Mitogenome information in cattle breeding and conservation genetics: Developments and possibilities of the SNP chip

V. Brajkovic [a,1,*], D. Hršak [b,1,*], L. Bradić [a], K. Turkalj [a], D. Novosel [c], S. Ristov [b], P. Ajmone-Marsan [d], L. Colli [e], V. Cubric-Curik [a], J. Sölkner [f], I. Curik [a,1,*]

[a] University of Zagreb, Faculty of Agriculture, Svetošimunska 25, 10000, Zagreb, Croatia
[b] Ruđer Bošković Institute, Bijenička 54, 10000, Zagreb, Croatia
[c] Croatian Veterinary Institute, Savska cesta 143, 10000, Zagreb, Croatia
[d] Universita Cattolica del Sacro Cuore, Dipartimento di Scienze Animali, della Nutrizione e degli Alimenti, via Emilia Parmense 84, 29122, Piacenza (PC), Italy
[e] Universita Cattolica del Sacro Cuore, BioDNA Centro di ricerca sulla Biodiversità e sul DNA Antico, via Emilia Parmense 84, 29122, Piacenza (PC), Italy
[f] University of Natural Resources & Life Sciences Vienna, Gregor-Mendel-Strasse 33, 1180, Vienna, Austria

HIGHLIGHTS

- Mitogenome SNPs are useful in animal breeding and conservation genetics.
- Mitogenome 310 SNPs are integrated into the commercial GGP Bovine 100K SNP Chip.
- Magellan v2.0 enables automated and efficient use of mitogenome information.
- Selected 310 SNPs with the Magellan v2.0 algorithm enable mitogenome classification.
- Mitogenome variations can have significant effects on cattle phenotypes (disorders).

ARTICLE INFO

ABSTRACT

In contrast to nuclear markers routinely used for genomic selection, mitogenome information has been underutilized for breeding and biodiversity management of cattle populations. Our main goal was to promote the efficient use of mitogenome SNPs contained in commercial high-throughput SNP arrays. In collaboration with NEOGEN Genomics (Lincoln, NE, USA), we integrated 310 SNPs into the commercial GGP Bovine 100K SNP array. In doing so, we demonstrated how mitogenome SNPs can be used in high-throughput arrays to (i) analyze population structure and diversity, (ii) classify bovine haplogroups and identify introgression and/or upgrading, (iii) screen and identify pedigree defects, (iv) impute mitogenome information on maternal lineages to increase statistical power in estimating the effects of mitogenome variation on quantitative production traits, and (v) identify deleterious mutations found in humans. In addition, we have developed protocols and pipelines integrated with the Magellan v2.0 software to enable efficient and routine use of mitogenome information in cattle breeding and genetic diversity management. Finally, we have highlighted some other interesting opportunities for the use of mitogenome information in the near future.

1. Introduction

Cattle were domesticated about 12,000 years ago when people recognized their ability to convert large amounts of roughage into high quality food, as well as other potential benefits (draft power, transportation, leather, manure as fertilizer, etc.) that were substantial (Sherratt, 1983). Today, the human population is growing exponentially and is projected to reach 10 billion by 2050 (FAO, 2015), requiring more efficient livestock production to meet the high demand for food with high nutritional value. Over the past 100 years, great success has already been achieved in increasing the productivity of cattle. For example, at the beginning of the twentieth century, the average milk yield of a dairy

---

cow was 2.6 to 3.2 tons per lactation (305 days), whereas today the best dairy cows can give more than 14.3 tons per lactation, with average production systems in Canada, the United Kingdom, and the United States yielding 8.0 to 11.0 tons per lactation (Britt et al., 2018; Britt et al., 2021).

Much of the high-performance success in modern cattle breeding has been achieved through the classical quantitative breeding approach based on pedigree and phenotype information but has been accelerated in recent decades by the application of genomic selection, even many times over for some traits (Boichard et al., 2015; Weigel et al., 2017; Cole and VanRaden, 2018). However, the approach used is based on the breeding value concept and additive polygenic inheritance of autosomal genes in the nucleus following Mendel's segregation, while neglecting the effects of other types of inheritance, such as those of the mitochondrial genome and sex chromosomes.

On the other hand, there are several hundred to several thousand mitochondria in each cell that play a critical role in the general functional aspects of mammalian cells, and their dysfunction has been associated with various metabolic phenotypes related to bioenergy (Wallace 2018). This is not surprising, because although their unit of inheritance (mitogenome) is a small circular and double-stranded molecule of approximately 16- to 17-thousand nucleotides in length, mitochondria are responsible for many cellular processes, beyond energy metabolism, such as signal transduction by reactive oxygen species (Li et al., 2013), regulation of membrane potentials (Vasan et al., 2022), apoptosis-programmed cell death (Green 1998), calcium signaling (Hajnoczky et al., 2006), regulation of cellular metabolism (McBride et al., 2006), certain heme synthesis reactions (Oh-Hama, 1997), and many others.

All these multifunctional bioenergetic aspects and the importance of mitochondria for cell biology strongly suggest that mitogenome variations, although not currently used in genomic selection and breeding, should have significant implications for energy-intensive modern cattle production, as shown by experimental analyses. For example, the potential impact of mitogenome variation on quantitative traits in cattle has long been questioned by cytoplasmic models in which maternal lineages in the pedigree explain up to 10% of phenotypic variability in dairy (Albuquerque et al., 1998; Boettcher and Gibson, 1997; Kennedy, 1986; Roughsedge et al., 1999) and beef (Mannen et al., 1998; Mezzadra et al., 2005; Pun et al., 2012) populations. The effects of molecular variation in the mitogenome, although limited to only a few hundred nucleotides, have been used in several studies (Boettcher et al., 1996; Brown et al., 1989; Qin et al., 2012; Schutz et al., 1994). Using complete mitogenome information traced back (imputed) to maternal lineages, Brajkovic (2019) confirmed that mitogenome polymorphism can explain a substantial part of the variation in milk (up to 10.2%), fat (up to 8.1%), and protein (up to 14.6%) performance. Recently, Fortuna et al. (2022) showed in a simple model through computer simulations that accounting for mitogenome variation can increase the accuracy of estimated breeding values from 1% to 4%, although given that mitochondrial DNA is passed through the maternal line, it is clear that genetic improvement only affects the selection pathway from dam to daughter. For more on incorporating cytoplasmic (mitogenome) information into dairy cattle breeding programs, see Gibson et al., 1997. Biological evidence for the importance of the mitogenome in bovine production performance has been provided by gene expression analyses based on RNA sequencing which linked mitochondrial genes variation to energy metabolism and feed efficiency in dairy cows (Dorji et al., 2020, 2021). The association of DNA copy number (an indicator variable for mitochondrial efficiency) and mitogenome variation (SNP and haplogroups) with growth and carcass traits in cattle was confirmed by Sanglard et al., (2022b). Evidence for the effects of the mitogenome on phenotype was also provided by Novosel et al. (2022), who showed that mutations known to be deleterious in humans can have negative effects similar to those of Leber's hereditary optic neuropathy when they occur in cattle, whereas deleterious mitogenome mutations have also been observed in dogs (Baranowska et al., 2009) and mice (Wredenberg et al., 2002; Trifunovic et al., 2005).

In addition to their functional and physiological effects on phenotypic variation, which is important for high productivity, mitochondrial DNA information can also be used for a variety of tasks, including genealogical verification with error detection in the pedigree (Oliehoek and Bijma 2009; Čačić et al., 2014), understanding domestication and phylogenetics (Bradley et al., 1996; Loftus et al., 1994; Verdugo et al., 2019), providing evidence for ancient introgression (Achili et al., 2008; Cubric-Curik et al., 2022; Schibler et al., 2014; Verdugo et al., 2019), estimating maternal demographics (Achili et al., 2008; Cubric-Curik et al., 2022; Mannen et al., 2020; Olivieri et al., 2015), estimating maternal lineage diversity with analysis of population structure (Ginja et al., 2019; Lorenzo et al., 2018) and several other topics.

Despite their tremendous applicability, mitogenome information contained in high-throughput SNP arrays, unlike nuclear markers, has not been adequately exploited for breeding and biodiversity management of cattle populations. Therefore, the main objective of this study was to change this state of affairs by enabling and promoting the routine use of mitogenome SNPs. To achieve our goal, we collaborated with Neogen Genomics (Lincoln, NE, USA) to select and verify 310 mitochondrial SNPs that were included in the new NEOGEN GGP Bovine 100K SNP Chip. This is the most comprehensive and systematic incorporation of mitogenome information into a SNP array that we are aware of to date. We have also developed pipelines and analysis tools, all implemented in the beta version of the improved Magellan software (Ristov et al., 2016), enabling numerous mitogenome analyses, including i) automatic readout of raw genotyping data with formation of mitogenome haplotypes in Fasta formated files, ii) classification of haplotypes based on 310 verified SNPs, iii) genealogical verification with identification of pedigree errors, iv) imputation of haplotypes by maternal lineages - necessary to reduce genotyping costs and improve the performance of quantitative genetic mitogenome analyses, v) identification of some hereditary deleterious mutations, and more. Finally, exemplary analyses were performed on a dataset of 264 genotyped animals (NEOGEN GGP Bovine 100K SNP Chip) to illustrate how mitogenome information can be efficiently used in cattle breeding and biodiversity management.

2. Materials and methods

2.1. Selection and verification of informative SNP markers

The selection of informative SNP markers included in the NEOGEN GGP Bovine 100K SNP Chip was based on our previous large-scale mitogenome study that included 736 complete mitogenome sequences from over 150 cattle breeds (Cubric-Curik et al., 2022), with 164 mitogenomes added (all listed in Table S1). Our original reference dataset was aligned to the bovine reference mitogenome sequence (GenBank accession number V00654, 16338bp) using the Clustal Omega program (Sievers et al., 2011), visualized using MEGA7 software (Kumar et al., 2016), and set to a length of 16388bp without indels. All mitogenomes in our reference dataset were classified respect to haplogroups using both MitoToolPy (Peng et al., 2015) and BEAST algorithms according to the procedures described in Cubric-Curik et al. (2022).

Once the reference dataset was created, the selection was based on several criteria. First, we selected all discriminative SNPs that would provide a good classification of mitogenomes according to known haplogroups. Second, we selected those SNPs that represented high diversity metrics in our large meta-population dataset of over 150 breeds. In addition, we focused on selecting SNPs that could promote the use of mitogenomes in animal breeding, i.e., that would provide the haplotype variation required for quantitative genetic mitogenome analyses. Third, we selected a small number of SNPs that have been shown to cause mitochondrial disease in humans (Wallace 2007; Wallace et al., 1988,

2018). This decision was motivated by previous experience where we detected different putatively detrimental mutations in our reference dataset that originated from several breeds and were associated with LHON-like phenotypes in several breeds (Novosel et al., 2022). Intuitively, the observed prevalence was much more disproportionate than originally expected, which deserves further attention. Overall, our selection included 331 SNPs among which 70, 219 and 42 SNPs were chosen with respect to haplogroup classification, diversity and their potential impact on the health status, respectively.

Validation of the selected mitogenome SNPs was carried out by high-throughput genotyping with the NEOGEN GGP Bovine 100K SNP Chip of 264 individuals belonging to 19 cattle breeds sampled in Austria, Croatia, Greece, and Italy. Complete mitogenome sequences were previously determined for 129 animals, and the remaining 135 animals were sampled from the same breeds. To ensure reliable SNP verification and haplogroup classification, our data set included animals assigned to all other established haplogroups (R, P, Q, T_1, T_2, T_3, T_4, and T_5), with the exception of I, the haplogroup characteristic of Indicine cattle.

2.2. Computerized enhancement of mitogenome analyses

Mitogenome analyses require additional knowledge due to their specific nature of haplotype maternal inheritance, which differs from classical breeding approaches and the usual handling of high-throughput SNP arrays. For example, to identify haplotypes from the set of mitogenome SNPs of high-throughput genotyping arrays and correctly classify them into correct haplogroups, one would have to go through a rather time-consuming protocol that could be quite daunting for a non-specialized user. To address this problem, we have developed pipelines and analysis tools, all implemented in the publicly available beta version of the Magellan v2.0 software (https://angen.agr.hr/en/group/77/MaGelLAn±2.0 or https://github.com/sristov/magellan), that allow convenient and efficient execution of numerous mitogenome analyses. Using the 264 animal dataset described above, we further explained how to perform a number of useful mitogenome analyses using Magellan v2.0 software. Please note that the current beta version of Magellan v2.0 is not final, as there are a number of opportunities for improvement beyond this study that represent our future challenge.

2.2.1. Automatic readout of mitogenome SNPs from the raw genotyping data of the NEOGEN GGP Bovine 100K SNP chip

The Magellan v2.0 release includes a new module mag_snp.py ("SNP analysis" in the GUI version) that can accept as input the NEOGEN GGP Bovine 100K final report in TSV (tabulator separated values) format, along with three additional mandatory files (Map file, SNP list, and population list). A Map file is a TSV formated file that describes the SNPs present in the chip and contains the following mandatory columns: Name, Chromosome, Position and GenTrain Score. The SNP list is a TXT formated file containing the list of the relevant mitochondrial SNPs. Finally, the population list is a TSV formated file containing the list of analyzed individuals along with the corresponding population. The population list file requires two mandatory columns (Sample ID and Population), while the functions of the module can be used to analyze the input and prepare the .lgen, .fam and .map files for further analysis, for example for further use in PLINK software, and to output a FASTA formated file containing the sequences (310 SNP long haplotypes) of the relevant SNPs from the map file for all individuals in the population. An example of a final report can be found in the example inputs in Magellan v2.0 as "1. Example_of_Final_report_of_Neogen_GGP_Bovine 100K_SNP_array".

2.2.2. Biodiversity, classification of haplotypes, and phylogenetic analyses

We have developed another method that provides meaningful classification of bovine mitogenome haplogroups based only on "reduced haplogroups", i.e., containing 310 verified SNPs previously stored in the FASTA-formated file. In Magellan v2.0 software, classification is performed in mag_snp.py and/or "SNP analysis" (highlight "Fasta", run "Run classification", and select the FASTA input file). The classification algorithm was trained on our highly informative dataset of 900 sequences representing all known haplogroups (see **Table S1**). Effectively, our dataset included 436 haplotypes, as we removed duplicates that occurred when reducing the complete mitogenome to 310 informative SNPs. The Magellan v2.0 software uses two different approaches for assigning "reduced haplotypes" to haplogroups, depending on the type of classification required. For classification into highly divergent haplogroups (R, I, P, Q, T), the software uses a multilayer machine learning model called eXtreme Gradient Boosting (Xgboost) (Chen et al., 2016), while for classification between T haplogroups (T_1, T_2, T_3, T_4, or T_5), the Siamese Dense Neural Network (SDNN) approach is used (Bromly et al., 1993). The classification results were provided in a textual output file "OutputSNP_Classification.txt".

Please note that any other biodiversity or/and phylogenetic analysis can be performed on the selected FASTA file with the reduced 310 SNPs of the informative haplogroups. Here, biodiversity analysis was performed using the DnaSp program (Rozas et al., 2017) on our sample of 264 haplotypes enriched with additional 248 "reduced haplotypes" derived from the 310 SNP information extracted from publicly available complete mitogenomes with previously classified haplogroups (Cubric-Curik et al., 2022) and 288 mitogenomes from GenBank. To confirm the results obtained with the classification algorithm, we also performed median-joining network analyses (Bandelt et al., 1999) on a selected dataset visualised with PopART (Leigh and Bryant, 2015).

2.2.3. Genealogical verification with identification of pedigree errors

Maintaining an accurate and informative pedigree has been an important component of animal breeding for over 100 years and has lost none of its applicability today with the use of the single-step genomic BLUP model that considers phenotypic, pedigree, and genomic information simultaneously (Legarra et al., 2009; Aquilar et al., 2010; Kudinov et al., 2022). Although not yet fully recognized in practice, mitogenome information, if typed for large numbers of animals and integrated into commercial high-throughput SNP arrays, could be a powerful tool to test the accuracy of pedigrees and find misaligned individuals deep in the pedigree (Čačić et al., 2014; Ristov et al., 2016). Using a pedigree of Croatian Busha cattle with 2444 triplets (animal, sire, dam) and 25 genotyped individuals, we demonstrated the power of mitogenome information that was integrated into the SNP chip and Magellan v2.0 algorithms (mag_ver.py) to identify and visualized pedigree errors. The example file (BushaPedErrors.txt) with recoded partial pedigree and mitogenome information for Busha cattle is included in the example inputs in Magellan v2.0 as "2. Example_of_Busa_pedigree_with_haplotypes".

2.2.4. Imputation of haplotypes by maternal lineages

The specific mitochondrial inheritance pattern allows "imputation" of haplotypes across maternal lineages. Magellan v2.0 provides several options in mag_verif ("Object of verification" highlight "SNP Sequence" and run "Impute haplotype") that allow imputation with varying degrees of confidence depending on the assumptions made (e.g., that the pedigrees are correct and that no mutations occur within certain lineages), which is demonstrated here using the Croatian Busha cattle dataset mentioned above.

2.2.5. Identification of deleterious mutations

As mentioned earlier, we also selected 40 SNP positions from the 310 SNPs that are associated with various pathological effects (diseases) in humans, such as Leber Hereditary Optic Neuropathy, Mitochondrial Encephalopathy, Lactic Acidosis, Stroke-like Episodes, and others (see **Table S3** for more information on the selected potentially deleterious variants their positions). The large number of selected SNP positions was motivated by the relatively high frequency of supposedly deleterious mutations found in our previous study (Cubric-Curik et al., 2022;

Novosel et al., 2022). Here, we used Busha cattle as an example to show how the Magellan v2.0 program (with the "Check deleterious mutations" option) uses mitogenome information from the NEOGEN GGP Bovine 100K SNP Chip to identify potentially deleterious mutations. For example, Magellan v2.0 provides the output "MUTATIONS_ALERT.TXT" which includes individual identification, SNP position in cattle, SNP position in humans, and type of associated diseases. Please use "1. Example_of_Final_report_of_Neogen_GGP_Bovine_100K_SNP_array" from the example inputs in Magellan v2.0 to check for the occurrence of deleterious mutations.

3. Results and discussion

3.1. Validation of mitochondrial SNPs from NEOGEN GGP Bovine 100K array

In our validation analyses, only 310 SNPs included in the NEOGEN GGP Bovine 100K SNP Chip matched 100% with previously known mitogenome sequences, whereas 21 SNPs were excluded from further analyses as unreliable.

Unfortunately, some alternative mutations within the selected SNPs were not verified directly, although the match to their wild types was confirmed, because we did not have access to samples with all known mutations appearing among selected 331 SNP positions. Visualization of the verified mitogenome SNPs is shown in Fig. 1, whereas the list of mitogenome SNPs included in the analysis contains additional

information (**Table S2**).

3.2. Biodiversity and classification of haplotypes

Mitogenome diversity and haplotype classification for the 264 animals genotyped with the NEOGEN GGP Bovine 100K SNP chip and the 288 mitogenomes downloaded from the GenBank is presented in Table 1. The results clearly demonstrate that the mitogenome diversity explained by our "reduced haplotypes" (310 mitogenome SNPs) is sufficient to be effectively used in numerous applications such as biodiversity monitoring, pedigree verification, breed identification, quantitative genetic analysis, etc., even when applied to extremely small-sized indigenous cattle breeds such as Croatian Busha cattle, Slavonian-Syrmian Podolian cattle, and Katerinis-Outras cattle. For example, we observed haplotype diversity values that frequently exceeded 0.85 (see **Table 2**), even though some samples were intentionally taken from the same maternal lineages (this was required for pedigree verification), which is quite high a value considering that haplotype diversity of the selected world breeds complete mitogenome dataset ranged from 0.93 to 1.00 (Dorji et al., 2022). We analyzed an additional 288 mitogenomes from GenBank to test how well our selected SNPs describe diversity in the very distant breeds not included in our training (SNP selection) dataset. The results of the diversity analysis of these breeds were consistent with those of the breeds of the training set. For example, haplotype diversity ranged from 0.68 to 0.94, with most populations exceeding 0.91.
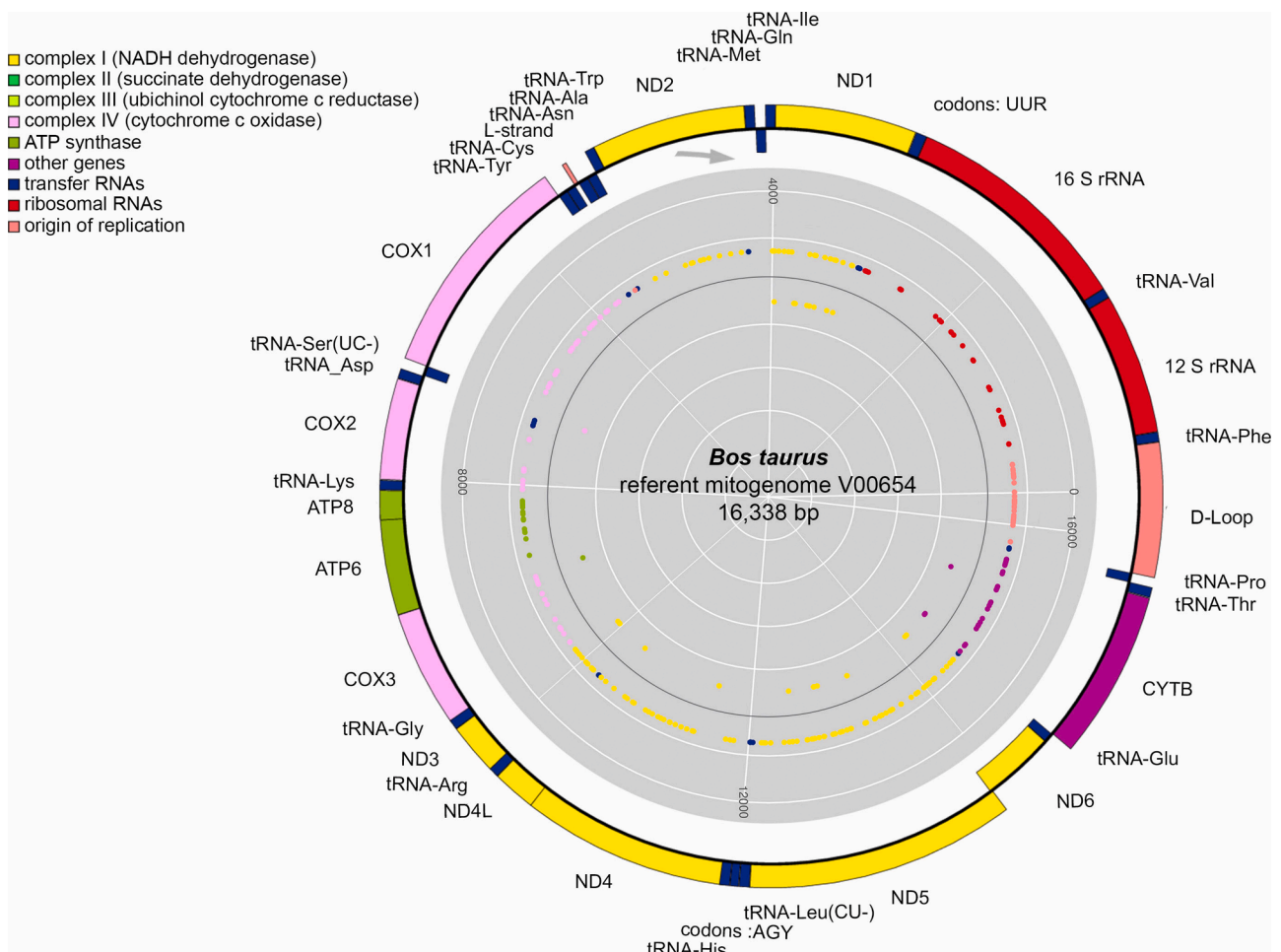

Fig. 1. Visualization of the verified 310 mitogenome SNPs with their position colored according to their function. SNP positions associated with deleterious mutations in humans are shown in the inner circle. Genes inside the circle are transcribed clockwise, genes outside the circle are transcribed counterclockwise. Solar plot was generated by modifying R code (https://github.com/stephenturner/solarplot) and OrganellarGenomeDRAW (Greiner et al., 2019).

Table 1

Mitogenome diversity and haplotype classification for 264 cattle genotyped by the NEOGEN GGP Bovine 100K SNP Chip and for 362 mitogenomes from the GenBank.

Breed	N	nML	nH	HD	VS	T_1	T_2	T_3	T_4	T_5	Q	P	R	I
Breeds for which mitogenome information was provided by the NEOGEN GGP Bovine 100K SNP Chip														
CHF	85	38	30	0.87	39	7	4	73		1				
CIP	30	12	9	0.86	10	2		28						
CB	25	16	18	0.97	25	6	8	11						
CSSP	24	8	5	0.77	9			18		6				
AMB	24		14	0.93	56		2	17				5		
GKO	10		6	0.78	22		1	8			1			
GKT	10		3	0.60	8		6	4						
GSP	12		6	0.85	17	1		8			3			
GSR	13		6	0.86	19	1		10			2			
OTH	31		17	0.89	72		1	26						4
Breeds for which mitogenome information was obtained from GenBank														
Gir	28		17	0.88	29									28
Holstein	26		9	0.83	10	1		25						
Kankrej	20		11	0.92	23									20
Ladahki	11		7	0.91	260			1						10
Lakhimi	12		7	0.83	206	2								10
Longlin	21		8	0.68	206			1	2					18
Mewati	10		7	0.91	17									10
Mongolian	26		13	0.94	212	3	7	11	2	2				1
Tibetan	48		18	0.92	34	8	1	34	2		3			
Yunling	160		35	0.92	226	30	5	47	7					71

Number of genotyped animals (N); Number of maternal lineages derived from the pedigree (nML and was not provided for breeds where pedigree was not available); Number of haplotypes (nH); Haplotype diversity (HD); VS – variable sites; Pedigree not available (PNA); Croatian Holstein (CHF), Istrian Cattle - Boskarin (CIP), Croatian Busha (CB), Slavonian Syrmian Podolian (CSSP), Murbodner (AMB), Katerinis Outras (GKO), Katerinis Tsiandouris (GKT), Shorthorn Pindos (GSP), and Shorthorn Rodopi (GSR), while other breeds genotyped in this study but represented with less than 10 animals were pooled (OTH). Breeds for which mitogenome sequences were taken from the GenBank are labelled with their full names, while the same mitogenomes were also used in Kinoshita et al. (2018) for Holstein, Xia et al. (2019) for Yunglin, Xia et al. (2020) for Longlin, Xia et al. (2021) for Mongolian and Tibetan cattle and Manisha et al. (unpublished) for Gir, Kankrej, Ladakhi, Lakhimi and Mewati breeds (accession numbers can be found in the Supplemental Table S1). All analyses have been performed on the 310 selected and verified SNPs.

At the same time, the classification algorithm implemented in Magellan v2.0 was able to correctly define haplogroups according to the classification of complete mitogenomes obtained in Cubric-Curik et al. (2022), as 100% empirical agreement (T_1, T_2, T_3, T_5, Q, P, and R haplotypes) was found for all samples that were additionally sequenced or that were in the same lineage as the verified sequenced samples. In classifying the "reduced haplotypes" from the non-training dataset from GenBank, we identified a wide range of haplogroups, including haplogroups T_4 and I, which were not present in our genotyped samples. These haplotypes were 97% correctly assigned by the algorithm implemented in MitoToolPy according to the classification of complete mitogenomes, further highlighting the efficiency of our classification algorithm and SNP selection. In this way, we were able to confirm the efficiency of the classification for all existing haplogroups.

The performance of the Magellan algorithm in correctly assigning the "reduced haplotypes" to the haplogroups is further supported by visualization of the Median-Joining network analysis (Fig. 2), in which, for example, four Chianina and five Murbodner haplotypes genotyped in this study (in gray) were correctly assigned to highly divergent haplogroups R and P, respectively. Agreement with the full mitogenome classification was also observed for other "reduced haplotypes" containing only 177 SNPs due to missing information of ancient samples obtained from the SNP chip (in gray). The Median-Joining network shown in Fig. 2 allows highlighting the relationships between representative haplotypes, including ancient aurochs sequences, and their positioning within the haplogroups. Please note that genetic relationship of mitogenomes can also be analyzed and visualized using phylogenetic trees and other clustering approaches, which has been presented in more detail in in Dorji et al. (2022).

Proper classification of haplotypes is necessary if we are to trace the long-term history of introgressions and the origin of cattle breeds. Ancient introgressions of different "ecotypes" of aurochs into modern cattle are a good example, as shown by the presence of haplogroups, as e.g. R and P, that diverged long before domestication (for more information, see Achilli et al., 2008; Verdugo et al., 2019; Cubric-Curik et al.,

2022). High maternal lineage diversity and/or the presence of rare haplogroups could be the result of upgrading or deliberate continuous crossing of low production breeds (cows) with bulls representing high production breeds.

This could be the case in the CHF breed investigated in this study, where the presence of rare haplotypes (T_1, T_2 and T_5) that are not specific to the Holstein-Frisian breed most likely indicates a maternal background resulting from the upgrading of local indigenous cattle in which a higher frequency of T_1, T_2 and T_5 was observed (CB and CSSP), see Table 1. Another example is the presence of haplogroup I, which indicates the maternal introgression of Bos indicus into Mongolian cattle (Table 1).

Our selection of 310 mitogenome SNPs allows analysis of maternal relatedness between different cattle breeds, as shown by the two-dimensional PCA plot of the pairwise calculated Wrights F_{ST} distance matrix between nine cattle breeds (Fig. 3B). Although theoretically higher in mammals, the estimated F_{ST} between breeds from mitogenome are expected to correlate with nuclear SNP estimates (Larson et al., 2009; Petit and Escoffier, 2009). To allow comparison with autosomal relatedness, the two-dimensional PCA representation of the pairwise Wright F_{ST} distance matrix is shown in Fig. 3C. When we regressed 36 pairwise nuclear F_{ST} values between nine breeds ($n > 9$) genotyped with the NEOGEN GGP Bovine 100K SNP Chip on estimates from the mitogenome (310 SNPs), the coefficient of determination ($R^2 = 0.48$) indicated that substantial variation in nuclear F_{ST} values can be explained by mitogenome estimates. Nevertheless, F_{ST} estimates from nuclear SNPs and mitogenomes can provide complementary information for population differentiation, and the degree of correlation between the two estimates should be interpreted contextually (Hedrick et al., 2013). For example, the strong sex-biased dispersal or gene flow between populations will lower correlation between F_{ST} estimates from nuclear SNPs and those from mitogenome (Petit and Escoffier, 2009). The same approach was repeated for 10 breeds with complete mitogenomes (362) from GenBank (Table 1). Surprisingly, the coefficient of determination of the regression of the pairwise F_{ST} values estimated from the complete
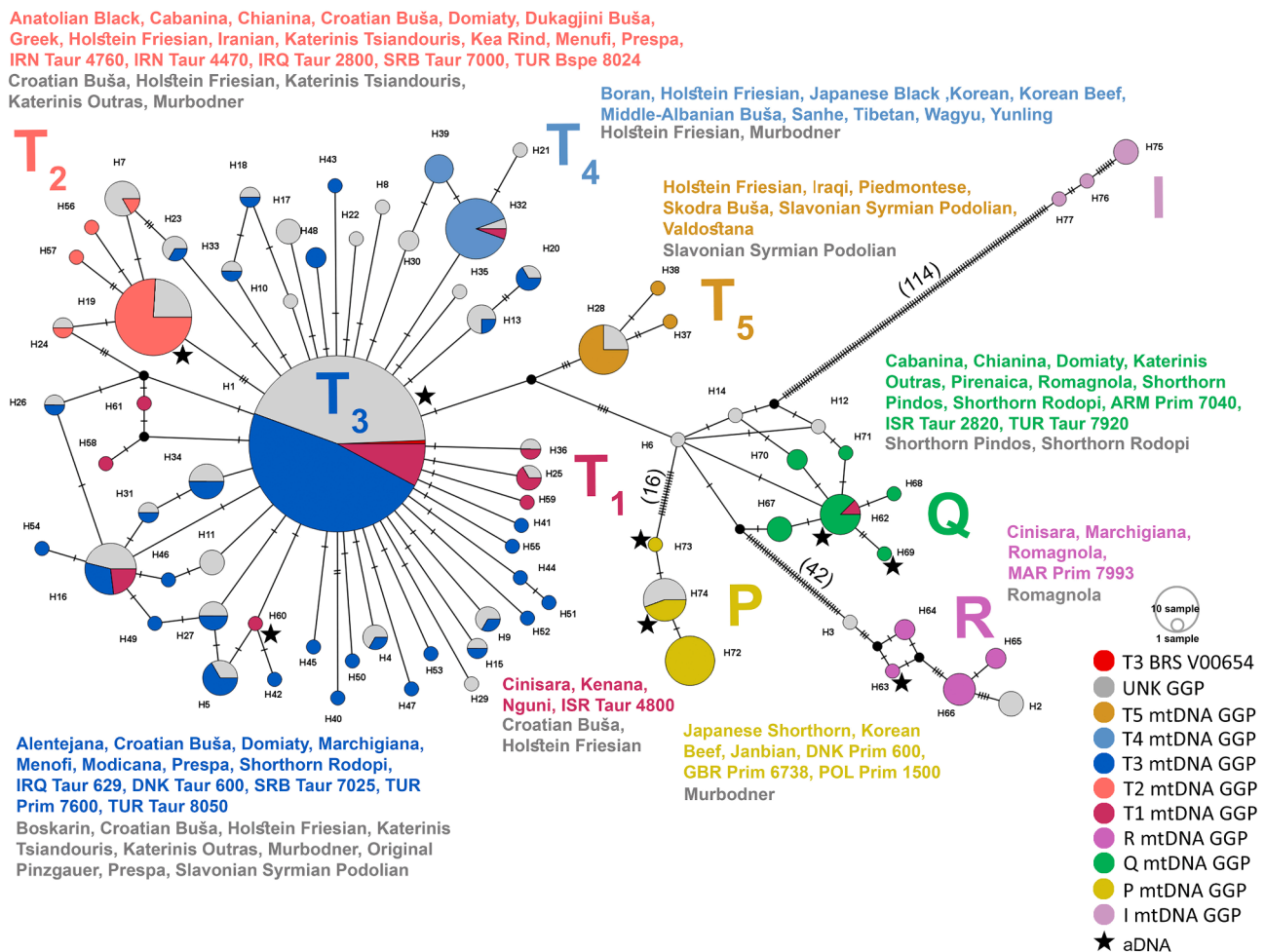
Fig. 2. The Median-Joining network shows the relationships among 77 "reduced haplotypes" (177 SNPs) found in 383 representative cattle samples. For aDNA cattle samples, the number in the name indicates the year before present, while aurochs are indicated by a black star.

mitogenomes on the estimates from the "reduced" mitogenomes (310 SNPs) was very high (R^2 = 0.98), demonstrating the power of the selected 310 SNPs in estimating maternal population relatedness. The regression analyses performed are shown in Fig. 4.

3.3. Genealogical verification with identification of pedigree errors

While verifying the correctness of Croatian Busha cattle pedigree, we found five cases where the maternal haplotypes did not match. One such example can be seen on the small portion of the Croatian Busha pedigree in Fig. 5, where cow 2308 with the alternative haplotype, here labelled A, did not match two other mitogenome haplotypes, here labelled G, identified in cows 2708 and 2191. The genealogical verification of pedigrees by mitogenome segregation can be very informative when the genotypes of a large number of individuals are available, especially when it complements the segregation of nuclear SNPs. For more information on the principles of the pedigree verification algorithm, see Čačić et al. (2014) and Ristov et al. (2016). Unfortunately, the number of SNPs included in the SNP array did not provide satisfactory discriminatory power to assign individual animals to specific breeds because some haplotypes occurred in multiple breeds (information in the Supplement Table S5). However, given the suitability of mitochondrial DNA for human forensic analyses with large numbers of genotyped animals, the implemented 310 SNPs could be used as useful prescreening information prior to more informative full sequencing of the mitogenome. By combining the information from the SNP array and full mitogenome sequencing, it may be possible to improve the accuracy and resolution of

breed assignment in animals, similar to how mtDNA analysis is used in forensic analyses in humans. For more details on the revolution of animal genomics in forensic sciences, see Cardinali et al. (2023).

3.4. Imputation of haplotypes by maternal lineages

Imputation of haplotypes across maternal lineages, while easily integrable into a software, is almost a prerequisite for performing quantitative genetic analyses because they require large numbers of genotyped animals. No less importantly, imputation across maternal lineages can be useful for detecting individuals with specific or deleterious mutations. For example, we have used the Magellan algorithm to trace Murbodner cattle with the P haplotype specific to aurochs and in tracing Cika cattle with the deleterious LHON mutation (for more information, see Cubric-Curik et al. (2022) and Novosel et al. (2022), respectively). However, the quality of the imputation and of any further analysis depends on a correct pedigree. Therefore, we strongly advise to check the pedigree before imputation. Depending on the reliability of the imputation, the new version of Magellan software offers two options. In the less reliable option, all individuals in the genotyped maternal lineages that do not turn out to be false are imputed. In the example pedigree shown in Fig. 5, a less reliable imputation would list all animals with letters (all animals with green or orange frames), specifically those labelled 1525, 1828, 2191, 2708, 2713, 2516, 3045, 3284, and 3360 representing haplotype G, and the directly genotyped cow "2308" with haplotype A (the only reliable animal in this part of the pedigree). In contrast, a very reliable imputation would list only animals with green
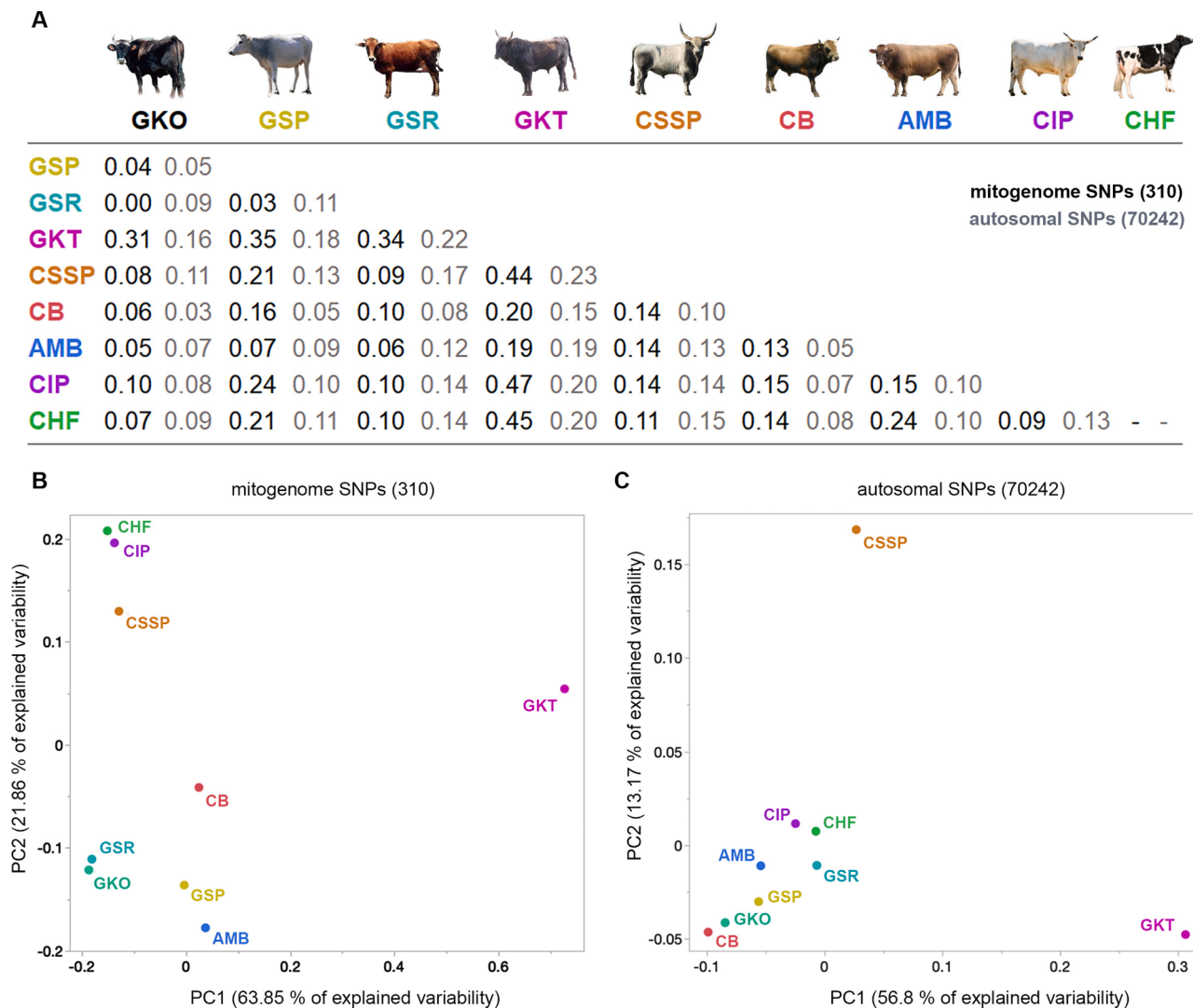
A

	GKO	GSP	GSR	GKT	CSSP	CB	AMB	CIP	CHF
GSP	0.04 0.05								
GSR	0.00 0.09	0.03 0.11							
GKT	0.31 0.16	0.35 0.18	0.34 0.22						
CSSP	0.08 0.11	0.21 0.13	0.09 0.17	0.44 0.23					
CB	0.06 0.03	0.16 0.05	0.10 0.08	0.20 0.15	0.14 0.10				
AMB	0.05 0.07	0.07 0.09	0.06 0.12	0.19 0.19	0.14 0.13	0.13 0.05			
CIP	0.10 0.08	0.24 0.10	0.10 0.14	0.47 0.20	0.14 0.14	0.15 0.07	0.15 0.10		
CHF	0.07 0.09	0.21 0.11	0.10 0.14	0.45 0.20	0.11 0.15	0.14 0.08	0.24 0.10	0.09 0.13	- -

mitogenome SNPs (310)
autosomal SNPs (70242)

B — mitogenome SNPs (310)

PC1 (63.85 % of explained variability) vs PC2 (21.86 % of explained variability)

C — autosomal SNPs (70242)

PC1 (56.8 % of explained variability) vs PC2 (13.17 % of explained variability)

Fig. 3. Illustration of the genetic relatedness between genotyped cattle breeds. A. Pairwise plot of F_{ST} values (estimated using 310 mitogenome SNPs versus 70,242 autosomal SNPs) between cattle breeds (breed represented with only more than nine animals and genotyped with NEOGEN GGP Bovine 100K SNP Chip). B. Principal component plot from F_{ST} distance matrix based on 310 mitogenome SNPs and C. Principal component plot from F_{ST} distance matrix based on 70,242 autosomal SNPs. The concept has been applied to the following breeds of cattle: Holstein (CHF), Istrian Cattle - Boskarin (CIP), Croatian Busha (CB), Slavonian Syrmian Podolian (CSSP), Murbodner (AMB), Katerinis Outras (GKO), Katerinis Tsiandouris (GKT), Shorthorn Pindos (GSP), and Shorthorn Rodopi (GSR).

frames, more precisely the animals numbered 1525, 1828, 2191 and 2708 with haplotype G and the cow 2308 with haplotype A (genotyped), since these are the only individuals with verified haplotypes.

3.5. Identification of some deleterious mutations

The presence of deleterious mutation, either placed on autosomes or X chromosomes, are well documented in cattle (Fritz et al., 2013; Sonstegard et al., 2013; VanRaden et al., 2011; Jenko et al., 2019) and according to Bosse et al. (2019) everyone is expected to possess a certain number of such mutations in its genome. In contrast to humans, where deleterious mitogenome mutations are well documented (Wallace 2007; Wallace et al., 1988, 2018), only one deleterious mitogenome mutation has been shown to cause pathological changes in cattle (Novosel et al., 2022), whereas the deleterious effects of mutation C3965A (C4171A mutation in humans) found in several cattle breeds (GenBank accession numbers Croatian Busha - MZ901457; Croatian Holstein - MZ901512; two Slovenian Cika - MZ901663, MZ901664; Korean cattle - DQ124399; Kosovo Busha - MZ901737) and reported by Novosel et al. (2019) remain to be confirmed. Among the cattle genotyped with the NEOGEN GGP Bovine 100K SNP chip, we detected the presence of the C3965A

mutation in four animals (one Croatian Busha and three Croatian Holstein cattle) and the presence of the T10432C mutation (T10663C mutation in humans) in the Cika cow "Gavtraža" (GenBank accession number MZ901664). Our analysis of the SNP array data revealed three additional Croatian Holstein animals with the C3965A mutation. Automatic reporting of putatively deleterious mutations was done via the Magellan v2.0 output file "MUTATIONS_ALERT.TXT" (see **Figure S1**). In general, we were surprised by the high frequency of these supposedly deleterious mutations in cattle, a problem that requires further assessment. When referring to deleterious mitogenome mutations, we have considered changes occurring in the mtDNA sequence and not the nuclear mutations that disrupt normal mitochondrial function and lead to dysfunction, such as the mutation in the MFN2 gene found in Tyrolean grey cattle and associated with degenerative axonopathy (for more information, see Drögemüller et al., 2011). The inclusion of many known deleterious human mitogenome mutations in high-throughput arrays routinely used for genomic selection, as is the case with the NEOGEN GGP Bovine 100K SNP Chip, will certainly help to clarify this issue.
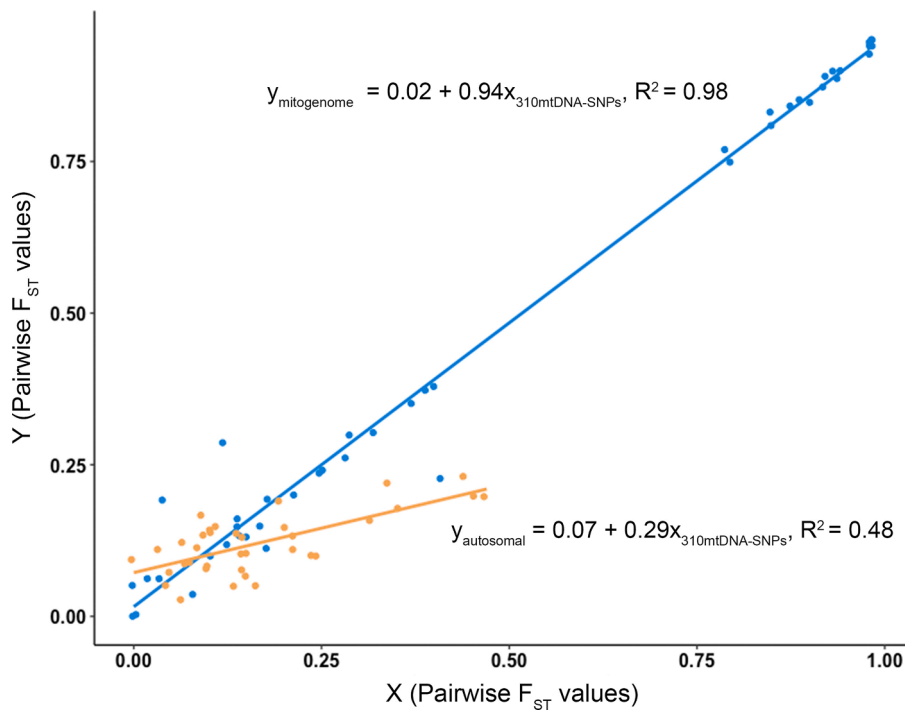
$$y_{mitogenome} = 0.02 + 0.94x_{310mtDNA-SNPs}, R^2 = 0.98$$

$$y_{autosomal} = 0.07 + 0.29x_{310mtDNA-SNPs}, R^2 = 0.48$$

Fig. 4. Linear regression analyses of the relationship between pairwise estimates of genetic relatedness among cattle breeds. The relationship between complete (16,338 SNPs) and "reduced" (310 SNPs) mitogenome information is shown in blue (10 breeds), whereas the relationship between autosomal (70,242 SNPs) and "reduced" mitogenome (310 SNPs) estimates is shown in orange (nine breeds).
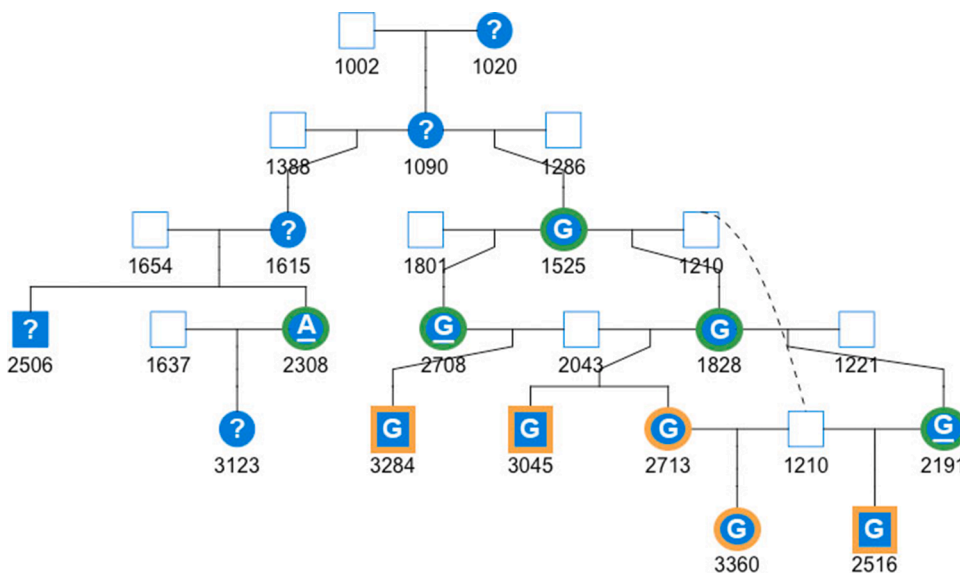

Fig. 5. Illustration of an example of verification and imputation performed on the pedigree of the Croatian Busha. The inheritance segregation pattern of haplotype A in cow 2308 does not match the haplotypes G identified in cows 2708 and 2191 (only the haplotypes of directly genotyped animals are underlined), indicating an error in the pedigree. All individuals with less reliable imputed haplotypes G are outlined in the orange, while highly reliable imputed (verified) or directly genotyped haplotypes are outlined in green.

4. Future prospects and possibilities

Besides the research applications of mitogenome information that we presented here, there are several additional usages we have not addressed in this study but whose potential we would like to highlight. Certainly, of great importance is the concept of relating tissue-specific mitochondrial content to various production-related phenotypes, as reported for chickens by Reverter et al. (2017) or for cattle by Sanglard et al. (2022b). This is of particular interest because mitochondrial DNA copy number, a good indicator of mitochondrial activity and a proxy for the number of mitochondria in cells, is heritable and can be readily determined by low-pass whole genome sequencing (Sanglard et al.,

2022a). In humans, heteroplasmic mitogenome mutations detectable by whole-genome sequencing (Young et al., 2022; Battle et al., 2022) are known to cause the prevalence of various disorders, particularly above a certain threshold (Ye et al., 2014; Davis et al., 2022). In addition, the Magellan v2.0 software provides some features that were not described in this study. A good example is the "optimal sample selection" feature, which is particularly useful for populations with informative and correct pedigree information when our financial resources are limited, and/or for diversity studies (the implemented algorithm searches for the sample set with the largest sum of mutual distances in a pedigree) to achieve the best coverage of sample variation for the available sample resources (the algorithm is described in Hršak et al., 2022b). We will continue to work

on improving the Magellan software (for some recent changes see Hršak et al. (2022a) and we take advantage of this opportunity to announce our plans for the near future. The most obvious improvement to be included in the Magellan software is the automatic reading of other high-throughput arrays containing mitogenome information. As mentioned earlier, low-pass whole genome sequencing is becoming more common in animal breeding, providing additional opportunities such as retrieving complete mitogenome sequence information and/or estimating mitochondrial DNA copy number. Therefore, in the future, we plan to integrate some automation pipelines into the Magellan software to enable complete mitogenome identification, heteroplasmy quantification, and mitochondrial DNA copy number estimation (for more information, see Sanglard et al., 2022a) to expand and further promote routine use.

5. Conclusion

Mitogenome information contained in high-throughput SNP arrays, unlike nuclear markers, has not been adequately exploited for breeding and biodiversity management of cattle populations. Here, we demonstrated the broad potential of using mitogenome information contained in high-throughput arrays to: (i) analyze population structure and diversity, (ii) classify mitogenome haplogroups in cattle and identify introgression or upgrading, (iii) screen and identify pedigree errors, (iv) impute mitogenome information on maternal lineages to analyze the effects of mitogenome variation on quantitative production traits, (v) identify suspected deleterious mutations. While full use of mitogenome information can be time consuming and requires additional specific knowledge, our enhancements built into the Magellan v2.0 software enable automation of several required analyses and thus the routine use by a broader scientific community. Overall, we hope that this study will foster use of mitogenome information in cattle breeding and genetic diversity monitoring.

Author statement

The authors state that no experiments with human subjects were performed in this study.

CRediT authorship contribution statement

V. Brajkovic: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Validation, Visualization, Writing – original draft, Writing – review & editing. **D. Hršak:** Investigation, Methodology, Software, Validation, Writing – original draft, Writing – review & editing. **L. Bradić:** Methodology, Formal analysis, Validation. **K. Turkalj:** Methodology, Formal analysis, Validation. **D. Novosel:** Formal analysis, Methodology, Validation. **S. Ristov:** Software. **P. Ajmone-Marsan:** Investigation, Resources. **L. Colli:** Investigation, Resources, Writing – original draft. **V. Cubric-Curik:** Funding acquisition, Investigation, Resources. **J. Sölkner:** Investigation, Resources. **I. Curik:** Conceptualization, Formal analysis, Methodology, Supervision, Validation, Writing – original draft, Writing – review & editing.

Declaration of Competing Interest

The authors declare that there is no conflict of interests regarding this study.

Acknowledgements and Funding

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.livsci.2023.105299.

References

Achilli, A., Olivieri, A., Pellecchia, M., Uboldi, C., Colli, L., Al-Zahery, N., Accetturo, M., Pala, M., Kashani, B.H., Perego, U.A., Battaglia, V., Fornarino, S., Kalamati, J., Houshmand, M., Negrini, R., Semino, O., Richards, M., Macaulay, V., Ferretti, L., Bandelt, H.J., Ajmone-Marsan, P., Torroni, A., 2008. Mitochondrial genomes of extinct aurochs survive in domestic cattle. Curr. Biol. 18, R157–R158. https://doi.org/10.1016/j.cub.2008.01.019.

Aguilar, I., Misztal, I., Johnson, D.L., Legarra, A., Tsuruta, S., Lawlor, T.J., 2010. Hot topic: a unified approach to utilize phenotypic, full pedigree, and genomic information for genetic evaluation of Holstein final score. J. Dairy Sci. 93, 743–752. https://doi.org/10.3168/jds.2009-2730.

Albuquerque, L.G., Keown, J.F., Van Vleck, L.D., 1998. Variances of direct genetic effects, maternal genetic effects, and cytoplasmic inheritance effects for milk yield, fat yield, and fat percentage. J. Dairy Sci. 81, 544–549. https://doi.org/10.3168/jds.S0022-0302(98)75606-1.

Bandelt, H.J., Forster, P., Röhl, A., 1999. Median-joining networks for inferring intraspecific phylogenies. Mol. Biol. Evol. 16, 37–48. https://doi.org/10.1093/oxfordjournals.molbev.a026036.

Baranowska, I., Jäderlund, K.H., Nennesmo, I., Holmqvist, E., Heidrich, N., Larsson, N.G., Andersson, G., Wagner, E.G.H., Hedhammar, Å., Wibom, R., Andersson, L., 2009. Sensory ataxic neuropathy in golden retriever dogs is caused by a deletion in the mitochondrial tRNATyr gene. PLoS Genet. 5, e1000499 https://doi.org/10.1371/journal.pgen.1000499.

Battle, S.L., Puiu, D., Verlouw, J., Broer, L., Boerwinkle, E., Taylor, K.D., Rotter, J.I., Rich, S.S., Grove, M.L., Pankratz, N., Fetterman, J.L., Liu, C., Arking, D.E., 2022. A bioinformatics pipeline for estimating mitochondrial DNA copy number and heteroplasmy levels from whole genome sequencing data. NAR Genom. Bioinforma 4, lqac034. https://doi.org/10.1093/nargab/lqac034.

Boettcher, P.J., Freeman, A.E., Johnston, S.D., Smith, R.K., Beitz, D.C., Mcdaniel, B.T., 1996. Relationships between polymorphism for mitochondrial deoxyribonucleic acid and yield traits of Holstein cows. J. Dairy Sci. 79, 647–654. https://doi.org/10.3168/jds.S0022-0302(96)76410-X.

Boettcher, P.J., Gibson, J.P., 1997. Estimation of variance of maternal lineage effects among Canadian Holsteins. J. Dairy Sci. 80, 2167–2176. https://doi.org/10.3168/jds.S0022-0302(97)76164-2.

Boichard, D., Ducrocq, V., Fritz, S., 2015. Sustainable dairy cattle selection in the genomic era. J. Anim. Breed. Genet. 132, 135–143. https://doi.org/10.1111/jbg.12150.

Bosse, M., Megens, H.J., Derks, M.F.L., de Cara, Á.M.R., Groenen, M.A.M., 2019. Deleterious alleles in the context of domestication, inbreeding, and selection. Evol. Appl. 12, 6–17. https://doi.org/10.1111/eva.12691.

Bradley, D.G., Machugh, D.E., Cunningham, P., Loftus, R.T., 1996. Mitochondrial diversity and the origins of African and European cattle. Proc. Natl. Acad. Sci. U. S. A. 93, 5131–5135. https://doi.org/10.1073/pnas.93.10.5131.

Brajković, V., 2019. Impact of Mitogenome On Milk Traits in Cattle. University of Zagreb Faculty of Agriculture. PhD thesis. https://dr.nsk.hr/en/islandora/object/agr%3A1290.

Britt, J.H., Cushman, R.A., Dechow, C.D., Dobson, H., Humblot, P., Hutjens, M.F., Jones, G.A., Mitloehner, F.M., Ruegg, P.L., Sheldon, I.M., Stevenson, J.S., 2021. Review: perspective on high-performing dairy cows and herds. Animal 15, 100298. https://doi.org/10.1016/j.animal.2021.100298.

Britt, J.H., Cushman, R.A., Dechow, C.D., Dobson, H., Humblot, P., Hutjens, M.F., Jones, G.A., Ruegg, P.S., Sheldon, I.M., Stevenson, J.S., 2018. Invited review: learning from the future—a vision for dairy farms and cows in 2067. J. Dairy Sci. 101, 3722–3741. https://doi.org/10.3168/jds.2017-14025.

Bromley, J., Bentz, J.W., Bottou, L., Guyon, I., Lecun, Y., Moore, C., Sackinger, E., Shah, R., 1993. Signature verification using a "Siamese" time delay neural network. Int. J. Pattern Recognit. Artif. Intell. 07, 669–688. https://doi.org/10.1142/s0218001493000339.

Brown, D.R., Koehler, C.M., Lindberg, G.L., Freeman, A.E., Mayfield, J.E., Myers, A.M., Schutz, M.M., Beitz, D.C., 1989. Molecular analysis of cytoplasmic genetic variation in Holstein cows. J. Anim. Sci. 67, 1926–1932. https://doi.org/10.2527/jas1989.6781926x.

Čačić, M., Cubric-Curik, V., Ristov, S., Curik, I., 2014. Computational approach to utilisation of mitochondrial DNA in the verification of complex pedigree errors. Livest. Sci. 169, 42–47. https://doi.org/10.1016/j.livsci.2014.09.009.

Cardinali, I., Tancredi, D., Lancioni, H., 2023. The revolution of animal genomics in forensic sciences. Int. J. Mol. Sci. 24, 8821. https://doi.org/10.3390/ijms24108821.

Chen, T., Guestrin, C., 2016. XGBoost: a scalable tree boosting system. In: Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 785–794. https://doi.org/10.1145/2939672.2939785.

Cole, J.B., VanRaden, P.M., 2018. Symposium review: possibilities in an age of genomics: the future of selection indices1. J. Dairy Sci. 101, 3686–3701. https://doi.org/10.3168/jds.2017-13335.

Cubric-Curik, V., Novosel, D., Brajkovic, V., Rota Stabelli, O., Krebs, S., Sölkner, J., Šalamon, D., Ristov, S., Berger, B., Trivizaki, S., Bizelis, I., Ferenčaković, M., Rothammer, S., Kunz, E., Simčič, M., Dovč, P., Bunevski, G., Bytyqi, H., Marković, B., Brka, M., Kume, K., Stojanović, S., Nikolov, V., Zinovieva, N., Schönherz, A.A., Guldbrandtsen, B., Čačić, M., Radović, S., Miracle, P., Vernesi, C., Curik, I., Medugorac, I., 2022. Large-scale mitogenome sequencing reveals consecutive expansions of domestic taurine cattle and supports sporadic aurochs introgression. Evol. Appl. 15, 663–678. https://doi.org/10.1111/eva.13315.

Davis, R.L., Kumar, K.R., Puttick, C., Liang, C., Ahmad, K.E., Edema-Hildebrand, F., Park, J.S., Minoche, A.E., Gayevskiy, V., Mallawaarachchi, A.C., Christodoulou, J., Schofield, D., Dinger, M.E., Cowley, M.J., Sue, C.M., 2022. Use of whole-genome sequencing for mitochondrial disease diagnosis. Neurology 99, E730–E742. https://doi.org/10.1212/WNL.0000000000200745.

Dorji, J., MacLeod, I.M., Chamberlain, A.J., Vander Jagt, C.J., Ho, P.N., Khansefid, M., Mason, B.A., Prowse-Wilkins, C.P., Marett, L.C., Wales, W.J., Cocks, B.G., Pryce, J.E., Daetwyler, H.D., 2021. Mitochondrial protein gene expression and the oxidative phosphorylation pathway associated with feed efficiency and energy balance in dairy cattle. J. Dairy Sci. 104, 575–587. https://doi.org/10.3168/jds.2020-18503.

Dorji, J., Vander Jagt, C.J., Chamberlain, A.J., Cocks, B.G., MacLeod, I.M., Daetwyler, H. D., 2022. Recovery of mitogenomes from whole genome sequences to infer maternal diversity in 1883 modern taurine and indicine cattle. Sci. Rep. 12, 5582. https://doi.org/10.1038/s41598-022-09427-y.

Dorji, J., Vander Jagt, C.J., Garner, J.B., Marett, L.C., Mason, B.A., Reich, C.M., Xiang, R., Clark, E.L., Cocks, B.G., Chamberlain, A.J., MacLeod, I.M., Daetwyler, H.D., 2020. Expression of mitochondrial protein genes encoded by nuclear and mitochondrial genomes correlate with energy metabolism in dairy cattle. BMC Genomics 21, 1–17. https://doi.org/10.1186/s12864-020-07018-7.

Drögemüller, C., Reichart, U., Seuberlich, T., Oevermann, A., Baumgartner, M., Boghenbor, K.K., Stoffel, M.H., Syring, C., Meylan, M., Müller, S., Müller, M., Gredler, B., Sölkner, J., Leeb, T., 2011. An unusual splice defect in the mitofusin 2 gene (mfn2) is associated with degenerative axonopathy in tyrolean grey cattle. PLoS ONE 6, e18931. https://doi.org/10.1371/journal.pone.0018931.

edited by FAO, 2015. In: Scherf, B.D., Pilling, D. (Eds.), The Second Report On the State of the World's Animal Genetic Resources For Food and Agriculture. FAO Commission on Genetic Resources for Food and Agriculture Assessments, Rome. edited by. http://www.fao.org/3/a-i4787e/index.html.

Fortuna, G.M., Zumbach, B.J., Johnsson, M., Pocrnic, I., Gorjanc, G., 2022. Accounting for nuclear- and mito-genome in genetic evaluation and breeding of dairy cattle. In: Proceedings of 12th World Congress on Genetics Applied to Livestock Production (WCGALP) Technical and species orientated innovations in animal breeding, and contribution of genetics to solving societal challenges, 301, pp. 1266–1269. https://doi.org/10.3920/978-90-8686-940-4_301.

Fritz, S., Capitan, A., Djari, A., Rodriguez, S.C., Barbat, A., Baur, A., Grohs, C., Weiss, B., Boussaha, M., Esquerré, D., Klopp, C., Rocha, D., Boichard, D., 2013. Detection of haplotypes associated with prenatal death in dairy cattle and identification of deleterious mutations in GART, SHBG and SLC37A2. PLoS ONE 8, e65550. https://doi.org/10.1371/journal.pone.0065550.

Gibson, J.P., Freeman, A.E., Boettcher, P.J., 1997. Cytoplasmic and mitochondrial inheritance of economic traits in cattle. Livest. Prod. Sci. 47, 115–124. https://doi.org/10.1016/S0301-6226(96)00023-1.

Ginja, C., Gama, L.T., Cortés, O., Burriel, I.M., Vega-Pla, J.L., Penedo, C., Sponenberg, P., Cañón, J., Sanz, A., do Egito, A.A., Alvarez, L.A., Giovambattista, G., Agha, S., Rogberg-Muñoz, A., Lara, M.A.C., Afonso, S., Aguirre, L., Armstrong, E., Vallejo, M. E.C., Canales, A., Cassamá, B., Contreras, G., Cordeiro, J.M.M., Dunner, S., Elbeltagy, A., Fioravanti, M.C.S., Carpio, M.G., Gómez, M., Hernández, A., Hernandez, D., Juliano, R.S., Landi, V., Marques, R., Martínez, R.D., Martínez, O.R., Melucci, L., Flores, B.M., Mújica, F., Parés i Casanova, P.M., Quiroz, J., Rodellar, C., Tjon, G., Adebambo, T., Uffo, O., Vargas, J.C., Villalobos, A., Zaragoza, P., Delgado, J.V., Martinez, A., 2019. The genetic ancestry of American Creole cattle inferred from uniparental and autosomal genetic markers. Sci. Rep. 9, 1–16. https://doi.org/10.1038/s41598-019-47636-0.

Green, D.R., 1998. Apoptic pathways: the roads to ruin. Cell 94, 695–698. https://doi.org/10.1016/S0092-8674(00)81728-6.

Greiner, S., Lehwark, P., Bock, R., 2019. OrganellarGenomeDRAW (OGDRAW) version 1.3.1: expanded toolkit for the graphical visualization of organellar genomes. Nucleic Acids Res. 47, W59–W64. https://doi.org/10.1093/nar/gkz238.

Hajnóczky, G., Csordás, G., Das, S., Garcia-Perez, C., Saotome, M., Sinha Roy, S., Yi, M., 2006. Mitochondrial calcium signalling and cell death: approaches for assessing the role of mitochondrial Ca2+ uptake in apoptosis. Cell Calcium 40, 553–560. https://doi.org/10.1016/j.ceca.2006.08.016.

Hedrick, P.W., Allendorf, F.W., Baker, C.S., 2013. Estimation of male gene flow from measures of nuclear and female genetic differentiation. J. Hered. 104, 713–717. https://doi.org/10.1093/jhered/est047.

Hršak, D., Brajkovic, V., Cubric-Curik, V., Curik, I., Ristov, S., 2022a. Update on Magellan, the software for the analysis of a single gender lineage in pedigrees. Int. Conf. Electr. Comput. Energy Technol. ICECET 2022, 1–3. https://doi.org/10.1109/ICECET55527.2022.9872932.

Hršak, D., Katanić, I., Ristov, S., 2022b. A fast method for the selection of samples in populations with available genealogical data. Diversity (Basel) 14, 150. https://doi.org/10.3390/d14020150.

Jenko, J., McClure, M.C., Matthews, D., McClure, J., Johnsson, M., Gorjanc, G., Hickey, J.M., 2019. Analysis of a large dataset reveals haplotypes carrying putatively recessive lethal and semi-lethal alleles with pleiotropic effects on economically

important traits in beef cattle. Genet. Sel. Evol. 51, 1–14. https://doi.org/10.1186/s12711-019-0452-z.

Kennedy, B.W., 1986. A further look at evidence for cytoplasmic inheritance of production traits in dairy cattle. J. Dairy Sci. 69, 3100–3105. https://doi.org/10.3168/jds.S0022-0302(86)80773-1.

Kinoshita, A., Kenéz, Á., Hasselmann, M., Dänicke, S., Huber, K., 2018. Inter-individual variation in adaptive capacity at onset of lactation: linking metabolic phenotype with mitochondrial DNA haplotype in Holstein dairy cows. Sci. Rep. 8, 15439. https://doi.org/10.1038/s41598-018-33853-6.

Kudinov, A.A., Koivula, M., Aamand, G.P., Strandén, I., Mäntysaari, E.A., 2022. Single-step genomic BLUP with many metafounders. Front. Genet. 13, 3266. https://doi.org/10.3389/fgene.2022.1012205.

Kumar, S., Stecher, G., Tamura, K., 2016. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol. Biol. Evol. 33, 1870–1874. https://doi.org/10.1093/MOLBEV/MSW054.

Larsson, L.C., Charlier, J., Laikre, L., Ryman, N., 2009. Statistical power for detecting genetic divergence-organelle versus nuclear markers. Conserv. Genet. 10, 1255–1264. https://doi.org/10.1007/s10592-008-9693-z.

Legarra, A., Aguilar, I., Misztal, I., 2009. A relationship matrix including full pedigree and genomic information. J. Dairy Sci. 92, 4656–4663. https://doi.org/10.3168/jds.2009-2061.

Leigh, J.W., Bryant, D., 2015. POPART: full-feature software for haplotype network construction. Methods Ecol. Evol. 6, 1110–1116. https://doi.org/10.1111/2041-210X.12410.

Li, X., Fang, P., Mai, J., Choi, E.T., Wang, H., Yang, X.F., 2013. Targeting mitochondrial reactive oxygen species as novel therapy for inflammatory diseases and cancers. J. Hematol. Oncol. 6, 1–19. https://doi.org/10.1186/1756-8722-6-19.

Loftus, R.T., MacHugh, D.E., Bradley, D.G., Sharp, P.M., Cunningham, P., 1994. Evidence for two independent domestications of cattle. Proc. Natl. Acad. Sci. U. S. A. 91 https://doi.org/10.1073/pnas.91.7.2757.

Lorenzo, P.Di, Lancioni, H., Ceccobelli, S., Colli, L., Cardinali, I., Karsli, T., Capodiferro, M.R., Sahin, E., Ferretti, L., Marsan, P.A., Sarti, F.M., Lasagna, E., Panella, F., Achilli, A., 2018. Mitochondrial DNA variants of Podolian cattle breeds testify for a dual maternal origin. PLoS ONE 13, e0192567. https://doi.org/10.1371/journal.pone.0192567.

Mannen, H., Kojima, T., Oyama, K., Mukai, F., Ishida, T., Tsuji, S., 1998. Effect of mitochondrial DNA variation on carcass traits of Japanese black cattle. J. Anim. Sci. 76, 36–41. https://doi.org/10.2527/1998.76136x.

Mannen, H., Yonezawa, T., Murata, K., Noda, A., Kawaguchi, F., Sasazaki, S., Olivieri, A., Achilli, A., Torroni, A., 2020. Cattle mitogenome variation reveals a post-glacial expansion of haplogroup P and an early incorporation into northeast Asian domestic herds. Sci. Rep. 10, 20842. https://doi.org/10.1038/s41598-020-78040-8.

McBride, H.M., Neuspiel, M., Wasiak, S., 2006. Mitochondria: more than just a powerhouse. Curr. Biol. 16, 551–560. https://doi.org/10.1016/j.cub.2006.06.054.

Mezzadra, C.A., Melucci, L.M., Corva, P.M., Valiente, S.L., Rípoli, M.V., Lirón, P., Giovambattista, G., 2005. Effects of cytoplasmic inheritance on preweaning traits of Hereford cattle. Genet. Mol. Biol. 28, 357–362. https://doi.org/10.1590/S1415-47572005000300003.

Novosel, D., Brajković, V., Simčič, M., Zorc, M., Svara, T., Cakanic, K.B., Jungić, A., Logar, B., Cubric-curik, V., Dovc, P., Curik, I., 2022. The consequences of mitochondrial T10432C mutation in Cika Cattle: a "Potential" model for Leber's hereditary optic neuropathy. Int. J. Mol. Sci. 23, 6335. https://doi.org/10.3390/ijms23116335.

Novosel, D., Cubric-Curik, V., Simic, M., Dovc, P., Curik, I., 2019. Two detrimental mutations in cattle mitogenome indicate the presence of leber's hereditary optic neuropathy. J. Cent. Eur. Agric. 20, 19–24. https://doi.org/10.5513/JCEA01/20.1.2413.

Oh-Hama, T., 1997. Evolutionary consideration on 5-aminolevulinate synthase in nature. Orig. Life Evol. Biosph. 27, 405–412. https://doi.org/10.1023/A:1006583601341.

Oliehoek, P.A., Bijma, P., 2009. Effects of pedigree errors on the efficiency of conservation decisions. Genet. Sel. Evol. 41, 1–11. https://doi.org/10.1186/1297-9686-41-9.

Olivieri, A., Gandini, F., Achilli, A., Fichera, A., Rizzi, E., Bonfiglio, S., Battaglia, V., Brandini, S., De Gaetano, A., El-Beltagi, A., Lancioni, H., Agha, S., Semino, O., Ferretti, L., Torroni, A., 2015. Mitogenomes from Egyptian cattle breeds: new clues on the origin of haplogroup Q and the early spread of Bos taurus from the Near East. PLoS ONE 10, e0141170. https://doi.org/10.1371/journal.pone.0141170.

Peng, M.S., Fan, L., Shi, N.N., Ning, T., Yao, Y.G., Murphy, R.W., Wang, W.Z., Zhang, Y. P., 2015. DomeTree: a canonical toolkit for mitochondrial DNA analyses in domesticated animals. Mol. Ecol. Resour. 15, 1238–1242. https://doi.org/10.1111/1755-0998.12386.

Petit, R.J., Excoffier, L., 2009. Gene flow and species delimitation. Trends Ecol. Evol. 24, 386–393. https://doi.org/10.1016/j.tree.2009.02.011.

Pun, A., Goyache, F., Cervantes, I., Gutiérrez, J.P., 2012. Cytoplasmic line effects for birth weight and preweaning growth traits in the Asturiana de los Valles beef cattle breed. Livest. Sci. 143, 177–183. https://doi.org/10.1016/j.livsci.2011.09.008.

Qin, Y.H., Chen, S.Y., Lai, S.J., 2012. Polymorphisms of mitochondrial ATPASE 8/6 genes and association with milk production traits in holstein cows. Anim. Biotechnol. 23, 204–212. https://doi.org/10.1080/10495398.2012.686468.

Reverter, A., Okimoto, R., Sapp, R., Bottje, W.G., Hawken, R., Hudson, N.J., 2017. Chicken muscle mitochondrial content appears co-ordinately regulated and is associated with performance phenotypes. Biol. Open 6, 50–58. https://doi.org/10.1242/bio.022772.

Ristov, S., Brajkovic, V., Cubric-Curik, V., Michieli, I., Curik, I., 2016. MaGelLAn 1.0: a software to facilitate quantitative and population genetic analysis of maternal

inheritance by combination of molecular and pedigree information. Genet. Sel. Evol. 48, 1–10. https://doi.org/10.1186/s12711-016-0242-9.

Roughsedge, T., Brotherstone, S., Visscher, P.M., 1999. Estimation of variance of maternal lineage effects at the Langhill dairy herd. Anim. Sci. 68, 79–86. https://doi.org/10.1017/s1357729800050104.

Rozas, J., Ferrer-Mata, A., Sanchez-DelBarrio, J.C., Guirao-Rico, S., Librado, P., Ramos-Onsins, S.E., Sanchez-Gracia, A., 2017. DnaSP 6: DNA sequence polymorphism analysis of large data sets. Mol. Biol. Evol. 34, 3299–3302. https://doi.org/10.1093/molbev/msx248.

Sangland, L.P., Kuehn, L.A., Snelling, W.M., Spangler, M.L., 2022a. Influence of environmental factors and genetic variation on mitochondrial DNA copy number. J. Anim. Sci. 100 https://doi.org/10.1093/jas/skac059.

Sangland, L.P., Snelling, W.M., Kuehn, L.A., Thallman, R.M., Freetly, H.C., Wheeler, T.L., Shackelford, S.D., King, D.A., Spangler, M.L., 2022b. Genetic and phenotypic associations of mitochondrial DNA copy number, SNP, and haplogroups with growth and carcass traits in beef cattle. J. Anim. Sci. 101 https://doi.org/10.1093/jas/skac415.

Schibler, J., Elsner, J., Schlumbaum, A., 2014. Incorporation of aurochs into a cattle herd in Neolithic Europe: single event or breeding? Sci. Rep. 4, 1–6. https://doi.org/10.1038/srep05798.

Schutz, M.M., Freeman, A.E., Lindberg, G.L., Koehler, C.M., Beitz, D.C., 1994. The effect of mitochondrial DNA on milk production and health of dairy cattle. Livest. Prod. Sci. 37, 283–295. https://doi.org/10.1016/0301-6226(94)90123-6.

Sherratt, A., 1983. The secondary exploitation of animals in the old world. World Archaeol. 15, 90–104. https://doi.org/10.1080/00438243.1983.9979887.

Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., Thompson, J.D., Higgins, D.G., 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol. Syst. Biol. 7, 539. https://doi.org/10.1038/msb.2011.75.

Sonstegard, T.S., Cole, J.B., VanRaden, P.M., van Tassell, C.P., Null, D.J., Schroeder, S.G., Bickhart, D., McClure, M.C., 2013. Identification of a nonsense mutation in CWC15 associated with decreased reproductive efficiency in Jersey cattle. PLoS ONE 8, e54872. https://doi.org/10.1371/journal.pone.0054872.

Trifunovic, A., Hansson, A., Wredenberg, A., Rovio, A.T., Dufour, E., Khvorostov, I., Spelbrink, J.N., Wibom, R., Jacobs, H.T., Larsson, N.G., 2005. Somatic mtDNA mutations cause aging phenotypes without affecting reactive oxygen species production. Proc. Natl. Acad. Sci. U. S. A. 102, 17993–17998. https://doi.org/10.1073/pnas.0508886102.

VanRaden, P.M., Olson, K.M., Null, D.J., Hutchison, J.L., 2011. Harmful recessive effects on fertility detected by absence of homozygous haplotypes. J. Dairy Sci. 94, 6153–6161. https://doi.org/10.3168/jds.2011-4624.

Vasan, K., Clutter, M., Fernandez Dunne, S., George, M.D., Luan, C.H., Chandel, N.S., Martínez-Reyes, I., 2022. Genes involved in maintaining mitochondrial membrane potential upon electron transport chain disruption. Front. Cell Dev. Biol. 10, 154. https://doi.org/10.3389/fcell.2022.781558.

Verdugo, M.P., Mullin, V.E., Scheu, A., Mattiangeli, V., Daly, K.G., Delser, P.M., Hare, A. J., Burger, J., Collins, M.J., Kehati, R., Hesse, P., Fulton, D., Sauer, E.W., Mohaseb, F. A., Davoudi, H., Khazaeli, R., Lhuillier, J., Rapin, C., Ebrahimi, S., Khasanov, M., Farhad Vahidi, S.M., MacHugh, D.E., Ertuğrul, O., Koukouli-Chrysanthaki, C., Sampson, A., Kazantzis, G., Kontopoulos, I., Bulatovic, J., Stojanović, I., Mikdad, A., Benecke, N., Linstädter, J., Sablin, M., Bendrey, R., Gourichon, L., Arbuckle, B.S., Mashkour, M., Orton, D., Horwitz, L.K., Teasdale, M.D., Bradley, D.G., 2019. Ancient cattle genomics, origins, and rapid turnover in the Fertile Crescent. Science 365, 173–176. https://doi.org/10.1126/science.aav1002.

Wallace, D.C., 2018. Mitochondrial genetic medicine. Nat. Genet. 50, 1642–1649. https://doi.org/10.1038/s41588-018-0264-z.

Wallace, D.C., 2007. Why do we still have a maternally inherited mitochondrial DNA? insights from evolutionary medicine. Annu. Rev. Biochem. https://doi.org/10.1146/annurev.biochem.76.081205.150955.

Wallace, D.C., Singh, G., Lott, M.T., Hodge, J.A., Schurr, T.G., Lezza, A.M.S., Elsas, L.J., Nikoskelainen, E.K., 1988. Mitochondrial DNA mutation associated with Leber's hereditary optic neuropathy. Science 242, 1427–1430. https://doi.org/10.1126/science.3201231.

Weigel, K.A., VanRaden, P.M., Norman, H.D., Grosu, H., 2017. A 100-year review: methods and impact of genetic selection in dairy cattle—From daughter–dam comparisons to deep learning algorithms. J. Dairy Sci. 100, 10234–10250. https://doi.org/10.3168/jds.2017-12954.

Wredenberg, A., Wibom, R., Wilhelmsson, H., Graff, C., Wiener, H.H., Burden, S.J., Oldfors, A., Westerblad, H., Larsson, N.G., 2002. Increased mitochondrial mass in mitochondrial myopathy mice. Proc. Natl. Acad. Sci. U. S. A. 99, 15066–15071. https://doi.org/10.1073/pnas.232591499.

Xia, X., Huang, G., Wang, Z., Sun, J., Wu, Z., Chen, N., Lei, C., Hanif, Q., 2020. Mitogenome diversity and maternal origins of Guangxi cattle breeds. Animals 10, 19. https://doi.org/10.3390/ani10010019.

Xia, X., Qu, K., Li, F., Jia, P., Chen, Q., Chen, N., Zhang, J., Chen, H., Huang, B., Lei, C., 2019. Abundant genetic diversity of yunling cattle based on mitochondrial genome. Animals 9, 641. https://doi.org/10.3390/ani9090641.

Xia, X.T., Achilli, A., Lenstra, J.A., Tong, B., Ma, Y., Huang, Y.Z., Han, J.L., Sun, Z.Y., Chen, H., Lei, C.Z., Hu, S.M., Chen, N.B., 2021. Mitochondrial genomes from modern and ancient Turano-Mongolian cattle reveal an ancient diversity of taurine maternal lineages in East Asia. Heredity (Edinb) 126, 1000–1008. https://doi.org/10.1038/s41437-021-00428-7.

Ye, K., Lu, J., Ma, F., Keinan, A., Gu, Z., 2014. Extensive pathogenicity of mitochondrial heteroplasmy in healthy human individuals. Proc. Natl. Acad. Sci. U. S. A. 111, 10654–10659. https://doi.org/10.1073/pnas.1403521111.

Young, M.J., Sachidanandam, R., Hales, D.B., Brard, L., Robinson, K., Rahman, M.M., Khadka, P., Groesch, K., Young, C.K.J., 2022. Identification of somatic mitochondrial DNA mutations, heteroplasmy, and increased levels of catenanes in tumor specimens obtained from three endometrial cancer patients. Life 12, 562. https://doi.org/10.3390/life12040562.