LIÈGE université
Gembloux
Agro-Bio Tech

EAST-BELGIAN RED & WHITE (EBRW): ASSESSING GENETIC AND PHENOTYPIC DIVERSITY, AND DEVELOPING PRACTICAL RECOMMENDATIONS FOR THE MANAGEMENT OF GENETIC DIVERSITY AT MATING

THOMAS DEMONTY

TRAVAIL DE FIN D'ETUDES PRESENTE EN VUE DE L'OBTENTION DU DIPLOME DE MASTER BIOINGENIEUR EN SCIENCES AGRONOMIQUES

ANNEE ACADEMIQUE 2020-2021

CO-PROMOTEURS: PR. NICOLAS GENGLER ET PR. JOHANN SÖLKNER

LIÈGE université
Gembloux
Agro-Bio Tech

EAST-BELGIAN RED & WHITE (EBRW): ASSESSING GENETIC AND PHENOTYPIC DIVERSITY, AND DEVELOPING PRACTICAL RECOMMENDATIONS FOR THE MANAGEMENT OF GENETIC DIVERSITY AT MATING

THOMAS DEMONTY

TRAVAIL DE FIN D'ETUDES PRESENTE EN VUE DE L'OBTENTION DU DIPLOME DE MASTER BIOINGENIEUR EN SCIENCES AGRONOMIQUES

ANNEE ACADEMIQUE 2020-2021

CO-PROMOTEURS: PR. NICOLAS GENGLER ET PR. JOHANN SÖLKNER

Acknowledgments

First, I would like to express my gratitude toward my supervisors, Pr. Gengler and Pr. Sölkner for their advice and their time. From the beginning of my studies, their teachings made me learn so much about animal genetics that I am really pleased that they accepted to help me in this master thesis.

Mrs Sylvie Vanderick, Mrs Hélène Wilmot, I would like to thank you for all your advice, your help and your support. Thank you for your infinite patience explaining me further genetic concepts. Thank you for your skills in solving coding issues. Thank you for your continued coaching and all the time you gave to me.

I thank Elevéo for sharing pedigree and genomic data used in this work. A special thank you goes to Mr Glorieux who kindly answered all my questions.

I would like to extend my gratitude to Mrs Lassois and all the jury members that accepted to be warrant of my final step here in GxABT. That means a lot to me.

My thanks also go to all the professors and teaching assistants of GxABT and BOKU that I met during these five years. Their work, but first of all their passion, have always impressed me. With the same passion, they made me discover new topics, new concepts, new tools and new applications that gave bio-engineering a bigger dimension and that made these five years a thrilling journey.

All along the way, friends have always been there. Léna, Marie, François, Pierre and Gilles you were the best fellow travellers that I could have hoped for. I would really like to thank you for all the happiness we shared in moments of joy but also for all the support you offered me in more difficult times. However, five years cannot be summarised with five friends. Thanks to every single person I met, here or in Vienna, because every moment and talk we shared made me grow as a person.

Finally, I do not want to miss the opportunity to address a very special thought to my family that was by my side in every step of my life.

Sophie, without you, this master's thesis would simply not have been completed but the worst thing would have been that our paths would not have crossed one year ago.

Abstract

The objective of this study is to assess the genetic and phenotypic diversity of the East-Belgian Red & White (EBRW) cattle breed. First, pedigree and genomic data were used to have an insight of population structure and to the within-breed genetic diversity by estimating heterozygosity and inbreeding and relationship coefficients. Then, records from milk recording were used in parallel with pseudo-phenotypes, i.e. deregressed estimated breeding values, to display the (pseudo-)phenotypic diversity existing between animals and herds reflecting differences in breeding objectives. This was achieved through principal component analysis. Based on the limited pedigree information, 2% of the EBRW population studied appear to be inbred with an average pedigree inbreeding of 2.3%. The pedigree information also confirmed to use of introgression in the past decades. Turning to genomic data, the observed heterozygosity was 0.358 and more than 40% of the genotyped population was inbred with average inbreeding estimates ranging from 2.9 to 3.5%. Phenotypical variation mostly occurs in regards to the milk yield what might reflect different breeding objectives. This tendency was also confirmed in pseudo-phenotypes. To conclude, practical recommendations were also formulated to manage the EBRW genetic diversity through mating with a focus on the 13 EBRW bulls available for artificial insemination. Finally, the EBRW genetic diversity seems to be relatively high while presenting limited inbreeding level and the phenotypic diversity seems to reflect divergent breeding objectives.

Résumé

L'objectif de ce travail est d'évaluer la diversité génétique et phénotypique de la race bovine Rouge-Pie de l'Est de la Belgique. Premièrement, les pédigrees et l'information génomique ont été utilisés pour obtenir un aperçu de la structure de la population et de la diversité génétique grâce à l'estimation de l'hétérozygotie et des coefficients de consanguinité et de parenté. Ensuite, les phénotypes résultant du contrôle laitier ont été utilisés en parallèle des pseudo-phénotypes, i.e. valeurs d'élevage dérégressées, pour représenter la diversité (pseudo-)phénotypique existant parmi les animaux ainsi qu'entre les troupeaux, ce qui pourrait refléter des objectifs de sélection différents. Cela a été réalisé au moyen d'analyses en composantes principales. Sur base des pédigrees disponibles, 2% de la population étudiée était consanguine avec un coefficient de consanguinité moyen de 2,3%. Les pédigrees ont également permis de confirmer des événements d'introgression survenus dans le passé. Sur base des données génomiques, l'hétérozygotie observée est de 0.358 et plus de 40% de la population génotypée est consanguine avec des valeurs moyennes allant de 2,9 à 3,5% en fonction des estimations. La variation phénotypique est principalement basée sur des différences de production de lait reflétant certainement des objectifs de sélection différents. Cette tendance a été confirmée sur base de pseudo-phénotypes. Enfin, des recommandations pratiques ont été rédigées pour gérer aux mieux les accouplements tout en préservant la diversité génétique de la race, avec un accent sur les 13 taureaux disponibles pour l'insémination artificielle. En conclusion, la diversité génétique au sein de la RPE semble relativement préservée et le taux de consanguinité reste limité.

Table of contents

General introduction 1

Chapter I – Literature review 2

1. EAST-BELGIAN RED & WHITE 2
   1.1. Historical context 2
   1.2. Current status 3
   1.3. Genetic distinctness 3
   1.4. Phenotypic description 4
   1.5. Herd-Book management 5
2. GENETIC DIVERSITY 7
   2.1. Assessing genetic diversity 8
3. PHENOTYPIC DIVERSITY 10
   3.1. Assessing phenotypic diversity 10
4. MANAGEMENT OF DIVERSITY AT MATING 12
   4.1. Contribution of individuals 12
   4.2. Mate allocation 13
5. CONCLUSION 14

Chapter II – Data, material & methods 15

1. PEDIGREE ANALYSIS 15
   1.1. Herd-Book and pedigree data 15
   1.2. Analyses 16
      1.2.1. Pedigree structure and demographic parameters 16
      1.2.2. Control of pedigree completeness 16
      1.2.3. Genetic diversity parameters 18
2. GENOTYPES ANALYSIS 19
   2.1. Data 19
   2.2. Data editing 19
   2.3. Analyses 20
      2.3.1. Quality control 20
      2.3.2. Genetic diversity parameters 21
3. PHENOTYPES ANALYSIS 23
   3.1. Data 23

3.1.1. Phenotypes 23
3.1.2. Estimated Breeding Values 23
3.1.3. Pseudo-phenotypes 24
3.2. Generalized Linear Mixed Model 24
3.3. Principal Component Analysis 25
Chapter III – Results & discussion 26
1. PEDIGREE ANALYSIS 26
1.1. Overview of the population structure 26
1.2. Pedigree quality 28
1.3. Inbreeding coefficient 30
1.4. Pedigree relationship matrix 32
2. GENOTYPES ANALYSIS 33
2.1. Pre-processing and quality control 33
2.2. Heterozygosity 34
2.3. Genomic relationship matrix 34
2.4. Inbreeding coefficients 35
3. PHENOTYPES ANALYSIS 37
3.1. Generalized Linear Mixed Model 37
3.1.1. Descriptive statistics 37
3.1.2. Principal Component Analysis 38
3.2. Pseudo-phenotypes 40
3.2.1. Descriptive statistics 41
3.2.2. Principal Component Analysis 42
4. RECOMMENDATIONS TO PRESERVE THE EBRW DIVERSITY 46
4.1. Maintaining the within breed genetic diversity 46
4.1.1. Reducing genetic drift 46
4.1.2. Using allele flow carefully 47
4.2. Prevent increase in inbreeding 48
4.3. Mating advice based on AI bulls EBVs 50
General conclusions & perspectives 53
References 56
Appendices 66

List of figures

Figure 1. East-Belgian Red & White cow …… 2

Figure 2. Unrooted neighbour-joining consensus tree constructed using Nei's genetic distance of 12 breeds …… 4

Figure 3. Number of births per year in refALL and refHB populations …… 27

Figure 4. Average pedigree completeness over six generation by Herd-Book section …… 29

Figure 5. Schematic representation of mating leading to, a. 0.25 and b. 0.125 pedigree inbreeding coefficient in individual D …… 31

Figure 6. Number of SNPs removed from the SNP panel with different SNP call-rate values …… 33

Figure 7. Correlation graph of production traits based on GLMM herd estimates …… 37

Figure 8. Circle of correlations of the original variables with principal components (PC) 1 and 2 …… 39

Figure 9. Distribution graph of owner herds in the first two Principal Components (PC) space …… 40

Figure 10. Correlation graph of production traits based on pseudo-phenotypes …… 42

Figure 11. Circle of correlations of the original variables with principal components (PC) 1 and 2 …… 43

Figure 12. Distribution graph of cows in the first two Principal Components (PC) space…… 44

Figure 13. Distribution graph of owner herds averages in the first two Principal Components (PC) space …… 45

Figure 14. Breeding circle for six herds …… 48

List of tables

Table 1.	Number of animals registered in the different Herd-Book sections ……	26
Table 2.	Distribution of refHB, ref0.6 and ref0.8 animals within pedigree inbreeding coefficient intervals ……	30
Table 3.	Average pedigree inbreeding in the inbred animals by reference population …….	30
Table 4.	Inbreeding coefficients distribution for the 214 genotyped individuals ……	35
Table 5.	Averaged inbreeding coefficients in the inbred individuals ……..	35
Table 6.	Eigenvalue, percentage of variance and cumulative variance of the principal components (PC) ……...	38
Table 7.	Correlations between variables and the first two principal components (PC) and their respective p-values ……	39
Table 8.	Contributions (%) of variables to the first two principal components (PC) …….	39
Table 9.	Eigenvalue, percentage of variance and cumulative variance of the principal components (PC) ……...	43
Table 10.	Correlations between variables and the first two principal components (PC) and their respective p-values ……	44
Table 11.	Contributions (%) of variables to the first two principal components (PC) …….	44
Table 12.	Number of offspring per EBRW AI bull ……..	47
Table 13.	Pedigree relationship matrix (A matrix) for the 13 AI bulls ……	49
Table 14.	Genomic relationship matrix (GRM) for the 13 AI bulls ……	49
Table 15.	Estimated Breeding Values (EBVs) for the production traits and SCS, expressed in the EBRW publication base ……..	51
Table 16.	EBVs reliability for production traits and SCS for the AI bulls ……..	51

List of abbreviations

AI	Artificial Insemination
A matrix	Numerator relationship matrix
AnGR	Farm Animal Genetic Resource
CompGen	number of Complete Generation
DIM	Days In Milk
DNA	Deoxyribonucleic Acid
DP	Dual-Purpose
DPBB	Dual-Purpose Belgian Blue
DPSF	Dual-Purpose Simmental-Fleckvieh
EABR	European Animal Breeding Regulation
EBRW	East-Belgian Red & White
EBV	Estimated Breeding Value
Elevéo	Walloon breeding association
EU	European Union
FAO	Food & Agriculture Organisation
F_{GRM}	Inbreeding estimate based on the Genomic Relationship Matrix
F_{HET}	Inbreeding estimate based on the excess of homozygosity
F_{PED}	Inbreeding estimate based on the pedigree
F_{UNI}	Inbreeding estimate based on the correlation between uniting gametes
GLMM	Generalized Linear Mixed Model
GRM	Genomic Relationship Matrix
HB	Herd-Book
HD	High Density
H_e	Expected Heterozygosity
H_o	Observed Heterozygosity
HWE	Hardy-Weinberg Equilibrium
IBD	Identical By Descent
ID	Identification
KEM	Campine breed

LD	Low Density
MAF	Minor Allele Frequency
MaxGen	Maximum number of Generation
MD	Mid Density
MR	Milk Recording
MRY	Maas-Rijn-Yssel
MS	Mate Selection
OC	Optimum Contribution
OCS	Optimum Contribution Selection
PC	Principal Component
PCA	Principal Component Analysis
PCI	Pedigree Completeness Index
RDN	Rotbunte Doppelnutzung
ref0.6	Reference population made of refHB animals with a PCI > 0.6
ref0.8	Reference population made of refHB animals with a PCI > 0.8
refALL	Reference population made of animals registered in the EBRW HB and their ancestors
refHB	Reference population made of animals registered in the EBRW HB
RPO	Red-Pied of Ösling
SCC	Somatic Cell Count
SCS	Somatic Cell Score
SNP	Single Nucleotide Polymorphism

General introduction

Currently, the livestock sector plays very important roles economically, socially and environmentally. Worldwide, it generates up to 40% of agricultural gross domestic product and provides incomes and livelihood to more than a billion of people (ILRI, 2021). Socially, communities around the world have always attributed an historical, social and cultural value to their livestock as they have been part of the agricultural development of their region and of the social life of rural populations, while preserving ancient local traditions (Gandini & Villa, 2003). Finally, the livestock sector can have a positive impact on the environment and biodiversity, contributing to landscape protection and to the global agro-biodiversity (Baltussen et al., 2017; Sturaro et al., 2013).

However, several changes, mostly led by new societal expectations, already started to occur in the livestock sector in Western societies. These citizens expectations mostly address the environmental impact of animal production systems and modernisation at the expense of animal's natural behaviour (e.g. zero-grazing) (Boogaard et al., 2011; Dockès et al., 2011). In the case of dairy farming, the general public is pleading for grassland based systems with reduced concentrate feeding and allowing more space for the animals to move, what corresponds to small scale organic farming in the public's mind (Christoph-Schulz et al., 2015).

In that direction, local and dual-purpose breeds seem to show some advantages compared to mainstream breeds such as Holstein. Traditional production system breeding more local breeds are considered as more sustainable regarding the agro-biodiversity but also in terms of carbon sequestration and landscape management (Sturaro et al., 2013). Breeders of local breeds are aware of the positive public perception of their activity and are demanding more promotion effort for their breed from authorities and breeders' association (Gandini et al., 2010). Moreover, it seems that local breeds benefit from a renewed attention as well as an increased awareness from public authorities in regards to their endangerment status (ERFP, 2019; EuReCa, 2010). However, to preserve these breeds, sometimes suffering from loss of genetic diversity due to evolutionary process, efforts are needed both for the breed promotion and the genetic management of remaining populations.

Therefore, this master thesis had two main objectives: (i) to assess the genetic and phenotypic diversity within the East-Belgian Red & White (EBRW) population using different data and (ii), based on what is found previously, to make practical recommendations to preserve the EBRW genetic and phenotypic diversity. The first chapter of this work reviews the literature about genetic diversity and methods for its conservation in livestock and present the EBRW and the context surrounding endangered local breeds. The second chapter presents the data used and computation analyses. Third, the Results & Discussion chapter presents obtained results and their interpretation in terms of genetic diversity conservation for the EBRW.

Chapter I – Literature review

1. EAST-BELGIAN RED & WHITE

The East-Belgian Red & White (EBRW) (**Figure 1.**) is a Belgian cattle (*Bos primigenius taurus* L.) breed located in the eastern part of Wallonia (**Appendix 1.**), especially in the German speaking part, called *Ostbelgien* (East-Belgium).

Figure 1.: East-Belgian Red & White cow (Demonty, 2021).

1.1. Historical context

According to Felius (2016), in the 18th century, the red colour, either pied or not, was preferred in western Europe (Felius, 2016). This may have led to the formation of a red-pied cluster of cattle around 1850 (**Appendix 2.**) (Bouffioux, 2014; Mastrangelo et al., 2020). In the 20th century, when selection really started for Red and White cattle in Belgium, demand for both meat and milk promoted the dual-purpose (DP) breed development (Bay et al., 2009). In 1924, the complete standards of the Belgian Red & White cattle, which were DP animals, were established (EuReCa, 2010). However, two types were coexisting: the one originating from Flanders, the Campine breed (KEM), and the second from East-Belgium, the EBRW (Bay et al., 2009; Colinet et al., 2015). In 1972, the two previously separated Herd-Books (HB) (KEM and EBRW) were merged into a single Belgian Red and White HB (Colinet et al., 2015; François et al., 2017). In the 1990's, holsteinization lead to the disappearance of the original

animals (Bay et al., 2009; Colinet et al., 2015). In addition, in Belgium a large proportion of cattle was not registered in HB, especially in East-Belgium were involvement on breeders in breeder association was impeded by the linguistic differences. As a result of the combination of these two main factors, both EBRW and KEM were considered as extinct.

1.2. Current status

In 2011, a group of few breeders still keeping original EBRW animals and concerted efforts from Elevéo, i.e. the Walloon breeding association, scientists and regional authorities, allowed to re-establish an EBRW HB and the breed was officially recognized in April 2015 (Colinet et al., 2015). Currently, the EBRW is considered as "endangered" according to the European Union (EU) threshold (**Appendix 3.**) as only 590 cows were under Milk Recording (MR) in 2019 (Elevéo, 2019). Moreover in regards to its distribution area, restricted to Belgium only, this breed could be considered as local according to the Food & Agriculture Organisation (FAO) definition (**Appendix 4.**) (FAO, 2007b).

The use of Artificial Insemination (AI) technologies has also been implemented for the EBRW and, until now, 13 EBRW AI bulls were collected.

1.3. Genetic distinctness

The official recognition of the EBRW breed was preceded by a genetic characterisation study which demonstrated the genomic distinctness of EBRW from Holstein and other "sister breeds", i.e. other close red and white breeds (**Figure 2.**), such as the KEM, the Red-Pied of Ösling (RPO) from Luxemburg, the Dutch *Maas-Rijn-Yssel* (MRY) and the German *Rotbunte Doppelnutzung* (RDN) (Colinet et al., 2015). The **Figure 2.**, was achieved by computing the Nei's genetic distances using Single Nucleotide Polymorphism (SNP) (**Appendix 5.**) data from a set of genotypes from 65 EBRW animals (Colinet et al., 2015). Similar results in François et al. (2017) supported the relative distinctness between EBRW and these other breeds (**Appendix 6.**).

Since then, genotyping efforts are going on and, nowadays, more than 200 EBRW have been genotyped. Only males are routinely genotyped as the genomic adherence to the breed genomic pool is a condition to their HB acceptance. Indeed, a genomic assignment tool, required to allow EBRW bulls of unknown origin to be included in the HB, was developed by Wilmot et al. (2021) based on distances of the genotype to be tested to reference genotypes of different breeds.

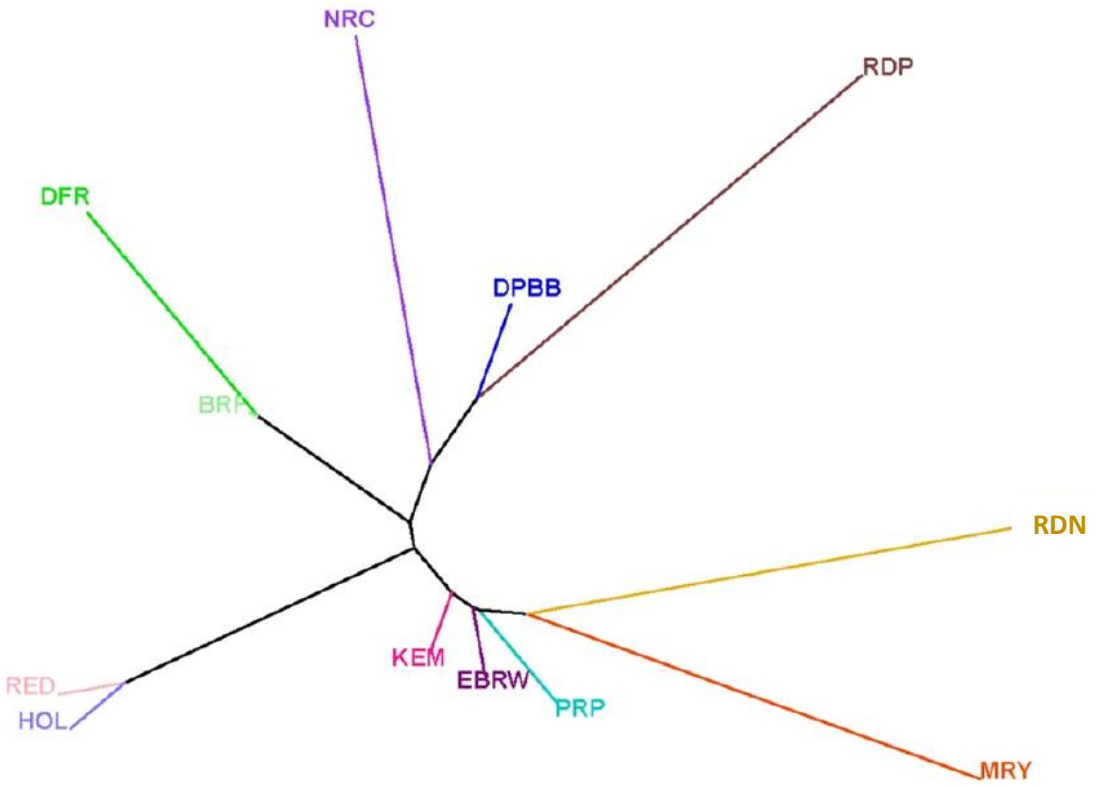


Figure 2.: Unrooted neighbour-joining consensus tree constructed using Nei's genetic distance of 12 breeds. EBRW: East-Belgian Red & White; KEM: Campine breed; PRP: French Red-Pied Lowland; MRY: Maas-Rijn-Yssel; RDN: Rotbunte Doppelnutzung; RDP: Maine-Anjou; DPBB; Dual-Purpose Belgian Blue; NRC: Norwegian Red Cattle; BRF: British Friesian; DFR: Dutch Friesian; RED: Red Holstein; HOL: Holstein. (Colinet et al., 2015)

1.4. Phenotypic description

The EBRW is a DP breed (Colinet et al., 2015), meaning it is kept for both milk and meat production (Syrstad, 1993). Avoiding extremes, DP breeds show intermediate phenotypes compared to purely dairy or beef breeds. Therefore, the milk yield of DP animals is lower than for dairy breeds (Dillon et al., 2003; Haiger & Knaus, 2010; Horn et al., 2013) but often, the milk quality (fat and protein content) is higher, depending on the breed (Kaptijn, 2016). Balancing the lower milk yield, DP animals show a superior carcass quality compared to dairy breeds (De Winter et al., 2010) which leads to higher payments at the slaughterhouse. Breeders also get a better price for calves thanks to higher daily gain compared to dairy calves (Geuder et al., 2012). Another characteristic of DP breeds is that they show superior fitness traits such as fertility, health or longevity compared to Holstein (Evans et al., 2004;

Kaptijn, 2016). Poor information were found on calving ease in DP breeds compared to beef breed, but it seems to be a selection criteria in Dual-Purpose Belgian Blue (DPBB) and Dual-Purpose Simmental-Fleckvieh (DPSF) (Mota et al., 2017; Sölkner et al., 2000)

Complete EBRW breed standards are fully reported in the **Appendix 7.** but the most important points are developed in the next sentences. Height at withers should be between 1m35 and 1m40 for adult cows. The head should be small and short with relatively large muzzle and short horns slightly curved to the front and the inside (**Appendix 8a.**). The coat should be mostly red with a little of white. Both colours are clearly separated. The head should be red: a white spot of variable size is allowed but a completely white head is forbidden. The underbelly and end legs should be white, as well as the tip of tail. The "Burned red" colour is allowed (**Appendix 8b.**). Shoulders, back and hindquarters should be broad and muscular. Legs should be short, robust and muscled.

As EBRW are now milk recorded, statistics on milk yields are available. In 2015, 548 cows were registered in the Walloon MR program. The mean milk production of those cows was 5,061 kg per lactation with 4.16 % fat and 3.47 % proteins (Servais, 2016). Herd means were ranging from 4,500 kg to 6,000 kg milk per cow per year (Servais, 2016). In 2019, 590 EBRW cows were recorded in the MR program with a mean milk production of 5,154 kg per lactation and fat and protein percentage of 4.14% and 3.43%, respectively. As expected, milk yields of the EBRW are clearly lower compared to Holstein but are higher than the DPBB (**Appendix 9.**).

Very poor information is available about the meat production of the EBRW: only a liveweight indication of 600-750 kg for adult cows (Commission Herdbook Rouge-Pie de l'Est, 2017). However, better prices for EBRW calves and culled cows than for dairy animals, resulting from a higher liveweight, were confirmed by an Elevéo agent (F. Thomas, personal communication, July 24, 2020) and a breeder (B. Darimont, personal communication, August 8, 2021).

1.5. Herd-Book management

Since 2015, an EBRW HB has also been implemented and its management scheme is reported under the **Appendix 10.**. As for every breed, the HB is divided into principal and annex sections according to the European Animal Breeding Regulation (EABR) (EU, 2016). However, the EBRW HB benefits from eased access and upgrading – from annex sections to the main section – procedures due to the endangered status of the breed (EU, 2016).

The first noticeable exception is that the EBRW Book B, for animals with unknown origin, is open to males (**Appendix 10.**), meaning that a bull can enter the Book B even when its parents are unknown, if he meets the breed standards and succeeds a breed assignment genomic evaluation (Elevéo, 2021). This exception is aimed to ease the reconstruction of breeds in serious danger of extinction and to avoid the further deterioration of their genetic diversity (EU, 2016).

Second, the upgrading of offspring to the main HB section is usually only allowed through the female line for non-endangered breeds (**Appendix 11.**). However, in endangered breeds (**Appendix 10**.), upgrading can be authorized for progeny from parents (male and female lines) recorded in the main or supplementary sections of the HB (Elevéo, 2021; EU, 2016). This rule is especially useful for breeds in which the number of purebred breeding bulls is low (EU, 2016).

If an animal is accepted in the Book A, a zootechnical certificate can be created in agreement with the EABR (**Appendix 12.**) (EU, 2016).

2. GENETIC DIVERSITY

One of the main objectives of conservation programs for endangered breeds is the preservation of Farm Animal Genetic Resources (AnGR) (EU, 2016; FAO, 2007a). Therefore, in parallel of setting up conservation programs, it is important to make a comprehensive description of genetic diversity (SanCristobal et al., 2006) defined as differences in allele combinations and allele frequencies within and between livestock breeds (Ellegren & Galtier, 2016; Kantanen et al., 2015).

Genetic diversity is continuously increased by mutational process, creating new alleles, and by recombination events, mixing parts of the homologous chromosomes during meiosis, resulting in new allele combinations (Andrews, 2010; Upadhyay, 2019).

Allele frequencies in a population might vary over the following generations due to genetic drift, selection and gene flow (Andrews, 2010; Ellegren & Galtier, 2016), defined below:

Genetic drift is caused by the random sampling of gametes contributing to the next generation in a finite population (Leffler et al., 2012). This may end up with the loss of several alleles and the fixation of others, therefore lowering genetic diversity (de Rochambeau et al., 2000). The smaller the population, the higher the risk of genetic drift, especially for rare alleles. As endangered breeds are small populations, genetic drift might be strong, which could reduce the genetic diversity as well as the fitness of the population (Kohn et al., 2006).

Selection allows a higher reproduction rate of individuals showing specific traits. Alleles carried by individuals more likely to reproduce will therefore be transmitted more heavily in the next generation. Therefore, frequency of the favoured alleles will increase over time at the expense of others leading to the reduction of genetic diversity (Leffler et al., 2012).

Allele flow is the movement of alleles into or out of a population (Andrews, 2010). It can increase genetic diversity of an almost isolated population as it brings new alleles in the population. In a breeding context, allele flow can be achieved through crossbreeding, what might affect the breed integrity, or by increasing exchange of breeding animals between breeders (Windig & Kaal, 2008).

2.1. Assessing genetic diversity

To investigate the genetic diversity of populations, increased or decreased by the process explained in the previous section, several methods are available. Below are presented three of them frequently used in animal breeding studies: (i) heterozygosity, (ii) inbreeding and (iii) genetic relationship.

Heterozygosity is defined as the proportion of loci, in the genome, being heterozygous, i.e. for which alleles carried by homologous chromosomes are different (Altshuler, 2012; Griffiths et al., 2008). It is therefore a direct measure of the genetic variation in individuals that can be averaged across the population. The expected heterozygosity (H_e), also called gene diversity, defined in Nei (1973), is a widely used parameter to measure diversity within populations (Abebe et al., 2015; Eusebi et al., 2020; Toro & Caballero, 2005). Under Hardy-Weinberg Equilibrium (HWE) conditions (no mutation, no selection, no migration, infinite population and random mating), expected proportions of heterozygous genotypes can be statistically determined as $H_e = 1 - \sum_{i=1}^{I} p_i^2$ with I the number of different alleles at a locus, and p_i , with $(i = 1, 2, \dots, I)$, is the expected frequency of allele i in the population (Harris & DeGiorgio, 2017). In contrast, for real populations that are not under the HWE conditions, the observed heterozygosity (H_o) is often estimated as the average heterozygosity across a panel of genetic markers, such as SNPs (**Appendix 5.**) (Coulon, 2010). Knowing the homozygous or heterozygous state of each SNP, H_o can be estimated as the proportion heterozygous SNPs over the total number of genotyped SNPs (Salazar et al., 2015).

Inbreeding was first defined as the correlation between homologous alleles at a random locus within a diploid individual (Wang, 2016; Wright, 1922) but it can also be defined as the probability for two alleles at the same locus to be Identical By Descent (IBD) i.e. derived from a common ancestor (de Rochambeau et al., 2000; Malécot, 1948; J. Wang, 2016). Both alleles at this locus are therefore in a homozygous state. As a consequence, inbred animals, resulting from the mating of related parents, are further decreasing genetic variation of the population. Inbreeding can originate from a low number of founders or from a current small population size (de Rochambeau et al., 2000) and these two factors can be combined, which might particularly be the case for endangered local breeds. However, inbreeding can also occur in breeds, like Holstein, with very large population size when only a small number of elite animals are heavily used (Miglior & Burnside, 1995). Besides a reduction in genetic diversity, inbreeding also causes a reduction of average individual fitness and performances called inbreeding depression (Kareiva & Floberg, 2008). A high inbreeding level can lower the long-term fitness of a population (Senczuk et al., 2020). It is therefore important to assess and monitor the inbreeding coefficient of a population for conservation purposes (Senczuk et al., 2020). The different

F estimates therefore gives information about the genetic diversity within a breed (Makina et al., 2014) and can be based on pedigree or genomic data (Bérénos et al., 2016). Even if the inbreeding coefficient was traditionally calculated based on pedigree data (F_{PED}), the genomic SNP-based estimates show better results (Gazal et al., 2014; Kardos et al., 2015). Indeed, F_{PED} does not account for recombination events and suffers from pedigree incompleteness, low depth and recording errors (Cassell et al., 2003; J. Wang, 2016).

As explained above, an animal is inbred when its parents are related. Therefore, knowing, a priori, the degree of relatedness between individuals allows to predict and prevent inbreeding and the resulting decrease in genetic diversity (Li et al., 2011; Strandén & Peura, 2007). Individuals sharing common ancestors are genetically related as they inherited alleles from these ancestors (Weir et al., 2006). Genetic relationship was first defined as the correlation between homologous alleles randomly chosen from each of two individuals (Wang, 2016; Wright, 1922) but it can also be defined as the probability of IBD alleles shared between two individuals because of their common ancestors (Bourrat, 2021; Malécot, 1948; Weir et al., 2006). The relationship between two animals can be measured by the relationship coefficient either based on pedigree or genomic data (Li et al., 2011; Strandén & Peura, 2007; Wang, 2016). It can also be given at population scale by averaging the pairwise coefficients across the whole set of animals (Addo et al., 2019; Carolino et al., 2020; Strandén & Peura, 2007).

3. PHENOTYPIC DIVERSITY

The phenotypes are quantitative or qualitative observable characteristics of an individual, determined by its genotype and by environmental influences (Baye et al., 2011; Griffiths et al., 2008). The phenotypic diversity therefore refers to the variation of phenotypes occurring between and within breeds (Andersson, 2001; Rahim et al., 2008). Globally, livestock breeds show a high phenotypic diversity resulting from artificial selection, previously only based on phenotypes, in different environments and for different objectives (Andersson, 2001). In local DP breeds, such as EBRW, this phenotypic diversity might be especially high as clear phenotypic breed standards have been missing for years, as well as global breeding objectives.

3.1. Assessing phenotypic diversity

To assess the phenotypic diversity, phenotypes are needed and might include production as well as morphometric measures. Globally, recording for phenotypic traits is very important in animal breeding as it is a mandatory step for any genetic improvement programs (Andersson, 2001; Cole et al., 2020). However, in many local breeds as in the EBRW, phenotypic data are scarce as this breed suffered from poor interest from breeders' association and low involvement of breeders into data recording services before the breed recognition in 2015. Therefore, MR measures (milk yield (kg/day), fat percentage, protein percentage and Somatic Cell Count (SCC)) are often the only phenotypes available for such breeds. However, additional data may be available in terms of calving, fertility, or longevity traits. A critical element in DP breeds is the availability of morphological data. This would be an important further step because it can help to select for better udders, but also better DP adapted morphology. Unfortunately, appropriate linear classification grids, such as the one developed for the DPBB (**Appendix 13.**), are not common. In DP breeds, beef traits are mostly missing as fattening of animals is done outside of the herds where animals are born. Therefore, future advances in the access to slaughter data may also allow to assess the phenotypical diversity.

As phenotypic traits might be numerous, multivariate analysis is a powerful tool to assess the phenotypic diversity of a breed (de Barros Nascimento de Medeiros et al., 2020; Gomes Arandas et al., 2017). Indeed, it gives a data overview considering all variables together allowing an easier exploration of data structures and patterns by eliminating variables that contribute less to variation (de Barros Nascimento de Medeiros et al., 2020; Gomes Arandas et al., 2017).

The multivariate analysis of phenotypic diversity can be used for precise breed phenotypic characterisation which is a first important step in conservation of local breeds for which precise breed standards might have been missing (FAO, 2011; González Ariza et al., 2021).

Phenotypic diversity and phenotypes recording is also required for selection and breed improvement programs (**Appendix 14.**) (Cole et al., 2020; Suhardi et al., 2020; van de Graaf, 2015) which might be key elements for the long-term survival of small local breeds as stated in Biscarini et al. (2015).

4. MANAGEMENT OF DIVERSITY AT MATING

Mating is a cornerstone for the diversity conservation. Indeed, preserving diversity is all about transmitting alleles to the next generation through mating.

4.1. Contribution of individuals

The first step in population management, is to choose individuals to reproduce and set their contribution to the next generation, i.e. the number offspring they will produce (Fernández et al., 2011; Yoshida et al., 2020).

The Optimum Contribution (OC) strategy determines the contributions of individuals in regards to their relationship coefficient with each other, which has to be minimized (Fernández et al., 2011; Meuwissen & Sonesson, 1998). The idea is to penalise individuals that are closely related to the rest of the population so that their contribution to the next generation will be lowered (Fernández et al., 2011). However, this strategy has not as objective to improve the population on a phenotypic level. Indeed, as the OC strategy does not account for genetic gain, it will not help improve the low production level of local breeds that is discriminating them against high-yielding international breeds (Biscarini et al., 2015; Gandini et al., 2010).

The Optimum Contribution Selection (OCS) maximizes the genetic gain of the next generation while limiting inbreeding to a chosen rate from one generation to the next (Fernández et al., 2011). While the OCS can be used either to improve breeds or for conservation purposes, its effectiveness in the context of the breeding of small endangered breeds was proven (Nielsen & Kargo, 2020; Schierenbeck et al., 2011; Solé et al., 2013). However, if introgression of alleles from another breed resulting from crossbreeding occurred in a local breed, the use of OCS for upgrading purposes could lead to an increased selection on introgressed alleles as exogenous alleles increase the rate of genetic gain and decrease the average relationship of the population.

By adding new constraints on migrant contribution and native founders' alleles, the Advanced OCS enables to recover the genetic originality of an admixed breed by reducing migrant contribution (Kohl et al., 2020; Y. Wang et al., 2017).

4.2. Mate allocation

The second step in population management is to choose the mating scheme, i.e. the way selected animal are to be mated together (Fernández et al., 2011; Yoshida et al., 2020). For example, OCS is usually applied in combination with mating strategies such as "minimum relationship" but before addressing this particular strategy, (i) random, (ii) hierarchical and (iii) factorial mate allocation strategies will be briefly explained.

Random mating means mating randomly individuals within the population (APS, 2004). In this situation, which is one of the HWE conditions, no change occurs in the inbreeding coefficient and H_o should tend to H_e (APS, 2004). However, random mating rarely occurs in livestock breeding because of breeding objectives.

In most breeding schemes, mating is hierarchical, meaning that one sire is mated to multiple dams, one or several times, which is convenient for breeders (Fernández et al., 2011). However, hierarchical mating can lead to a high increase in inbreeding coefficient if several sibs are selected from best families (Fernández et al., 2011; Koenig & Simianer, 2006; Stachowicz et al., 2011).

Factorial mating schemes, where parents of both sexes are randomly mated to several different individuals of the other sex, produce a higher proportion of half-sibs compared to hierarchical design (Fernández et al., 2011; Woolliams, 1989). Indeed, as attention is paid to mate cows to different bulls, there is less chance for a cow to be mated several times with the same bull and thus a lower proportion of full-sibs. Therefore, the chance to select full-sibs is lowered which leads to lower inbreeding compared to hierarchical mating (Woolliams, 1989).

The "minimum relationship" scheme directly avoids inbreeding by mating the least related individuals together using the relationship coefficient (Fernández et al., 2011). The idea is then to avoid mating of individuals showing a relationship coefficient higher than a selected threshold (Fernández et al., 2011).

Finally, Mate Selection (MS), provides an option to merge the selection and contribution step with the mate allocation step (Fernández et al., 2011; Yoshida et al., 2020). Using MS requires a complex objective function that might account for technical, logistical and cost issues as well as genetic gain, genetic diversity, progeny inbreeding and use of reproductive technologies (Yoshida et al., 2020). In the end, MS provides a ranked mating list suggesting pairs for mating that would maximize the objective function (Kinghorn, 2011). The MS strategy is of high interest in AnGR management as it is a powerful tool to minimize rate of inbreeding in conservation programs of critical and endangered breeds (Kremer et al., 2010).

5. CONCLUSION

As the EBRW has been presented as well as the context surrounding the breed, it appears that the EBRW population is still endangered even though efforts were made and are still going on for its conservation. In that direction, assessing the genetic and phenotypic diversity of the breed might be a relevant action for the further preservation of the EBRW. The different genetic diversity estimates that were presented in this chapter applied to the EBRW population could give information about its current genetic diversity level. The multivariate analysis of EBRW phenotypic data could also give an insight of the breed phenotypic diversity and of breeding objectives of owners. In the end, some conservation measures could be formulated based on the new information available.

Chapter II – Data, material & methods

1. PEDIGREE ANALYSIS

Through pedigree analysis two groups of information can be reached. First, it allows to have an overview of the population structure (Carolino et al., 2020). Then, if pedigree data are sufficiently complete, pedigree analysis allows to have an insight of the genetic diversity of the population using, e.g.: inbreeding coefficient, number of founders or relationship estimators (Wellmann, 2021b).

1.1. Herd-Book and pedigree data

A list of 3,426 animals registered in the EBRW HB containing animals' IDs and their attributed HB section was provided by Elevéo. These HB animals have been chosen to set a reference population used for further pedigree analysis in this work and will be referenced as the "refHB" population.

The pedigree information (animal's ID, ID of their sire and dam) of these EBRW animals was available from the all-breed pedigree file shared by Elevéo with Gembloux Agro-Bio Tech for the genetic evaluations. This pedigree file also contains, among others, two breed indications, sex and birthdate of the animals. From the two breed indications, called "breed composition" and "breeding activity", the first one gives the proportion of different breeds entering in the composition of the individual. It is half the breed composition of each parent. The second breed indication, the breeding activity, is a single code corresponding to the breed accounting for more than 50% of the breed composition of the animal. If no breed in the breed composition reaches a proportion higher than 50%, the breed referred by the breeder in the birth declaration is taken as the 'breeding activity'.

Information about breeders and current owners of the animals, were also added to the pedigree information.

Then, an upward pedigree extraction was performed for the 3426 refHB animals. This procedure extracts from the all-breed pedigree file all the recorded ancestors refHB. The refHB population and their extracted ancestors will be referenced as the "refALL". For refALL, a file containing IDs of the animals, ID of their sires and dam, year of birth, breed composition and breeding activity, country of origin and HB section was created to have an insight in the population structure.

1.2. Analyses

1.2.1. Pedigree structure and demographic parameters

The refALL pedigree was used to investigate the pedigree structure. Indeed, even if the official EBRW breed is quite new, genealogies records are not. Therefore, looking at EBRW ancestors might be helpful to understand some aspect of the breed origin, what will also influence the genetic diversity. Simple demographic parameters (number of animals, number of founders, sex ratio) were computed for refALL with CFC v1.0 software (Sargolzaei et al., 2006) and were displayed by year of birth and country of origin.

1.2.2. Control of pedigree completeness

As a prior to genetic diversity indicators computation, such as inbreeding and relationship coefficients, the quality of the refHB individual's pedigree will be checked through the computation of different pedigree completeness parameters presented below. To compute these parameters for refHB individuals, refALL is used as it contains refHB animals plus their ancestors.

The Maximum number of Generation (MaxGen) is the number of generations traced back between an individual and its further known ancestor (Gebremariam, 2013; Wellmann, 2021b). For example, if the MaxGen value of an animal equals 10, it means that 10 generations separate this animal from its older known ancestor.

The number of Complete Generation (CompGen) is the number of generations traced back with no ancestors missing in any upward generation (Gebremariam, 2013; Wellmann, 2021b).

The pedigree completeness is computed as the averaged proportion of known ancestors in each ancestral generation over a defined number of generation (MacCluer et al., 1983; Wellmann, 2021b). In the second generation, for example, if the 4 grand-parents are known, the completeness for this generation would be 1. If only four of the great-grand-parents were known out of eight, the completeness of the third generation would equal 0.5. Here, a limit of six generations was set. As some EBRW ancestors are American and Canadian Holsteins, this limit is relevant to reduce bias in the computation of EBRW pedigree completeness due to very complete American and Canadian Holstein

pedigree information. The pedigree completeness will be displayed as a mean within HB sections in order to see the relationship between the HB management and the quality of the pedigree recording.

The Pedigree Completeness Index (PCI), is described in (Wellmann, 2021b) as a useful parameter to identify individuals with insufficient pedigree information to estimate inbreeding. Indeed, inbreeding can only be detected if both maternal and paternal ancestries are known (Wellmann, 2021b). Therefore, the PCI is computed as the harmonic mean of the contribution of the parents to give more weight to less complete parts pedigree (Groeneveld et al., 2009; MacCluer et al., 1983; Wellmann, 2021a). The formula is first given in (MacCluer et al., 1983) as follow:

$$PCI_d = \frac{4 \cdot C_{d_{pat}} \cdot C_{d_{mat}}}{C_{d_{pat}} + C_{d_{mat}}}$$

where $C_{d_{pat}}$ and $C_{d_{mat}}$ are contributions from the paternal and maternal lines respectively:

$$C_{d_{par}} = \frac{1}{d} \sum_{i=1}^{d} a_i$$

Where, par is either maternal (mat) or paternal (pat), a_i is the proportion of known ancestors in generation i and d is the number of generations considered in the calculation of pedigree completeness. For example, if the paternal contribution $C_{d_{pat}}$ of an animal equals zero because the sire is unknown, the PCI value for this animal will be zero. Therefore, the further inbreeding estimation will be really weak as no common ancestors could be detected between the maternal and paternal lines.

Two PCI threshold will be set at 0.6 – as a compromise between a better inbreeding coefficient estimate while keeping enough animals (Li et al., 2011) – and 0.8 to have a more reliable inbreeding estimation (Gautason et al., 2021). The two subpopulations drawn from refHB when PCI thresholds were set will be referred as ref0.6 and ref0.8. The same computations will also be run in parallel with the whole refHB population with no PCI threshold.

The pedigree quality of the refHB population can then be assessed through the average of those parameters, computed using the OptiSel v2.0.5 R package (Wellmann, 2021a) in RStudio v1.4.1106 (RStudio Team, 2021).

1.2.3. Genetic diversity parameters

To have an insight into the EBRW genetic diversity, two parameters will be computed for each of the three sets refHB, ref0.6 and ref0.8 using OptiSel v2.0.5 (Wellmann, 2021a) and AGHmatrix R packages (Amadeu et al., 2016).

First, the inbreeding coefficient, a relevant indicator of the genetic diversity of a population was computed. Indeed, as inbreeding reduces the heterozygosity level in a population, the higher the inbreeding coefficient, the lower genetic diversity. Therefore, the pedigree based inbreeding coefficient (F_{PED}) will be computed individually and as average of the inbreds.

Second, the numerator relationship matrix (A) was computed (Amadeu et al., 2016). This symmetric matrix contains additive genetic relationships between all pairs of individuals in a population. The additive genetic relationship between two individuals reflects the proportion of IBD alleles shared between two individuals. Relationships between individuals is a precious information for mating in a conservation context.

2. GENOTYPES ANALYSIS

A genotype is the DNA sequence of an individual (Austin, 2021). Genotyping several animals using SNP markers provides information on DNA variations at different specific locus of the genome within the population. Computations, using specific programs, can then be performed based on genotypes to estimates genetic diversity of the animals.

2.1. Data

Until now, EBRW animals have been genotyped in the context of the breed revival (Colinet et al., 2015), the breed assignment program (Wilmot et al., 2021) and paternity testing (Elevéo, 2019). A total of 317 animals were sent by Elevéo for genotype analysis but only 254, recorded in EBRW HB, were kept in this work to be sure of their breed status.

For this breed, genotyping is performed either with Mid Density (MD) SNP chips (BovineSNP50 v2 (Illumina, 2010); BovineSNP50 v3 (Illumina, 2020); EuroG MD (EuroGenomics, 2019)) for breed assignment or with Low Density (LD) SNP chips (EuroG 10K v7 (EuroGenomics, 2019)) for paternity testing, respectively. The MD SNP chips genotype around 50k SNPs evenly spaced on the genome while LD SNP chips only genotype 10k SNPs which lower the information quality. A total of 56,059 SNPs were genotyped by these two chips.

2.2. Data editing

Two data files were used for this work, a genetic map of the SNPs, containing the ordered list of SNPs detected by the chips and a genotype file, containing a numerical SNP code (**Appendix 15.**) for each of the 56,059 SNPs genotyped. The SNP code correspond to the state – homozygous, heterozygous –of the SNPs.

However, these raw data needed some editing steps to be suitable for further analysis. First, EBRW individuals genotyped with the 10k chip were discarded as the SNP density was judged insufficient in regard to the analysis that will be performed later on (Mészáros et al., 2015; Zhang et al., 2015). Therefore, out of the 254 EBRW genotypes available, only 216 genotyped with MD chip were kept for

further analysis. Then, for the input files to be used in PLINK v.1.9 (Chang, 2021; Chang et al., 2015; Purcell et al., 2007), specific .ped and .map file formats are required. Therefore, the numerical SNP codes were transposed into alphanumerical AB allele codes (**Appendix 15.**). The .ped text file was, finally, transformed into a binary file with the --make-bed PLINK function and undergone a quality control with PLINK.

2.3. Analyses

2.3.1. Quality control

First, the available .map file contains all genotyped SNPs from both LD and MD chips. However, samples are only genotyped with one or the other type of SNP chip and therefore should only get a value for the SNP covered by the particular SNP chip used. Unfortunately, following PLINK requirement, the genotype file needed a value for every SNP of the map file (SNPs genotyped by both LD and MD chips). Therefore, missing values were attributed to SNPs not covered by the chip used for the sample of interest. This induces a bias in the percentage of missing SNPs per sample. To solve this problem, knowing that only samples genotyped with MD chips were kept, SNPs covered only by the LD chip needed to be removed of the SNP panel of the map file. To remove SNPs covered only by the LD chip a first SNP call-rate was applied.

The call-rate for a given SNP is the proportion of genotypes for which the corresponding SNP information is not missing (Reed et al., 2015). Different call-rate thresholds were tested, from 0 to 100% in 5% increments, to best fit to our dataset and eliminate LD SNPs. Next, to discard animals with poor genotype quality, samples were filtered with an individual call-rate, which is the proportion of SNPs for which information is not missing (Reed et al., 2015). Individuals with a proportion of missing or unreliable genotypes above the fixed threshold are removed (Reed et al., 2015). A second SNP call-rate filtering was performed, after removing unreliable animals, to discard remaining genotype errors that occurred within the MD chip SNP panel.

A filtering on Minor Allele Frequency (MAF) was performed with a value of 0.01 with the PLINK "--maf" function. This threshold allows to remove genotyping errors in SNPs that are monomorphic in the population (Anderson et al., 2010; Trujano-chavez et al., 2021) while keeping as much genetic variation as possible which is important for genetic diversity studies.

Finally, alleles can be correlated due to both inbreeding and linkage disequilibrium. Therefore, to avoid taking account for linkage disequilibrium based homozygosity in the inbreeding estimates, Chang (2021) recommends to perform a linkage disequilibrium pruning before using the "--het " and "--ibc" functions with the "--indep-pairwise" function. The following settings, also used for the KEM breed by François et al. (2017), were applied: window size of 50 SNPs, step size of 5 , pairwise r^2 of 0.2. A correlation (r^2) between two SNPs is considered as linkage disequilibrium if higher than 0.2.

2.3.2. Genetic diversity parameters

The samples and SNPs that passed the quality control were used for further analysis with the PLINK software (Chang, 2021; Chang et al., 2015; Purcell et al., 2007) and AGHmatrix R package (Amadeu et al., 2016).

First indicators of the within breed genetic diversity that were presented in the literature review are H_e and H_o. The "-- het" PLINK function computes the observed homozygosity for each sample as the proportion of homozygous SNPs over the whole set of SNPs analysed. The expected homozygosity is also given by this function and is estimated based on loaded MAFs through the "--read-freq" function. As an allele is either in homozygous or heterozygous state, H_o and H_e are therefore accessed as the observed or expected homozygosity minus one, respectively.

Second, the Genomic Relationship Matrix (GRM) is similar to the A matrix (cf. section 1.2.3 of this chapter), but based on genomic data. The method proposed by VanRaden (2008), based on estimated allele frequencies from SNP data, was chosen in the AGHmatrix R package (Amadeu et al., 2016) as relationship estimates are scaled to allow comparison with the pedigree relationships.

Third, inbreeding was assessed through three estimates. The first inbreeding coefficient estimate is based on the excess of homozygosity in the genomes of individuals that can reflect inbreeding. The "--het" function calculates a first inbreeding coefficient F_{HOM} as:

$$F_{HOM} = \frac{O - E}{N - E}$$

with O, the observed number of homozygous markers of the individual; E, the expected number of homozygous markers under HWE conditions calculated from the allele frequencies estimated on the sample; and N, the total number of markers without missing values for the individual (Gazal et al., 2014).

Then, through the "--ibc" PLINK function, two other individual inbreeding coefficients estimates were produced: Fhat1 and Fhat3. The Fhat1 estimate (renamed F_{GRM}) is based on the GRM (Chang, 2021). As diagonal elements of the GRM equal $1 + F_i$ where F_i is the inbreeding coefficient of the animal i, F_{GRM} is individual's relationship to itself minus 1 (Zhang et al., 2015). The Fhat3 coefficient (renamed F_{UNI}) is based on the correlation between uniting gametes (Chang, 2021) and is defined in Zhang et al. (2015) as:

$$F_{UNI} = \frac{x_i^2 - (1 + 2p_i)x_i + 2p_i^2}{2p_i(1 - p_i)}$$

Where x_i is the number of copies of the reference allele for the ith SNP and p_i is the observed population-wide allele frequency of the reference allele (i.e., the allele whose homozygous genotype was coded as "0") of the ith SNP.

3. PHENOTYPES ANALYSIS

The diversity within the EBRW breed was also investigated through phenotypes as they partly reflect genetic background of the animals using the model P=G+E where P is the estimated phenotypic value, G is the part of P explained by the genotype, E is the deviation of P from observed phenotypic values explained by the environment (Baye et al., 2011).

3.1. Data

3.1.1. Phenotypes

The first data used are observed phenotypes for 6 different traits available for the 1647 milk recorded refHB cows. The traits used in this study are (i) milk yield (kg/day), (ii) fat percentage (%), (iii) protein percentage (%) and (iv) SCC (cells/mL). From these values, the (v) fat and (vi) protein quantities (g/day) were also drawn.

3.1.2. Estimated Breeding Values

Estimated Breeding Values (EBV) are estimates of G from the P=G+E model, i.e. it estimates the genetic potential of animals. In Wallonia, EBVs are based on a random regression test-day model and are given as an equivalent production on a 305 Days In Milk (DIM) basis, averaged over the first three lactations (Vanderick et al., 2020).

Usually, EBVs are expressed relative to the average of a reference population which was settled, here, as all HB registered EBRW cows born in 2015 with records for production traits. For these animals, milk, fat and protein quantities EBVs were averaged and the mean was subtracted from raw EBVs of every animal. For fat and protein percentages, the following conversion equations were used (Elevéo, 2021):

$$EBV_{\%FAT} = \frac{(EBV_{KGFAT} \times 100) - (EBV_{KGMILK} \times \%P_{FAT})}{(EBV_{KGMILK} + P_{KGMILK})}$$

$$EBV_{\%PROT} = \frac{(EBV_{KGPROT} \times 100) - (EBV_{KGMILK} \times \%P_{PROT})}{(EBV_{KGMILK} + P_{KGMILK})}$$

Where $EBV_{\%FAT}$ and $EBV_{\%PROT}$ are EBVs for fat and protein percentages, respectively, of the new base, EBV_{KGMILK} , EBV_{KGFAT} and EBV_{KGPROT} are production EBVs of the new base and P_{KGMILK} , $\%P_{FAT}$ and $\%P_{PROT}$ are phenotypic means for milk production and percentage of fat and protein.

For somatic cells, the SCC was further transposed in Somatic Cell Score (SCS) using the following equation (Vanderick et al., 2020) before computing EBVs for SCS using the same procedure as for production EBVs:

$$SCS = [log_2(SCC \ / \ 100.000)] + 3$$

3.1.3. Pseudo-phenotypes

To display distribution of animals reflecting their phenotypes, it was chosen to work with deregressed EBVs, called pseudo-phenotypes. Indeed, phenotypes were not directly usable because several measures were taken for each animal and because those different measures were taken at different moment of the lactation curve from one cow to the other and are therefore not comparable. On the other hand, the regressed nature of the EBVs might introduce bias resulting from differences in reliabilities. Therefore, EBVs were divided by their reliabilities, i.e. deregressed, to produce pseudo-phenotypes to work with. Pseudo-phenotypes were available for six traits: (i) milk quantity (kg), (ii) fat quantity (kg), (iii) fat percentage, (iv) protein quantity (kg), (v) protein percentage and (vi) SCS. Pseudo-phenotypes that were out of a three standard deviations from the mean range were considered as outliers and discarded (Niero et al., 2016).

3.2. Generalized Linear Mixed Model

Based on phenotypes (cf. section 3.1.1. of this chapter), Generalized Linear Mixed Models (GLMM) were set to investigate the factors affecting phenotypes. The idea was to determine if significative differences arose between animals from different owners as animals were taught to reflect the breeding objectives of their owner.

The data used were 18,065 test-day records for milk (kg), fat (g), fat percentage, protein (g), protein percentage and SCS of the refHB cows registered at MR. These production variables were used as the dependant variables. The animal was set as a random factor while the fixed effects were (i) owner (23 levels), (ii) the year of the test (from 2010 to 2020; years 2010 to 2013 with few records were merged into 2014), (iii) the month of the test, (iv) the interaction between the lactation (ranging from 1 to 3) and the DIM classes (12 classes of 30 DIM from 0 to 365 DIM). The SAS proc MIXED was used as it seems to be a very efficient procedure for repeated measures (Hamer & Simpson, 2000). To run this model, only cows with more than one observation and owners with more than 20 cows and 100 records were kept (Ray et al., 1992) because owners with more cows are the most active in EBRW conservation and selection.

3.3. Principal Component Analysis

Two Principal Component Analyses (PCA) were performed using the R package FactoMineR v2.4 (Husson et al., 2020). The first one used the GLMM estimates for the 6 production traits, corrected for fixed factors (see previous section) and averaged by owner. Trying to get rid of the environmental factors affecting phenotypes, pseudo-phenotypes for the same six production traits were used for a second PCA. In both cases, data were scaled because of different units and standard deviation between variables. The goal of these analyses was to visualize the distribution of animals that might reflect difference in breeding objectives of the owners. Therefore, graph would display individuals' distribution as well as centroids of the most important owners. Pertinence of these analyses was assessed, a priori, by computing correlations between the different variables. Strong positive correlations between certain variables would justify the use of a PCA (Gomes Arandas et al., 2017). The number of dimensions was chosen based on eigenvalues explaining a minimum of 70% of cumulated variance (Gomes Arandas et al., 2017; Kaiser, 1961).

Results & discussion

1. PEDIGREE ANALYSIS

1.1. Overview of the population structure

The list of 3,426 animals registered in the HB of EBRW, provided by Elevéo, was used as the reference population for pedigree analysis. The number of animals registered in each HB section is presented in **Table 1.**.

Table 1.: Number of animals registered in the different Herd-Book sections.

Herd-book section	Count
A	67
B	2,456
B1	859
B2	6
C	37
C1	1
Total	**3,426**

Few animals are registered in the main HB section (book A) what was expected considering the HB management procedures and the short period of time since the HB creation (cf. Chapter I, section 1.2.). Most of the animals are therefore registered in the HB annex sections (Book B and C). These sections can contain animals with less EBRW breed percentage as they might have unknown origins or even ancestors from other breeds such as Holstein, DPSF or Improved Red. In order to check if a filtering based on breed composition would allow to select only "purebred" HB animals, the EBRW breed composition percentages were checked (**Appendix 16.**). However, it rapidly appears that the breed composition was not reliable to select purebred animals.

The refALL population contains 17,266 animals (the 3,426 refHB animals and their 13,840 ancestors) born between 1910 and 2020. The number of births per year since 1960 is presented in the **Figure 3.** with distinction for refHB animals. Births before 1960 are not displayed because it is mainly foreign animals and because some birth dates were, at some point, attributed by default to animals with unknown birth date.

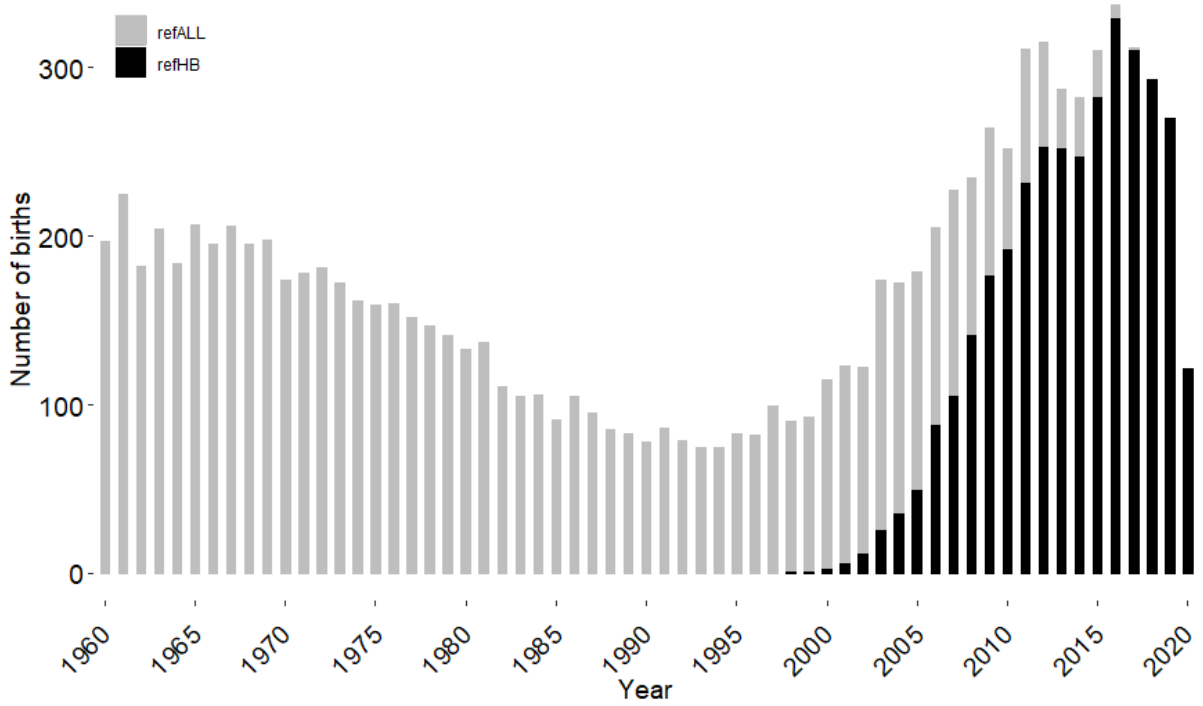


Figure 3.: Number of births per year in refALL and refHB populations.

Globally, after a constant decrease, the number of births rose since 1995 to reach a peak in 2016. In the 1990's, EBRW was considered extinct and the few remaining EBRW breeders were most of the time out of any breeding association. Therefore, it was expected to see a reduced number of registered births at that time. Between 1990 and 1995, less than 50% of the EBRW recorded ancestors were born in Belgium, with births registered mainly in the Netherlands and in Germany. This might indicate importations of German and Dutch breeds such as the RDN and MRY which are also DP red and white breeds. However, François et al. (2017) did not found the expected proximity between EBRW and MRY but would instead suggest crosses with Improved Red.

The first refHB animal was born in 1998, seventeen years before the HB creation in 2015. This cow was genotyped in 2014 to investigate the genomic differentiation of the EBRW relative to other breeds, what lead to the HB creation. The maximum number of HB animals was reached in 2016 with 329 births what is consistent with the HB creation and the revival campaign that started in 2014. However, since then, the number of births is decreasing by, 20 births each year on average. This number of births for the last years shouldn't be considered as static because some animals are potentially not registered in the HB yet. Notwithstanding, this decrease might reflect a downturn in breeder's involvement in the conservation of the EBRW. Another explanation is the ageing breeder's population. Indeed, some of them are close to retirement and might already reduce the size of their herd.

The "breeding activity" of the refALL animals was checked to investigate the occurrence of other breeds in the EBRW genealogies (**Appendix 17.**). Ancestors of EBRW HB animals appeared to be registered as very different breeds from which Holstein is the most frequent one. The other well represented breeding activities were Simmental, Red & White (that includes RDN and MRY animals). This confirms the use of different breeds in the EBRW genealogies and therefore, before accurately quantifying the genetic diversity, we can already state that part of it is due to the composite origin of the EBRW breed.

The refHB contains 3,257 cows and only 169 bulls leading to very unbalanced breeding sex-ratio of 0.051. This number means that within the refHB only few bulls are potentially available for mating. Moreover, this imbalance is also due to less stringent rules for a cow to enter the HB than for a bull for which a genomic test is required. Making the hypothesis that only HB bulls will be used to breed the next EBRW generations (what will happen when the HB will be closed, i.e. the transition period will end), the little number of bulls available will increase the size of full- or half-sib families. Therefore, if full- and half-sib relationship are more frequent, the average relationship of the population will increase which might lead to an increase in inbreeding, reducing genetic diversity.

From the 17,622 refALL animals, 4,709 were considered as founders, i.e. individuals with no registered parents. The vast majority of these founders were registered as Holstein (n=2,107), unkwown (n=916), or Simmental (n=470). Within these 4,709 animals, 53 belong to the refHB population. In other words, out of the 3,426 refHB animals, 53 have no registered parents what is more than expected for a revived breed as the EBRW. These 53 animals are all registered in HB section B, what is consistent with HB management procedure. They are Belgian cows born between 1998 and 2013 (6 animals with unknown birth date) except one German bull (not used for AI) born in 2018.

1.2. Pedigree quality

Through the pedigree quality parameters, it quickly appeared that the EBRW pedigree quality was very low. Indeed, even if, for some animals, up to 33 generations were recorded, the MaxGen average for refHB animals equals 5.6 generations. However, these generations are most of the time far from complete. This is evidenced by the really low CompGen average value of 0.49 indicating that, in average, even both parents of the animals are not known as less than one generation is complete. In that direction, pedigree completeness, i.e. averaged proportion of ancestors known in each ancestral

generation, for the different HB section is depicted in **Figure 4.**. A limit of six generations was set to avoid taking account for pedigree of other breeds in the EBRW PC computation.

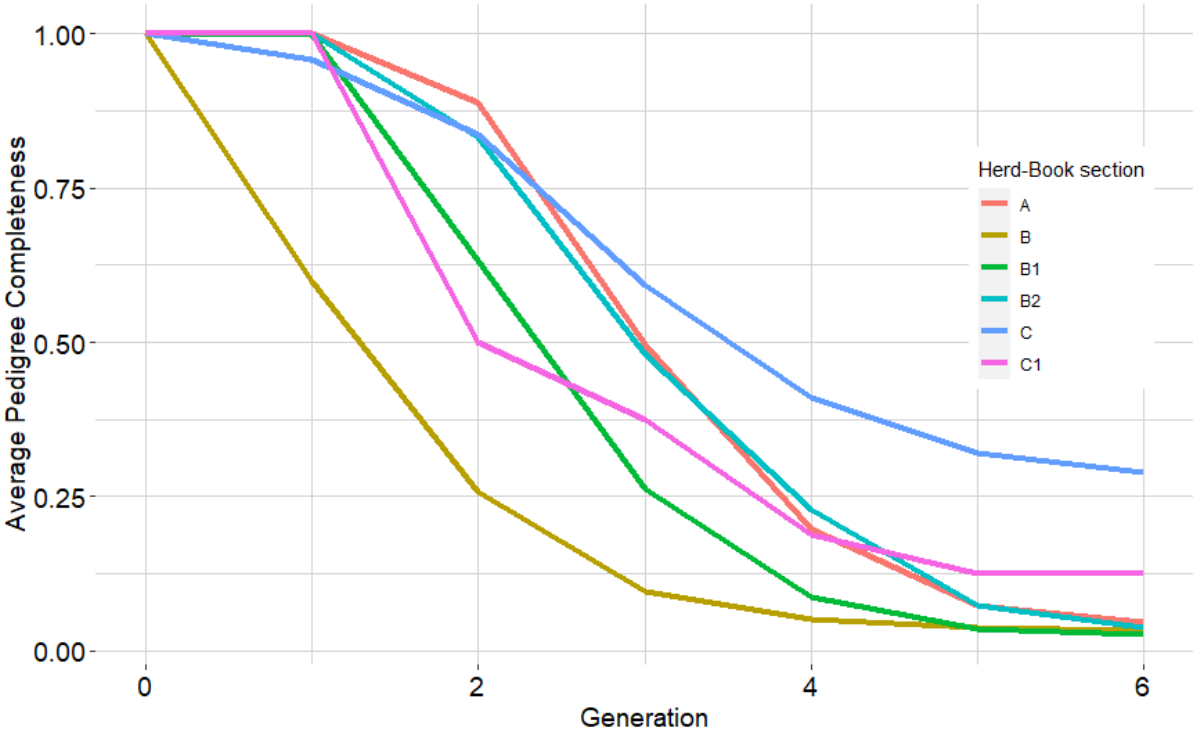


Figure 4.: Average pedigree completeness over six generation by Herd-Book section.

From the **Figure 4.**, it appeared that the pedigree completeness decreased rapidly to low values, reflecting a poor pedigree information. For almost every HB section, the pedigree completeness felt under 0.5 after only three generations. In the first generation, Book A, B1, B2 and C1 animals have a pedigree completeness value of one, meaning that both parents are known. This was expected because to enter these sections, these animals must have both parents registered in the EBRW HB. Average pedigree completeness of book A animals shows the best value in the second generation (grand-parental generation) with 0.89 but falls to 0.50 in the third generation. This value of 0.5 means that, in average, for book A animals, half of the great-grand-parents are recorded in the pedigree file. Those values reflect a better pedigree recording directly linked to the HB management. The HB creation was therefore a very relevant action for the EBRW conservation and need to be continued and promoted in the years to come. In contrast, Book B animals that can enter the EBRW HB without known ancestries, get the worst pedigree completeness values over most of the ancestral generations. For generations 3 to 6, Book C animals obtained the higher pedigree completeness value. This can be explained by the fact that these animals had ancestors registered as other breeds (Holstein, DPSF, RDN, Improved Red) for which it seems that pedigree were better recorded. Therefore, the book C animals joined the EBRW HB with already very complete ancestries. In the end, this low pedigree quality is a

problematic issue to compute pedigree based genetic diversity parameter such as inbreeding. Indeed, both paternal and maternal lines are needed to detect common ancestors leading to inbreeding.

The mean PCI for the 3,426 HB animals over six generations is only 0.15 what is really insufficient for a reliable inbreeding coefficient estimate therefore 0.6 and 0.8 PCI thresholds were applied to discard individuals with the less complete pedigree. When selecting for a PCI value over six generation higher than 0.6 and 0.8, it raised 39 and 5 animals respectively.

1.3. Inbreeding coefficient

Distribution of animals within pedigree inbreeding coefficient (F_{PED}) intervals and average inbreeding values are presented for refHB, ref0.6 and ref0.8 populations in **Table 2.** and **Table 3.**.

Table 2.: Distribution of refHB, ref0.6 and ref0.8 animals within pedigree inbreeding coefficient intervals.

Population		**refHB**	**ref0.6**	**ref0.8**
F_{PED} **value**	**0**	3 361	12	0
	]0;0.01]	47	19	1
	]0.01;0.05]	9	6	4
	]0.05;0.10]	2	1	0
	]0.10;0.15]	5	0	0
	]0.15;0.20]	0	0	0
	]0.20;0.25]	1	0	0
	]0.25;0.30]	1	1	0
Total		3 426	39	5
Total inbred		65	27	5

Table 3.: Average pedigree inbreeding in the inbred animals by reference population.

Population	**refHB**	**ref0.6**	**ref0.8**
Average F_{PED}	0.023	0.024	0.020

From **Table 2.**, it appears that when applying the different PCI thresholds, the number of inbred individuals decrease as the threshold increase. This means that even if the inbreeding estimates are more reliable with a higher PCI threshold, many inbred individuals are discarded, especially individuals with low inbreeding values. These low values reflect old inbreeding, that happened several generations ago and that were reduced through generations. These inbreeding can appear in one parental line

only, meaning that individuals with only one parental line known (PCI = 0) can get such inbreeding value. However, every relationship between individuals from the two parental lines can not be accounted for in this situation. That is why inbreeding coefficient computed for animals with low PCI values can only be underestimated and why inbreeding coefficient is more reliable with high PCI values reflecting a more complete pedigree. Another issue is that individuals with high inbreeding estimate are also discarded when the PCI threshold increase. Indeed, in refHB, two animals with inbreeding coefficient higher than 0.20 had one of their parents as their own grand-parent (**Figure 5a.**). This leads to an inbreeding coefficient estimates of at least 0.25. Moreover, five refHB animals had an inbreeding value of 0.125 reflecting the existence of a common grand-parent in both paternal and maternal lines as illustrated in **Figure 5b.**.

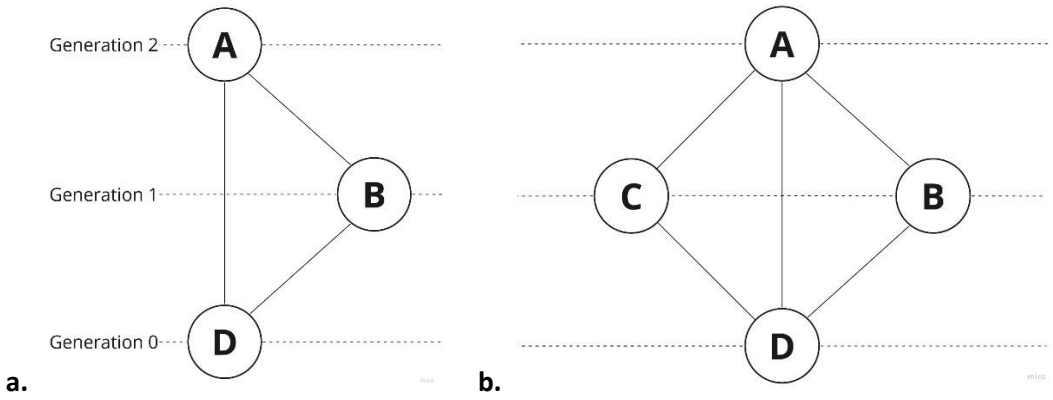


Figure 5.: Schematic representation of mating leading to, **a.** 0.25 and **b.** 0.125 pedigree inbreeding coefficient in individual D.

As these relations occurs in generations 1 and 2, if the generations 3 to 6 are completely unknown, the PCI value will be lower than the threshold which occurs for one of the 0.25 inbreeding animal. In our situation, the 0.6 PCI threshold over six generation might be too stringent because it discards animals with recent inbreeding happening in the very first generations. However, the average inbreeding values for the three reference populations (**Table 3.**) are quite similar and it is difficult to state which best represent the true inbreeding situation knowing issues discussed hereabove. Nevertheless, these values are similar to what was found in Baumung & Sölkner (2002) for three Austrian endangered breeds in similar pedigree completeness situation. Unfortunately, these values remain underestimates of the real inbreeding in the population due to incomplete pedigree information (Forutan et al., 2018; Keller et al., 2011).

1.4. Pedigree relationship matrix

Three A matrices were built for ref0.6 and ref0.8 populations. Once self-relationships removed, the average relationship coefficients equal 0.028 and 0.032, for the two reference populations respectively. The refHB was not used because its too low PC would lead to unexploitable results. We see here that the better the PC, the higher the relationship coefficient. However, for those animals with more known pedigree (ref0.6 and ref0.8), large part of their pedigree was composed of Holstein animals. Therefore, the value obtained here should be taken carefully and might not reflect the average relationship of the whole EBRW population. In the end, these values reflect that, in average, considering ref0.6 and ref0.8, around 2.8% and 3.2% of alleles, shared between two randomly chosen animals in ref0.6 and ref0.8 respectively, are IBD. Making the hypothesis that these values can be transposed to the whole EBRW population, an average relationship of around 3% depicts a low level of relatedness between individuals within the breed. If animals are weakly related, it can be supposed that they are coming from different origins and that they bring with them lots of variations creating a gene pool containing a lot of genetic diversity.

The values obtained for ref0.6 and ref0.8 fall into the range of what was found in other endangered cattle breeds (Addo et al., 2019; Carolino et al., 2020; Carolino & Gama, 2007; Mrode et al., 2009).

2. GENOTYPES ANALYSIS

2.1. Pre-processing and quality control

To remove SNPs only genotyped by LD chips, the SNP call rate threshold was set to 80% as this threshold is the end of the plateau depicted in **Figure 6.**. From this call-rate value and in regards to the number of SNPs removed, we can assume that are removed (i) SNPs only genotyped by LD chips, (ii) SNPs not shared between the different MD chips but also (iii) SNPs with missing values for most animals.

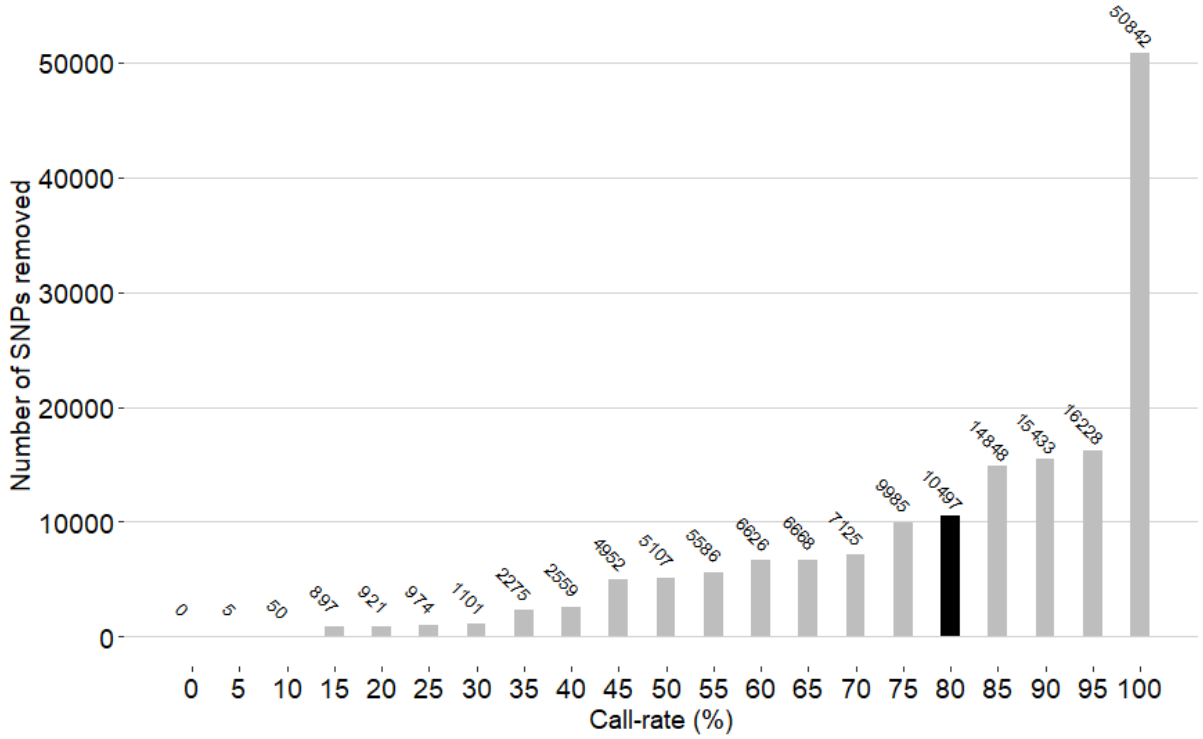


Figure 6.: Number of SNPs removed from the SNP panel with different SNP call-rate values.

Then, two animals were discarded when applying an 90% threshold for individual call-rate. This value, commonly found in the literature is also relevant here as it allows to keep the 13 AI bull genotypes which are of main interest. For the second SNP call-rate, a 95% threshold was applied. This more stringent threshold, removing 5,574 more SNPs, is aimed to remove low quality SNP probably due to genotyping errors. The MAF filtering further removed 418 SNPs also assumed to be related to genotyping errors. From the LD pruning, 3,591 SNP were discarded. In the end, 35,979 SNPs and 214 animals passed filters and quality control.

2.2. Heterozygosity

From these genomic data, the expected and observed homozygosity, computed by PLINK for every sample and then averaged over the whole population, equals 0.641 and 0.642, respectively. Therefore, the average H_e and H_o equals 0.359 and 0.358, respectively. The H_o based on a sample of SNPs is only an estimate of the complete genome heterozygosity level. However, even with MD chips, the H_o estimate is frequently used as a within-breed indicator of genetic diversity. Compared to values found in Browett et al. (2018) for 32 cattle breeds with H_o ranging from 0.21 to 0.32 the observed heterozygosity of the EBRW was the highest. However, the 0.358 observed heterozygosity value of the EBRW is very close to values found for the RDN (0.356) (Addo et al., 2019) which is a German breed similar to the EBRW. This level of heterozygosity is relatively high possibly reflecting a preserved genetic diversity. However, part of that diversity can be due to the verified introgression of Holstein genes and crossbreeding which was also observed in the RDN (Addo, 2020). On a conservation perspective, an individual being heterozygote for a particular SNP is able to transmit to his offspring two different versions of the SNP. More genetic diversity can therefore be transmitted to the next generations with a higher H_o.

2.3. Genomic relationship matrix

The VanRaden (2008) method for computing the GRM matrix, relatedness are estimated based on allele frequencies. However, in the current study, allele frequencies are estimated from the same genotypes from which relationship between individuals were calculated. This leads to an average relationship estimate equal to zero. Therefore, as explained in Wang (2014), these relationships can no longer be interpreted as the proportion of IBD alleles shared between the two individuals, as for the pedigree-based relationship.

Instead, these genomic relationship estimates can be interpreted as correlations of homologous alleles between individuals due to shared ancestry (Wang, 2014). A negative genomic correlation for a single SNP between two individuals means that a specific allele tends to be sampled from one individual, but usually not from both (Ackerman et al., 2017). A negative genomic relationship between two individuals for a full set of SNPs would mean that in average detecting an allele in one individual makes it less likely to detect the allele in the other individual. Therefore, these two animals are less related than the average. In the end, rather than giving an averaged relationship coefficient over the full

populations, the genomic relationships between the 13 AI bulls will be compared to their pedigree relationships given in the A matrix into section 4.2. of this chapter.

2.4. Inbreeding coefficients

Distribution of animals within inbreeding coefficient intervals and average inbreeding values in the inbred are presented for F_{HOM}, F_{GRM} and F_{UNI} in **Table 4.** and **Table 5.**. The negative inbreeding values in **Table 4.** Might be related to estimations of alleles frequencies.

Table 4.: Inbreeding coefficients distribution for the 214 genotyped individuals.

Inbreeding coefficent estimate		F_{HOM}	F_{GRM}	F_{UNI}
Inbreeding value	**[-0.15;-0.1[**	1	1	1
	[-0.1;-0.05[	1	19	1
	[-0.05;0[	125	80	120
	[0;0.05[	66	89	73
	[0.05-0.10[	18	20	15
	[0.10;0.15[	1	4	2
	[0.15;0.20[			1
	[0.20;0.25]	2	1	1

Table 5.: Averaged inbreeding coefficients in the inbred individuals.

Inbreeding coefficient estimate	F_{HOM}	F_{GRM}	F_{UNI}
Averaged inbreeing coefficient	0.033	0.035	0.029

Negative F_{HOM} values can result from a lower observed number of homozygous SNPs than expected. This can be explained by outbreeding, i.e. mating of non-relatives (Vilà et al., 2003), that occurred through crossbreeding for example. For F_{GRM} and F_{UNI} , the same issue occurs than for GRM computation and results should be interpretated as correlations, where negative values denote less relatedness between individuals than positive ones.

When looking at averaged inbreeding values (**Table 5.**) the F_{UNI} value is clearly lower than the F_{HOM} and F_{GRM} values. A property of F_{GRM} and F_{UNI} estimates is to give more weight to homozygosity at rare alleles while F_{HOM} weights alleles equally (Alemu et al., 2021; Keller et al., 2011). Therefore, the low F_{UNI} estimate might reflect lower inbreeding at rare alleles what would be positive as F_{UNI} seems

to be positively correlated with inbreeding depression (Alemu et al., 2021; Keller et al., 2011). However, the relatively large difference between F_{GRM} and F_{UNI} estimates is difficult to explain but could be due to the use of MD SNP chip data that provides less precise estimates than HD SNP chip data. The averaged genomic inbreeding coefficients in the inbred (**Table 5.**) are all superior to the F_{PED} values obtained in the previous section reflecting the relevance of genomic inbreeding estimates as they can better capture the inbreeding level of a population.

In the end, even if the interpretation of genomic inbreeding estimate is more complex than the pedigree-based estimate, genomic inbreeding estimates can still give precious information. Indeed, it allows to plot highly inbred individuals that had no pedigree information.

3. PHENOTYPES ANALYSIS

3.1. Generalized Linear Mixed Model

To investigate phenotypic differentiations between animals grouped in owner herds, GLMMs were built. Six GLMMs were run in parallel with each of the 6 production variables (milk quantity (kg/day), fat quantity (g/day), fat percentage (%), protein quantity (g/day), protein percentage (%) and SCS). Every fixed factor had a significant effect on the dependent variables except for the year effect on protein percentage that had a p-value of 0.051.

3.1.1. Descriptive statistics

Strong positive correlations appeared between GLMM herd estimates for milk, fat and protein quantities and were displayed through the correlation graph (**Figure 7.**). These correlations were expected because the more milk is produced, the more milk constituents are produced. The observed strong positive correlations between phenotypic traits justified the use of a PCA.

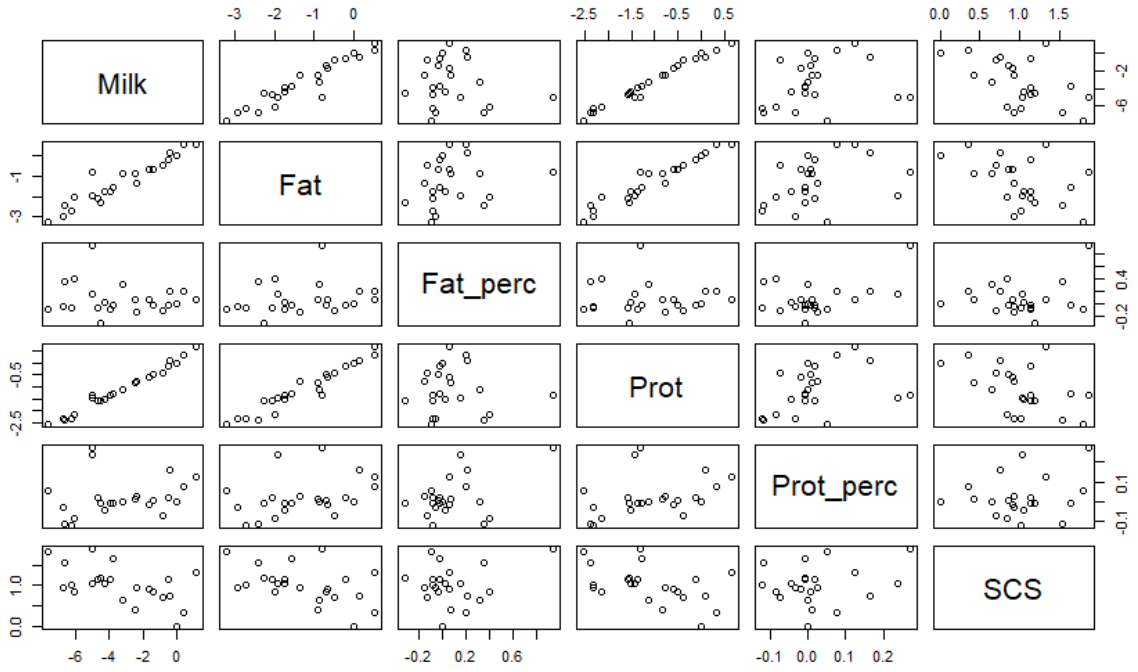


Figure 7.: Correlation graph of production traits based on GLMM herd estimates. Milk: daily milk yield (kg); Fat: daily fat yield (g); Fat_perc: daily fat percentage; Prot: daily protein yield (g); Prot_perc: daily protein percentage; SCS: Somatic Cell Score.

3.1.2. Principal Component Analysis

The first two components, explaining 56.66% and 22.17% of variance respectively (**Table 6.**), reached a cumulative variance of 77.83% what was fitting the 70% threshold. In addition, the eigenvalues of these two components are superior to 1 which is also a common criterion for component setting (Gomes Arandas et al., 2017; Kaiser, 1961). Correlations between the original variables and the two principal component (PC), are displayed in **Figure 8.** (Abdi & Williams, 2010). These correlations are reported in **Table 7.** along with their p-values, whereas **Table 8.** shows the contributions of these variables to the first two PC. Finally, **Figure 9.** presents the distribution of owner herds in the first two PC space.

Table 6.:
Eigenvalue, percentage of variance and cumulative variance of the principal components (PC).

PC	Eigenvalue	Variance (%)	Cumulative variance (%)
1	3.33	55.57	55.57
2	1.62	26.94	82.51
3	0.66	16.03	93.54
4	0.39	6.43	99.96
5	0.0017	0.028	99.99
6	0.00049	0.0082	100

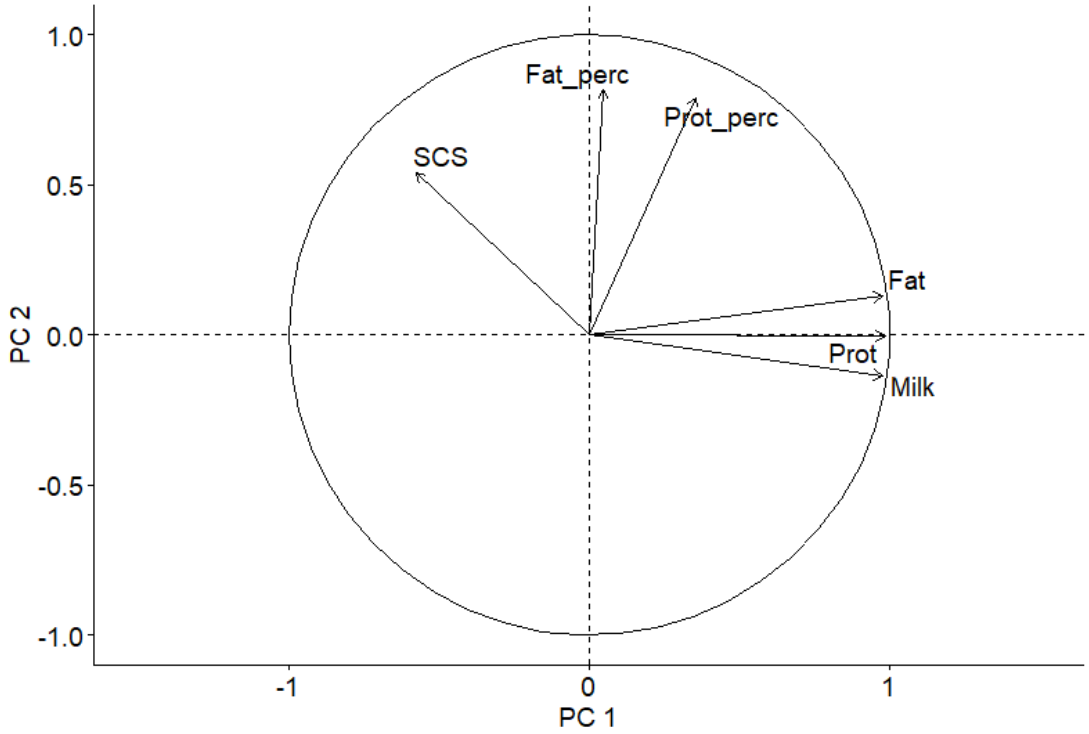


Figure 8.: Circle of correlations of the original variables with principal components (PC) 1 and 2. Milk: daily milk yield (kg); Fat: daily fat yield (g); Fat_perc: daily fat percentage; Prot: daily protein yield (g); Prot_perc: daily protein percentage; SCS: Somatic Cell Score.

Table 7.: Correlations between variables and the first two principal components (PC) and their respective p-values.

	PC 1		PC 2	
Variable	**Correlation**	**p-value**	**Correlation**	**p-value**
Milk	0.98	$2.38.10^{-15}$ (***)	-0.13	$5.37.10^{-1}$
Fat	0.97	$3.72.10^{-15}$ (***)	0.12	$5.55.10^{-1}$
%Fat	0.04	$8.30.10^{-1}$	0.81	$2.11.10^{-6}$ (***)
Protein	0.99	$1.64.10^{-17}$ (***)	-0.003	$9.87.10^{-1}$
%Protein	0.36	$9.53.10^{-2}$	0.79	$7.22.10^{-6}$ (***)
SCS	-0.58	$3.97.10^{-3}$ (**)	0.53	$7.90.10^{-3}$ (**)

(**): very significant correlation; (***): highly significant correlation

Table 8.: Contributions (%) of variables to the first two principal components (PC).

Variable	**PC 1**	**PC 2**
Milk	28.56	1.14
Fat	28.49	1.04
%Fat	0.07	41.17
Protein	29.10	0.0007
%Protein	3.81	38.65
SCS	9.97	18.00

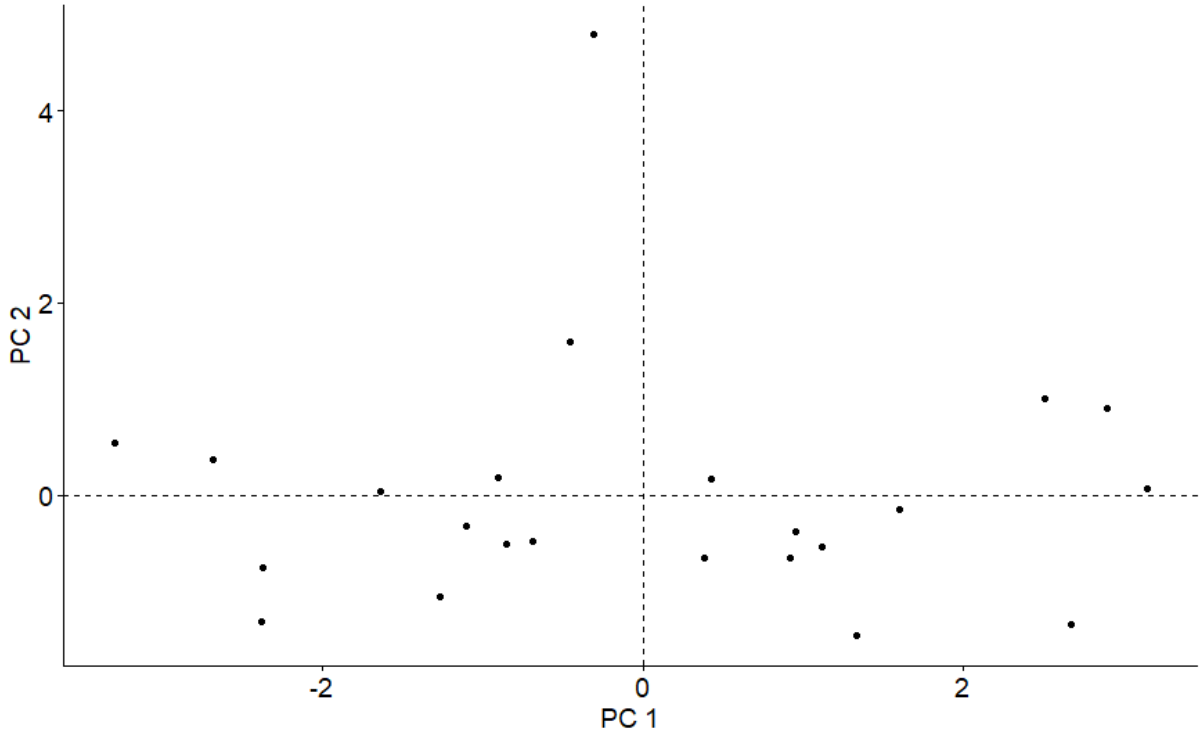


Figure 9.: Distribution graph of owner herds in the first two Principal Components (PC) space.

In **Figure 9.**, the owner herds clearly appear distributed along the first PC axis and grouped around the origin of the second PC axis. The first axis being highly correlated with the three quantity traits the PCA indicates that the most differences between owner herds are based on milk, fat and protein production differences. This result was expected. Indeed, as fat and protein percentages differences between cows of a same breed are often reduced, the greatest differences were supposed to be on quantities. Even if genetic factors probably explain part of that phenotypic variation, the environmental factors such as feeding or management practices also have a big effect on milk production. One herd lie far above the others on the second PC axis. This herd had the highest fat and protein percentages estimates from the GLMM as well as the highest SCS estimates but had only 23 cows. Therefore, the estimates might be less precise or it results truly from the breeding objective of the owner.

3.2. Pseudo-phenotypes

Pseudo-phenotypes are based on EBVS that are estimates of the genetic part affecting phenotypes. After outliers' removal, the six traits pseudo-phenotypes of 1,595 cows distributed over 40 owners were kept for the PCA. A strong variation in the numbers of cows per owner was observed (from 1 cow to a maximum of 192 cows per owner). These differences in the number of cows per owner with

pseudo-phenotypes, i.e. registered in the MR program, can reflect the herd size but also depends on how prone are the owners to pay for the MR service.

3.2.1. Descriptive statistics

From the correlation graph (**Figure 10.**) some strong positive correlations clearly appear between deregressed milk, fat and protein quantity variables as explained before. Compared to **Figure 7.**, the correlation trends in **Figure 10.** are more precise because of the higher number of observations and because of the better reliability of the pseudo-phenotypes compared to GLMM estimates. Therefore, some new negative correlation patterns appear between protein percentage and milk, fat as well as protein quantities. This might reflect the dilution effect between milk yield and milk constituents percentages that can be explained by the negative genetic correlations between those traits (Linn, 1988). Surprisingly, no such negative correlations appear between fat percentage and milk quantities while negative genetic correlations was reported in Linn (1988). This absence of negative correlation between pseudo-phenotypes could be due to the influence of environmental factors such as feeding or due to the EBV/pseudo-phenotypes estimation method. The strong positive correlations and the negative correlations observed between pseudo-phenotypes justified the implementation of a PCA.

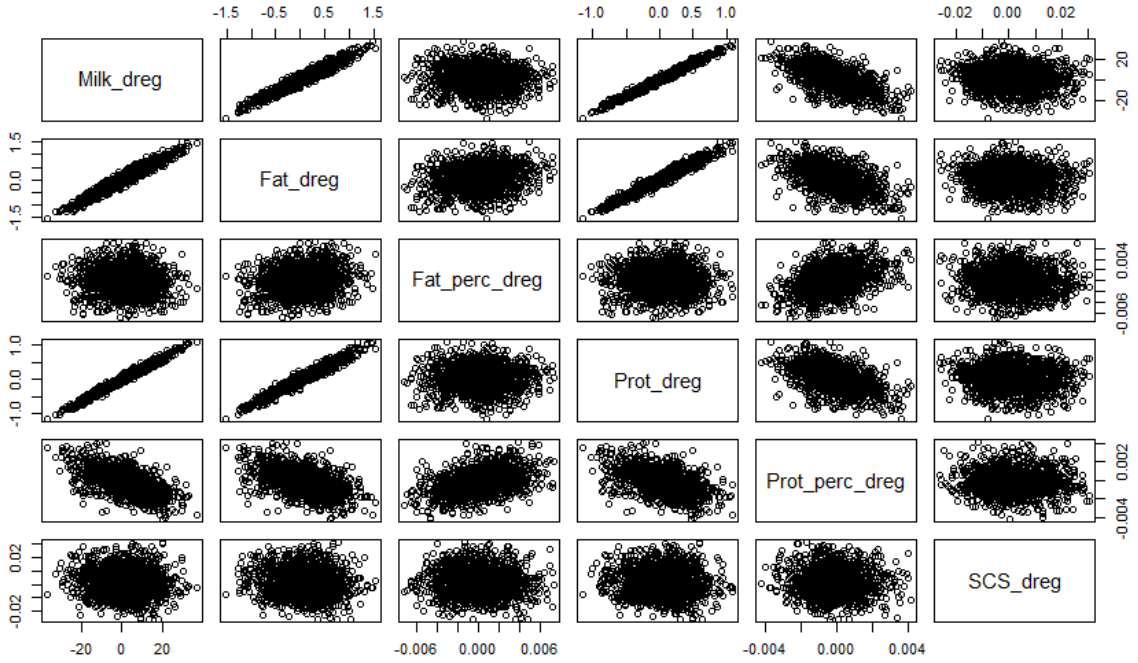


Figure 10.: Correlation graph of production traits based on pseudo-phenotypes. Milk_dreg: deregressed EBV for 305DIM milk production; Fat_dreg: deregressed EBV for fat production; Fat_perc_dereg: dregressed EBV for fat percentage; Prot_dreg: deregressed EBV for protein production; Prot_perc_dereg: dregressed EBV for protein percentage; SCS_dereg: dregressed EBV for SCS.

3.2.2. Principal Component Analysis

The first two components, explaining 56.66% and 22.17% of variance respectively (**Table 9.**), reached a cumulative variance of 77.83% what was fitting the 70% threshold. In addition, the eigenvalues of these two components are superior to 1 which is also a common criterion for component setting (Gomes Arandas et al., 2017; Kaiser, 1961). Correlations between the original variables and the two PC, are displayed in **Figure 11.** (Abdi & Williams, 2010). These correlations are reported in **Table 10.** along with their p-values, whereas **Table 11.** shows the contributions of these variables to the first two PC. **Figure 12.** presents the distribution of cows in the first two PC space. For the same 23 owner herds used for the GLMM, pseudo-phenotypes were averaged and plotted as supplementary points (red points) in **Figure 12.**. For easier comparison with **Figure 9.**, only the 23 owner herd averages are plotted in **Figure 13.**.

Table 9.: Eigenvalue, percentage of variance and cumulative variance of the principal components (PC).

PC	Eigenvalue	Variance (%)	Cumulative variance (%)
1	3.34	55.66	56.66
2	1.33	22.17	77.83
3	0.97	16.11	93.94
4	0.36	6.05	99.996
5	0.00015	0.0026	99.998
6	0.00009	0.0015	100

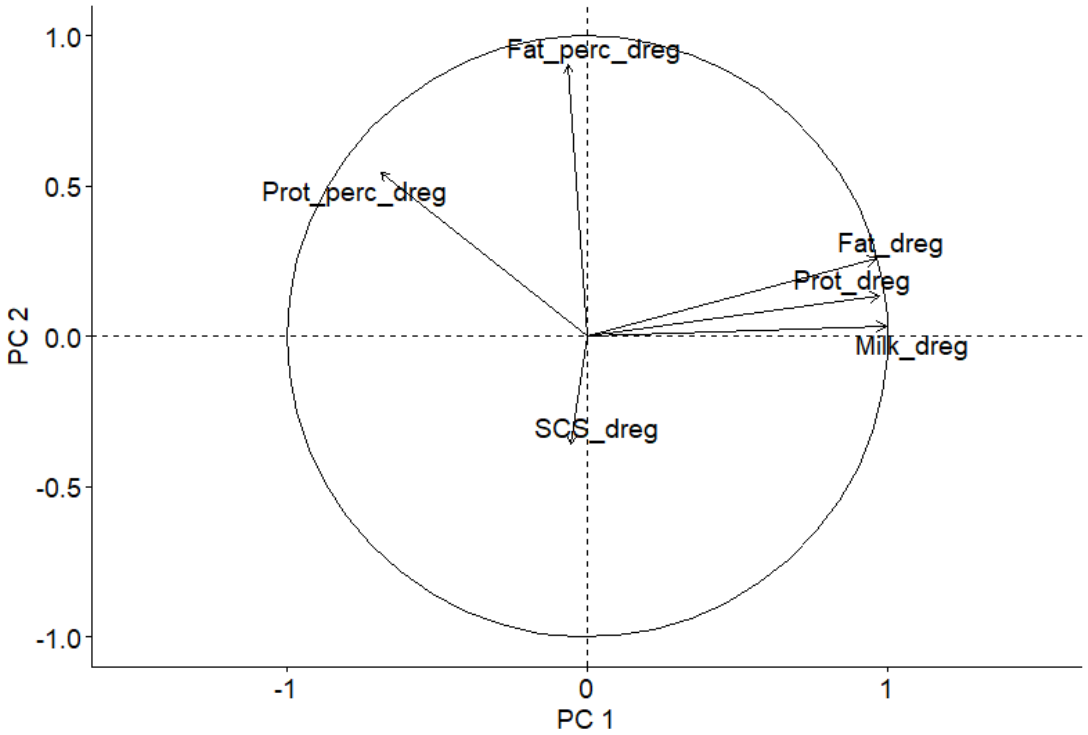


Figure 11. : Circle of correlations of the original variables with principal components (PC) 1 and 2. Milk_dreg: deregressed EBV for 305DIM milk production; Fat_dreg: deregressed EBV for fat production; %Fat_dereg: dregressed EBV for fat percentage; Prot_dreg: deregressed EBV for protein production; %Prot_dereg: dregressed EBV for protein percentage; SCS_dereg: dregressed EBV for SCS.

Table 10.: Correlations between variables and the first two principal components (PC) and their respective p-values.

Variable	PC 1		PC 2	
	Correlation	p-value	Correlation	p-value
Deregressed Milk	0.99	$< 1.10^{-300}$ (***)	0.03	$1.69.10^{-1}$
Deregressed Fat	0.96	$< 1.10^{-300}$ (***)	0.26	$7.57.10^{-26}$ (***)
Deregressed %Fat	-0.06	$9.94.10^{-3}$ (**)	0.90	$< 1.10^{-300}$ (***)
Deregressed Protein	0.97	$< 1.10^{-300}$ (***)	0.14	$6.09.10^{-8}$ (***)
Deregressed %Protein	-0.69	$3.25.10^{-225}$ (***)	0.55	$2.50.10^{-124}$ (***)
Deregressed SCS	-0.06	$2.65.10^{-2}$ (*)	-0.36	$3.20.10^{-49}$ (***)

(*): significant correlation; (**): very significant correlation; (***): highly significant correlation

Table 11.: Contributions (%) of variables to the first two principal components (PC).

Variable	PC 1	PC 2
Deregressed Milk	29.60	0.09
Deregressed Fat	27.70	5.03
Deregressed %Fat	0.12	61.54
Deregressed Protein	28.26	1.37
Deregressed %Protein	14.23	22.36
Deregressed SCS	0.09	9.60

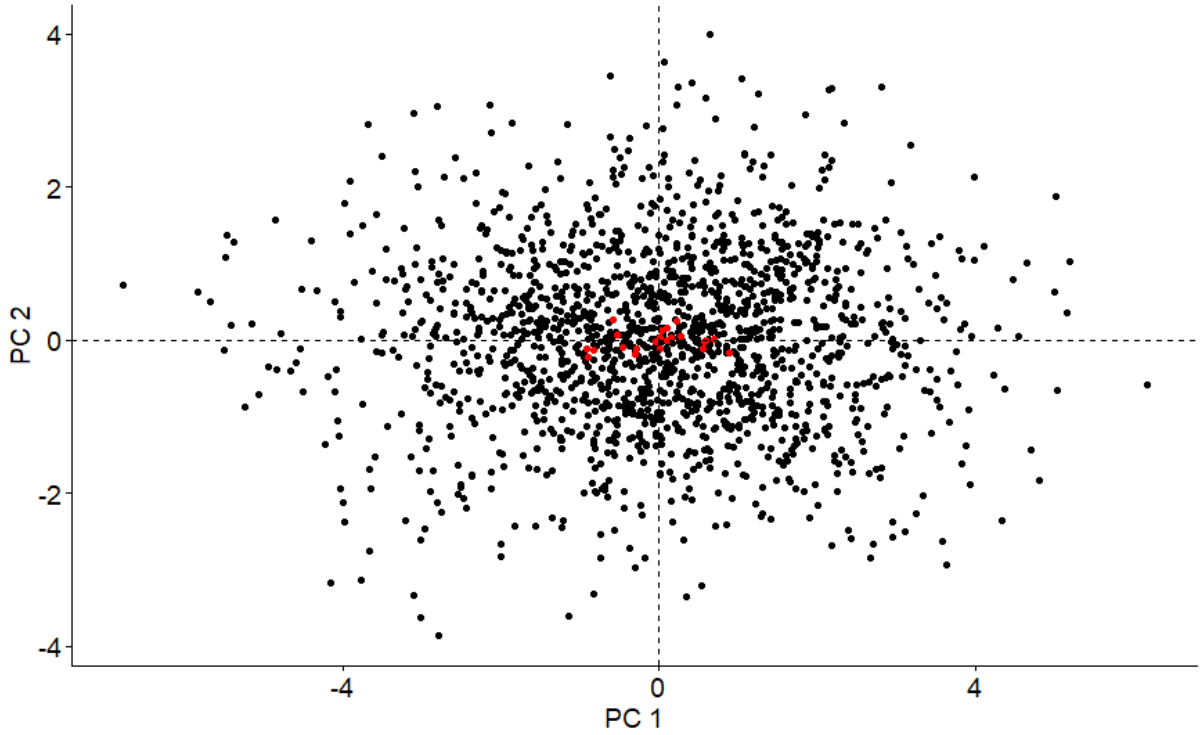


Figure 12.: Distribution graph of cows in the first two Principal Components (PC) space. The 23 owner herds averages are coloured in red.

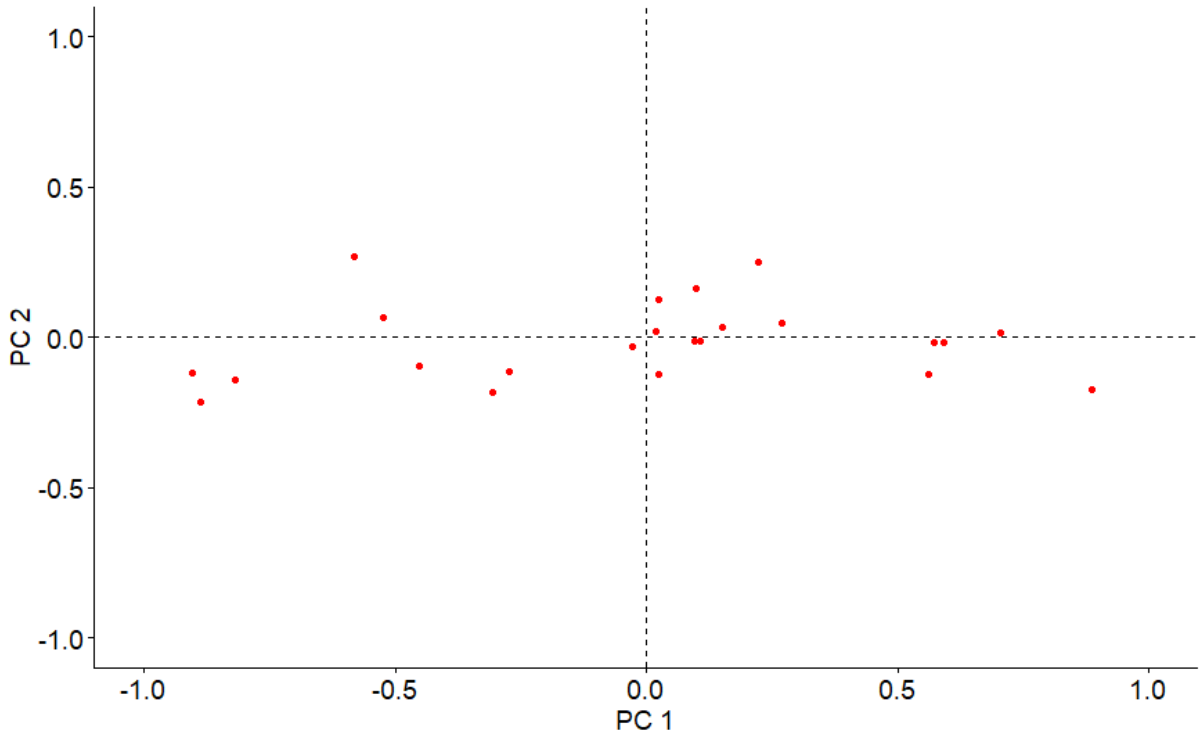


Figure 13.: Distribution graph of owner herds averages in the first two Principal Components (PC) space.

From the **Figure 12.**, the individuals appear grouped together around the centre of the two-dimensional space and evenly distributed over the four quadrants. Still, a slight elongation of the point cloud can be seen on the PC 1 axis. This spread over the PC 1 is even more clear when looking at the owner herds averages (**Figure 13.**) and support what has been found with PCA based on GLMM herd estimates (**Figure 9.**).

Mazza et al. (2016) and Sartori et al. (2018) demonstrate the negative genetic correlation between morphology and test-day milk yield traits in endangered DP breeds. Therefore, animals and owners on the left end of the point cloud, probably have better muscularity and conformation traits, leaving aside differences in milk production due to feeding differences.

4. RECOMMENDATIONS TO PRESERVE THE EBRW DIVERSITY

4.1. Maintaining the within breed genetic diversity

From the previous analyses, the within-breed diversity existing in the EBRW seems quite preserved. However, in endangered breeds, due to their small finite population size this within- breed genetic diversity is mostly threatened by genetic drift (Kristensen et al., 2015; Simianer, 2005). Therefore, conservation measures should be taken to reduce genetic drift.

4.1.1. Reducing genetic drift

To reduce genetic drift, several points can be considered.

First is to increase or, at least, stabilize the population size. Indeed, reduction or strong variations in the number of individuals per generation might lead to a genetic bottleneck that can considerably change the alleles frequencies in the next generation with risks to lose or fix some alleles, lowering genetic diversity (Choudhuri, 2014; Vilà et al., 2003; Willoughby et al., 2015). The decreasing number of births in the EBRW population since 2016 put the existing genetic diversity at risk through increased genetic drift and increased inbreeding. Therefore, continuous efforts should be made to promote the breed and get new breeders involved in the breed conservation.

Then balancing the breeding sex ratio can also slow down genetic drift (Rosche et al., 2018; Simianer, 2005; Willoughby et al., 2015). In our situation, with a potential unbalanced breeding sex ratio, a low number of sires contributes to half of the alleles of the offspring. Therefore, the alleles frequencies can vary due to the sampling process and might lead to a reduced genetic diversity. In the EBRW situation, breeders should be encouraged to register more bulls in the HB to offer a more balanced breeding sex ratio.

Third, balancing individuals contributions to the next generation will also lower the genetic drift (Simianer, 2005; Willoughby et al., 2015). Indeed, if an animal is used much more times than another, its alleles will be transmitted more heavily, changing the allele frequency in the next generation. Therefore, in an ideal situation each EBRW animal should produce the same number of claves. In that direction, mating could be planned to use AI bulls as evenly as possible, for example, as it seems not to be the case yet. Indeed, when looking at **Table 12.**, the number of offspring per AI bull is very unbalanced. More worryingly, the proportion of offspring registered in the EBRW HB per bull never

exceeds one third of the total number of offspring for each bull. This might reflect a broad use of EBRW semen for crossbreeding. Another option could be to try to cull cows after a fixed number of calving.

Table 12.: Number of offspring per EBRW AI bull.

		Offspring	Daughters	Sons	EBRW HB registered offspring	% of EBRW HB registered offspring	MR registered daughters
Bull	**Becco**	34	18	16	0	0.0	0
	Eddy	318	156	162	64	20.1	0
	Galleo	122	52	70	26	21.3	15
	Hilar	231	126	105	37	16.0	16
	James-RPE	9	3	6	0	0.0	0
	Johann	138	58	80	14	10.1	9
	Manuel ORB	155	75	80	40	25.8	26
	Medello	66	30	36	1	1.5	0
	Praeses	90	43	47	14	15.6	0
	Rudi-RPE	21	9	12	0	0.0	0
	Tom	211	129	82	51	24.2	15
	Willy	220	117	103	59	26.8	58
	Yannick	158	90	68	45	28.5	38

4.1.2. Using allele flow carefully

Allele migration can counteract genetic drift and increase genetic diversity by bringing new alleles into the genetic pool of the breed but also by preventing the formation of sub-population within the breed where genetic drift will occur more severely as the size of those sub-population will be even lower. Therefore, animals from other breeds, could be used to breed with EBRW in compliance with the HB rules that accept animals of other breeds, but still meeting the EBRW breed standards, through the Book C. In addition, transfer of animals among breeders should be encouraged and eased through the promotion and development of AI.

However, to preserve the genetic uniqueness of the EBRW breed, introgression of alleles from other breeds should be limited and the use of animals with a high percentage of exogenous alleles transmitted by ancestors belonging to other breeds should be avoided. As it is now known that the EBRW breed suffered from Holstein introgression, the percentage of Holstein genetic material in the AI bulls was verified using an “in-house” program. For Eddy and Yannick up to 12.5% of the alleles might have been inherited from Holstein ancestors and should be used for AI in limited amount.

In the end, implementing a breeding circle mating scheme (**Figure 14.**), where herds are organised in a circle with each herd being a bull donor herd for another herd in static donor–recipient combination from year to year (Windig & Kaal, 2008), could be another step toward the conservation of the EBRW diversity. Indeed, according to some authors, it seems to be an effective genetic diversity management tool for small populations with low pedigrees as it reduces the increase in inbreeding in the population (Windig et al., 2019; Windig & Kaal, 2008).

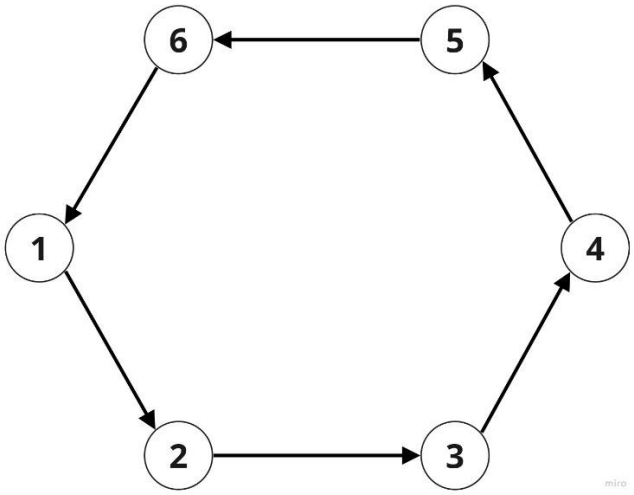


Figure 14.: Breeding circle for six herds. Herd 1 providing breeding bulls to herd 2, herd 2 to herd 3, etc. (Windig & Kaal, 2008).

4.2. Prevent increase in inbreeding

A critical point to avoid loss of genetic diversity is preventing the increase in homozygosity due to inbreeding. To do so, mating related individuals should be avoided. Therefore, information about the relatedness between the individuals need to be available. In this work, the relatedness information is given by the relationship matrices, both pedigree and genomic, computed and displayed for the 13 AI bulls under **Table 13.** and **Table 14.**.

Table 13.: Pedigree relationship matrix (A matrix) for the 13 AI bulls.

	Becco	Eddy	Galleo	Hilar	James-RPE	Johann	Manuel ORB	Medello	Praeses	Rudi-RPE	Tom	Willy	Yannick
Becco	1												
Eddy	0	1											
Galleo	0	0	1										
Hilar	0	0	0	1									
James-RPE	0	0	0	0	1								
Johann	0	0	0	0	0	1							
Manuel ORB	0	0	0	0	0	0	1						
Medello	0	0	0	0	0.25	0	0	1					
Praeses	0.0625	0	0	0	0	0	0	0	1				
Rudi-RPE	0	0	0	0	0	0	0	0	0.25	1			
Tom	0	0	0	0	0	0	0	0	0	0	1		
Willy	0	0	0	0	0	0	0	0	0	0	0	1	
Yannick	0	$6.01.10^{-05}$	0	0	0	0	0	0	0	0	0	0	1

Table 14.: Genomic relationship matrix (GRM) for the 13 AI bulls.

	Becco	Eddy	Galleo	Hilar	James-RPE	Johann	Manuel ORB	Medello	Praeses	Rudi-RPE	Tom	Willy	Yannick
Becco	1.00												
Eddy	-0.01	0.98											
Galleo	-0.01	-0.01	1.07										
Hilar	-0.02	-0.02	0.00	1.00									
James-RPE	0.02	-0.01	-0.00	-0.02	0.99								
Johann	-0.02	-0.01	0.01	-0.02	-0.01	0.99							
Manuel ORB	-0.05	-0.02	-0.01	-0.04	-0.05	-0.05	0.95						
Medello	-0.00	-0.02	-0.01	-0.03	0.16	-0.03	-0.06	0.93					
Praeses	0.05	-0.04	-0.02	-0.04	-0.05	-0.04	-0.06	-0.03	0.97				
Rudi-RPE	-0.01	-0.02	-0.02	-0.05	-0.04	-0.03	-0.07	-0.02	0.20	0.98			
Tom	-0.03	-0.01	-0.00	-0.01	-0.02	-0.00	-0.04	-0.03	-0.05	-0.05	1.05		
Willy	-0.01	-0.02	-0.00	-0.01	-0.03	-0.01	-0.05	-0.02	-0.05	-0.02	0.01	1.00	
Yannick	-0.03	0.01	-0.00	-0.03	0.04	-0.04	-0.05	0.04	-0.03	-0.02	-0.03	-0.01	0.96

From the A matrix (**Table 13.**), four relationships were detected between pairs of bulls. The 0.25 relationships reflect the presence of the same father for the two pairs of bulls. The 0.0625 value results from a half-great-uncle relationship between Becco and Praeses. Yannick and Eddy, having deep Holstein pedigree, their $6.01.10^{-05}$ relationship coefficient is due to common Holstein ancestors.

From the GRM (**Table 14.**), the same four aforementioned relationships were detected, validating the genomic information. However, a downward bias seems to be present, also reflected by the negative relationship values. This is probably due to the computation process estimating allele frequencies on the sample itself (cf. section 2.3. of this chapter). Moreover, new relationships were unveiled using the genomic information, making genomic information relevant here as it brings new information that was not captured by the pedigree due to its low depth. In that direction, a 0.4 relationship coefficient was unveiled between Yannick and Medello that were born in the same herd, which might explain the presence of a common ancestor.

To prevent increase in inbreeding, using heavily inbred animals for mating should also be avoided. Indeed, these inbred animals, carrying more homozygous alleles are therefore less prone to transmit genetic diversity to the next generations. Looking at the diagonal elements of the A matrix, it appears that no pedigree inbreeding was detected in these 13 AI bulls as none of the diagonal element are superior to 1. However, when looking at the genomic matrix, Galleo and Tom have diagonal values indicating an inbreeding coefficient of 0.07 and 0.05 respectively which is not extremely high but still need to be considered.

Practically, it can be recommended not to use James-RPE semen on Medello's daughters or the opposite what would result in a 0.0625 pedigree inbreeding.

4.3. Mating advice based on AI bulls EBVs

As selection can be a powerful tool to improve the profitability of local endangered breed and therefore their survival (Biscarini et al., 2015; Gandini et al., 2010), the production EBVs of the AI bulls, presented in the **Table 15**. allow to make some recommendations for selection.

Table 15.: Estimated Breeding Values (EBVs) for the production traits and SCS, expressed in the EBRW publication base.

Variable		Milk	Fat	%Fat	Prot	%Prot	SCS
Bull	**James-RPE**	306	9	-0.07	9	-0.03	-3.05
	Tom	268	-2	-0.25	6	-0.06	-3.05
	Rudi-RPE	266	18	0.13	9	-0.01	3.17
	Becco	132	4	-0.03	4	-0.01	-3.24
	Johann	114	3	-0.04	3	-0.02	3.14
	Medello	84	2	-0.02	2	-0.02	3.01
	Praeses	-109	-2.	0.04	-3	0.02	3.15
	Manuel ORB	-135	-6	-0.01	-2	0.06	-3.31
	Eddy	-164	-3	0.09	-4	0.03	3.09
	Yannick	-167	-12	-0.10	-6	-0.01	3.26
	Willy	-558	-21	0.05	-16	0.07	-3.02
	Hilar	-681	-31	-0.04	-22	0.05	-3.44
	Galleo	-1054	-47	-0.06	-30	0.17	3.05

Table 16.: EBVs reliability for production traits and SCS for the AI bulls.

Reliability		Production	SCS
Bull	**James-RPE**	14	12
	Tom	49	41
	Rudi-RPE	26	23
	Becco	10	8
	Johann	34	20
	Medello	10	7
	Praeses	17	15
	Manuel ORB	63	55
	Eddy	11	10
	Yannick	71	59
	Willy	80	74
	Hilar	51	42
	Galleo	49	39

Bulls are ranked in **Table 15.** by Milk EBV values. Therefore, compared to the EBRW publication base (all refHB females born in 2015 = 206 animals), daughters of James-RPE for example are expected to produce on average half of James EBV kg of milk on a 305 DIM lactation basis due to the genetic superiority of James-RPE. If the breeding objective is to increase milk production bulls with the most positive milk EBVs should be used in preference. However, as it can be seen from **Table 16.**, EBVs reliability might be very low due to the inexistent or low number milk recorded daughters (**Table 12.**). Therefore, the very superior milk EBV of James-RPE should be considered with extreme caution.

If the selection objective is to increase milk contents, Rudi and Galleo should be preferred according to the fat and protein percentages EBVs respectively. For the udder health, as low SCS reflect a better udder health, negative EBVs are to be preferred. Therefore, Hilar seems to be a good bull to use to get offspring with improved udder health, always considering the reliability.

Finally, if the selection objective is the meat production there is a lack of information. Ones might rely on the negative genetic correlation between milk yield and muscularity and choose bulls with the lower milk EBVs to serve their cows but it seems a very inaccurate method. In this case, the best option might be to choose bulls on their appearance as no phenotypic records exist for the muscularity.

General conclusions & perspectives

The EBRW, considered as being extinct in the 1990's, have been re-established. Studies started in 2011 leading to the creation of a new EBRW HB in 2015. The EBRW cattle are today considered as one of the two endangered local breeds in Wallonia. However, their very low population size as well as the particular risk status of this breed call for an informed management of the EBRW genetic diversity.

Therefore, the objectives of this master thesis were to assess the current within breed genetic diversity level using pedigree and genomic information. In addition, the breed phenotypic diversity was investigated through the use of phenotype records and pseudo-phenotypes obtained from deregressed EBVs. Finally, practical recommendations were formulated for the management of the genetic diversity at mating.

Despite low knowledge of pedigree, its analysis confirmed a certain degree of crossbreeding with other breeds resulting in exogenous alleles admixture in the EBRW. The inbreeding and relationship estimates based on pedigree were low. However, the low depth and completeness of the pedigree imply that these parameters are underestimated.

The genomic information allowed to improve the inbreeding and relationship estimates as they added new information not covered by the pedigree data. Still, the inbreeding estimates ranging between 2.9% and 3.5% were relatively low and did not seem to be an issue in the studied population, suggesting that breeders despite being out of the official HB keeping did a reasonable job avoiding inbreeding. Furthermore, the observed heterozygosity value of 0.359 seems to be very similar to that found in the RDN and higher than in other breeds (Addo et al., 2019; Browett et al., 2018). This heterozygosity value reflects a preserved genetic diversity but it might also be increased by exogenous alleles admixture that occurred through crossbreeding, occurring in most local breeds.

The phenotypic diversity was displayed through two PCA using GLMM estimates and pseudo-phenotypes. Clear differences between herds were unveiled relating to the milk production, what might reflect large phenotypic diversity within the EBRW breed as well as differences in breeding objectives of specific breeders. This last point must be considered in future definition of global breeding goals for EBRW.

For the management of genetic diversity, recommendations are mostly aimed at reducing the genetic drift by (i) increasing the number of animals and avoid fluctuations, (ii) reducing the imbalance in sex-ratio and (iii) balancing animal's contributions. Allele flow within the breed could be increased by promoting transfer of reproduction animals between breeders. The current efforts to identify suitable AI bulls that are then distributed throughout the entire EBRW population is therefore, considering the current status of this breed, a positive measure to increase allele flow. However, the use of these sires must be carefully monitored to ensure their balanced use.

An important conclusion and perspective, established by this study and required for the global management of the EBRW breed, is the need to pursue, enhance and promote data acquisition from the field. Indeed, increasing the data quantity, and maintaining their high quality, would allow to perform more accurate and diverse (e.g., other important traits for DP cattle) analyses. Deeper knowledge of pedigree will also allow the reduction of the level of underestimation of pedigree inbreeding. Registering more animals to the HB will enable to broaden the breeding population and the reference population for further studies. The amount of genotyping should also be increased. Genotyping some cows (e.g., bull mothers before their use) may allow to maximise chances that the genotypes of their sons remain acceptable as EBRW. The "sister breed" RPO also requires genotyping for all animals to acquire HB status. The EBRW could adopt a similar policy for new animals in the future. This would also give a better sampling of the population as, for now, mainly bulls are genotyped. The DPBB in Belgium is currently pursuing a strategy of genotyping nearly all cows and sires. The objective is not to validate HB status but to allow genomic genetic evaluations in this breed. Moreover, the quality of the genomic data could also be increased by using High-Density (HD) SNP chips for example. Data from HD SNP chips would also allow to impute currently available genomic data to this higher density. This would result in more accurate heterozygosity, relationship and inbreeding estimates. In addition, based on HD genomic data, new methods could be used to assess the genetic diversity, such as the ROH inbreeding coefficient that showed very poor preliminary results when computed with the mid-density data currently available. Data from HD SNP chips would also facilitate comparison and joint analyses across different but related breeds.

For phenotypes, current rules for EBRW require milk recording to allow access to EU subsidies. Even if this creates a barrier for some breeders to join the EBRW conservation program, it also allows a continuous flow of precious phenotypic information. On a rather short-term basis, the generation of relevant extra data could be obtained by implementing a linear classification for morphological traits as it was done with success for the DPBB. In the context of DP breed, as the EBRW breed, the acquisition of phenotypes for dairy and, at least indirectly (i.e., classification data), beef production is necessary to switch from managing its strict breed conservation to breed conservation and improvement.

Finally, a critical element for each endangered breed, as is the EBRW, is the increase of its population size and the genomic and phenotypic characterisation of this population. This would be beneficial for the assessment and management of the EBRW genetic diversity. Moreover, the survival of this breed is not yet assured in regards to the low number of animals and the ageing breeders. Therefore, an increase of the population and of the number of EBRW breeders should be the first main objective for the years to come. It could be achieved through increased efforts for the breed promotion based on

its economic, environmental, and social advantages. The status of EBRW as a locally rooted breed in its region is here clearly an advantage. The current strategies to perform genomic and phenotypic characterisation of its population are well established but, as previously explained, could still be strengthened in order to acquire the relevant data to manage and to improve the EBRW breed in the future.

References

Abdi, H., & Williams, L. J. (2010). Principal component analysis. *Wiley Interdisciplinary Reviews: Computational Statistics*, *2*(4), 433–459. https://doi.org/10.1002/wics.101

Abebe, A. S., Mikko, S., & Johansson, A. M. (2015). Genetic diversity of five local Swedish chicken breeds detected by microsatellite markers. *PLoS ONE*, *10*(4), 1–13. https://doi.org/10.1371/journal.pone.0120580

Ackerman, M., Johri, P., Spitze, K., Xu, S., Doak, T., Young, K., & Lynch, M. (2017). Estimating Seven Coefficients of Pairwise Relatedness. *Genetics*, *206*, 105–118. https://doi.org/10.1534/genetics.116.190660/-/DC1.1

Addo, S. (2020). *Assessment of genetic diversity in local breeds*. University of Kassel, Germany.

Addo, S., Klingel, S., Hinrichs, D., & Thaller, G. (2019). Runs of Homozygosity and NetView analyses provide new insight into the genome-wide diversity and admixture of three German cattle breeds. *PLoS ONE*, *14*(12), 1–20. https://doi.org/10.1371/journal.pone.0225847

Alemu, S. W., Kadri, N. K., Harland, C., Faux, P., Charlier, C., Caballero, A., & Druet, T. (2021). An evaluation of inbreeding measures using a whole-genome sequenced cattle pedigree. *Heredity*, *126*, 410–423. https://doi.org/10.1038/s41437-020-00383-9

Altshuler, D. (2012). The Inherited Basis of Common Diseases. In L. Goldman & A. I. Schafer (Eds.), *Goldman's Cecil Medicine* (24th ed., Vol. 1, pp. 195–198). Elsevier Inc. https://doi.org/10.1016/B978-1-4377-1604-7.00041-5

Amadeu, R. R., Cellon, C., Olmstead, J. W., Garcia, A. A. F., Resende, M. F. R., & Muñoz, P. R. (2016). AGHmatrix: R Package to Construct Relationship Matrices for Autotetraploid and Diploid Species: A Blueberry Example. *The Plant Genome*, *9*(3), 1–8. https://doi.org/10.3835/plantgenome2016.01.0009

Anderson, C. A., Pettersson, F. H., Clarke, G. M., Cardon, L. R., Morris, A. P., & Zondervan, K. T. (2010). Data quality control in genetic case-control association studies. *Nature Protocols*, *5*(9), 1564–1573. https://doi.org/10.1038/nprot.2010.116.

Andersson, L. (2001). Genetic dissection of phenotypic diversity in farm animals. *Nature Reviews Genetics*, *2*, 130–138. https://doi.org/10.1038/35052563

Andrews, C. A. (2010). Natural Selection, Genetic Drift, and Gene Flow Do Not Act in Isolation in Natural Populations. *Nature Education Knowledge*, *3*(10). https://www.nature.com/scitable/knowledge/library/natural-selection-genetic-drift-and-gene-flow-15186648/, (2021-07-24).

APS. (2004). *Reproductive/Mating Systems*. Population Genetics of Plant Pathogens. https://www.apsnet.org/edcenter/disimpactmngmnt/topc/PopGenetics/Pages/ReproductiveMatingSystems.aspx (2021-07-10).

Austin, C. P. (2021). *Genotype*. National Human Genome Research Institute. https://www.genome.gov/genetics-glossary/genotype, (2021-07-10).

Baltussen, W., Arets, E., De Blaeij, A., Galgani, P., De Groot-Ruiz, A., & Vellinga, T. (2017). Environmental and Social Effects of Livestock Systems: Poultry, Beef and Dairy. *Proceedings of the 21st International Farm Management Congres*, *1*, 1–13. www.ifmaonline.org

Baumung, R., & Sölkner, J. (2002). Analysis of pedigrees of Tux-Zillertal, Carinthian Blond and Original Pinzgau cattle population in Austria. *Journal of Animal Breeding and Genetics*, *119*(3), 175–181. https://doi.org/10.1046/j.1439-0388.2002.00332.x

Bay, E., Colinet, F., Hick, C., & Gengler, N. (2009). The dual purpose Red and White. *Gembloux Agro-*

Bio Tech University of Liège.

Baye, T. M., Abebe, T., & Wilke, R. A. (2011). Genotype-environment interactions and their translational implications. *Personalized Medicine*, *8*(1), 59–70. https://doi.org/10.2217/pme.10.75

Bérénos, C., Ellis, P. A., Pilkington, J. G., & Pemberton, J. M. (2016). Genomic analysis reveals depression due to both individual and maternal inbreeding in a free-living mammal population. *Molecular Ecology*, *25*(13), 3152–3168. https://doi.org/10.1111/mec.13681

Biscarini, F., Nicolazzi, E., Alessandra, S., Boettcher, P., & Gandini, G. (2015). Challenges and opportunities in genetic improvement of local livestock breeds. *Frontiers in Genetics*, *5*, 1–16. https://doi.org/10.3389/fgene.2015.00033

Boogaard, B. K., Oosting, S. J., Bock, B. B., & Wiskerke, J. S. C. (2011). The sociocultural sustainability of livestock farming: An inquiry into social perceptions of dairy farming. *Animal*, *5*(9), 1458–1466. https://doi.org/10.1017/S1751731111000371

Bouffioux, A. (2014). *Utilisation de l'information génomique dans la caractérisation d'une race bovine Wallonne menacée d'extinction : la Rouge-Pie de l'Est de la Belgique*. ULiège - Gembloux Agro-Bio Tech.

Bourrat, P. (2021). Genetic Relatedness. In T. K. Shackelford & V. A. Weekes-Shackelford (Eds.), *Encyclopedia of Evolutionary Psychological Science* (pp. 3401–3404). Springer Nature Switzerland. https://doi.org/10.1007/978-3-319-19650-3_1358

Browett, S., McHugo, G., Richardson, I. W., Magee, D. A., Park, S. D. E., Fahey, A. G., Kearney, J. F., Correia, C. N., Randhawa, I. A. S., & MacHugh, D. E. (2018). Genomic characterisation of the indigenous Irish Kerry cattle breed. *Frontiers in Genetics*, *9*(51). https://doi.org/10.3389/fgene.2018.00051

Camp, K. M., & Trujillo, E. (2014). Position of the academy of nutrition and dietetics: Nutritional genomics. *Journal of the Academy of Nutrition and Dietetics*, *114*(2), 299–312. https://doi.org/10.1016/j.jand.2013.12.001

Carolino, N., & Gama, L. T. (2007). Indicators of genetic erosion in an endangered population: The Alentejana cattle breed in Portugal. *Journal of Animal Science*, *86*(1), 47–56. https://doi.org/10.2527/jas.2007-0148

Carolino, N., Vitorino, A., Carolino, I., Pais, J., Henriques, N., Silveira, M., & Vicente, A. (2020). Genetic diversity in the Portuguese mertolenga cattle breed assessed by pedigree analysis. *Animals*, *10*(11), 1–23. https://doi.org/10.3390/ani10111990

Cassell, B. G., Adamec, V., & Pearson, R. E. (2003). Effect of incomplete pedigrees on estimates of inbreeding and inbreeding depression for days to first service and summit milk yield in Holsteins and Jerseys. *Journal of Dairy Science*, *86*, 2967–2976. https://doi.org/10.3168/jds.S0022-0302(03)73894-6

Chang, C. C. (2021). *PLINK 1.90 beta*. https://www.cog-genomics.org/plink/1.9/, (2021-06-30).

Chang, C. C., Chow, C. C., Tellier, L. C. A. M., Vattikuti, S., Purcell, S. M., & Lee, J. J. (2015). Second-generation PLINK: Rising to the challenge of larger and richer datasets. *GigaScience*, *4*(7), 1–16. https://doi.org/10.1186/s13742-015-0047-8

Choudhuri, S. (2014). Fundamentals of Molecular Evolution. In S. Choudhuri (Ed.), *Bioinformatics for Beginners* (pp. 27–53). Academic Press.

Christoph-Schulz, I., Salamon, P., & Weible, D. (2015). What is the benefit of organically-reared dairy

cattle? Societal perception towards conventional and organic dairy farming. *International Journal on Food System Dynamics*, *6*(3), 139–146. https://doi.org/10.18461/ijfsd.v6i3.632

Cole, J. B., Eaglen, S. A. E., Maltecca, C., Mulder, H. A., & Pryce, J. E. (2020). The future of phenomics in dairy cattle breeding. *Animal Frontiers*, *10*(2), 37–44. https://doi.org/10.1093/af/vfaa007

Colinet, F. G., Bouffioux, A., Mayeres, P., Malzahn, M., François, L., Janssens, S., Buys, N., Hiemstra, S. J., Windig, J. J., & Gengler, N. (2015). Genetic heritage of the Eastern Belgium Red and White breed, an endangered local breed. *EAAP Annual Meeting, Warsaw, Poland, August 31st – September 4th 2015 Genetic*, *2*, 5590.

Commission Herdbook Rouge-Pie de l'Est. (2017). *Rouge-Pie de l'Est*. http://www.rougepiedelest.be/index.php/language/fr/, (2021-04-09).

Coulon, A. (2010). Genhet: An easy-to-use R function to estimate individual heterozygosity. *Molecular Ecology Resources*, *10*, 167–169. https://doi.org/10.1111/j.1755-0998.2009.02731.x

de Barros Nascimento de Medeiros, R., Gomes Arandas, J. K., Silva Cavalcante, P. O., Vieira da Silva, N. M., Vieira de Oliveira, J. C., & Norma Ribeiro, M. (2020). Is multivariate analysis a useful tool to assess the morphometric profile of endangered goats? *Small Ruminant Research*, *190*. https://doi.org/https://doi.org/10.1016/j.smallrumres.2020.106175

de Rochambeau, H., Fournet-Hanocq, F., & Vu Tien Khang, J. (2000). Measuring and managing genetic variability in small populations. *Annales de Zootechnie*, *49*(2), 77–93. https://doi.org/10.1051/animres:2000109

De Winter, M. A., Vogelzang, T. A., & Van Schaick, J. (2010). *De blaarkop : ouderwets goed*.

Demonty, T. (2021). *East-Belgian Red & White cow*.

Dillon, P., Buckley, F., O'Connor, P., Hegarty, D., & Rath, M. (2003). A comparison of different dairy cow breeds on a seasonal grass-based system of milk production. 1. Milk production, live weight, body condition score and DM intake. *Livestock Production Science*, *83*(1), 21–33. https://doi.org/10.1016/S0301-6226(03)00041-1

Dockès, A. C., Magdelaine, P., Daridan, D., Guillaumin, A., Rémondet, M., Selmi, A., Gilbert, H., Mignon-Grasteau, S., & Phocas, F. (2011). Attentes en matière d'élevage des acteurs de la sélection animale, des filières de l'agroalimentaire et des associations. *INRA Productions Animales*, *24*(4), 285–296. https://doi.org/10.20870/productions-animales.2011.24.1.3233

Elevéo. (2018a). *PdS BBB Mixte*. https://www.awenet.be/awe/commun/asbl/communication/pds.php?espece=bovin&race=bbb mixte, (2021-08-04).

Elevéo. (2018b). *Programme de sélection Holstein (reproducteurs de race pure de l'espèce bovine)*. https://www.awenet.be/awe/commun/asbl/communication/pds.php?espece=bovin&race=hols tein, (2021-08-04).

Elevéo. (2019a). *Moyennes des productions par race et par lactation des animaux taris dans l'année*.

Elevéo. (2019b). *Règlement d'enregistrement des Généalogies Bovines chez Elevéo*.

Elevéo. (2021a). *PdS Rouge-Pie de l'Est*. https://www.awenet.be/awe/commun/asbl/communication/pds.php?espece=bovin&race=RPE , (2021-08-04).

Elevéo. (2021b). *references bases for combined interbull international evaluations and walloon evaluation of sires - disclaimer*. http://www.elinfo.be/indexEN.html, (2021-08-04)

Ellegren, H., & Galtier, N. (2016). Determinants of genetic diversity. *Nature Reviews Genetics*, *17*(7), 422–433. https://doi.org/10.1038/nrg.2016.58

ERFP. (2019). *Terms of Reference (TOR) for the European Regional Focal Point (ERFP) for Animal Genetic Resources*.

EU. (2016). *Regulation (EU) 2016/1012 of the European Parliament and of the Council of 8 June 2016 on zootechnical and genealogical conditions for the breeding, trade in and entry into the Union of purebred breeding animals, hybrid breeding pigs and the germinal*.

EuReCa. (2010). Local cattle breeds in Europe. In S. J. Hiemstra, Y. de Haas, A. Mäki-Tanila, & G. Gandini (Eds.), *Local cattle breeds in Europe*. Wageningen Academic Publishers. https://doi.org/10.3920/978-90-8686-697-7

EuroGenomics. (2019). *Launch of improved EuroGenomics genotyping beadchip : EuroG MD*.

Eusebi, P. G., Martinez, A., & Cortes, O. (2020). Genomic tools for effective conservation of livestock breed diversity. *Diversity*, *12*(1). https://doi.org/10.3390/d12010008

Evans, R. D., Dillon, P., Shalloo, L., Wallace, M., & Garrick, D. J. (2004). An economic comparison of dual-purpose and Holstein-Friesian cow breeds in a seasonal grass-based system under different milk production scenarios. *Irish Journal of Agricultural and Food Research*, *43*(1), 1–16. https://www.jstor.org/stable/25562501

FAO. (2007a). *Global plan of action for animal genetic resources and the Interlaken declaration*. http://www.journals.cambridge.org/abstract_S0043933909000245

FAO. (2007b). *The State of the World's Animal Genetic Ressources for Food and Agriculture - in brief* (D. Pilling & B. Rischkowsky (Eds.)). Food and Agriculture Organization of the United Nations.

FAO. (2011). Draft Guidelines on Phenotypic Characterization of Animal Genetic Resources. *COMMISSION ON GENETIC RESOURCES FOR FOOD AND AGRICULTURE - Thirteenth Regular Session*.

Felius, M. (2016). *On the breeds of cattle Their history, classification and conservation*. Universiteit Utrecht.

Fernández, J., Meuwissen, T. H. E., Toro, M. A., & Mäki-Tanila, A. (2011). Management of genetic diversity in small farm animal populations. *Animal*, *5*(11), 1684–1698. https://doi.org/10.1017/S1751731111000930

Forutan, M., Ansari Mahyari, S., Baes, C., Melzer, N., Schenkel, F. S., & Sargolzaei, M. (2018). Inbreeding and runs of homozygosity before and after genomic selection in North American Holstein cattle. *BMC Genomics*, *19*(98), 1–12. https://doi.org/10.1186/s12864-018-4453-z

François, L., Wijnrocx, K., Colinet, F. G., Gengler, N., Hulsegge, B., Windig, J. J., Buys, N., & Janssens, S. (2017). Genomics of a revived breed: Case study of the Belgian campine cattle. *PLoS ONE*, *12*(4), 1–14. https://doi.org/10.1371/journal.pone.0175916

Gandini, G., Avon, L., Bohte-Wilhelmus, D., Bay, E., Colinet, F. ., Choroszy, Z., Díaz, C., Duclos, D., Fernández, J., Gengler, N., Hoving-Bolink, R., Kearney, F., Lilja, T., Mäki-Tanila, A., Martín-Collado, D., Maurice-van Eijndhoven, M., Musella, M., Pizzi, F., Soini, K., … Hiemstra, S. . (2010). Motives and values in farming local cattle breeds in Europe: a survey on 15 breeds. *Animal Genetic Resources/Ressources Génétiques Animales/Recursos Genéticos Animales*, *47*, 45–58. https://doi.org/10.1017/s2078633610000901

Gandini, G., & Villa, E. (2003). Analysis of the cultural value of local livestock breeds: A methodology. *Journal of Animal Breeding and Genetics*, *120*(1), 1–11. https://doi.org/10.1046/j.1439-

0388.2003.00365.x

Gautason, E., Schönherz, A. A., Sahana, G., & Guldbrandtsen, B. (2021). Genomic inbreeding and selection signatures in the local dairy breed Icelandic Cattle. *Animal Genetics*, *52*, 251–262. https://doi.org/10.1111/age.13058

Gazal, S., Sahbatou, M., Perdry, H., Letort, S., Génin, E., & Leutenegger, A. L. (2014). Inbreeding coefficient estimation with dense SNP data: Comparison of strategies and application to HapMap III. *Human Heredity*, *77*(1–4), 49–62. https://doi.org/10.1159/000358224

Gebremariam, W. (2013). *Characterization of the Global Brown Swiss Cattle Population Structure*. Swedish University of Agricultural Sciences.

Geuder, U., Pickl, M., Scheidler, M., Schuster, M., & Goetz, K.-U. (2012). Growth performance, carcass traits and meat quality of Bavarian cattle breeds. *Zuchtungskunde*, *84*(6), 485–499.

Gomes Arandas, J. K., Vieira da Silva, N. M., Nascimento, R. de B., Filho, E. C. P., Brasil, L. H. de A., & Ribeiro, M. N. (2017). Multivariate analysis as a tool for phenotypic characterization of an endangered breed. *Journal of Applied Animal Research*, *45*(1), 152–158. https://doi.org/10.1080/09712119.2015.1125353

González Ariza, A., Arando Arbulu, A., León Jurado, J. M., Navas González, F. J., Delgado Bermejo, J. V., & Camacho Vallejo, M. E. (2021). Discriminant Canonical Tool for Differential Biometric Characterization of Multivariety Endangered Hen Breeds. *Animals*, *11*, 2211. https://doi.org/10.3390/ani11082211

Griffiths, A., Wessler, S., Lewontin, R., & Carroll, S. (2008). *Introduction to Genetic Analysis* (9th ed.). W.H. Freeman and Co.

Groeneveld, E., Westhuizen, B. D., Maiwashe, A., Voordewind, F., & Ferraz, J. B. (2009). POPREP: a generic report for population management. *Genetics and Molecular Research : GMR*, *8*(3), 1158–1178. https://doi.org/10.4238/vol8-3gmr648

Haiger, A., & Knaus, W. (2010). A comparison of dual-purpose Simmental and Holstein Friesian dairy cows in milk and meat production: 1st comm. Milk production without concentrates. *Züchtungskunde*, *82*(2), 131–143.

Hamer, R., & Simpson, P. M. (2000). Mixed-Up Mixed Models: Things That Look Like They Should Work But Don't, and Things That Look Like They Shouldn't Work But Do. *Proceedings of the Twenty-Fifth Annual SAS® Users Group International Conference*, 20–25. http://www2.sas.com/proceedings/sugi25/25/aa/25p020.pdf

Harris, A. M., & DeGiorgio, M. (2017). An unbiased estimator of gene diversity with improved variance for samples containing related and inbred individuals of any ploidy. *G3: Genes, Genomes, Genetics*, *7*(2), 671–691. https://doi.org/10.1534/g3.116.037168

Horn, M., Steinwidder, A., Gasteiner, J., Podstatzky, L., Haiger, A., & Zollitsch, W. (2013). Suitability of different dairy cow types for an Alpine organic and low-input milk production system. *Livestock Science*, *153*(1–3), 135–146. https://doi.org/10.1016/j.livsci.2013.01.011

Husson, A. F., Josse, J., Le, S., & Mazet, J. (2020). *Package ' FactoMineR '* (2.4).

Illumina. (2010). *BovineSNP50 Genotyping BeadChip*.

Illumina. (2020). *BovineSNP50 v3 BeadChip*.

ILRI. (2021). *Prosperity*. Research Themes. https://www.ilri.org/research/themes/prosperity, (2021-08-03).

Kaiser, H. F. (1961). A Note on Guttman's Lower Bound for the Number of Common Factors. *British Journal of Statistical Psychology*, *14*, 1–2.

Kantanen, J., Løvendahl, P., Strandberg, E., Eythorsdottir, E., Li, M. H., Kettunen-Praebel, A., Berg, P., & Meuwissen, T. (2015). Utilization of farm animal genetic resources in a changing agroecological environment in the Nordic countries. *Frontiers in Genetics*, *5*(FEB), 1–9. https://doi.org/10.3389/fgene.2015.00052

Kaptijn, G. (2016). Evaluation of the performance of dual-purpose cows in European pasture-based systems. In *Farming Systems Ecology Group*. University of Wageningen

Kardos, M., Luikart, G., & Allendorf, F. W. (2015). Measuring individual inbreeding in the age of genomics: Marker-based measures are better than pedigrees. *Heredity*, *115*(1), 63–72. https://doi.org/10.1038/hdy.2015.17

Kareiva, P., & Floberg, J. (2008). Endangered Species. In S. E. Jørgensen & B. D. Fath (Eds.), *Encyclopedia of Ecology* (1st ed., pp. 1246–1253). Elsvier B.V.

Keller, M. C., Visscher, P. M., & Goddard, M. E. (2011). Quantification of inbreeding due to distant ancestors and its detection using dense single nucleotide polymorphism data. *Genetics*, *189*, 237–249. https://doi.org/10.1534/genetics.111.130922

Kinghorn, B. P. (2011). An algorithm for efficient constrained mate selection. *Genetics Selection Evolution*, *43*(1), 1–9. https://doi.org/10.1186/1297-9686-43-4

Koenig, S., & Simianer, H. (2006). Approaches to the management of inbreeding and relationship in the German Holstein dairy cattle population. *Livestock Science*, *103*, 40–53. https://doi.org/10.1016/j.livsci.2005.12.009

Kohl, S., Wellmann, R., & Herold, P. (2020). Advanced optimum contribution selection as a tool to improve regional cattle breeds: A feasibility study for Vorderwald cattle. *Animal*, *14*(1), 1–12. https://doi.org/10.1017/S1751731119001484

Kohn, M. H., Murphy, W. J., Ostrander, E. A., & Wayne, R. K. (2006). Genomics and conservation genetics. *Trends in Ecology and Evolution*, *21*(11), 629–637. https://doi.org/10.1016/j.tree.2006.08.001

Kremer, V. D., Newman, S., Wilson, E. R., & Kinghorn, B. P. (2010). Mate Selection For Sustained Genetic Improvement In Small Populations. *Proceedings of the 9th World Congress on Genetics Applied to Livestock Production*, Paper 0536.

Kristensen, T. N., Hoffmann, A. A., Pertoldi, C., & Stronen, A. V. (2015). What can livestock breeders learn from conservation genetics and vice versa? *Frontiers in Genetics*, *5*, 1–12. https://doi.org/10.3389/fgene.2015.00038

Leffler, E. M., Bullaughey, K., Matute, D. R., Meyer, W. K., Ségurel, L., Venkat, A., Andolfatto, P., & Przeworski, M. (2012). Revisiting an Old Riddle: What Determines Genetic Diversity Levels within Species? *PLoS Biology*, *10*(9). https://doi.org/10.1371/journal.pbio.1001388

Li, M. H., Strandén, I., Tiirikka, T., Sevón-Aimonen, M. L., & Kantanen, J. (2011). A comparison of approaches to estimate the inbreeding coefficient and pairwise relatedness using genomic and pedigree data in a sheep population. *PLoS ONE*, *6*(11). https://doi.org/10.1371/journal.pone.0026256

Linn, J. G. (1988). FACTORS AFFECTING THE COMPOSITION OF MILK FROM DAIRY COWS. In Committee on the Technological Options to Improve the Nutritional Attributes of Animal Products Board on Agriculture National Research Council (Ed.), *Designing Foods: Animal Prodcut*

Options in the Marketplace (pp. 224–241). National Academy Press.

MacCluer, J. W., Boyce, A. J., Dyke, B., Weitkamp, L. R., Pfennig, D. W., & Parsons, C. J. (1983). Inbreeding and pedigree structure in Standardbred horses. *Journal of Heredity*, *74*(6), 394–399.

Makina, S. O., Muchadeyi, F. C., van Marle-Köster, E., MacNeil, M. D., & Maiwashe, A. (2014). Genetic diversity and population structure among six cattle breeds in South Africa using a whole genome SNP panel. *Frontiers in Genetics*, *5*(9), 1–7. https://doi.org/10.3389/fgene.2014.00333

Malécot, G. (1948). *Les Mathématiques de l'hérédité*. Masson et Cie.

Mastrangelo, S., Tolone, M., Ben Jemaa, S., Sottile, G., Di Gerlando, R., Cortés, O., Senczuk, G., Portolano, B., Pilla, F., & Ciani, E. (2020). Refining the genetic structure and relationships of European cattle breeds through meta-analysis of worldwide genomic SNP data, focusing on Italian cattle. *Scientific Reports*, *10*(1), 1–13. https://doi.org/10.1038/s41598-020-71375-2

Mazza, S., Guzzo, N., Sartori, C., & Mantovani, R. (2016). Genetic correlations between type and test-day milk yield in small dual-purpose cattle populations: The Aosta Red Pied breed as a case study. *Journal of Dairy Science*, *99*, 8127–8136. https://doi.org/10.3168/jds.2016-11116

Mészáros, G., Boison, S. A., Pérez O'Brien, A. M., Ferenčaković, M., Curik, I., Da Silva, M. V. B., Utsunomiya, Y. T., Garcia, J. F., & Sölkner, J. (2015). Genomic analysis for managing small and endangered populations: A case study in Tyrol Grey cattle. *Frontiers in Genetics*, *6*, 173. https://doi.org/10.3389/fgene.2015.00173

Meuwissen, T. H. E., & Sonesson, A. K. (1998). Maximizing the Response of Selection with a Predefined Rate of Inbreeding: Overlapping Generations. *Journal of Animal Science*, *76*(10), 2575–2583. https://doi.org/10.2527/1998.76102575x

Miglior, F., & Burnside, E. B. (1995). Inbreeding of Canadian Holstein Cattle. *Journal of Dairy Science*, *78*(5), 1163–1167. https://doi.org/10.3168/jds.S0022-0302(95)76733-9

Mota, R. R., Mayeres, P., Bastin, C., Glorieux, G., Bertozzi, C., Vanderick, S., Hammami, H., Colinet, F. G., & Gengler, N. (2017). Genetic evaluation for birth and conformation traits in dual-purpose Belgian Blue cattle using a mixed inheritance model. *Journal of Animal Science*, *95*, 4288–4299. https://doi.org/10.2527/jas2017.1748

Mrode, R., Kearney, J. F., Biffani, S., Coffey, M., & Canavesi, F. (2009). Short communication: Genetic relationships between the Holstein cow populations of three European dairy countries. *Journal of Dairy Science*, *92*, 5760–5764. https://doi.org/10.3168/jds.2008-1931

Nei, M. (1973). Analysis of gene diversity in subdivided populations. *Proceedings of the National Academy of Sciences of the United States of America*, *70*(12 (I)), 3321–3323. https://doi.org/10.1073/pnas.70.12.3321

Nielsen, H. M., & Kargo, M. (2020). An endangered horse breed can be conserved by using optimum contribution selection and preselection of stallions. *Acta Agriculturae Scandinavica A: Animal Sciences*, *69*(1–2), 127–130. https://doi.org/10.1080/09064702.2020.1728370

Niero, G., Visentin, G., Ton, S., De Marchi, M., Penasa, M., & Cassandro, M. (2016). Phenotypic characterisation of milk technological traits,protein fractions,and major mineral and fatty acid composition of Burlina cattle breed. *Italian Journal of Animal Science*, *15*(4), 576–583. https://doi.org/10.1080/1828051X.2016.1250128

Ostbelgien. (n.d.). Wikipedia. https://de.wikipedia.org/wiki/Ostbelgien, (2021-08-03)

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., Maller, J., Sklar, P., De Bakker, P. I. W., Daly, M. J., & Sham, P. C. (2007). PLINK: A tool set for whole-genome

association and population-based linkage analyses. *American Journal of Human Genetics*, *81*(3), 559–575. https://doi.org/10.1086/519795

Rahim, N. G., Harismendy, O., Topol, E. J., & Frazer, K. A. (2008). Genetic determinants of phenotypic diversity in humans. *Genome Biology*, *9*(4). https://doi.org/10.1186/gb-2008-9-4-215

Ray, D. E., Halbach, T. J., & Armstrong, D. V. (1992). Season and Lactation Number Effects on Milk Production and Reproduction of Dairy Cattle in Arizona. *Journal of Dairy Science*, *75*, 2976–2983. https://doi.org/10.3168/jds.S0022-0302(92)78061-8

Reed, E., Nunez, S., Kulp, D., Qian, J., Reilly, M. P., & Foulkes, A. S. (2015). A guide to genome-wide association analysis and post-analytic interrogation. *Statistics in Medicine*, *34*, 3769–3792. https://doi.org/10.1002/sim.6605

Rosche, C., Schrieber, K., Lachmuth, S., Durka, W., Hirsch, H., Wagner, V., Schleuning, M., & Hensen, I. (2018). Sex ratio rather than population size affects genetic diversity in Antennaria dioica. *Plant Biology*, *20*, 789–796. https://doi.org/10.1111/plb.12716

RStudio Team. (2021). *RStudio: Integrated Development Environment for R*. RStudio, PBC. http://www.rstudio.com/, (2021-07-05).

Salazar, J. A., Rubio, M., Ruiz, D., Tartarini, S., Martínez-Gómez, P., & Dondini, L. (2015). SNP development for genetic diversity analysis in apricot. *Tree Genetics and Genomes*, *11*(1). https://doi.org/10.1007/s11295-015-0845-2

SanCristobal, M., Chevalet, C., Haley, C. S., Joosten, R., Rattink, A. P., Harlizius, B., Groenen, M. A. M., Amigues, Y., Boscher, M. Y., Russell, G., Law, A., Davoli, R., Russo, V., Désautés, C., Alderson, L., Fimland, E., Bagga, M., Delgado, J. V., Vega-Pla, J. L., … Cardellino, R. (2006). Genetic diversity within and between European pig breeds using microsatellite markers. *Animal Genetics*, *37*(3), 189–198. https://doi.org/10.1111/j.1365-2052.2005.01385.x

Sargolzaei, M., Iwaisaki, H., & Colleau, J. J. (2006). *CFC: Coancestry, Inbreeding (F), Contribution* (1.0).

Sartori, C., Guzzo, N., Mazza, S., & Mantovani, R. (2018). Genetic correlations among milk yield, morphology, performance test traits and somatic cells in dual-purpose Rendena breed. *Animal*, *12*(5), 906–914. https://doi.org/10.1017/S1751731117002543

Schierenbeck, S., Pimentel, E. C. G., Tietze, M., Körte, J., Reents, R., Reinhardt, F., Simianer, H., & König, S. (2011). Controlling inbreeding and maximizing genetic gain using semi-definite programming with pedigree-based and genomic relationships. *Journal of Dairy Science*, *94*(12), 6143–6152. https://doi.org/10.3168/jds.2011-4574

Senczuk, G., Mastrangelo, S., Ciani, E., Battaglini, L., Cendron, F., Ciampolini, R., Crepaldi, P., Mantovani, R., Bongioni, G., Pagnacco, G., Portolano, B., Rossoni, A., Pilla, F., & Cassandro, M. (2020). The genetic heritage of Alpine local cattle breeds using genomic SNP data. *Genetics Selection Evolution*, *52*(1), 1–12. https://doi.org/10.1186/s12711-020-00559-1

Servais, L. (2016). La Rouge-Pie de l'Est Retour en grâce d 'une mixte oubliée. *Wallonie Elevages*, 8–11.

Simianer, H. (2005). Using expected allele number as objective function to design between and within breed conservation of farm animal biodiversity. *Journal of Animal Breeding and Genetics*, *122*(3), 177–187. https://doi.org/10.1111/j.1439-0388.2005.00523.x

Solé, M., Valera, M., Gómez, M. D., Cervantes, I., & Fernández, J. (2013). Implementation of Optimum Contributions Selection in endangered local breeds: The case of the Menorca Horse population. *Journal of Animal Breeding and Genetics*, *130*(3), 218–226. https://doi.org/10.1111/jbg.12023

Sölkner, J., Miesenberger, J., Willam, A., Fuerst, C., & Baumung, R. (2000). Total merit indices in dual purpose cattle. *Archives Animal Breeding*, *43*(6), 597–608. https://doi.org/10.5194/aab-43-597-2000

Stachowicz, K., Sargolzaei, M., Miglior, F., & Schenkel, F. S. (2011). Rates of inbreeding and genetic diversity in Canadian Holstein and Jersey cattle. *Journal of Dairy Science*, *94*, 5160–5175. https://doi.org/10.3168/jds.2010-3308

Strandén, I., & Peura, J. (2007). Inbreeding and relationship coefficients in the Finnish blue fox population. *Agricultural and Food Science*, *16*, 147–156. https://doi.org/10.2137/145960607782219319

Sturaro, E., Marchiori, E., Cocca, G., Penasa, M., Ramanzin, M., & Bittante, G. (2013). Dairy systems in mountainous areas: Farm animal biodiversity, milk production and destination, and land use. *Livestock Science*, *158*(1–3), 157–168. https://doi.org/10.1016/j.livsci.2013.09.011

Suhardi, Summpunn, P., Duangjinda, M., & Wuthisuthimethavee, S. (2020). Phenotypic diversity characterization of Kalang and Thale Noi buffalo (Bubalus bubalis) in Indonesia and Thailand: Perspectives for the buffalo breeding development. *Biodiversitas*, *21*(11), 5128–5137. https://doi.org/10.13057/biodiv/d211118

Syrstad, O. (1993). Evaluation of dual-purpose (milk and meat) animals. *World Animal Review*, *77*, 56–59.

Toro, M. A., & Caballero, A. (2005). Characterization and conservation of genetic diversity in subdivided populations. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *360*(1459), 1367–1378. https://doi.org/10.1098/rstb.2005.1680

Trujano-chavez, M. Z., Valerio-Hernández, J. E., López-Ordaz, R., & Ruíz-Flores, A. (2021). Minor allele frequency in genomic prediction for growth traits in Braunvieh cattle. *Revista Bio Ciencias*, *8*. https://doi.org/https://doi.org/10.15741/revbio.08.e1

Upadhyay, M. (2019). *Genomic variation across European cattle: contribut ion of gene flow*. Swedish University of Agricultural Sciences. Wageningen University.

van de Graaf, A. (2015). *Textbook Animal Breeding and Genetics*. Groen Kennisnet. https://wiki.groenkennisnet.nl/display/TAB/Textbook+Animal+Breeding+and+Genetics, (2021-08-03).

Vanderick, S., Reis Mota, R., Wijnrocx, K., & Gengler, N. (2020). *Description of the Genetic Evaluation Systems used in the Walloon Region of Belgium*. http://www.elinfo.be/indexEN.html, (2021-07-03).

VanRaden, P. M. (2008). Efficient methods to compute genomic predictions. *Journal of Dairy Science*, *91*, 4414–4423. https://doi.org/10.3168/jds.2007-0980

Vilà, C., Sundqvist, A. K., Flagstad, Ø., Seddon, J., Björnerfeldt, S., Kojola, I., Casulli, A., Sand, H., Wabakken, P., & Ellegren, H. (2003). Rescue of a severely bottlenecked wolf (Canis lupus) population by a single immigrant. *Proceedings of the Royal Society B: Biological Sciences*, *270*, 91–97. https://doi.org/10.1098/rspb.2002.2184

Wang, J. (2014). Marker-based estimates of relatedness and inbreeding coefficients: An assessment of current methods. *Journal of Evolutionary Biology*, *27*(3), 518–530. https://doi.org/10.1111/jeb.12315

Wang, Jinliang. (2016). Pedigrees or markers: Which are better in estimating relatedness and inbreeding coefficient? *Theoretical Population Biology*, *107*, 4–13.

https://doi.org/10.1016/j.tpb.2015.08.006

Wang, Y., Bennewitz, J., & Wellmann, R. (2017). Novel optimum contribution selection methods accounting for conflicting objectives in breeding programs for livestock breeds with historical migration. *Genetics Selection Evolution*, *49*(1), 1–12. https://doi.org/10.1186/s12711-017-0320-7

Weir, B. S., Anderson, A. D., & Hepler, A. B. (2006). Genetic relatedness analysis: Modern data and new challenges. *Nature Reviews Genetics*, *7*, 771–780. https://doi.org/10.1038/nrg1960

Wellmann, R. (2021a). *Package ' optiSel '* (2.0.5).

Wellmann, R. (2021b). *Pedigree-based Evaluations*. https://cran.r-project.org/web/packages/optiSel/vignettes/ped-vignette.html, (2021-06-30).

Willoughby, J. R., Fernandez, N. B., Lamb, M. C., Ivy, J. A., Lacy, R. C., & Dewoody, J. A. (2015). The impacts of inbreeding, drift and selection on genetic diversity in captive breeding populations. *Molecular Ecology*, *24*, 98–110. https://doi.org/10.1111/mec.13020

Wilmot, H., Bormann, J., Soyeurt, H., Hubin, X., Glorieux, G., Mayeres, P., Bertozzi, C., & Gengler, N. (2021). Development of a genomic tool for breed assignment by comparison of different classification models - Application to three local cattle breeds. *Journal of Animal Breeding and Genetics*. Under Press

Windig, J. J., & Kaal, L. (2008). An effective rotational mating scheme for inbreeding reduction in captive populations illustrated by the rare sheep breed Kempisch Heideschaap. *Animal*, *2*(12), 1733–1741. https://doi.org/10.1017/S1751731108003029

Windig, J. J., Verweij, M. J. W., & Oldenbroek, J. K. (2019). Reducing inbreeding rates with a breeding circle: Theory and practice in Veluws Heideschaap. *Journal of Animal Breeding and Genetics*, *136*, 51–62. https://doi.org/10.1111/jbg.12371

Woolliams, J. (1989). Modifications to MOET nucleus breeding schemes to improve rates of genetic progress and decrease rates of inbreeding in dairy cattle. *Animal Science*, *49*(1), 1–14. https://doi.org/10.1017/S0003356100004190

Wright, S. (1922). Coefficients of Inbreeding and Relationship. *The American Naturalist*, *56*, 330–338.

Yoshida, G. M., Yáñez, J. M., de Queiroz, S. A., & Carvalheiro, R. (2020). Mate selection provides similar genetic progress and average inbreeding than optimum contribution selection in the long-term. *Aquaculture*, *526*, 1–7. https://doi.org/10.1016/j.aquaculture.2020.735376

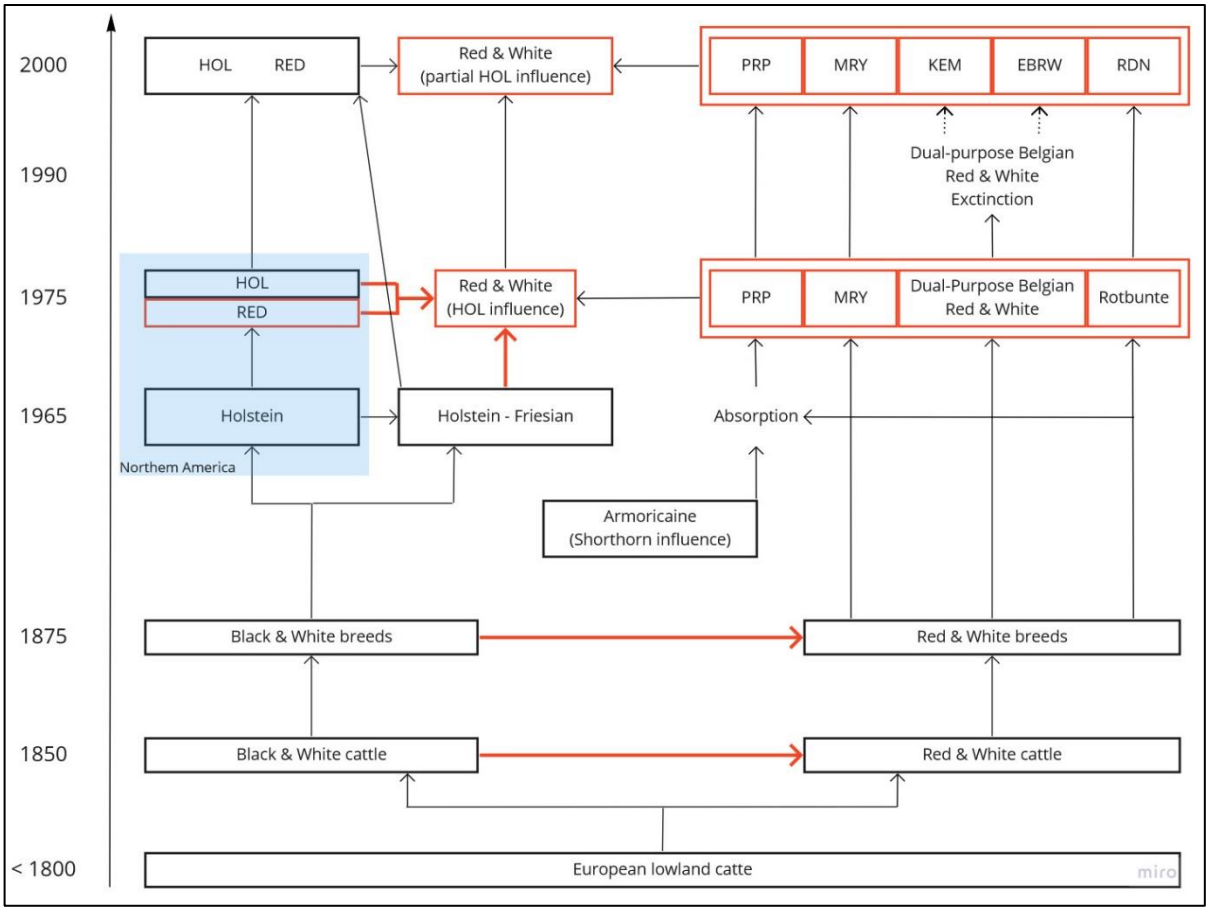
Zhang, Q., Calus, M. P. L., Guldbrandtsen, B., Lund, M. S., & Sahana, G. (2015). Estimation of inbreeding using pedigree, 50k SNP chip genotypes and full sequence data in three cattle breeds. *BMC Genetics*, *16*, 1–11. https://doi.org/10.1186/s12863-015-0227-7

Appendices

Appendix 1.: Political regions of Belgium and East-Belgium (Ger.: *Ostbelgien*).

Adapted from: "*Ostbelgien*" (n.d.).

Appendix 2.: Historical background of some European cattle breeds.



: red factor Source: Bouffioux (2014)

EBRW: East-Belgian Red & White; KEM: Campine breed; HOL: Holstein; RED: Red Holstein; PRP: French Red-Pied Lowland; MRY: Maas-Rijn-Yssel; RDN: Rotbunte Doppelnutzung.

Appendix 3.: EU thresholds for endangered breeds of different livestock species based on the number of breeding females.

Eligible farm animal species	Thresholds under which a local breed is considered as endangered (number of breeding females (*))
Cattle	7,500
Sheep	10,000
Goat	10,000
Equidae	5,000
Pigs	15,000
Avian	25,000
(*) Number, calculated for all Member States, of breeding females of the same breed available for pure-bred reproduction registered in a herd-book kept by an approved breeding organisation recognised by the Member State in accordance with Community zootechnical legislation.	

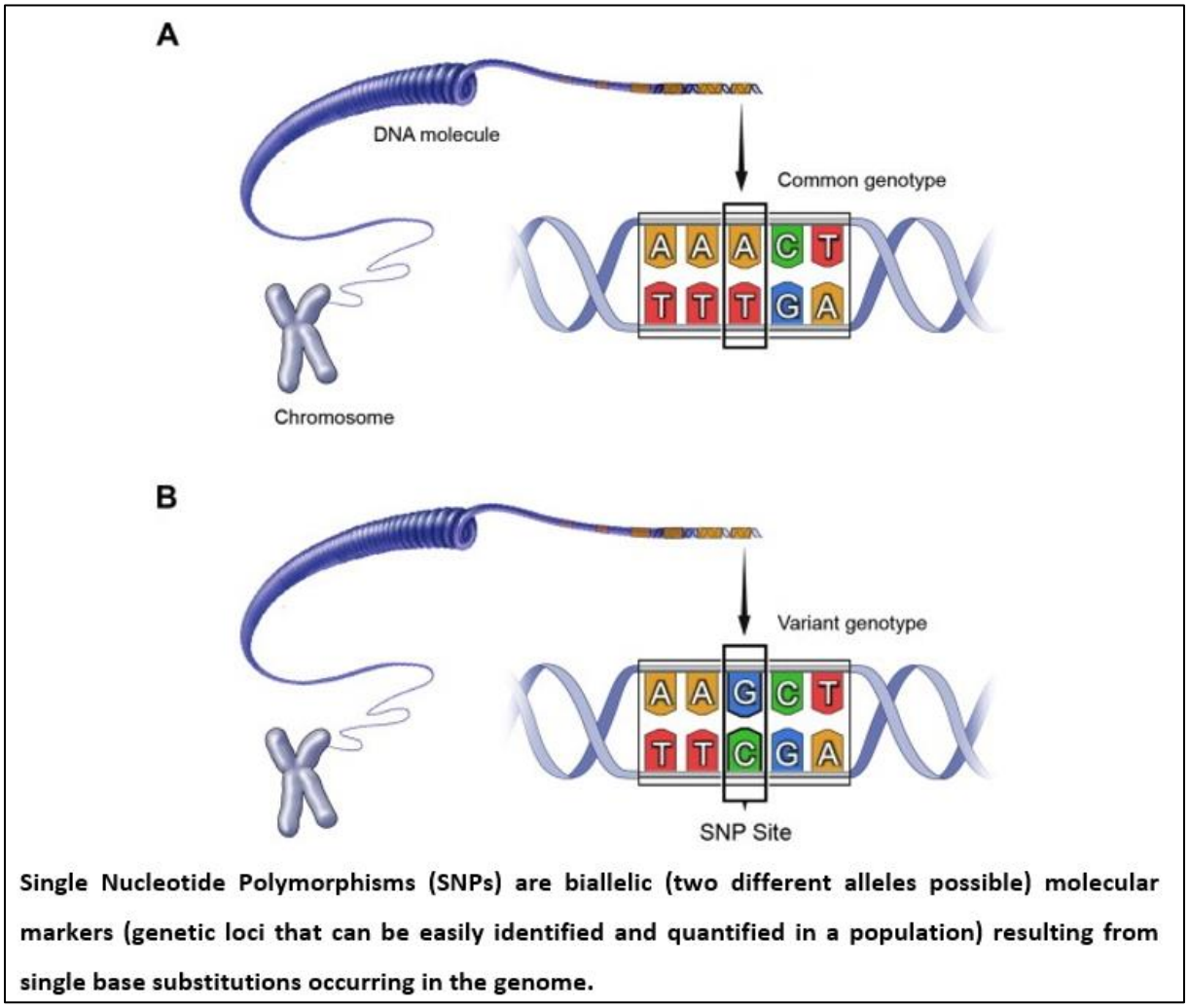
Source: EU (2016)

Appendix 4.: FAO new classification system for breed populations.

Under the new system of breed classification developed for The State of the World's Animal Genetic Resources for Food and Agriculture, the primary distinction is between breeds that occur in only one country, which are referred to as "local" breeds, and those that occur in more than one country, which are referred to as "transboundary" breeds. Within the transboundary breed category, a further distinction is drawn between "regional" transboundary breeds – those that occur in more than one country within a single region, and "international" transboundary breeds – those that occur in more than one region. The decision as to which national-level breed populations should be considered as belonging to a transboundary breed was taken on the basis of expert judgment and reviewed by National Coordinators for the Management of Animal Genetic Resources from the relevant countries. Although some refinements are still required, the new classification has proved to be very useful as a framework for assessing breed diversity at global and regional levels.

Source: FAO (2007)

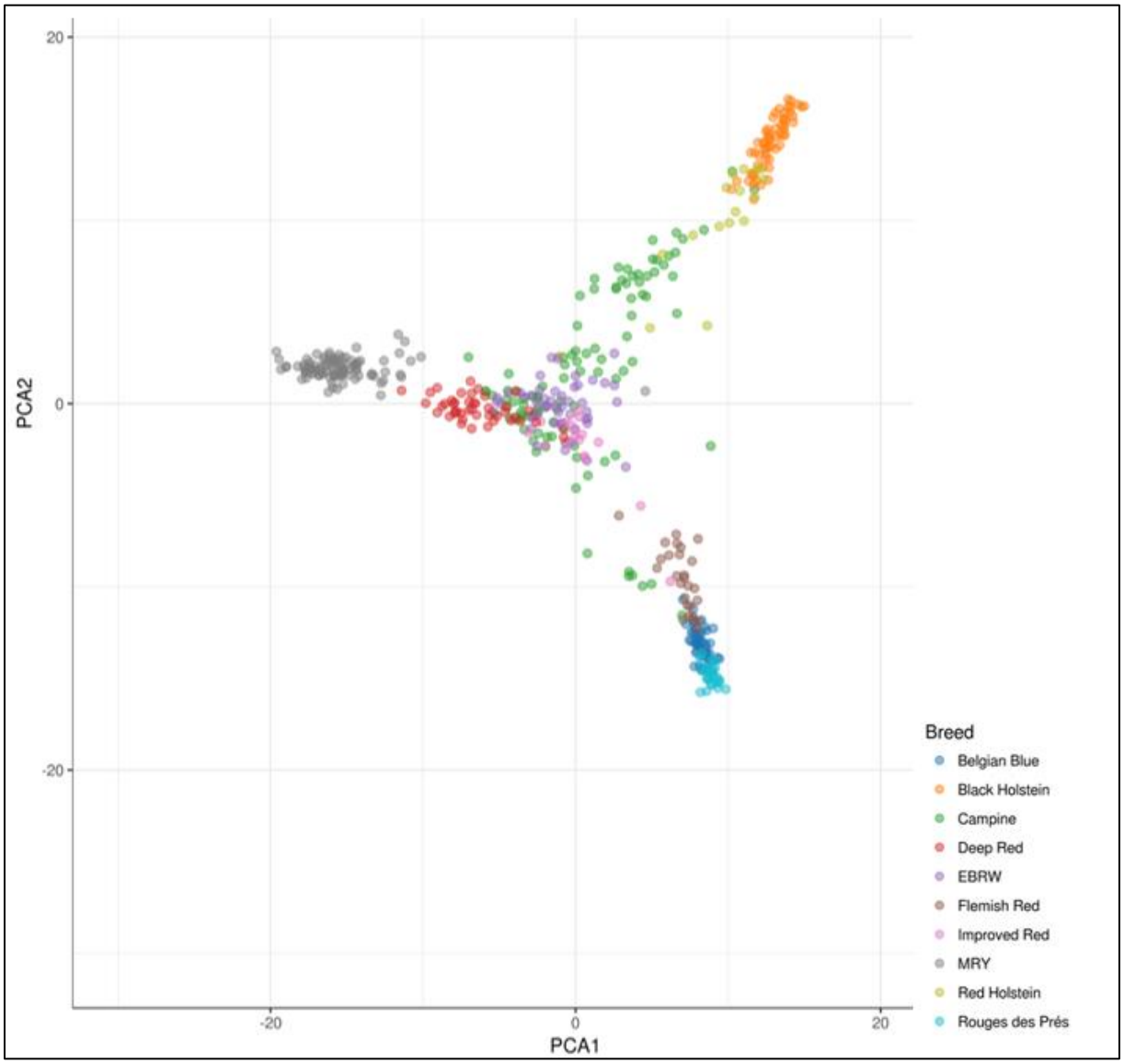
Appendix 5.: Single Nucleotide Polymorphism **A**: common genotype **B**: variant genotype.



Single Nucleotide Polymorphisms (SNPs) are biallelic (two different alleles possible) molecular markers (genetic loci that can be easily identified and quantified in a population) resulting from single base substitutions occurring in the genome.

Adapted from: Camp & Trujillo (2014)

Appendix 6.: Principal component analysis showing the genomic relationship between ten cattle breeds.



Source: François et al. (2017)

Appendix 7.: EBRW breed standards.

La vache Rouge-Pie de l'Est est une vache de type mixte (à deux fins) de taille moyenne, précoce et robuste ayant de bonnes caractéristiques laitières et des qualités viandeuse accusées (facile à engraisser).

Elle a un caractère calme et « facile à l'entretien », est très rustique et présente une bonne longévité (10-12 ans).

La tête est petite, assez courte et large avec un mufle assez large. Les cornes sont assez petites, légèrement courbés vers l'avant et vers l'intérieur.

Robe : Sa robe est rouge et blanche, les différentes tâches de couleur sont clairement séparées, la tête rousse peut avoir une tache blanche de taille variable (mais sans occuper le visage entier, comme chez le Fleckvieh). Les parties en dessous du ventre et les pattes doivent être blanches ; la pointe de la queue est blanche. La couleur « rouge brulée » (vers le noir) est admise.

Morphologie :

- Les épaules sont obliques, longues, larges, bien musclées et fermes. Le garrot est assez large.
- Sa poitrine est large pour offrir la place pour le cœur, les poumons, et ainsi assurer une bonne rusticité.
- Son corps est profond et large avec des côtes arquées, lui permettant de bien valoriser les fourrages grossiers.
- Le dos est également large et d'une longueur suffisante, bien musclé, fort et ferme.
- Les hanches doivent être bien larges, fortes, bien musclées.
- L'arrière-train, de même que les hanches sont les principaux fournisseurs de viande de bonne qualité. Dès lors, il joue un rôle important dans l'appréciation des caractéristiques viandeuses.
- La croupe, en premier lieu, est assez large, notamment en vue du vêlage, de la formation de viande et de l'implantation du pis. Elle est bien musclée.
- La culotte typique est légèrement rebondie vers l'arrière et sur les côtés, et est profonde. Ainsi, on a généralement un bassin large et fort dont la viande est de première qualité.
- Le bassin, légèrement incliné, permet des vêlages faciles.
- La queue doit être solidement implantée, légèrement saillante et assez fine.
- Les membres doivent être courts, forts, secs et assez musclés.

La hauteur au garrot des vaches adultes se situe entre 1,35 m et 1,40 m, pour un poids allant de 600 à 750 kg. La production laitière des vaches avoisine 5000 kg de lait, avec des taux de MG et de protéines avoisinant 4.20 et 3,50% respectivement.

Source: Commission Herdbook Rouge-Pie de l'Est (2017)
French: http://www.rougepiedelest.be/index.php/language/fr/presentation/standard-racial/
German: http://www.rougepiedelest.be/index.php/language/de/vorstellung/rassestandard/

Appendix 8.: East-Belgian Red & White cows.

a. Head

Source: Demonty (2021)

b. The “burned red” colour.

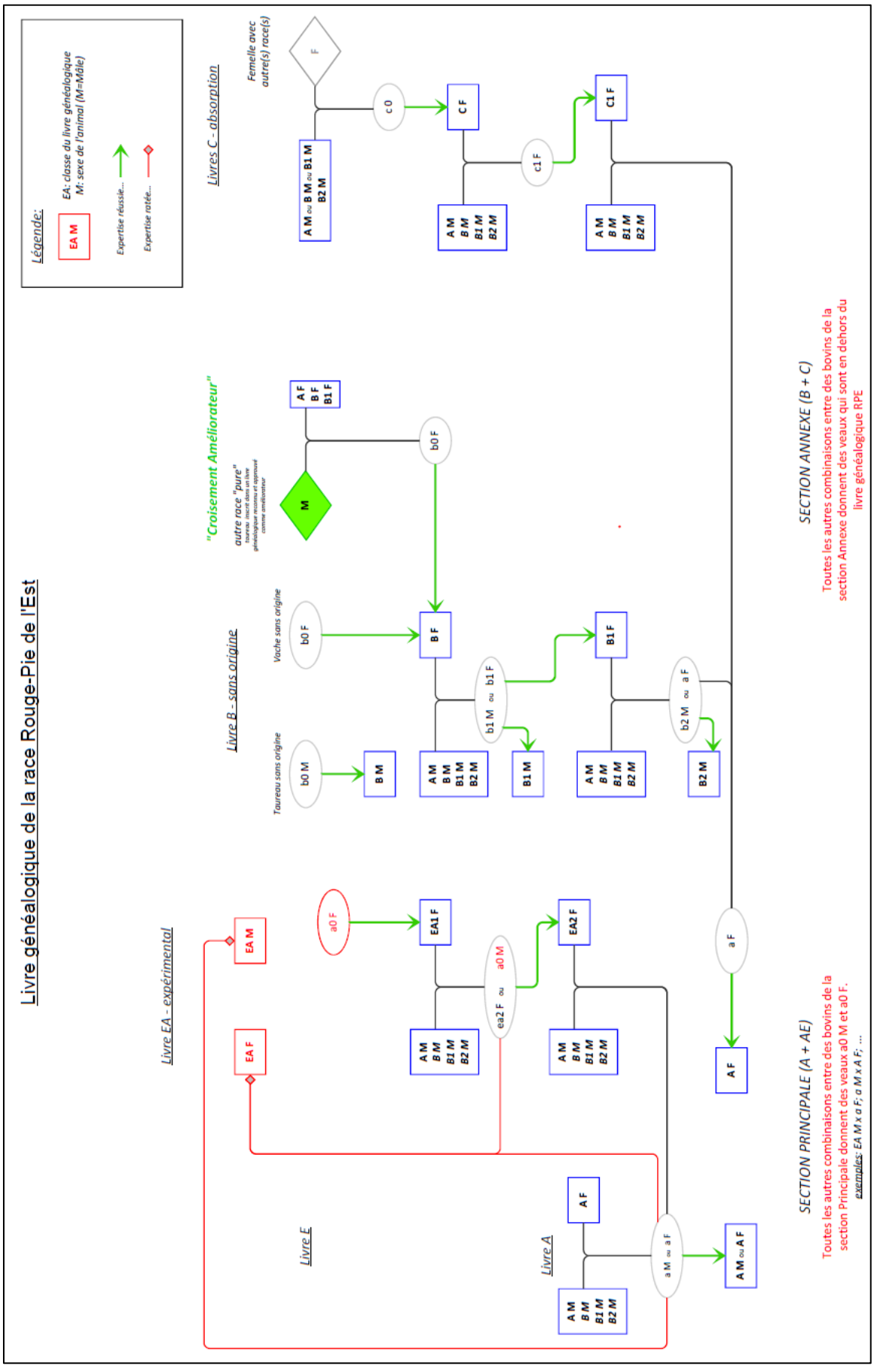
Source: Demonty (2021)

Appendix 9.: Walloon milk recording results for EBRW and 4 other breeds in 2019.

Breed	Number of animals	Milk yield (kg)	Fat yield (kg)	Fat percentage	Protein yield (kg)	Protein percentage
EBRW	590	5,154	213	4.14	177	3.43
DPBB	2,244	4,433	164	3.69	147	3.31
Holstein	35,954	9,391	380	4.04	319	3.39
Montbéliarde	911	7,878	312	3.96	274	3.47
Normande	238	6,824	285	4.17	244	3.57

Source: Elevéo (2019)

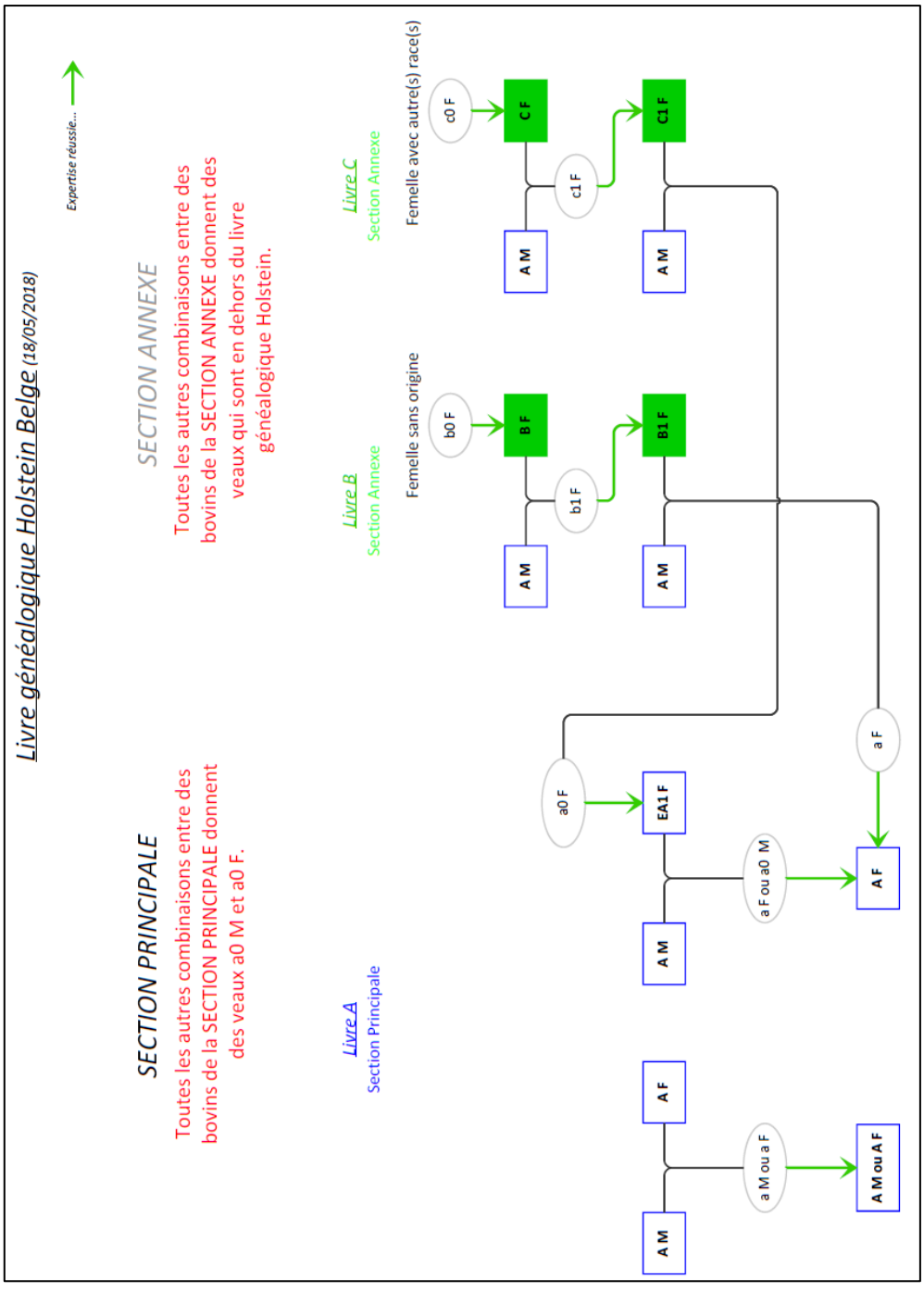
Appendix 10.: EBRW herd-book management scheme.



Book A: main section, purebred animals; Book B: animals without known origin; Book C: animals with ancestors from other breeds (together, books B and C are called the annex section of the herdbook); Book E: animals failing the phenotypical inspection and/or genomic breed assignment. The B1 and B2 subsections are attributed to offspring of a female being registered in subsections B and B1 respectively. The C1 subsection is dedicated for the offspring of females from C subsection.

Source: Elevéo (2021)

Appendix 11.: Belgian Holstein herd-book management scheme.



Source: Elevéo (2018b)

Appendix 12.: Zootechnical certificate for animals registered in the main EBRW HB section.

Certificat Zootechnique Rouge Pie de l'Est

1 Elevéo asbl
Rue des Champs Elysées 4 – 5590 Ciney (Belgique) – Tél : +32 (0) 83 23 06 11

Ce document contient les informations requises, pour un certificat zootechnique, par le règlement (UE) 2016/1012, pour les échanges de reproducteurs de race pure de l'espèce bovine (Bos taurus, Bos indicus, Bubalus bubalis).

10 Naisseur : DUPONT Jean-Baptiste Rue du chêne 12 6440 Boussu-lez-Walcourt	11 Propriétaire : DUPONT Jean-Baptiste Rue du chêne 12 6440 Boussu-lez-Walcourt

6. n° inscription dans le LG : BE-826139550 7.1 Système : conforme au règlement 1760/2000 7.3 Boucle saumon : BE 8-26139550 7.4 Nom : Pauline de Waleffes 9 Date de naissance : 24/08/2015 – pays : Belgique	2. LG : Rouge Pie de l'Est 3. Race : Rouge Pie de l'Est de la Belgique 4 Section principale, livre A (2018) 5 Sexe : Femelle Robe : Rouge et Blanche 8 Vérification identité : 8.1 : ADN 8.2 : père/parenté confirmé

12.1 Père Notair de Waleffes BE-757546963 Rouge et blanche SP – livre A Parenté confirmée	12.1.1 Grand-père paternel Notair de Waleffes BE-757546963 Rouge et blanche SP – livre A	Notair de Waleffes BE-757546963
		Notair de Waleffes BE-757546963
	12.1.2 Grand-mère paternelle Notair de Waleffes BE-757546963 Rouge et blanche SP – livre A	Notair de Waleffes BE-757546963
		Notair de Waleffes BE-757546963
12.2 Mère Notaire de Waleffes BE-757546964 Rouge et blanche SA – livre B1 Père confirmé	12.2.1 Grand-père maternel Notair de Waleffes BE-757546963 Rouge et blanche SP – livre A	Notair de Waleffes BE-757546963
		Notair de Waleffes BE-757546963
	12.2.2 Grand-mère maternelle Notair de Waleffes BE-757546963 Rouge et blanche SA – livre B	Notair de Waleffes BE-757546963

14 Insémination 14.1 01/05/2018 14.2.1 BE- 156726062 14.2.2 MEDELLO 14.2.3 ADN – confirmé père	13 Informations complémentaires 13.1 voir verso 13.2 voir verso 13.3 voir verso 13.4 voir verso

15.1 Fait à Ciney 15.2 le : 26/06/2018
N° émission : 1

15.3 P. MAYERES, DIRECTEUR DES SERVICES A L'ELEVAGE

B DARIMONT, PRESIDENT DU HERDBOOK

15.4

CERTIFICAT DE PERFORMANCES Date d'édition : 26/12/2018

Association Wallonne de l'Elevage asbl,
rue des Champs Elysées 4 5590 Ciney (Belgique)
www.awenet.be

IDENTIFICATION

6. n° inscription dans le LG : BE-826139550	12.1. P : GBR 582093 PICSTON SHAKER
7.3 Boucle saumon : BE 8-26139550	
7.4 Nom : Pauline de Waleffes	
9 Date de naissance : 24/08/2015	12.2. M : BE 853732466 BACARA
5 Sexe : Femelle Robe : rouge et blanc	

LIVRE GENEALOGIQUE Rouge Pie de l'Est livre A (2017)

13.1. CONTRÔLE DES PERFORMANCES

Lactations en 305 jours animal

	N°	Age	Kg L	%MG	KgMG	%Prot	KgProt
First	1	2a 11m	6710	3.73	250	3.37	226
	2	4a 0m	7121	3.57	254	3.23	230
	3	5a 2m	8883	3.53	314	3.25	288
	4	6a 10m	8905	3.55	316	3.28	292
	5	8a 10m	9725	3.66	356	3.18	309
Ø5		8a 10m	8269	3.6	298	3.25	269
Best	5		9725	3.66	356	3.18	309

Lactations en 305 jours mère (12.2)

	N°	Age	Kg L	%MG	KgMG	%Prot	KgProt
First	1	2a 5m	7635	3.48	265	3.15	241
Ø1			7635	3.48	265	3.15	241
Best	1	2a 5m	7635	3.48	265	3.15	241

Lactations en 305 jours Grand-mère mat (12.2.2)

First	1	2a 2m	7754	3.45	267	3.26	253
Ø4			9138	3.76	344	3.32	303
Best	3	4a 3m	10164	3.61	367	3.33	339

Dernière lactation

N°	Date	Kg L	% MG	% Prot	N°	Date	Kg L	% MG	% Prot	N°	Date	Kg L	% MG	% Prot
1	18/11/17	41.9	3.20	3.28	6	14/04/18	30.8	3.47	3.01	11	14/10/18			
2	23/12/17	33.4	2.78	2.90	7	18/05/18	32.7	3.99	3.22	12	17/11/18			
3	15/01/18	32.4	3.00	2.86	8	16/06/18	24.8	3.36	3.11					
4	17/02/18	31.3	3.51	2.89	9	18/08/18	24.2	2.41	3.16					
5	13/03/18	36.5	3.47	3.01	10	18/09/18	25.4	3.32	3.73					

Production Laitière totale (Lactations terminées)

	N°	J Lact	Kg L	%MG	%Prot	Kg MU		N°	J Lact	Kg L	%MG	%Prot	Kg MU
Animal	5	1088	51200	3.65	3.25	1400	Mère 12.2	5	1088	51200	3.65	3.25	1400

Adéquation au standard racial

Date d'expertise : 24/12/2018 Expertise réussie
Taille : 134cm, le 24/12/20108 Conformation : 3

13.2. EVALUATION GENETIQUE

Index	animal	Père (12.1)	Mère (12.2)	Gestation	Index	animal	Père (12.1)	Mère (12.2)	Gestation
Production (date 12/18)									
V€L	182				Kg Mg				
Rep	*93*				Kg Prot				
Kg Lait	866								
% MG	-0.03				Cellules	3.73			
% Prot	0.14								

13.3. CARACTERISTIQUES GENETIQUES

Appartenance au pool génétique de la race : confirmée

Source: Elevéo (2021)

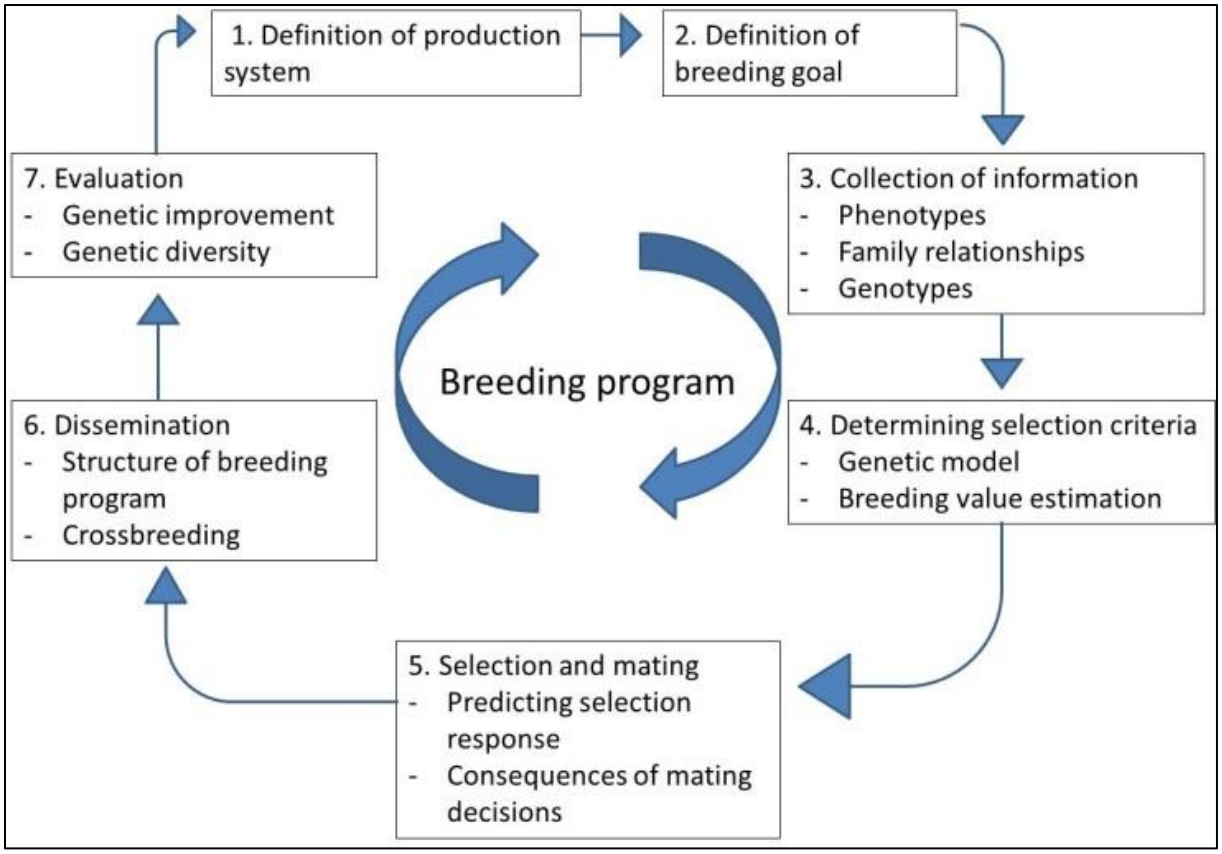
Appendix 13.: Morphological traits evaluated by the linear classification grid in DPBB.

General Synthesis	
Udder Synthesis	**Muscularity Synthesis**
Udder balance	Shoulder muscularity
Udder depth	Back muscularity
Udder ligament	Buttock muscularity
Front Teat position	**Pelvis Synthesis**
Back Teat position	Pelvis width
Teat length	Pelvis inclination
Development Synthesis	Tail head
Sacrum size (cm)	**Other**
Chest depth	Bones
Aplomb Synthesis	Skin
Hock	Back line
Feet angle	

A specific linear classification grid has been developed in partnership with the French *Bleue du Nord* Breeding Association. The linear classification is a service offered to breeders, without obligation, or restriction of access to certain classes of the Dual-Purpose Belgian Blue Herd-Book.The previous table shows all the assessed traits (or measured for height) on lactating cows. It is divided into 5 synthetic scores (Udder, Development, Aplomb, Muscularity, Pelvis) which are combined into a General Score.

Source: Elevéo (2018a)

Appendix 14.: Breeding program structure.



Source: van de Graaf (2015)

Appendix 15.: Conversion table between SNP and allele codes.

Description	SNP code "0,1,2,3,5"	SNP code "0,1,2,-9"	Allele code "AB"
Homozygous for major allele	0	0	AA
Heterozygous	1	1	AB
Heterozygous	2	1	BA
Homozygous for minor allele	3	2	BB
Error	5	-9	00

Appendix 16.: Distribution of HB animals per interval of EBRW breed composition percentage.

Percentage sections of EBRW breed composition:		[0;25[	[25;50[	[50;75[	[75;87.5[	[87.5;100]	Total
Herd-book section:	A	4	20	17	12	9	62
	B	5	56	85		1996	2142
	B1	20	139	375	8	256	798
	B2	1	1	2		1	5
	C		6	15			21
	C1			1			1
Total		30	222	495	20	2262	3029

Appendix 17.: Occurrence and frequency of breeding activities in the refALL population.

Breeding activity	Count	Frequency (%)
Crossing	330	1.91
Belgian Blue	781	4.52
Red & White	1959	11.35
Red	66	0.38
Holstein	7045	40.8
Jersey	2	0.01
Red Holstein *	13	0.08
Improved Red	21	0.12
Ayrshire	621	3.6
Brown Swiss	1	0.01
Simmental	1256	7.27
Fleckvieh	154	0.89
East Belgian Red & White	**3429**	**19.86**
Red Danish	4	0.02
Swedish Red & White	1	0.01
Angler	1	0.01
Montbéliarde	133	0.77
Rouge des Prés	93	0.54
Dikbil	18	0.1
Dual-purpose crossing	5	0.03
Beef crossing	2	0.01
Dairy breed	11	0.06
Dual-purpose breed	57	0.33
Beef breed	1	0.01
Unknown breed	1262	7.31
Total	17266	100

* Obsolete denomination, now recorded as Holstein