Conservation priorities for the different lines of Dutch Red and White Friesian cattle change when relationships with other breeds are taken into account

Hulsegge, B., Calus, M. P. L., Oldenbroek, J. K., & Windig, J. J.

Conservation priorities for the different lines of Dutch Red and White Friesian cattle change when relationships with other breeds are taken into account.

B. Hulsegge[*], M.P.L. Calus[*], J.K. Oldenbroek[†], J.J. Windig[*]

[*] Animal Breeding and Genomics Centre, Wageningen UR Livestock Research, P.O. Box 338, 6700 AH, Wageningen, The Netherlands.

[†]Centre for Genetic Resources The Netherlands, PO Box 16, 6700 AA Wageningen, The Netherlands

Corresponding Author:

Ina Hulsegge, Animal Breeding and Genomics Centre, Wageningen UR Livestock Research, P.O. Box 338, 6700 AH, Wageningen, The Netherlands. Tel: +31-317-480513; E-mail: ina.hulsegge@wur.nl

Summary

From a genetic point of view the selection of breeds and animals within breeds for conservation in a national genepool can be based on a maximum diversity strategy. This implies that priority is given to conservation of breeds and animals that diverge most and overlap of conserved diversity is minimised. This study investigated the genetic diversity in the Dutch Red and White Friesian (DFR) cattle breed and its contribution to the total genetic diversity in the pool of the Dutch dairy breeds. All Dutch cattle breeds are clearly distinct, except for Dutch Friesian breed (DF) and DFR, and have their own specific genetic identity. DFR has a small but unique contribution to the total genetic diversity of Dutch cattle breeds and is closely related to the Dutch Friesian breed. Seven different lines are distinguished within the DFR breed and all contribute to the diversity of the DFR breed. Two lines show the largest contributions to the genetic diversity in DFR. One of these lines comprises unique diversity both within the breed and across all cattle breeds. The other line comprises unique diversity for the DFR but overlaps with the Holstein Friesian breed. There seems to be no necessity to conserve the other 5 lines separately, because their level of differentiation is very low.

This study illustrates that, when taking conservation decisions for a breed, it is worthwhile to take into account the population structure of the breed itself and the relationships with other breeds.

Keywords: conservation**,** genetic diversity, population structure, relationships with other breeds

Introduction

Farm animal breeds are recognized for different values, with economic, social, historic and cultural aspects (Gandini and Oldenbroek 2007). Genetic diversity is the basis for the development and survival of animal breeds. However, many traditional, local, farm animal breeds have small (effective) population sizes, leading to a loss of their genetic diversity. It is, therefore, especially important to maintain genetic diversity in these small populations of farm animals (Fernandez, Meuwissen et al. 2011). Small populations of local breeds often comprise genetic variation with cultural, historical, sociological and environmental values (Hiemstra, De Haas et al. 2010) generally not present in the global highly productive breeds that dominate modern intensive livestock production systems. Genetic management of local breeds, is crucial for their own survival, and for maintaining diversity in the entire species, because the genetic diversity between breeds is a substantial part of the genetic diversity within the species (Wooliams and Toro 2007).

Maintaining high levels of within breed genetic diversity is the second important aim in conservation genetic diversity within the species. Traditionally, animal breeders quantify genetic diversity by analysing pedigrees, and estimating average kinships and inbreeding levels (Gutiérrez, Altarriba et al. 2003; Wooliams and Toro 2007). Pedigree analysis may not be adequate, since pedigrees are often not available in depth, so that a reliable quantification of within breed variation may not be possible. Moreover, pedigrees are generally only known since breed formation making analysis of between breed diversity by pedigree analysis impossible. Methods based on pedigree analysis can now be complemented with molecular genetic information facilitating analysis of diversity both within and across breeds (Boettcher, Tixier-Boichard et al. 2010).

Besides small effective population size, local breeds may be threatened by indiscriminate crossing with other breeds. Crossing may lead to increased genetic diversity in a population,

however at the expense of losing part or eventually all of the original genetic diversity in the population (FAO 2007). Thus both within and across breed variation need to be considered in order to preserve genetic diversity within species (Bennewitz, Eding et al. 2007; Wooliams and Toro 2007; Boettcher, Tixier-Boichard et al. 2010; Roberts and Lamberson 2015).

Eding *et al* (2002) provided a framework to quantify relative amounts of both within- and across population genetic diversity by using marker estimated kinships. In this method kinships are estimated with the help of markers and the genetic diversity within a breed is estimated as one minus the average kinship in that breed. The average kinship is also estimated across breeds, so that the genetic diversity of a set of breeds can be determined. Moreover, for each breed its contribution to the diversity of the total set can be quantified, thereby quantifying both its unique diversity and the overlap with other breeds.

After the study of Eding *et al*. (2002) progress in genotyping techniques has increased the number of available markers. The availability of dense molecular marker maps can provide a more precise picture of the genetic background of breeds (e.g. distances, uniqueness), which increase the capabilities for making decisions aimed at maintaining genetic diversity.

In this study the maximum diversity strategy was used to quantify the genetic diversity (Bennewitz, Eding et al. 2007). This strategy selects breeds that contribute in a significant way to the overall genetic diversity considering both within and across breeds diversity.

For local breeds, next to setting conservation priorities at breed level, a more detailed division into lines can be helpful to determine conservation priorities within the breed.

The objective of this study was to quantify the genetic diversity in a numerically small breed and its contribution to the total genetic diversity in other breeds of the same species in the same country. For these objectives, we used the Dutch Red and White Friesian cattle (DFR) and quantified the relationship with other Dutch dairy breeds. We assessed:

(a) The relationship of DFR with other Dutch dairy breeds and the contribution of the DFR to the total genetic diversity in Dutch dairy cattle breeds.

(b) The genetic differences between lines within the DFR

(c) The contribution of the within line genetic diversity to the total genetic diversity in the DFR, and to the gene pool of the Dutch dairy cattle breeds.

Materials and methods

Animals and Genotypes

A total of 68 Dutch Red and White Friesian cattle (DFR) animals (26 bulls and 42 cows) were sampled. The DFR is a local breed in the North of the Netherlands. Anecdotally and according to herdbook information it is closely related to the Dutch Friesian (DF) breed which is one of the founding breeds of the Holstein Friesian which is now the dominant dairy cattle breed in the world (Felius, Koolmees et al. 2011). Of the 68 sampled DFR animals, 48 animals were assigned to different lines, based on their ancestry from (founding) sires, within the breed by the Dutch herdbook "Stichting Roodbont Fries Vee" (Table 1). Two other groups consist of animals not (yet) registered in the herd book: one group from two farms with some Holstein Friesian (HF) blood and another group of isolated animals originating from the Dutch island Terschelling , from here on referred to as line 6 and 7 respectively.

Table 1. Number of samples per line of Dutch Red and White Friesian animals

Line	Name	#Bulls	#Cows	total
1	Jet	5	4	9
2	Marco-Kei	3	5	8
3	Koos	5	5	10
4	Reitsma	4	7	11
5	DF-line	8	2	10
6	Elsinga line		11	11
7	Terschelling	1	8	9
total		26	42	68

To obtain DNA we collected hair samples from the cows. From the bulls semen straws were provided by the Centre for Genetic Resources, the Netherlands (CGN). Samples were chosen, based on pedigree information of the herd book, so that they represent a wide variation in origin within a line. Samples were genotyped using the BovineSNP50 BeadChip (Illumina Inc., San Diego, CA, USA). All samples had a genotype call rate > 85%. During the quality check SNPs with a GenCall score ≤.0.20 and call rate ≤ 85% were deleted from the analyses (n=2,635). Missing genotypes were imputed using Beagle with 20 iterations (Browning and Browning 2009). The imputation was carried out for each chromosome independently. The mean r^2 value for the accuracy of imputation provided by Beagle was 0.98. After these editing steps 51,974 of the initial 54,609 SNPs remained.

Data from the DFR cattle were supplemented with data originating from studies with four other Dutch breeds (Maurice-van Eijndhoven 2014; Pryce, Johnston et al. 2014; Maurice-Van Eijndhoven, Bovenhuis et al. 2015). These data included 1,287 purebred cows; 989 were Holstein Friesian (HF), 97 Groningen White headed (GWH), 137 Meuse-Rhine-Yssel (MRY), and 64 Dutch Friesian (DF). Previously performed editing steps to remove uninformative SNP are described by Hulsegge *et al*. (2013). In short, Holstein Friesian animals were genotyped with a BovineSNP50 BeadChip and imputed to the BovineHD BeadChip using Beagle (Browning and Browning 2009). The mean Beagle r^2 was 0.96 across the imputed loci. Animals from the three other breeds (GWH, MRY and DF) were genotyped with the BovineHD BeadChip. The editing steps comprised deleting SNP with call rate < 95%, GenCall score ≤ 0.20 and GenTrain score ≤ 0.55. No MAF (minor allele frequency) thresholds were applied in the editing procedure. To investigate whether differences in results could arise with edits based on MAF, as is commonly done in other studies or applications, the impact of MAF threshold 0.02 was evaluated. The preliminary analyses indicated that our results and conclusions were hardly affected when not applying such editing step (results not

137 shown). After the editing steps 750,457 of the 777,962 SNP remained. These 750,457 SNP
138 contained 36,625 SNP that were also included in the DFR data after editing. For all animals,
139 genotypes on those 36,625 SNP were used in further analyses.
140
141 *Breed identity of DFR*
142 To investigate whether DFR is a breed with its own genetic identity, and to visualize the
143 relationship between DFR and the four other Dutch cattle breeds, principal component
144 analysis (PCA) was performed on the SNP genotypes (Price, Patterson et al. 2006) (Patterson,
145 Price et al. 2006) using the R-package Hierfstat (Goudet 2005). Genetic divergence between
146 each breed pair was quantified by calculating pairwise F_{ST} (Weir and Cockerham 1984) using
147 the R-package Hierfstat (Goudet 2005).
148
149 *Contribution of DFR to total genetic diversity in Dutch dairy cattle.*
150 To quantify the importance of DFR relative to the other breeds the marker estimated kinships
151 and the core set method of Eding *et al*. (2002) were used. In this method kinships are
152 estimated with the help of markers and the genetic diversity within a breed is estimated as one
153 minus the average kinship in that breed. The average kinship is also estimated across breeds,
154 so that the genetic diversity of the whole set can be determined. The total genetic diversity of
155 a set depends on the contribution of each breed to the total set. If all breeds contribute equally,
156 the total genetic diversity is equal to one minus the average within and across breed kinships.
157 Otherwise breed kinships have to be weighted by their contribution e.g.

158 $$g_{div} = 1 - \mathbf{c}'\mathbf{M}\mathbf{c},$$

159 with **c** being the vector with n (number of breeds) contributions of each breed (summing up to
160 1) and **M** being the n x n matrix with within and across breed kinships. So, if a relatively

uniform breed contributes more to the total set the genetic diversity of the total set will be lower compared to when a relatively diverse breed contributes more.

In the core set method of Eding *et al*. (2002) the contribution of each of the breeds that maximise the genetic diversity is estimated as

$$\mathbf{c_{max}} = \frac{\mathbf{M^{-1}1_n}}{\mathbf{1_n}'\mathbf{M^{-1}1_n}}$$

Where c_{max} is a vector with the contributions that maximises the diversity in the total set, and 1_n is a vector of n ones. The total diversity in the set is then estimated by:

$$Div_{set} = 1 - c_{max}'Mc_{max} = \frac{1}{1_n'M^{-1}1_n}$$

The contribution of each breed to this core set thus depends both on the between- and within-breed components of genetic diversity. However, not only the contribution determines the relative importance of a breed for the total genetic diversity. A breed may contribute a small amount to the core set (e.g. when their within breed kinship is high) but nevertheless increase the total genetic diversity considerably (e.g. when its across breed kinships are low). Therefore, the average kinship of the core set when the breed is included is compared to the average kinship of the core set when the breed is excluded (Eding, Crooijmans et al. 2002). The required kinships were obtained by first computing a genomic relationship matrix (G) according to Yang *et al* (2010) using the software Calc_grm (Calus 2013). Using those genomic relationships, average within and between breed kinships were computed across all pairwise relationships within and between breeds, including self-kinships.

Contribution of lines to genetic diversity within DFR

To visualize the separation of the different lines based on molecular genetic data, PCA was used. The core set method was used to determine the relative contribution of each line to the

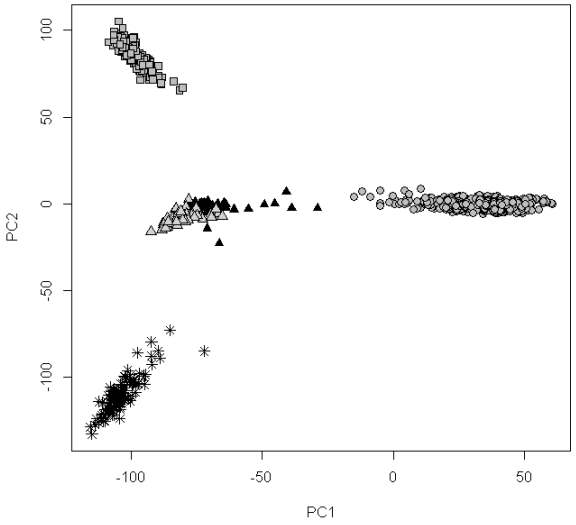
184 total genetic diversity in the DFR. The core set method was also performed with both the
185 DFR lines and the other breeds simultaneously, to determine the overlap of the contribution of
186 the individual DFR lines to the total genetic diversity with the contribution of other breeds.

187

188 **Results**

189 *Relationship of DFR cattle breed with other Dutch dairy breeds*

190 The combination of the first and second principal components (PC1 and PC2) separated
191 individual animals according to their breed (Fig. 1). PC1 distinguished the four local breeds
192 from the commercial breed HF. PC2 separated the local breeds MRY on the one hand and
193 GWH on the other hand from the Friesian breeds (DF, DFR and HF). Based on the first two
194 principal components overlap existed between the DF and DFR.



195

196 Figure 1 Principal component analysis (PCA) of five Dutch dairy cattle breeds based on 36
197 625 single-nucleotide polymorphisms (SNP's) [circle grey = Holstein Friesian (HF); star =
198 Groningen White headed (GWH); triangle grey = Dutch Friesian (DF); square grey = Meuse-
199 Rhine-Yssel (MRY); triangle black = Dutch Red and White Friesian (DFR)].

200

Genetic differentiation (pairwise F_{ST}) among breeds, confirmed that DFR is genetically closest to DF (F_{ST}=0.056) (Table 2). Pairwise F_{ST} values ranged from 0.056 (between DFR and DF) to 0.156 (between GWH and DF). The kinship values also indicated that DFR and DF were more related to each other than to the other breeds. DFR and DF had the highest average between-breed kinship (0.033) (Table 2). Average between-breed kinship ranged from -0.078 to 0.033.

Table 2. Estimated pairwise F_{ST} as a measure of genetic differentiation (below diagonal) and average genomic kinship (above diagonal) between five Dutch dairy cattle breeds.

	GWH	DF	MRY	HF	DFR
G	-	-0.078	-0.057	-0.053	-0.068
DF	0.156	-	-0.067	-0.056	0.033
MRY	0.155	0.135	-	-0.031	-0.050
HF	0.132	0.111	0.110	-	-0.036
DFR	0.136	0.056	0.111	0.088	-

DFR showed the lowest average within-breed kinship (0.106) and GWH the highest (0.248) (Table 3). The total diversity of the Dutch cattle breeds was 0.926. All five breeds contributed almost equal to the overall genetic diversity (varying from 19.55% to 20.64%). The highest unique genetic diversity was observed for GWH (0.015) and the lowest for DFR (0.006). Nevertheless, the DFR contains some unique genetic diversity not present in the other Dutch breeds, although it is less than the unique diversity of the other breeds (Table 3).

Table 3. Average genomic kinship (f) within breeds and contribution of breeds to a core set in which the diversity is maximised (= average f minimised). Unique diversity is measured as the increase in f when the core set is formed without a contribution of that breed.

	f	contribution	Unique diversity
DFR (all lines)	0.106	19.84%	0.006
GWH	0.248	19.93%	0.015
DF	0.155	19.55%	0.007
MRY	0.199	20.04%	0.012
HF	0.174	20.64%	0.010
Core set	0.074		-

Genetic differences between DFR lines

PCA distinguished DFR line 7 from the other lines by the first principal component (Fig. 2).

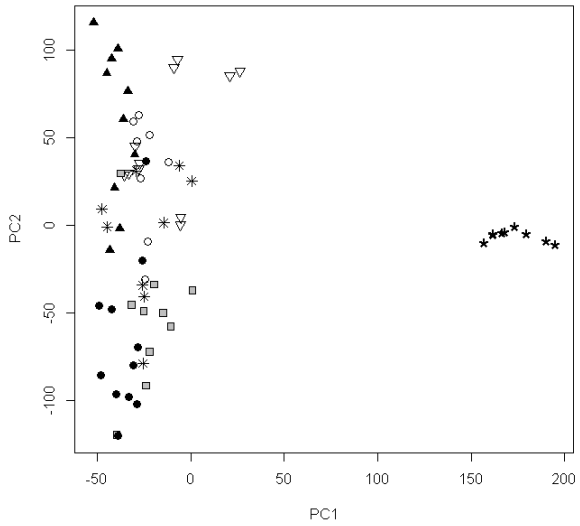


Figure 2 Principal component analysis (PCA) of seven Dutch Friesian Red cattle lines on 36 625 single-nucleotide polymorphisms (SNPs) [star = DFR line 1; circle white = DFR line 2; triangle point up black = DFR line 3; circle black = DFR line 4; square grey = DFR line 5; triangle point down = DFR line 6; asterisk = DFR line 7].

There was some differentiation among the other lines along the second principal component, but with a large overlap between the different lines. Genetic differentiation between the

234 different DFR lines was also confirmed by the pairwise F_{ST}, which varied between 0.012 and
235 0.190 (Table 4). Consistent with the PCA results, the F_{ST} values indicated that line 7 clearly
236 diverged from the other lines. Pairwise F_{ST} between DFR line 7 and the other six lines ranged
237 from 0.149 to 0.190, while the maximum pairwise F_{ST} between the lines 1 to 6 was 0.078
238 (between DFR lines 3 and 4). The F_{ST} values between DFR lines 1 to 6 were lower than the
239 F_{ST} values between breeds (Table 2), meaning that the DFR lines 1 to 6 were more related to
240 each other than the breeds were. The F_{ST} values between DFR line 7 and the other lines were
241 somewhat higher than the values found between the breeds as presented in Table 2.
242 The average kinships between-line and within-line of the DFR breed are presented in Table 4
243 and 5. Within-line kinship were higher (Table 5; varying between 0.131 and 0.478) compared
244 with the between-line kinship (Table 4; varying between 0.041 and 0.157). The lines 1 to 5
245 were more related to each other than to the lines 6 and 7. DFR line 7 showed the highest
246 within-line kinship (0.478) and the lowest between-lines kinship (ranging from 0.041 to
247 0.053). DFR line 6 had the lowest level of within-line and the second lowest level of between-
248 line kinship.

249

250 Table 4. Estimated pairwise F_{ST} as a measure of genetic differentiation (below diagonal) and
251 average genomic kinship (above diagonal) between 7 DFR lines.

	DFR line 1	DFR line 2	DFR line 3	DFR line 4	DFR line 5	DFR line 6	DFR line 7
DFR line 1	-	0.119	0.123	0.135	0.095	0.078	0.052
DFR line 2	0.042	-	0.140	0.114	0.091	0.080	0.053
DFR line 3	0.061	0.058	-	0.112	0.110	0.108	0.045
DFR line 4	0.036	0.056	0.078	-	0.157	0.069	0.043
DFR line 5	0.040	0.046	0.059	0.012	-	0.063	0.042
DFR line 6	0.048	0.049	0.057	0.063	0.046	-	0.041
DFR line 7	0.158	0.166	0.190	0.171	0.149	0.149	-

252

253

The contribution of each line (in %) to the DFR breed are shown in Table 5. All lines contributed to the diversity of the DFR breed. The highest contribution to the total diversity of the DFR breed was observed for line 6 (26.02 %), while line 3 showed the smallest contribution (6.81%). The total diversity of the DFR was 0.874. The largest part of diversity of most lines is represented in the other lines as well. The highest impact on the diversity was observed when line 6 or line 7 was removed, leading to a decrease in overall diversity of the DFR breed by about 2.3% and 1.6 %, respectively. Removing one of the lines 1 to 5 had only a small impact on the diversity. Apparently, the diversity contained in these lines is almost completely present in the other lines as well.

Table 5. Average genomic kinship (f) within lines and contribution of lines to a core set in which the diversity is maximised (= average f minimised). Unique diversity is measured as the increase in f when the core set is formed without a contribution of that breed/ line.

		DFR lines		All breeds/lines	
	f	contribution	Unique diversity	contribution	Unique diversity
DFR line 1	0.176	12.63%	0.005	13.26%	0.0002
DFR line 2	0.192	11.70%	0.004	11.90%	0.0002
DFR line 3	0.265	6.81%	0.001	15.37%	0.0003
DFR line 4	0.205	10.01%	0.002	16.35%	0.0002
DFR line 5	0.140	19.02%	0.008	14.97%	0.0002
DFR line 6	0.131	26.02%	0.020	14.82%	0.0002
DFR line 7	0.478	13.81%	0.014	13.21%	0.0004
Core set	0.126		-		

Contribution of the DFR lines to the total genetic diversity.

The average kinship between DFR lines and the Dutch cattle breeds are presented in Table 6. This kinship varied from -0.079 to 0.085. The highest values were estimated between DFR

lines and DF, while the lowest values were observed between DFR lines and GWH. Line 6 was the line most closely related to HF, and line 7 the line least related to DF.

Table 6. Average genomic kinship between Dutch cattle breeds and DFR lines.

	GWH	DF	MRY	HF
DFR line 1	-0.065	0.027	-0.046	-0.040
DFR line 2	-0.065	0.030	-0.048	-0.040
DFR line 3	-0.073	0.031	-0.054	-0.045
DFR line 4	-0.074	0.050	-0.058	-0.051
DFR line 5	-0.079	0.085	-0.065	-0.056
DFR line 6	-0.058	0.010	-0.046	-0.003
DFR line 7	-0.060	-0.007	-0.030	-0.017

Results of assessing the impact of removing one line from the DFR breed and calculating the contribution of each line (in %) to the pool of Dutch dairy cattle breeds with maximal genetic diversity are shown in Table 5. When considering all Dutch dairy cattle breeds, removing one of the DFR lines has a small impact on the diversity (loss of 0.0002 to 0.0004; Table 5). When considering all breeds the contribution of DFR line 6 was considerably smaller (14.82%) compared to DFR lines analysed in separation (26.02%). This was due to the inclusion of the HF breed, removing the HF breed increased the contribution of line 6 with 4,7% (results not shown). The contribution of DFR line 5 to the diversity across all breeds is also smaller (14.97%) compared to DFR lines only (19.02%). For DFR line 3 the contribution to the diversity across all breeds is larger (15.37%) compared to DFR lines only (6.81%). Removing DF increased the contribution of DFR, especially by the contribution of line 5. Thus analysing DFR in isolation of the other breeds suggests for some lines a larger proportion of unique diversity, while part of this diversity apparently is due to influences of the other breeds, in particular DF and HF, as revealed by the analysis including other breeds.

Discussion

Relationship of DFR cattle breed with other Dutch dairy breeds

Genetically, Dutch cattle breeds are clearly distinct from each other as shown by the PCA results, except for DF and DFR. As expected from breed history, the DFR breed is closely related to the DF breed (FAO 2007). These breeds were recorded as separate breeds for slightly more than 100 years. Red offspring of the DF breed, born out of the combination of two red-factor-carriers, could be incorporated in the DFR-breed. From 1970 DF and DFR became rare. (Porter 2002). Genetic differentiation between the breeds (pairwise F_{ST}) and the between breed kinship also indicated that DFR and DF were more related to each other than to the other Dutch breeds. In European cattle breeds, pairwise F_{ST} values have been reported i.e. ranging from 0.035 to 0.132 (Gautier, Faraut et al. 2007) and from 0.059 to 0.142 (Neuditschko 2011). The F_{ST} between DFR and DF of 0.056 is at the lower end of these ranges. DFR showed a reasonable contribution (19.84%) to the total genetic diversity of Dutch cattle breeds and contains a small amount of genetic diversity not present in the other Dutch breeds. This contribution is comparable to the contribution of each of the other breeds. Thus, although DFR and DF are closely related, the results of this study showed that DFR has its own genetic identity, containing some genetic diversity not present in other breeds.

Genetic management of lines within breeds

Management of breeds subdivided in lines implies a compromise of different factors: first, the maintenance of the highest possible levels of genetic diversity for the whole breed; second, the preservation of the genetic differentiation between lines; and third, the restriction of within-line diversity to acceptable levels, so inbreeding would not increase beyond these acceptable levels (Fernández, Toro et al. 2008). The results of our study revealed a high level of admixture between line 1 to 5. This reflects the similar origin of these lines. Consequently,

there seems to be no necessity to conserve these 5 lines separately, because their level of differentiation is very low. The line with the highest overall contribution to diversity in DFR is line 6. However, part of this diversity is due to some HF blood and therefore of lower conservation value.

The pairwise F_{ST} values indicated that DFR line 7 had a high level of genetic differentiation from other lines. This line has been bred for a considerable time in isolation from the other lines, and apparently conserved genetic diversity not present anymore in the rest of the population. However, this line showed high levels of inbreeding, and a low level of diversity.

Contribution of lines within breeds to the total genetic diversity across breeds.

A way to measure the influence of one line over the others in the DFR breed is to ascertain its genetic contribution to diversity by removing this line from the whole DFR breed and determining the remaining genetic diversity (Caballero and Toro 2002; Eding, Crooijmans et al. 2002). However, the results are different when relationships of other Dutch cattle breeds are taken into account. Some DFR lines contains a portion of genetic diversity which is also represented in the other Dutch cattle breeds. Maximizing genetic diversity within a breed is therefore not always the best strategy. Thus, our results demonstrate that when establishing conservation programs, it is necessary to take relationships with other breeds into account as well. Lenstra (2006) also indicated that for decisions on conservation priorities, the diversity of all local breeds related to the endangered population should be taken into account in order to assess their unique contribution to diversity.

Assessing contributions of lines without pedigree relationship to herdbook animals.

Previously, pedigree information was the most important information used for registration of animals in a herdbook.. Use of genome-wide SNP information, now provides a way to assess the relationship of animals without pedigree to animals registered in a herdbook. The Dutch

DFR herdbook “Stichting Roodbont Fries Vee” had assigned 48 sampled animals in this study to five different lines. Two additional DFR lines were defined consisting of animals that were not registered (DFR line 6 and 7). The lines might be considered as sub-populations, but there are no formal restrictions on pairing animals from different lines with each other, whereas, crosses between animals of different breeds are considered crossbreds and not registered as belonging to either breed. Consequently, in the context of diversity relationships between lines are generally much higher than relationships between breeds.

For the lines without an official pedigree the results of this study showed similarities and differences to the five lines (DFR line 1 to 5) with an official pedigree. This study indicated that line 6, a group with some HF blood, indeed represents part of the HF genetic diversity. Currently there seems to be no necessity to conserve DFR line 6. However, conserving line 6 *in situ* may be useful in practice for several reasons: first, this line consists of approximately 100 animals, while the total population size of DFR is 500; second, to increase the milk production of the DFR breed; and third, to increase the genetic diversity of DFR and consequently to decrease the chance of inbreeding. However, conserving line 6 should not be at the expense of other lines.

This study distinguished DFR line 7 from the other DFR lines. However, considering all Dutch cattle breeds line 7 is closely related to DFR and FH. This isolated group of animals will maximize the level of genetic diversity for the whole DFR breed and will increase genetic differentiation between lines, despite its high levels of inbreeding. Therefore, line 7 makes a unique contribution to the DFR cattle, and it is worthwhile to include this line without an official pedigree in the herdbook. The DFR herdbook and breeders are now considering the inclusion of line 6 and line 7 in the herdbook. It is often not possible, and may also not be desirable, to conserve all breeds/lines, mostly due to financial limitations (Bennewitz, Eding et al. 2007). As shown in this study, taking relationships with other breeds

368 into account can change conservation priorities within a breed, and thus may affect
369 conservation decisions made for this breed. This is not only applicable to the lines within a
370 breed in this study, but also for breeds within a species or in a gene pool of national breed as
371 in this study.
372 Conservation decisions also should take into account the degree of endangerment and costs of
373 conservations and economic, cultural and historic values of different characteristics of a breed
374 (Simianer, Marti et al. 2003) (Bennewitz, Eding et al. 2007). Endangerment of most DFR
375 lines is similar, however line 7 is clearly more endangered, since the owner has stopped active
376 farming. DFR line 6 had the highest overall contribution to diversity in DFR, however when
377 considering HF the contribution of DFR line 6 was considerably smaller, indicating that the
378 endangerment of line 6 is not really a threat for the DFR breed as a whole. Consequently,
379 conservation priorities based on genetic diversity coincides with priority based on degree of
380 endangerment.
381

382 **Acknowledgements**

383 Funding: This work was supported by Centre for Genetic Resources, The Netherlands, funded
384 by the Ministry of Economic Affairs, program “Kennisbasis Dier”, code KB-12-005.03.001
385 and program “WOT”, code WOT-03-003-056.

386 The RobustMilk project and the National Institute of Food and Agriculture (NIFA) are
387 acknowledged for providing the 50k genotypes of the HF cows, and the gDMI consortium is
388 acknowledged for imputing those to 777k genotypes. The Dutch Milk Genomics Initiative
389 and the project 'Melk op Maat', funded by Wageningen University (the Netherlands), the
390 Dutch Dairy Association (NZO, Zoetermeer, the Netherlands), the cooperative cattle
391 improvement organization CRV BV (Arnhem, the Netherlands), the Dutch Technology
392 Foundation (STW, Utrecht, the Netherlands), the Dutch Ministry of Economic Affairs (The

393 Hague, the Netherlands) and the Provinces of Gelderland and Overijssel (Arnhem, the
394 Netherlands), are thanked for providing the 777k genotypes of the GWH, FH and MRY cows.
395 The authors acknowledge Myrthe Maurice – van Eijndhoven for collecting the data of the
396 GWH, FH and MRY cows, and the herd owners for their help in collecting the data. The text
397 represent the author's views and does not necessary represent a position of the Ministry who
398 will not be liable for the use made of such information.

399

400 ## References

401 Bennewitz, J., H. Eding, et al. (2007). Selection of breeds for conservation. In, *Utilisation and Conservation of*
402 *Farm Animal Genetic Resources*. Wageningen Academic Publishers. pp. 131-146.
403 Boettcher, P. J., M. Tixier-Boichard, et al. (2010). Objectives, criteria and methods for using molecular genetic
404 data in priority setting for conservation of animal genetic resources. *Animal Genetics* **41**, 64-77.
405 Browning, B. L., S. R. Browning (2009). A unified approach to genotype imputation and haplotype-phase
406 inference for large data sets of trios and unrelated individuals. *American Journal of Human Genetics*
407 **84**, 210-223.
408 Caballero, A., M. A. Toro (2002). Analysis of genetic diversity for the management of conserved subdivided
409 populations. *Conservation Genetics* **3**, 289-299.
410 Calus, M. P. L. (2013). Calc_grm – a program to compute pedigree, genomic, and combined relationship
411 matrices. *ABGC, Wageningen UR Livestock Research.*
412 Eding, H., R. P. Crooijmans, et al. (2002). Assessing the contribution of breeds to genetic diversity in
413 conservation schemes. *Genet Sel Evol* **34**, 613-633.
414 FAO (2007). *The State of the World's Animal Genetic Resources for Food and Agriculture,* Commission on
415 Genetic Resources for Food and Agriculture Rome.
416 Felius, M., P. A. Koolmees, et al. (2011). On the Breeds of Cattle—Historic and Current Classifications. *Diversity*
417 **3**, 660-692.
418 Fernandez, J., T. H. Meuwissen, et al. (2011). Management of genetic diversity in small farm animal
419 populations. *Animal* **5**, 1684-1698.
420 Fernández, J., M. A. Toro, et al. (2008). Management of subdivided populations in conservation programs:
421 Development of a novel dynamic system. *Genetics* **179**, 683-692.
422 Gandini, G., K. Oldenbroek (2007). Strategies for moving from conservation to utilisation. In, *Utilisation and*
423 *Conservation of Farm Animal Genetic Resources*. Wageningen Academic Publishers. pp. 29-54.
424 Gautier, M., T. Faraut, et al. (2007). Genetic and haplotypic structure in 14 European and African cattle breeds.
425 *Genetics* **177**, 1059-1070.
426 Goudet, J. (2005). HIERFSTAT, a package for R to compute and test hierarchical F-statistics. *Molecular Ecology*
427 *Notes* **5**, 184-186.
428 Gutiérrez, J. P., J. Altarriba, et al. (2003). Pedigree analysis of eight Spanish beef cattle breeds. *Genetics*
429 *Selection Evolution* **35**, 43-63.
430 Hiemstra, S. J., Y. De Haas, et al. (2010). *Local cattle breeds in Europe: Development of policies and strategies*
431 *for self-sustaining breeds,* Wageningen Academic Publishers.
432 Hulsegge, B., M. P. Calus, et al. (2013). Selection of SNP from 50K and 777K arrays to predict breed of origin in
433 cattle. *J Anim Sci* **91**, 5128-5134.
434 Lenstra, J. A. (2006). Marker-assisted conservation of European cattle breeds: An evaluation. *Animal Genetics*
435 **37**, 475-481.
436 Maurice-Van Eijndhoven, M. H., H. Bovenhuis, et al. (2015). Overlap in genomic variation associated with milk
437 fat composition in Holstein Friesian and Dutch native dual-purpose breeds. *J Dairy Sci* **98**, 6510-6521.
438 Maurice-van Eijndhoven, M. H. T. (2014). Genetic variation of milk fatty acid composition between and within
439 dairy cattle breeds. In, *Animal Breeding and Genomics Centre*. Wageningen University, Wageningen.
440 Neuditschko, M. (2011). Eine Genomweite Populationsstrukturanalyse in Rinderrassen. In, *Faculty of Veterinary*
441 *Medicine* Ludwig-Maximilians-University, München.
442 Patterson, N., A. L. Price, et al. (2006). Population structure and eigenanalysis. *PLoS Genetics* **2**, 2074-2093.
443 Porter, V. (2002). *Mason's world dictionary of livestock breeds, types and varieties,* CABI Pub, Wallingford.
444 Price, A. L., N. J. Patterson, et al. (2006). Principal components analysis corrects for stratification in genome-
445 wide association studies. *Nature Genetics* **38**, 904-909.
446 Pryce, J. E., J. Johnston, et al. (2014). Imputation of genotypes from low density (50,000 markers) to high
447 density (700,000 markers) of cows from research herds in Europe, North America, and Australasia
448 using 2 reference populations. *Journal of Dairy Science* **97**, 1799-1811.

449 Roberts, K. S., W. R. Lamberson (2015). Relationships among and variation within rare breeds of swine. *Journal*
450 *of Animal Science* **93**, 3810-3813.
451 Simianer, H., S. B. Marti, et al. (2003). An approach to the optimal allocation of conservation funds to minimize
452 loss of genetic diversity between livestock breeds. *Ecological Economics* **45**, 377-392.
453 Weir, B. S., C. C. Cockerham (1984). Estimating F-statistics for the analysis of population structure. *Evolution*
454 **38**, 1358-1370.
455 Wooliams, J., M. Toro (2007). What is genetic diversity. In, *Utilisation and Conservation of Farm Animal Genetic*
456 *Resources*. Wageningen Academic Publishers. pp. 55-71.
457 Yang, J., B. Benyamin, et al. (2010). Common SNPs explain a large proportion of the heritability for human
458 height. *Nature Genetics* **42**, 565-569.

459

460