEB outgroup to root the neighbour-net tree (Figure 2). This network separated breeds from seven predefined groups (Figure 1a) into three main clusters.

The first cluster included the Buša and Podolian group as well as the Channel Island group. The resulting network demonstrated the common ancestry and genetic affinity within the Buša group in the form of short branches. The Albanian strains (PRE, LKB, IMB, DBB and SKB) and DGB from Kosovo[1] formed a small cluster that also included the Channel Island breeds (JSY, GNS). This was in accordance with the asymmetric gene flow from Jersey to the Albanian cattle populations (Table S2) and the results based on the individual relationship analysis (Table S4). The homogenous and inbred BHB strain (Table 1) showed the longest branch within the Buša group.

The second cluster of the neighbour-net (Figure 2) included the Alpine and West France groups, and the third cluster included the breeds from the North Sea, Great Britain and Iberian groups. The close relationship and long co-evolution of some breeds were demonstrated by a long, narrow net or a common branch that originated from the same basal node and finally diverged into two or three branches. These patterns could be observed for instance in Holstein [RHF-HF], Braunvieh [OBV-BBV], Angus [RAN-AAN-DAN] and two Iberian breeds [MAR-BAR]. Several breeds with a documented population bottleneck and/or a higher degree of inbreeding (HRP, BHB, JSY, GNS, HGL, SHR) showed characteristic longer branches indicating strong genetic drift or selection pressure (Figure 2).

3.4 | Assessment of population structure using heuristic and model-based methods

A two-dimensional diagram (2DD) that was generated from the D_{PS} distances among 1,828 individuals (Figure S1) demonstrated the existence of two clearly separated clusters consisting of the outlier breeds ZEB and NDA. These two outliers were placed near the Anatolian Black breed (ABB), which is known as an admixture with Zebu (Decker et al., 2014).

The two-dimensional genetic diversity space demonstrated a genetic gradient from Buša strains and Podolian breeds over the Alpine breeds to the West France and North Sea breeds. The breeds from the Great Britain group were genetically located on the opposite side of the Balkan Buša strains. The dispersed and scattered individuals from the Buša metapopulation occupied a large proportion of the diversity space that completely overlapped with the morphologically substantially different Podolian breeds.

Cross-validation of the ADMIXTURE analyses revealed that the most probable number of inferred populations was K = 45 (minimum cross-validation error, cv = 0.534). When the K value was increased from 11 to 45, breeds under strong artificial selection were progressively assigned to different groups with a very high membership proportion (q) (Figure S2). The Buša strains, however, showed low

TABLE 1 Parameters of the genetic diversity in 57 included cattle populations with 45,454 SNPs

Group	Breed	*N*	nA	$\overline{nA}$	mA	$\overline{mA}$	H_O	H_E	pA	$\overline{pA}$	spA	$\overline{spA}$	$\overline{fpA}$	AR	$\overline{AR}$	$H_{O(SNP)}$	$H_{E(SNP)}$	$A_{T(POOL)}$%	$A_{T(GROUP)}$%
Buša	RHS	24	41,245	37,157	7.85	7.61	0.737	0.742	225	181	255	218	0.024	6.78	6.47	0.334	0.337	5.0	52.6
	MKB	24	39,235		7.46		0.719	0.729	200		246		0.029	6.48		0.325	0.330	4.2	
	PRE	39	43,361		8.25		0.749	0.736	218		232		0.018	6.53		0.342	0.337	4.6	
	IMB	43	**46,179**		**8.79**		0.737	0.743	**290**		**381**		**0.013**	6.71		0.338	**0.341**	**7.0**	
	LKB	27	40,416		7.69		0.741	0.727	169		216		0.022	6.48		0.339	0.332	4.0	
	DBB	25	35,513		6.76		0.723	0.703	105		124		0.031	5.91		0.333	0.323	1.9	
	SKB	14	32,968		6.27		0.716	0.708	87		90		0.036	6.22		0.328	0.325	1.9	
	DGB	21	34,316		6.53		0.729	0.702	**71**		108		**0.040**	5.92		0.335	0.321	1.6	
	RMB	26	43,013		8.18		**0.754**	**0.744**	241		288		0.021	**6.90**		**0.343**	0.339	5.6	
	SHB	21	38,410		7.31		0.729	0.726	164		199		0.029	6.53		0.333	0.331	3.6	
	MNB	20	38,208		7.27		0.731	0.729	128		152		0.030	6.56		0.335	0.334	3.1	
	SRB	20	39,541		7.52		0.713	0.737	177		181		0.029	6.78		0.325	0.335	3.9	
	BHB	18	**29,688**		**5.65**		**0.685**	**0.646**	85		**80**		0.028	**5.27**		**0.311**	**0.291**	**1.4**	
	HRB	28	43,250		8.23		0.744	0.742	174		233		0.020	6.84		0.340	0.340	4.8	
Podolian	HRP	24	27,148	32,443	5.17	6.17	0.675	0.636	58	86	75	105	0.021	4.68	5.40	0.306	0.289	0.9	4.6
	HRI	30	37,799		7.19		0.713	0.705	113		141		0.027	5.98		0.325	0.321	2.2	
	RMG	21	30,843		5.87		0.675	0.660	80		89		0.024	5.40		0.306	0.300	1.5	
Alpine	SIC	26	36,118	34,836	6.87	6.63	0.730	0.712	69	69	97	83	0.019	6.00	5.44	0.336	0.328	1.6	12.5
	API	50	39,391		7.49		0.710	0.700	121		157		0.018	5.79		0.328	0.323	2.0	
	TGV	50	35,947		6.84		0.697	0.689	104		94		0.016	5.47		0.320	0.317	0.9	
	OBV	35	34,424		6.55		0.694	0.681	63		76		0.014	5.45		0.318	0.312	0.8	
	BBV	50	32,152		6.12		0.655	0.641	41		55		**0.010**	4.84		0.301	0.295	0.2	
	MWF	46	32,823		6.24		0.711	0.675	54		54		0.011	5.10		0.327	0.311	0.8	
	DFV	50	35,977		6.84		0.693	0.680	59		69		**0.010**	5.41		0.320	0.314	0.1	
	FGV	49	36,545		6.95		0.708	0.694	71		85		**0.010**	5.56		0.327	0.320	1.1	
	VOG	18	32,179		6.12		0.714	0.685	39		57		0.028	5.73		0.327	0.314	0.7	
	MON	28	29,323		5.58		0.688	0.646	23		41		0.018	4.88		0.316	0.297	0.3	
	ABO	22	31,202		5.94		0.703	0.665	36		47		0.023	5.32		0.324	0.305	0.4	
	TAR	37	35,144		6.69		0.690	0.680	75		99		0.014	5.45		0.317	0.312	1.1	
	PMT	22	37,323		7.10		0.733	0.716	107		118		0.029	6.33		0.336	0.328	2.5	

(Continues)

TABLE 1 (Continued)

Group	Breed	*N*	nA	$\overline{nA}$	mA	$\overline{mA}$	H_O	H_E	pA	$\overline{pA}$	spA	$\overline{spA}$	$\overline{fpA}$	AR	$\overline{AR}$	$H_{O(SNP)}$	$H_{E(SNP)}$	$A_{T(POOL)}$%	$A_{T(GROUP)}$%
West France	CHR	50	41,289	36,566	7.86	6.96	0.719	0.718	131	91	157	108	0.015	6.08	5.82	0.334	0.334	2.2	11.5
	LIM	46	38,822		7.39		0.709	0.702	109		115		0.016	5.86		0.327	0.323	1.6	
	AUB	22	33,790		6.43		0.688	0.681	68		76		0.023	5.72		0.315	0.311	1.1	
	SAL	26	33,442		6.36		0.664	0.665	62		77		0.019	5.46		0.304	0.304	0.8	
	MAN	20	32,198		6.13		0.684	0.665	47		62		0.025	5.57		0.320	0.312	0.7	
	PAR	17	33,534		6.38		0.721	0.697	78		82		0.029	6.01		0.332	0.319	1.7	
	BAQ	30	36,573		6.96		0.704	0.697	85		122		0.023	5.85		0.325	0.321	1.8	
	GAS	22	33,883		6.45		0.697	0.683	78		96		0.023	5.74		0.318	0.311	1.6	
Iberian	MAR	14	30,034	29,840	5.71	5.68	0.685	0.672	78	73	85	81	0.036	5.67	5.62	0.313	0.305	1.7	3.2
	BAR	14	29,645		5.64		0.657	0.663	67		77		0.036	5.57		0.301	0.302	1.5	
North Sea	RHF	50	35,913	28,414	6.83	5.41	0.704	0.699	49	55	55	66	**0.010**	5.49	5.48	0.336	0.334	0.7	9.2
	HF	50	34,667		6.60		0.698	0.687	42		68		**0.010**	5.31		0.334	0.329	0.3	
	LKF	20	28,964		5.51		0.675	0.656	29		43		0.025	5.10		0.315	0.306	0.6	
	BBB	45	37,053		7.05		0.697	0.691	100		130		0.026	5.57		0.328	0.325	1.5	
	NOR	30	32,729		6.23		0.693	0.665	88		90		0.017	5.30		0.319	0.305	1.3	
	FRP	22	35,475		6.75		0.729	0.709	72		79		0.023	6.03		0.341	0.332	1.8	
	BPN	15	31,392		5.97		0.722	0.688	75		86		0.033	5.82		0.333	0.317	1.6	
	NRC	32	33,858		6.44		0.700	0.685	81		82		0.016	5.46		0.331	0.323	1.4	
Channel Island	JSY	52	29,548	28,658	5.62		0.620	0.618	51		54		**0.010**	**4.52**	4.59	0.285	0.284	**0.0**	0.7
	GNS	16	25,764		4.90	5.45	0.631	0.622	34	47	48	53	0.031	4.80		0.290	0.287	0.7	
Great Britain	HER	39	32,819	32,479	6.24	6.18	0.664	0.670	79	62	119	78	0.028	5.17	5.2	0.325	0.328	1.5	5.7
	SHR	14	**24,851**		**4.73**		**0.575**	**0.587**	30		28		0.036	4.71		**0.272**	**0.279**	**0.0**	
	GLW	40	32,141		6.12		0.635	0.653	80		91		0.013	5.00		0.296	0.304	1.2	
	DAN	50	39,427		7.50		0.713	0.709	67		96		**0.010**	5.85		0.339	0.337	1.2	
	RAN	14	26,621		5.06		0.667	0.649	**15**		**20**		0.036	5.04		0.320	0.311	0.2	
	AAN	35	31,270		5.95		0.656	0.666	46		43		0.014	5.06		0.315	0.321	0.6	
	HGL	19	26,372		5.02		0.619	0.605	61		58		0.026	4.66		0.287	0.280	1.0	
	All	1,712	68,044		12.95		0.737	0.686	5,249		3,139			5.65		0.318	0.313	100	100

Population abbreviations are listed in Table S1. Minimum and maximum values within the Buša group and when all 57 breeds were taken into account are in bold typeface.

N—number of individuals; nA, $\overline{mA}$—number and mean number of alleles; H_O, H_E,—observed and expected heterozygosity; $\overline{H_O}$ and $\overline{H_E}$—average observed and expected heterozygosity; pA, $\overline{pA}$—number and mean of private alleles; spA, $\overline{spA}$—number and mean of semiprivate alleles; $\overline{fpA}$—frequency of private alleles; AR, $\overline{AR}$—number and mean of allelic richness; $H_{O[SNP]}$, $H_{E[SNP]}$—observed and expected heterozygosity by SNP; $A_{T(POOL)}$, $A_{T(GROUP)}$—contribution of each subpopulation to allelic diversity and its sum over breed group.

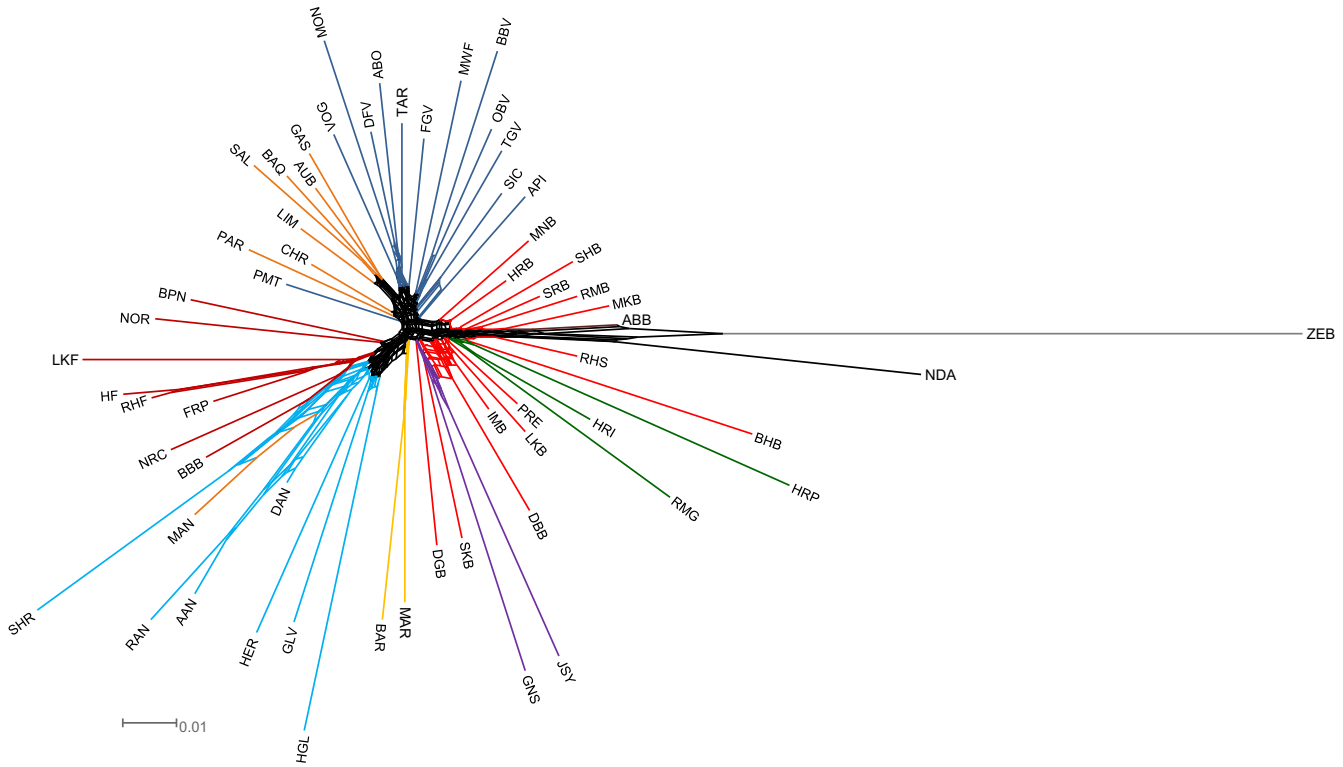

FIGURE 2 Neighbour-network based on pairwise Nei's D_A-genetic distances among 60 breeds with Zebu (ZEB) and N'Dama (NDA) serving as outliers. Population colours and abbreviations match the list in Table S1

differentiation and the highest admixture level even at $K = 45$ (<0.366; Figure 1b).

When the Bayesian procedure took only 14 Buša strains into account, a subdivision into six clusters was obtained. However, this clustering suggested a very low or absent differentiation among the majority of the Buša strains (the lowest cv = 6; Figure 1c). For the most reliable population structure inference, the ADMIXTURE analysis requires independent haplotypes (Alexander et al., 2009). However, here we sampled the last remains of the previously large native Buša cattle population and, in consequence, some of the investigated strains were represented by highly related and inbred animals. This is especially true for BHB and DBB. The population structure results of the Buša metapopulation (Figure 1c) could thus be affected by some familial structures in BHB, DBB and most possibly also in DGB and MKB.

3.5 | Effective population size (*Ne*)

For the time at which the majority of breeds were formed (50 generations, i.e., 250 years ago), a relatively high effective population size (Ne_{50}) was estimated with average values of 701 for the scarcely selected Buša strains and 444 for the highly and moderately selected breeds (Table 2). In the period from 50 to five generations ago, *Ne* slowly decreased in all included breeds. Unfortunately, *Ne* decreased 5.5-fold ($\overline{Ne_{50}} \sim 500$, $\overline{Ne_5} \sim 91$) over the last 45 generations (~225 years). The lowest estimated Ne_5 was determined in BHB (32) and the highest in CHR ($Ne_5 = 204$) (Table 2). Although the Buša strains showed the highest *Ne* 50 generations ago, only five of them (RHS, IMB, HRB, PRE and RMB) had a *Ne* greater than 100 five generations ago. The largest discrepancy became evident in the case of RMB in which *Ne* decreased 9.3-fold in the period from [Ne_{50}-Ne_5] compared to an only 2.2-fold decrease in HER. This resulted in comparable Ne_5 of 118.0 and 125.8 for RMB and HER, respectively (Table 2). Different processes could have been involved in the past reduction of *Ne* in breeds like RMB and HER. While the *Ne* in HER was primarily shaped by intensive selection in the nominally large population, RMB was characterized by fragmentation and nominal reduction of population size in the last generations.

3.6 | Conservation priorities at subpopulation level

As revealed by the METAPOP analyses, the Buša subpopulations contributed the largest proportion of total allelic diversity of all included European cattle ($\overline{A_T} = 0.125$). This is mainly due to the high diversity within ($A_S = 0.097$) and, to a lesser extent, between the subpopulations ($D_A = 0.028$) (Figure 3a). More precisely, four strains (RMB, HRB, RHS and SRB) had the highest allelic diversity ($A_T > 0.130$), while BHB had the lowest one ($A_T = 0.104$). Taking all breeds into account, the lowest allelic diversity was estimated for the two Channel Island subpopulations ($\overline{A_T} = 0.091$).

The METAPOP software was used to simulate a hypothetical genetic pool that consisted of European cattle breeds with maximum allelic diversity. In this analysis, the European *Bos taurus* pool was represented by the 57 breeds and strains (Table 1). The contribution (%) of each subpopulation to this pool was estimated in the form of

TABLE 2 Estimation of effective population size (*Ne*) of 57 cattle breeds from five to 50 generations (Ne_5-Ne_{50}), reduction of *Ne* (Ne_R) expressed as ratio for the period 5–50 generations

Breed	Ne_{50}	Ne_{45}	Ne_{40}	Ne_{35}	Ne_{30}	Ne_{25}	Ne_{20}	Ne_{15}	Ne_{10}	Ne_5	Ne_R
RHS	847.5	780.8	701.2	608.4	534.0	451.5	366.0	281.8	192.9	105.7	8.0
MKB	573.9	523.4	470.2	420.0	372.3	302.5	245.5	189.6	134.3	72.3	7.9
PRE	847.7	788.9	698.3	604.0	538.6	451.4	362.5	287.8	203.3	119.6	7.1
IMB	**1223.0**	**1116.7**	**992.4**	**901.0**	**784.7**	**664.5**	**541.5**	**430.7**	296.6	166.4	7.4
LKB	681.2	624.5	561.3	495.9	436.0	359.8	293.7	225.9	159.3	88.5	7.7
DBB	402.5	367.7	316.8	288.5	250.9	206.8	170.5	132.7	94.5	51.8	7.8
SKB	414.9	379.9	332.0	295.1	257.2	214.3	174.4	132.9	93.0	49.3	8.4
DGB	357.0	320.3	290.6	257.7	226.5	187.4	156.9	118.4	84.0	47.6	7.5
RMB	1095.3	982.5	884.2	788.0	659.1	576.6	457.8	343.1	233.7	118.0	**9.3**
SHB	586.4	532.6	469.4	421.5	365.9	303.2	247.8	192.5	132.1	70.9	8.3
MNB	663.7	621.2	538.4	468.9	405.4	343.8	278.0	212.7	144.5	77.5	8.6
SRB	718.4	676.1	576.9	519.5	445.9	384.5	302.8	236.3	162.9	87.2	8.2
BHB	210.1	197.0	170.9	150.7	134.7	110.4	92.6	72.4	**53.6**	**31.9**	6.6
HRB	1197.1	1058.5	939.8	822.3	731.8	598.7	482.3	377.6	250.5	130.4	9.2
HRP	**195.7**	**177.9**	**160.4**	**142.0**	**125.0**	**106.2**	**90.9**	**71.7**	54.9	34.5	5.7
HRI	506.1	471.7	429.4	374.8	329.5	283.1	230.2	187.0	136.3	81.4	6.2
RMG	300.1	282.8	249.5	230.6	203.2	177.6	149.8	126.8	98.7	64.2	4.7
SIC	594.4	561.5	488.0	432.2	376.3	332.2	269.6	205.3	144.8	81.4	7.3
API	565.0	520.5	468.4	423.5	381.6	325.7	277.7	225.9	175.0	121.1	4.7
TGV	464.9	428.3	390.3	353.3	315.9	269.5	231.1	186.9	141.6	95.3	4.9
OBV	485.2	455.3	405.4	372.8	333.3	294.1	254.9	215.4	163.1	104.7	4.6
BBV	277.2	254.3	232.4	215.5	193.7	172.1	151.4	132.6	110.0	88.6	3.1
MWF	314.2	290.2	263.1	231.0	204.4	173.2	146.9	114.7	86.4	55.4	5.7
DFV	605.4	549.4	532.3	483.7	443.6	393.3	341.8	291.2	224.0	143.7	4.2
FGV	506.4	475.2	422.8	392.7	344.5	309.0	263.9	220.5	169.3	117.1	4.3
VOG	445.6	400.8	361.7	323.8	286.6	245.1	196.2	154.7	108.6	58.8	7.6
MON	277.6	250.9	229.1	202.7	182.5	154.0	128.6	102.0	74.8	49.5	5.6
ABO	372.2	344.2	304.5	276.9	245.5	206.9	168.0	134.5	95.8	56.7	6.6
TAR	495.1	455.6	413.7	386.6	354.0	307.5	257.6	214.8	165.2	103.5	4.8
PMT	773.2	693.1	629.7	557.5	488.3	414.1	338.6	262.6	179.7	96.3	8.0
CHR	849.8	796.5	741.5	676.0	609.0	535.2	478.3	404.6	**318.1**	**204.3**	4.2
LIM	815.2	770.0	716.7	640.3	570.3	513.1	444.3	360.6	269.1	165.5	4.9
AUB	624.7	581.4	527.0	477.1	423.0	360.5	311.1	245.8	174.3	96.8	6.5
SAL	496.5	458.3	438.6	392.0	361.3	315.3	271.9	223.0	176.8	107.3	4.6
MAN	387.5	376.6	343.0	312.5	287.1	251.7	225.5	193.7	143.4	84.8	4.6
PAR	468.0	434.7	379.5	338.0	292.9	248.2	199.4	154.9	109.5	58.6	8.0
BAQ	673.5	633.5	556.0	514.4	452.1	382.4	324.9	253.0	186.4	111.2	6.1
GAS	539.3	505.0	466.3	420.9	375.6	321.9	270.7	220.8	157.2	90.1	6.0
MAR	401.2	372.7	339.1	300.6	274.8	237.9	196.5	165.3	118.4	63.9	6.3
BAR	390.6	358.7	327.4	296.1	258.7	232.1	196.8	161.1	121.5	66.1	5.9
RHF	428.9	400.0	368.3	332.6	303.4	268.2	227.5	197.0	155.5	116.2	3.7
HF	354.6	329.5	303.2	277.2	251.3	221.2	188.1	166.6	134.0	101.9	3.5
LKF	274.1	251.5	228.1	205.3	180.9	154.8	130.0	104.4	79.4	52.5	5.2
BBB	433.1	413.1	365.7	338.1	303.5	264.7	226.6	185.1	144.2	97.8	4.4
NOR	323.5	302.0	275.6	246.8	224.1	189.9	163.6	133.6	99.0	61.0	5.3

(Continues)

TABLE 2 (Continued)

Breed	Ne_{50}	Ne_{45}	Ne_{40}	Ne_{35}	Ne_{30}	Ne_{25}	Ne_{20}	Ne_{15}	Ne_{10}	Ne_{5}	Ne_{R}
FRP	592.2	545.5	486.9	437.6	391.4	333.9	273.3	213.1	151.8	83.7	7.1
BPN	377.0	347.1	301.7	272.8	244.9	204.3	158.4	125.8	87.1	48.6	7.8
NRC	360.5	335.2	307.0	274.2	250.2	213.1	188.4	157.7	122.6	82.7	4.4
JSY	259.5	244.6	234.6	217.9	200.6	180.8	163.0	148.7	130.3	106.5	2.4
GNS	251.1	236.9	214.2	199.3	178.8	155.0	134.3	114.2	87.2	53.7	4.7
HER	279.2	269.6	251.5	235.5	218.9	203.2	184.1	173.7	160.4	125.8	2.2
SHR	221.1	211.5	193.0	184.1	170.9	160.1	145.7	129.5	110.5	65.6	3.4
GLW	357.1	338.7	316.6	292.5	271.4	250.6	228.6	212.0	197.6	150.4	2.4
DAN	568.9	538.0	487.8	457.1	416.8	380.0	341.4	297.8	251.5	169.9	3.4
RAN	248.8	240.4	219.2	200.9	183.0	165.5	140.4	118.1	93.6	55.1	4.5
AAN	304.3	290.3	264.7	249.8	228.5	205.3	188.6	172.3	151.5	114.6	2.7
HGL	228.0	215.8	199.8	185.0	169.3	159.2	142.1	126.9	106.3	76.0	3.0
Mean	500.1	462.8	417.1	375.7	334.1	288.2	242.3	197.2	147.4	90.9	5.7

One generation interval is 5 years. Lines denote predefined group and breed abbreviations as listed in Table S1. Minimum and maximum values are in bold.

the $A_{T(POOL)}$ parameter and is presented in Table 1. The Buša metapopulation contributed more than half (52.6%) to the European *Bos taurus* pool with maximum allelic diversity, while the remaining 43 breeds contributed the rest (47.4%). There were only four breeds (HRI, API, PMT and CHR) that did not belong to the Buša strains with $A_{T(POOL)} > 2\%$ (Table 1). Nevertheless, each of these four breeds contributed less than 2.6%, while the Buša strains contributed 4.1% on average. Therefore, the first ten subpopulations that were prioritized for conservation belonged to the Buša strains, among which IMB had the largest contribution (7%; Table 1). From the other breeds, 24 contributed between 1.0% and 2.5% to $A_{T(POOL)}$ and only JSY and SHR were estimated with a contribution of 0% (Table 1).

Additionally, METAPOP was used to estimate the loss or gain of allelic diversity (A_T) after removal of each respective subpopulation (e.g., due to extinction) from the design. The essential differences in the allelic diversity of the Buša strains and the remaining European cattle breeds are unambiguously demonstrated in Figure 3b. The removal of the Buša strains (with the exception of the inbred BHB) caused a loss of allelic diversity, while the removal of the majority of the remaining European breeds had no effect or even caused a gain of allelic diversity (Figure 3b). A relative increase in overall allelic diversity by removal of a particular subpopulation in a pool is a paralogous feature to allelic richness defined by Petit, El Mousadik, and Pons (1998) and only indicates that the diversity of the respective subpopulation (e.g., 0.099 in JSY and GNS) was much lower than the overall mean (0.110). Therefore, this analysis suggested that the highly differentiated Channel Island breeds as well as most breeds from Great Britain contributed to the allelic diversity rather by variance between breeds (Figure 3b).

3.7 | The network analyses

The network analysis (NETVIEW) using the minimum spanning tree (MST) (Figure 4) largely agreed with the phylogenetic neighbour-net (Figure 2). The NETVIEW analysis misplaced both Jersey and Guernsey in the vicinity of the Buša cluster. The Jersey and Guernsey breeds are geographically (2,500 km) and ecologically well separated from the Buša strains. However, Jersey cattle served as gamete donors in some Albanian and Bulgarian Buša strains. Detailed analyses of the individual relationships (D_{PS} and UAR) and appropriate individual connections in the network (Figure 4b) revealed that a few or even one single admixed animal could lead to a phylogenetically unexpected placement of the whole population or group. These properties were known (Steinig et al., 2015) and made the method implemented in NETVIEW particularly suitable for the detection of familiar and fine-scale rather than large-scale phylogenetic structures. This NETVIEW property was helpful for the detection of admixture signatures in some members of the metapopulation (see below).

3.8 | Conservation priority at individual level

We estimated a matrix consisting of six different admixture signatures for every member of the metapopulation and used it for a multivariate outlier test (Filzmoser et al., 2005). Table S4 summarizes the input matrix (the first seven columns) as well as the results of the outlier test (next 3 columns). We designed an R function (Text S2) to visualize the results of the multivariate outlier test (Figure 5). The significance and intensity of admixture in some animals (Figure 5; Table S4) presented an objective criterion for conservation prioritization within strains. This information combined with the proportion of significantly admixed animals in a particular strain could support objective decisions in a cross-border conservation programme. The most important features of the outlier analysis and the overall roles applicable to any given metapopulation will be explained by two representative excerpts from Table S4.

First, the Middle-Albanian Buša animal IMB11 (Figure 5) was estimated as a significant outlier and, according to the here proposed

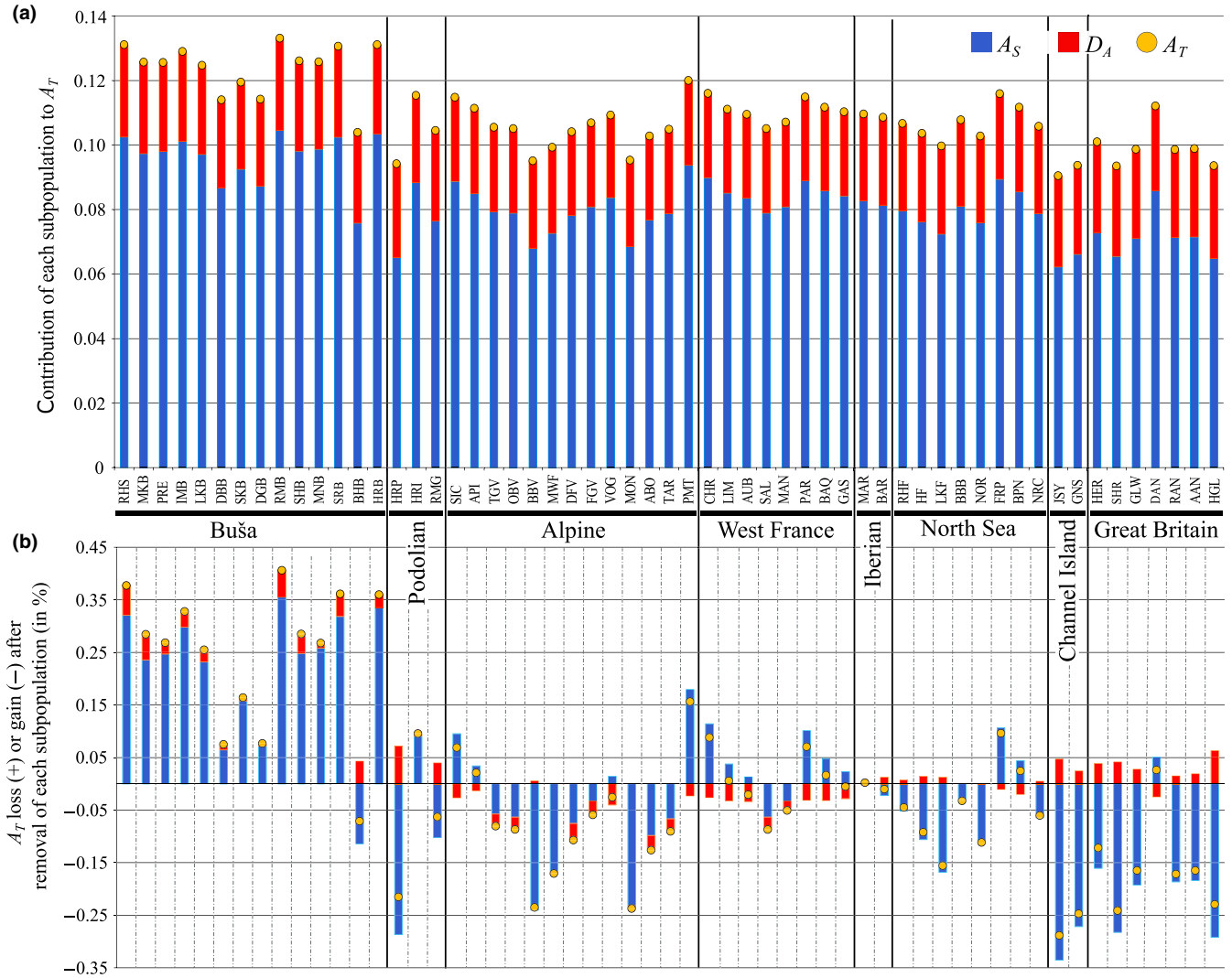

FIGURE 3 The total allelic diversity of taurine cattle breeds. (a) Within- (A_S) and between-subpopulation (D_A) contributions to the total allelic diversity (A_T; yellow circles), (b) Contributions of each subpopulation to the total allelic diversity loss (positive sign) or gain (negative sign) after removal of each respective subpopulation (in %, A_T). The vertical black line indicates the division of breeds (marked with abbreviations) into predefined groups (Table S1)

multivariate analysis, not suitable for a local (within its strain) and cross-border conservation programme. Detailed analyses of the individual relationships (D_{PS} and UAR) and connections in the network (Figure 4b) revealed IMB11 as highly related (UAR > 0.24) to some Jersey animals and also the whole Jersey breed (mean UAR = 0.215). IMB11 showed a high average distance ($D_{UAR(W)}$ = 3.27) to its own metapopulation and was connected with 52 animals outside of the metapopulation. Five of six admixture metrics showed high values (Table S4) and only a relative number of semiprivate alleles observed in IMB11 was not peculiar (nspAA = 0). A large proportion (29%) of Buša animals sampled in Albania (49 of 169) were detected as outliers, most of these cases due to admixture with Jersey cattle. This is consistent with historical data, that is in the 1970s and 1980s, the Albanian government imported semen from Jersey sires for replacement crossings with native Buša in regions with better infrastructure. After approximately 10 years of crossing, the government abandoned this breeding strategy due to the disappointingly low production level of the crossbred animals in their challenging environment.

Second, the most purebred Albanian Buša strain was the Prespa cattle strain. This population was administratively and geographically isolated in Albania (military zone at the border triangle between Albania, Macedonia and Greece) for a longer period of time during the last century. Some Prespa cattle were detected as outliers due to admixture with Alpine breeds, for example Braunvieh and Tyrolian Grey. The animal PRE04, for instance, showed a high distance ($D_{UAR(W)}$ = 3.2) to animals of the metapopulation and was connected with a large number of animals (55) outside of the metapopulation, while the remaining admixture parameters were not peculiar (Table S4). Similarly, most of the non-Albanian Buša animals detected as admixture outliers (e.g., in Croatian, Bosnian and Herzegovinian, Montenegrin, Kosovar[1] and Serbian Buša) showed similar patterns to PRE04.

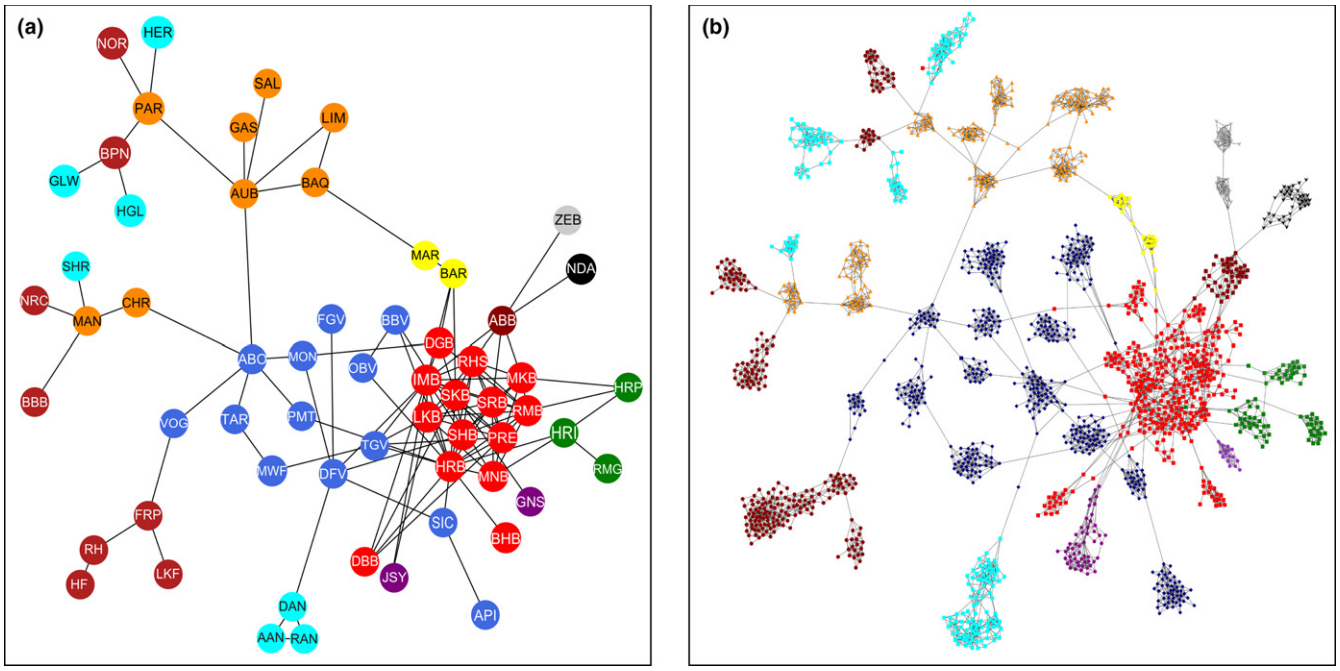

FIGURE 4 Organic network at $k = 10$ with minimum spanning tree (MST) for 60 cattle populations ($N = 1{,}828$) presented at (a) population and (b) individual level. The topology of the population in (b) is the same as in (a); therefore, abbreviations are not subsequently added

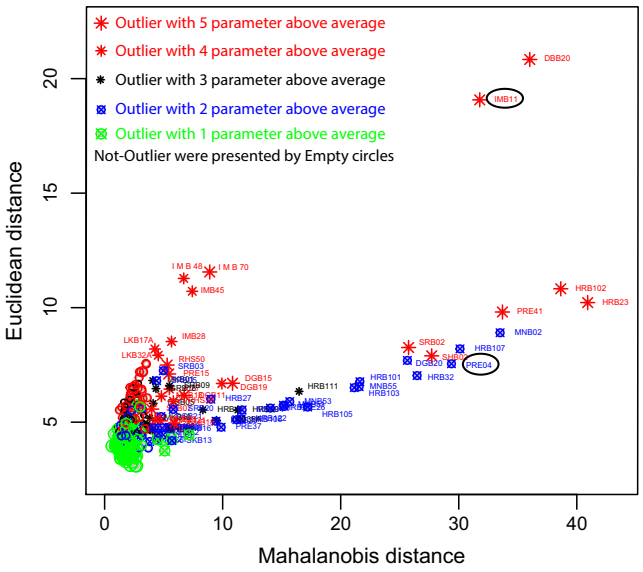

FIGURE 5 Two-dimensional presentation of the multivariate outlier test. The *y*-axis presents the Euclidean and the *x*-axis the robust Mahalanobis distance of the multivariate data (Filzmoser et al., 2005). Each animal is presented by a single symbol. The symbol size, type and colour depend upon the input and output data of the multivariate outlier test defined in the R function Admixture_OutLier_6v (Text S2) and the figure legend. Symbols of animals detected as outliers are marked with the respective Animal-ID. Two animals discussed in the text (IMB11 and PRE04) are emphasized by an ellipse

This is in accordance with historical records (Supporting information) documenting the long-term upgrading of local cattle populations with the grey-brown Alpine cattle breeds.

We performed the outlier test in two steps. In the first step, we used a more stringent threshold (quantile 0.75 and alpha 0.025) and detected 69 outliers. After exclusion of these outliers from the data set, we recalculated all parameters and performed an outlier analysis with the new matrix and a less stringent threshold (quantile 0.5 and alpha 0.05). The second step detected 62 additional animals as outliers. The conservation prioritization based on the second and additional rounds of outlier testing should be accompanied by a differentiated assessment of as complete information as possible, for example information on inbreeding, relationships, local availability of additional animals.

4 | DISCUSSION

The known breed history combined with comprehensive analyses of genetic diversity, differentiation and clustering suggested that the 14 investigated Buša strains were part of a single metapopulation with only diffuse barriers between strains. The genetic overlap with the morphologically different Podolian metapopulation is unexpected and should be investigated by analyses of additional strains (e.g., from Serbia, Hungary, Romania and Ukraine).

In this study, we prioritized the conservation value of breeds or strains by their total allelic diversity (A_T) and the relative contribution of each subpopulation to a theoretical pool of maximum allelic diversity ($A_{T(POOL)}$; Table 1) estimated with rarefaction using the program METAPOP version 2.0 (Pérez-Figueroa et al., 2009). Our previous analyses on field and simulated data sets (Medugorac et al., 2011) clearly underlined the general advantages of the allelic richness method (Caballero & Rodriguez-Ramilo, 2010; Pérez-Figueroa et al., 2009) in

species with high within-population diversity. Vilas et al. (2015) demonstrated that the allelic diversity of neutral markers served as the best indicator of a high adaptive potential and that maximized allelic diversity implied higher responses to selection in an unknown future. This long-term aspect of within-population diversity is primarily considered by methods based on allelic diversity ($\lambda = 1$, where λ = (within-population diversity)/(between-population diversity); Meuwissen, 2009), only partly by core set methods ($\lambda = 0.5$) but neglected by the Weitzman (1992) distance approach ($\lambda = 0$) (Medugorac et al., 2011; Meuwissen, 2009).

The METAPOP results (e.g., Figure 3a,b) recommended particular Buša strains as well as the whole metapopulation as the most important source of allelic diversity in the *Bos taurus* pool represented by 57 European breeds and strains. These results could partly be applied to the global cattle population as a large proportion of the globally widespread *Bos taurus* breeds are derived from European cattle. Of course, a consideration of the most complete data set, that is the inclusion of all breeds and metapopulations (e.g., Podolian), could result in a shift in some of the here estimated proportions. However, the assessment of the global conservation priorities of the Buša strains is strongly facilitated by genomewide phylogeny and diversity gradient analyses across Europe. The prioritization of breeds in a conservation programme is a complex and multifaceted decision-making process, and other factors besides genetics may need to be considered. However, this study relates only to the latter.

The here implemented methods and the use of multi-allelic markers have the potential to reduce the ascertainment bias of the SNP array (Auton et al., 2009; Boyko et al., 2010; Lohmueller, Bustamante, & Clark, 2009; Simčič et al., 2015). Furthermore, they allowed us to demonstrate how to benefit from the allelic diversity of multi-allelic markers as better indicators of future adaptation value (Caballero & Rodriguez-Ramilo, 2010; Vilas et al., 2015). Genotyping by sequencing as well as large-scale whole-genome sequencing could offer a definitive answer regarding the extent of the still remaining bias. This question, however, remains unanswered for the present since whole-genome sequencing initiatives like the 1,000 bull genomes project (Daetwyler et al., 2014; for breed structure in Run4 see, e.g., Kunz et al., 2016) exclusively consider commercial and highly selected breeds. Similar as for microsatellites in 1990s, there could be the need to define a common panel of target regions and thus make the studies based on genotyping by sequencing in different subpopulations comparable.

In accordance with the known replacement crossings with Jerseys (in Albania and Bulgaria) and some Alpine breeds in SE Europe, corresponding proportions of genetic admixture were observed. This admixture could have some impact on the high allelic diversity that was observed in the Buša strains. Nevertheless, even breeds with no (RMB) or very low levels of admixture (MKB, SHB, RHS, MNB and SRB) showed high allelic diversity. Amador, Hayes, and Daetwyler (2012) found that admixture necessarily reduced the number of private alleles in both involved breeds. Moreover, admixture of one breed in two or more strains will also reduce the number of semiprivate alleles. Therefore, the high level of allelic diversity in Buša, whether it was defined by nA, pA, spA, A_T, A_S or D_S, is not a result of admixture but an indicator of their unselected status and their high adaptive potential. Some of the true private alleles of the donor breeds like Jersey, on the other hand, could have been introgressed into Albanian Buša strains and thus were observed as semiprivate or common alleles. This could have led to a lower number of observed private alleles in Jersey. However, admixture did not affect the total number of observed alleles in the donor breed although a significantly lower nA was observed in the most important donor breed, Jersey (29548), compared with the un-admixed Buša strains, for example RMB (43013). The major proportion of the Buša allelic diversity thus is of native origin.

Several aspects of geographic and genetic space overlaps were investigated. First, the alleles of 5,255 short haplotype blocks that were observed in 1,712 animals suggested a decreasing gradient of genetic diversity from SE Europe towards Western and Northern Europe. Such a diversity gradient could be the consequence of the spread of agriculture in Europe by demic diffusion (Sokal, Oden, & Wilson, 1991) or a combination of the demic and the cultural model (Fort, 2012). Unfortunately, testing of this hypothesis is complicated by additional gradients. There is an overlapping gradient of artificial selection and managerial input that could be a more recent anthropogenic source of variation. This gradient also reaches from the scarcely selected animals in SE Europe over the mostly dual-purpose breeds in the Alpine area to the beef or dairy breeds with the highest degree of selection and specialization in Western and Northern Europe. Moreover, a gradient of nonanthropogenic sources like climate or environment could be observed from SE towards NW Europe as well. It could therefore be difficult to discriminate with certainty between the possible sources and reasons for the observed diversity gradient. However, this study suggests that the major proportion of the Buša genetic background is of native origin (see above) and probably reflects the genetic architecture of the similarly small (95–115 cm in withers height) short-horned cattle breeds that have been reported in SE Europe since the Late Bronze and Iron Ages (Becker, 1986; Sachenbacher-Palavestra, 1986). Through joint efforts of local partners in eight countries from SE Europe, samples of the last remains of the scarcely selected, relatively undifferentiated and previously large native Buša cattle population could be obtained. A combination of the here compiled collection of samples with the paleogenomes (e.g., Shapiro & Hofreiter, 2014) of ancient DNA from the Dinaric Alps, the Balkans and the Rhodope Mountains could help to infer the evolutionary processes and the genetic architecture of the most important domesticated species from the time of domestication until the present. This could improve the understanding and future testing of hypotheses related to demic and cultural diffusion in the Neolithic transition (e.g., Fort, 2012) and following eras in Europe.

To estimate the genetic distance to the own metapopulation and to detect animals that were highly related to foreign animals, we estimated the unified additive relationships (UAR) between animals, which are based on identity by descent (IBD) between corresponding

gametes (Powell, Visscher, & Goddard, 2010; Yang et al., 2010) (Table S4). Furthermore, we combined these estimates with the proportion of semiprivate alleles and the number of network connections to foreign animals and thus highlighted animals that were not suitable for a conservation programme. The here designed and applied composite test combines six metrics that each provide distinct information components about past admixture events and offers an objective, statistically valid and powerful tool for the detection of admixed outliers. The composite statistic (Filzmoser et al., 2005) as well as the six single metrics rely on well-established, broadly applied methods (UAR, Yang et al., 2010; Nearest Neighbour Graph, Steinig et al., 2015; and spA, Simčič et al., 2015 and this study) and were additionally confirmed by their consistency with historical records of local strains (Supporting information).

Export of gametes (semen), embryos and breeding animals from developed to developing countries represents the backbone for replacement crossings in local domestic populations. This progressive worldwide process is close to the final stage in many regions and species (FAO, 2015). Still present remains of such native domestic populations are highly fragmented (e.g., Ćurković et al., 2016; Ramljak et al., 2011) and will not be conserved without integrative large-scale action. As a fragmented metapopulation that is distributed across many countries is predetermined for the sequential extinction vortex, we suggest a cross-border conservation concept with limited exchange of animals or gametes chosen by an objective and transparent method. The here presented methods prioritize the conservation value of animals and gametes estimated as a most probably pure breed, that is not admixed with foreign, omnipresent breeds. These nonadmixed animals were suggested as optimal exchange candidates for a cooperative cross-border conservation programme and accepted as such by local partners. The list of simple decision rules (Text S3) could assist metapopulation-based conservation programmes for local domesticated strains. To ensure the integrity of the host population and to preserve the total diversity of the metapopulation itself, the exchange and reproduction of immigrant animals or gametes should be restricted and well monitored within the recipient strain (e.g., only one migrant per generation; Medugorac et al., 2011; Toro & Caballero, 2005). Taking a look beyond domestic species, there are various highly fragmented and partly admixed species (Medugorac et al., 2017) kept in numerous zoos across the globe, for example Bantengs and yaks. Therefore, the elaborated genome-assisted cross-border conservation concepts developed in this study could be easily customized to the needs of comparable conservation programmes for domesticated or other metapopulations bred and kept in captivity or under some other sort of human control.

ACKNOWLEDGEMENTS

Parts of the samples were collected and financed by the BUSHALIVE project (FAO, PO 299432). We thank all breeders and breeding organizations that provided samples. Genotyping of some reference populations and samples that were considered for reliable haplotyping but not directly included in this study was financially supported by the Tierzuchtforschung e.V. München. We thank Dr. M. Gautier and Dr. J. Decker who provided SNP genotypes of some valuable reference populations as indicated in Table S1. We thank Dr V. Cadavez who provided samples of two Iberian breeds.

DATA ACCESSIBILITY

SNP genotype data of 1,210 animals were deposited at DRYAD (https://doi.org/10.5061/dryad.q00k37c). SNP genotypes of 618 animals used in this study were produced by other authors (Table S1) and are accessible via Decker et al. (2014; https://doi.org/10.5061/dryad.th092) and/or Gautier et al. (2010; upon request, mathieu.gautier@inra.fr).

AUTHOR CONTRIBUTIONS

J.R. analysed data, wrote the manuscript, contributed samples and genotypes. G.B., H.B., B.M., M.B., A.I., K.K., S.S., V.N., H.P.G. and E.B. wrote the manuscript, contributed samples and genotypes. M.S., J.S. contributed samples and genotypes. E.K. and S.R. analysed data, wrote the manuscript. D.S. analysed data. W.K. designed the study. I.M. designed study, performed research, analysed data, contributed samples and genotypes and wrote the manuscript.

ORCID

Ivica Medugorac http://orcid.org/0000-0002-6891-3039

REFERENCES

Akçakaya, H. R., Mills, G., & Doncaster, C. P. (2007). The role of metapopulations in conservation. In D. W. Macdonald, & K. Service (Eds.), *Key topics in conservation biology* (pp. 64–84). Oxford: Blackwell Publishing.

Alexander, D. H., Novembre, J., & Lange, K. (2009). Fast model–based estimation of ancestry in unrelated animals. *Genome Research*, *19*, 1655–1664. https://doi.org/10.1101/gr.094052.109

Amador, C., Hayes, B. J., & Daetwyler, H. D. (2012). Genomic selection for recovery of original genetic background from hybrids of endangered and common breeds. *Evolutionary Applications*, *7*, 227–237.

Auton, A., Bryc, K., Boyko, A. R., Lohmueller, K. E., Novembre, J., Reynolds, A., … King, K. S. (2009). Global distribution of genomic diversity underscores rich complex history of continental human populations. *Genome Research*, *19*, 795–803. https://doi.org/10.1101/gr.088898.108

Barrett, J. C., Fry, B., Maller, J., & Daly, M. J. (2005). Haploview: Analysis and visualization of LD and haplotype maps. *Bioinformatics*, *21*, 263–265. https://doi.org/10.1093/bioinformatics/bth457

Becker, C. K. (1986). Ausgrabungen in einem Siedlungshügel der Bronze- und Eisenzeit Makedoniens 1975-1979. Die Tierknochenfunde. Prähistorische Archäologie in Südosteuropa 5. V. Spiess, Berlin.

Bowcock, A. M., Ruiz-Linares, A., Tomfohrde, J., Minch, E., Kidd, J. R., & Cavalli-Sforza, L. L. (1994). High resolution of human evolutionary trees with polymorphic microsatellites. *Nature*, *368*, 455–457. https://doi.org/10.1038/368455a0

Boyko, A. R., Quignon, P., Li, L., Schoenebeck, J. J., Degenhardt, J. D., Lohmueller, K. E., … Auton, A. (2010). A simple genetic architecture

underlines morphological variation in dogs. *PLoS Biology*, *8*, e1000451.

Browning, S., & Browning, B. (2007). Rapid and accurate haplotype phasing and missing–data inference for whole–genome association studies by use of localized haplotype clustering. *The American Journal of Human Genetics*, *81*, 1084–1097. https://doi.org/10.1086/521987

Burger, J. R., Allen, C. D., Brown, J. H., Burnside, W. R., Davidson, A. D., Fristoe, T. S., … Zuo, W. (2012). The Macroecology of Sustainability. *PLoS Biology*, *10*, e1001345. https://doi.org/10.1371/journal.pbio.1001345

Caballero, A., & Rodriguez-Ramilo, S. (2010). A new method for the partition of allelic diversity within and between subpopulations. *Conservation Genetics*, *11*, 2219–2229. https://doi.org/10.1007/s10592-010-0107-7

Ćurković, M., Ramljak, J., Ivanković, S., Mioč, B., Ivanković, A., Pavić, V., … Medugorac, I. (2016). The genetic diversity and structure of 18 sheep breeds exposed to isolation and selection. *Journal of Animal Breeding and Genetics*, *133*, 71–80.

Daetwyler, H. D., Capitan, A., Pausch, H., Stothard, P., Van Binsbergen, R., Brøndum, R. F., … Esquerré, D. (2014). Whole–genome sequencing of 234 bulls facilitates mapping of monogenic and complex traits in cattle. *Nature Genetics*, *46*, 858–865. https://doi.org/10.1038/ng.3034

Decker, J. E., McKay, S. D., Rolf, M. M., Kim, J., Alcalá, A. M., Sonstegard, T. S., … Babar, M. E. (2014). Worldwide patterns of ancestry, divergence, and admixture in domesticated cattle. *PLoS Genetics*, *10*, e1004254. https://doi.org/10.1371/journal.pgen.1004254

El Mousadik, A., & Petit, R. J. (1996). High level of genetic differentiation for allelic richness among populations of the argan tree [*Argania spinosa* (L.) Skeels] endemic to Morocco. *Theoretical and Applied Genetics*, *92*, 832–839. https://doi.org/10.1007/BF00221895

Falush, D., Stephens, M., & Pritchard, J. K. (2003). Inference of population structure using multilocus genotype data: Linked loci and correlated allele frequencies. *Genetics*, *164*, 1567–1587.

FAO. (2015). *The second report on the state of the world's Animal genetic resources for food and agriculture*. Rome: FAO Commission on Genetic Resources for Food and Agriculture Assessments. Available at http://www.fao.org/3/a-i4787e.pdf. Accessed 10 February 2018.

Felius, M. (2007). *Cattle breeds: An Encyclopedia*. Misset: Doetinchem.

Filzmoser, P., Garrett, R. G., & Reimann, C. (2005). Multivariate outlier detection in exploration geochemistry. *Computers & Geosciences*, *31*, 579–587. https://doi.org/10.1016/j.cageo.2004.11.013

Flori, L., Thevenon, S., Dayo, G. K., Senou, M., Sylla, S., Berthier, D., … Gautier, M. (2014). Adaptive admixture in the West African bovine hybrid zone: Insight from the Borgou population. *Molecular Ecology*, *23*, 3241–3257. https://doi.org/10.1111/mec.12816

Flury, C., Tapio, M., Sonstegard, T., Drögemüller, C., Leeb, T., Simianer, H., … Rieder, S. (2010). Effective population size of an indigenous Swiss cattle breed estimated from linkage disequilibrium. *Journal of Animal Breeding and Genetics*, *127*, 339–347. https://doi.org/10.1111/j.1439-0388.2010.00862.x

Fort, J. (2012). Synthesis between demic and cultural diffusion in the Neolithic transition in Europe. *Proceedings of the National Academy of Sciences of the United States of America*, *109*, 18669–18673. https://doi.org/10.1073/pnas.1200662109

Gautier, M., Laloë, D., & Moazami-Goudarzi, K. (2010). Insights into the genetic history of french cattle from dense SNP data on 47 worldwide breeds. *PLoS ONE*, *5*, e13038. https://doi.org/10.1371/journal.pone.0013038

Hayes, B., Visscher, P. M., McPartlan, H. C., & Goddard, M. E. (2003). Bovel multilocus measure of linkage disequilibrium to estimate past effective population size. *Genome Research*, *13*, 635–643. https://doi.org/10.1101/gr.387103

Huson, D. H., & Bryant, D. (2006). Application of phylogenetic networks in evolutionary studies. *Molecular Biology Evolution*, *23*, 254–267. https://doi.org/10.1093/molbev/msj030

Jost, L. (2007). Partitioning diversity into independent alpha and beta components. *Ecology*, *88*, 2427–2439. https://doi.org/10.1890/06-1736.1

Jost, L. (2008). G_{ST} and its relatives do not measure differentiation. *Molecular Ecology*, *18*, 4015–4026. https://doi.org/10.1111/j.1365-294X.2008.03887.x

Kunz, E., Rothammer, S., Pausch, H., Schwarzenbacher, H., Seefried, F. R., Matiasek, K., … Medugorac, I. (2016). Confirmation of a non-synonymous SNP in PNPLA8 as a candidate causal mutation for Weaver syndrome in Brown Swiss cattle. *Genetic Selection Evolution*, *48*, 21. https://doi.org/10.1186/s12711-016-0201-5

Lohmueller, K. E., Bustamante, C. D., & Clark, A. G. (2009). Methods for human demographic inference using haplotype patterns from genome wide single–nucleotide polymorphism data. *Genetics*, *182*, 217–231. https://doi.org/10.1534/genetics.108.099275

Mace, G. M. (2012). The limits to sustainability science: Ecological constraints or endless innovation? *PLoS Biology*, *10*, e1001343. https://doi.org/10.1371/journal.pbio.1001343

Mace, G. M., Reyers, B., Alkemade, R., Biggs, R., Chapin, F. S. III, Cornell, S. E., … Purvis, A. (2014). Approaches to defining a planetary boundary for biodiversity. *Global Environmental Change-Human and Policy Dimensions*, *28*, 289–297. https://doi.org/10.1016/j.gloenvcha.2014.07.009

Medugorac, I., Graf, A., Grohs, C., Rothammer, S., Zagdsuren, Y., Gladyr, E., … Eggen, A. (2017). Whole-genome analysis of introgressive hybridization and characterization of the bovine legacy of Mongolian yaks. *Nature Genetics*, *49*, 470–475. https://doi.org/10.1038/ng.3775

Medugorac, I., Medugorac, A., Russ, I., Veit-Kensch, C. E., Taberlet, P., Luntz, B., … Foerster, M. (2009). Genetic diversity of European cattle breeds highlights the conservation value of traditional unselected breeds with high effective population size. *Molecular Ecology*, *18*, 3394–3410. https://doi.org/10.1111/j.1365-294X.2009.04286.x

Medugorac, I., Veit-Kensch, C. E., Ramljak, J., Brka, M., Marković, B., Stojanović, S., … Bennewitz, J. (2011). Conservation priorities of genetic diversity in domesticated metapopulations: A study in taurine cattle breeds. *Ecology and Evolution*, *1*, 408–420. https://doi.org/10.1002/ece3.39

Meuwissen, T. (2009). Towards consensus on how to measure neutral genetic diversity?. *Journal of Animal Breeding and Genetics*, *126*, 333–334.

Nei, M. (1973). Analysis of gene diversity in subdivided populations. *Proceedings of the National Academy of Sciences of the USA*, *70*, 3321–3323. https://doi.org/10.1073/pnas.70.12.3321

Nei, M. (1987). *Molecular evolutionary genetics*. New York: Columbia University Press.

Nei, M., Tajima, F., & Tateno, Y. (1983). Accuracy of estimated phylogenetic trees from molecular data. *Journal of Molecular Evolution*, *19*, 153–170. https://doi.org/10.1007/BF02300753

Olden, J. D., LeRoy, Poff. N., Douglas, M. R., Douglas, M. E., & Fausch, K. D. (2004). Ecological and evolutionary consequences of biotic homogenization. *Trends in Ecology & Evolution*, *19*, 18–24. https://doi.org/10.1016/j.tree.2003.09.010

Pérez-Figueroa, A., Saura, M., Fernández, J., Toro, M., & Caballero, A. (2009). METAPOP—A software for the management and analysis of subdivided populations in conservation programs. *Conservation Genetics*, *10*, 1097–1099. https://doi.org/10.1007/s10592-008-9718-7

Petit, R. J., El Mousadik, A., & Pons, O. (1998). Identifying populations for conservation on the basis of genetic markers. *Conservation Biology*, *12*, 844–855. https://doi.org/10.1046/j.1523-1739.1998.96489.x

Powell, J., Visscher, P., & Goddard, M. (2010). Reconciling the analysis of IBD and IBS in complex trait studies. *Nature Reviews. Genetics*, *11*, 800–805. https://doi.org/10.1038/nrg2865

Pritchard, J. K., Stephens, M., & Donnely, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, *155*, 945–959.

R Development Core Team. (2008). R: A language and environment for statistical computing. Available at https://www.r-project.org. Accessed 10 February 2018.

Ramljak, J., Ivanković, A., Veit-Kensch, C., Förster, M., & Medugorac, I. (2011). Analysis of genetic and cultural conservation value of three indigenous Croatian cattle breeds in a local and global context. *Journal of Animal Breeding and Genetics*, *128*, 73–84. https://doi.org/10.1111/j.1439-0388.2010.00905.x

Sachenbacher-Palavestra, M. (1986). Tierknochenfunde aus Pod bei Bugojno, einer befestigten Siedlung der Bronze- und Eisenzeit in Zentralbosnien. Diss. med. vet., München.

Sattari, S. Z., Bouwman, A. F., Martinez Rodríguez, R., Beusen, A. H. W., & van Ittersum, M. K. (2016). Negative global phosphorus budgets challenge sustainable intensification of grasslands. *Nature Communications*, *7*, 10696. https://doi.org/10.1038/ncomms10696

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., … Ideker, T. (2003). Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Research*, *13*, 2498–2504. https://doi.org/10.1101/gr.1239303

Shapiro, B., & Hofreiter, M. (2014). A paleogenomic perspective on evolution and gene function: New insights from ancient DNA. *Science*, *343*, 1236573. https://doi.org/10.1126/science.1236573

Simčič, M., Smetko, A., Sölkner, J., Seichter, D., Gorjanc, G., Kompan, D., & Medugorac, I. (2015). Recovery of native genetic background in admixed populations using haplotypes, phenotypes, and pedigree information—using Cika Cattle as a case breed. *PLoS ONE*, *10*, e0123253.

Sokal, R. R., Oden, N. L., & Wilson, C. (1991). Genetic-evidence for the spread of agriculture in Europe by demic diffusion. *Nature*, *351*, 143–145. https://doi.org/10.1038/351143a0

Steffen, W., Richardson, K., Rockström, J., Cornell, S. E., Fetzer, I., Bennett, E. M., … Folke, C. (2015). Planetary boundaries: Guiding human development on a changing planet. *Science*, *347*, 1259855. https://doi.org/10.1126/science.1259855

Steinig, E. J., Neuditschko, M., Khatkar, M. S., Raadsma, H. W., & Zenger, K. R. (2015). NETVIEWP: A network visualization tool to unravel complex population structure using genome–wide SNPs. *Molecular Ecology Resources*, *16*, 216–227.

Taberlet, P., Valentini, A., Rezaei, H. R., Naderi, S., Pompanon, F., Negrini, R., & Ajmone-Marsan, P. (2008). Are cattle, sheep, and goats endangered species? *Molecular Ecology*, *17*, 275–284. https://doi.org/10.1111/j.1365-294X.2007.03475.x

The Bovine HapMap Consortium. (2009). Genome–wide survey of SNP variation uncovers the genetic structure of cattle breeds. *Science*, *324*, 528–532.

Toro, M. A., & Caballero, A. (2005). Characterization and conservation of genetic diversity in subdivided populations. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *360*, 1367–1378. https://doi.org/10.1098/rstb.2005.1680

Veit-Kensch, C., Medugorac, I., Jedrzejewski, W., Bunevich, A., & Foerster, M. (2007). A heuristic two–dimensional presentation of microsatellite–based data applied to dogs and wolves. *Genetics Selection Evolution*, *39*, 447. https://doi.org/10.1186/1297-9686-39-4-447

Vilas, A., Pérez-Figueroa, A., Quesada, H., & Caballero, A. (2015). Allelic diversity for neutral markers retains a higher adaptive potential for quantitative traits than expected heterozygosity. *Molecular Ecology*, *24*, 4419–4432. https://doi.org/10.1111/mec.13334

Waples, R. K., Larson, W. A., & Waples, R. S. (2016). Estimating contemporary effective population size in non-model species using linkage disequilibrium across thousands of loci. *Heredity*, *117*, 233–240. https://doi.org/10.1038/hdy.2016.60

Weitzman, M. L. (1992). On diversity. *Quarterly Journal of Economics*, *107*, 363–405.

Yang, J., Benyamin, B., McEvoy, B. P., Gordon, S., Henders, A. K., Nyholt, D. R., … Goddard, M. E. (2010). Common SNPs explain a large proportion of the heritability for human height. *Nature Genetics*, *42*, 565. https://doi.org/10.1038/ng.608

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: Ramljak J, Bunevski G, Bytyqi H, et al. Conservation of a domestic metapopulation structured into related and partly admixed strains. *Mol Ecol*. 2018;00:1–18. https://doi.org/10.1111/mec.14555