# Genetic variation and pedigree structure in two endangered Norwegian horse breeds

Hanne Fjerdingby Olsen[1], Gunnar Klemetsdal[1], John Ruane[1,2] and Turid Helfjord[3]

*[1]Department of Animal and Aquacultural Sciences, Agricultural University of Norway, P.O. Box 5003, N-1432 Ås, Norway.*

*[2]Food and Agriculture Organization of the United Nations (FAO), Viale delle Terme di Caracalla, 00100 Rome, Italy.*

*[3]Norwegian Equine Centre, Starum, N-2850 Lena, Norway.*

Corresponding author: Hanne Fjerdingby Olsen

Department of Animal and Aquacultural Sciences, Agricultural University of Norway, P.O. Box 5003, N-1432 Ås, Norway.

Tel.: +47 64948033

Fax: +47 64947960

E-mail: hanne.fjerdingby@iha.nlh.no

Key words: pedigree analysis, effective population size, effective number of founders, breed conservation.

**Abstract**

Artificial selection in small populations will reduce the effective population size and accelerate the genetic drift leading to loss of genetic variation and possible accumulation of recessives. To examine the genetic variance in these two endangered Norwegian horse-breeds, the genetic structures were investigated, and parameters as inbreeding coefficient, effective population size, effective number of founders, effective number of ancestors and effective number of founder genomes were calculated. The data consisted of 31,142 individuals of the Døle horse and 1,987 individuals of the Nordland horse, for which the complete generation equivalent was 10.5 and 7.3, respectively. The level of inbreeding was about 12% in both breeds, with a rather wavy pattern during the past 50 years in The Døle horse. Considering the last generations only, effective population size was found to be 158 in the Døle horse and 62 in the Nordland horse. In contrast to the Nordland, the effective population size in the Døle was shown sensitive to the time period considered, dropping to 12 individuals when including a period with steep increase of inbreeding in the 1970's. Loss of genetic variation was supported by all the parameters calculated, indicating a loss of genetic diversity in these two breeds, with needs for careful planning of future breeding.

## Introduction

Norway is a country that is relatively rich in domestic animal genetic resources (Ruane, 2000). However, for a variety of reasons, such as changes in farming practices, many of its breeds are in danger of extinction in the near future. Conservation of these resources is a challenge, requiring a multi-faceted approach, covering a range of aspects such as increasing public awareness about their importance or ensuring that the breeds are managed in a sustainable way so that problems of inbreeding and loss of genetic variation are minimised. Here, we focus on this latter aspect and report on the genetic structure and inbreeding levels in two endangered Norwegian horse breeds.

The populations of the Norwegian Nordland horse (also known as *Nordland-/Lyngshest*) and the Døle horse are small, and breeders also exercise artificial selection. Selection reduces the effective population size ($N_e$), because the number of breeding animals is reduced, and in an already small population it accelerates genetic drift. In such small populations genetic drift, in which allele frequencies at a locus change at random from one generation to the next, increases leading to the eventual fixation of alleles in the locus and reduction of additive genetic variation.

Changes in allele frequencies also result in changes of genotype frequencies. The frequencies of homozygotes increases and recessives accumulate in the population, a phenomenon known as inbreeding depression. The inbreeding coefficient is thus a measure of how far the process of genetic drift has progressed, relative to the base population, whereas the inbreeding rate describes the speed of the genetic drift. The objective of this study is to examine the genetic structure of these two endangered breeds and to calculate various parameters directly relevant for breed conservation efforts.

## Materials and methods

Data were received from Norwegian Equine Centre autumn 1999. The material contained all the animals ever registered in these two breeds; 30,712 registered animals of the Døle horse and 1,983 registered animals of the Nordland horse. The first registered Døle horse was a stallion born in 1846. In the Nordland horse the first registered animal was a mare born in 1906. In both breeds the last birth-year of registered animals was 1998. All the ID-numbers were renumbered, and the files were edited to remove obvious errors. For the analyses the reference populations were defined as horses born between 1990 and 1998. The founders of these populations were defined as horses with unknown parents. When an animal has only one parent known, the unknown parent was considered as a founder (Boichard et al., 1997).

All ancestors of the animals in the reference populations were traced back to the founders, individual by individual, to reveal the completeness of the data. At the end of this procedure, the total number of records generated for the two breeds were 31,142 and 1,987, respectively. From these files the maximum number of generations generated ($g_{max}$) were calculated, and so was the proportion of known ancestors per generation and in addition a 'complete generation equivalent' (CGE):

$$CGE = \frac{1}{N}\sum_{j=1}^{N}\sum_{i=1}^{n_j}\frac{1}{2^{g_{ij}}}$$

where N is the number of individuals in the reference population, $n_j$ is the number of ancestors generated for animal j and $g_{ij}$ is the number of generations between individual j and its ancestor i (Boichard et al., 1997).

To describe the genetic variation in a population, several measures can be used. To quantify the rate of genetic drift the rate of inbreeding is most frequently used (Boichard et al., 1997). Unfortunately, the coefficients of inbreeding are very sensitive to incomplete data. A complementary approach, which is more robust, is to analyse the probabilities of gene origin, as presented by Boichard et al., 1997. Parameters from probabilities of gene origin; effective number of founders ($f_e$), effective number of ancestors ($f_a$) and effective number of founder genomes ($f_g$), are also affected by incomplete pedigree information, but to a smaller extent than parameters derived from inbreeding coefficients (Boichard et al. 1997).

*Effective population size*

The calculations of the individual inbreeding coefficients were accomplished using the Quaas-Henderson algorithm (Quaas, 1976; Henderson, 1976). The effective population size is defined as the number of individuals in an idealised population which would give rise to the same inbreeding rate ( $\Delta F$ ) as observed in the real population (Falconer and Mackay, 1996). The effective population size was computed according to the classical formula:

$$\frac{1}{N_e} = 2\frac{F_t - F_{t-1}}{1 - F_t} = 2\Delta F$$

where $F_t$ is the average coefficient of inbreeding in the defined reference population, in generation t, and $F_{t-1}$ is the average inbreeding coefficient of parents of these individuals, in generation t-1.

Alternatively, to show the complexity of calculating a reliable rate of F and consequently the effective population size, an approximate value of $\Delta F$ was calculated by regressing individual inbreeding coefficients on year of birth of horses. The simple linear regression analysis was carried out for three periods, 1990-1998, 1980-1998 and 1970-1998, to examine the trend in rate of inbreeding. The effective population size was approximated by the following formula:

$$\frac{1}{N_e} = 2\Delta F_y L$$

where $\Delta F$ is the simple linear regression estimate of yearly rate of inbreeding and L is the calculated generation interval between parents and progeny, for the corresponding time period.

*Effective number of founders*

Effective number of founders was defined as the number of equally contributing founders that would produce the same genetic diversity as in the reference population under study (Lacy, 1989; Rochambeau et al., 1989):

$$f_e = 1 \Big/ \sum_{k=1}^{f} q_k^2$$

where $q_k$ is the genetic contribution of founder k to the population's gene pool. The balance of the founder contributions measures the preservation of the genetic diversity from the founders to the present population. If each founder contributes the same, the effective number of founders equals to the actual number of founders. In any other situation, the effective number is smaller than the actual number of founders. An important limitation with this approach is that it ignores potential bottlenecks in the pedigree (Boichard et al., 1997).

*Effective number of ancestors*

The effective number of ancestors, describes the minimum number of ancestors (they may or may not be founders) required to explain the complete genetic diversity of the population under study (Boichard et al., 1997):

$$f_a = 1 \Big/ \sum_{k=1}^{f} p_k^2$$

where $p_k$ is the marginal contribution of ancestor k to the population's gene pool, i.e. the contribution not explained by other ancestors. The exact computation of the marginal contribution is described in Boichard et al., 1997. In the Nordland all ancestors with a non-zero contribution to the reference population (horses born 1990-1998) were determined. In the Døle we stopped the procedure with the 100 highest contributing ancestors, as these individuals covered more than 97% of the contribution to the reference population. The effective number of ancestors accounts for potential bottlenecks, but still the probability of gene loss by drift needs to be considered.

*Effective number of founder genomes*

To account for loss of genetic variability by genetic drift, Lacy (1989, 1995) and Ballou and Lacy (1995) introduced the concept of effective number of founder genomes, or founder genome equivalent. This measures how many founder genes are maintained in the population for a given locus, and how balanced their frequency is. Originally, this parameter was calculated by probability calculations (Lacy, 1989) or by gene dropping analysis (MacCluer et al., 1986). In stead, Lacy (1995) proposed to stay with a definition of the founder genome equivalent as half the inverse of average coancestry:

$$f_g = \frac{1}{2\bar{f}_t}$$

This is supported, especially from a practical point of view, by Caballero & Toro (2000) and Zechner et al. (2002).

## Results

In the Døle horse the reference populations held a total of 1,535 horses, while the corresponding number for the Nordland horse was 1,050 horses (Table 1). In the reference populations, the largest number of registered horses during a year was 215 in the Døle and 148 in the Nordland, with a trend in both breeds somewhat towards a larger number over time.

Table 2 shows that the data on Døle consisted of 31,142 horses, of which more than 85% were mares. The table also shows that about 30% of the stallions and almost 40% of the mares were founders. Correspondingly, the data on Nordland was made up of 1,987 horses, with more equal sex representation than in the Døle, and with a lower frequency of founders (less than 5% for both sexes).

In the Døle horse every animal in the reference population had an average number of 22,464 ancestors (Table 3). The corresponding number in the Nordland horse was 524. Still, both breeds had a pedigree completeness of more than 95% in generation 6 (Table 3). Including more generations in calculating a more complete pedigree was demonstrated in the Døle than in the Nordland. In table 3, the quality of the pedigree information is also indicated by the complete generation equivalent (CGE). In the Døle horse the pedigree can be traced back a maximum of 26 generations, where CGE is 10.5. While in the Nordland CGE was 7.3, with tracing of pedigree for a maximum of 14 generations.

Figure 1 shows that inbreeding levels in the two breeds evolved somewhat differently throughout the past 100 years. The average inbreeding coefficient for the

Døle horse increased to about 7 per cent in the 1940s, after which it fell significantly before rising again towards the end of the 1970s. During the past few years, the inbreeding coefficient has been around 12%. For the Nordland horse, the average inbreeding coefficient has been high throughout most of the period. At present, the average lies at the same level of inbreeding one gets with half-sib mating (12.5%).

In the reference population of the Døle, the average inbreeding coefficient was 11.75%, as a result of an increase of 0.28% over the last generation (Table 4). Similarly, the reference population of the Nordland had a level of inbreeding of 12.77%, with an increase over the last generation of 0.71%. The corresponding effective population sizes calculated by use of the classical formula were 158 in the Døle and 62 in the Nordland (Table 4). Table 5 presents the effective population size calculated with values from the linear regression. For the Døle horse the value of $N_e$ stretched from 12 to 57, while the Nordland horse had values of $N_e$ from 78 to • .

The effective number of founders, $f_e$, was 48 for the Døle horse and 14 for the Nordland horse (Table 4). The corresponding values for effective number of ancestors, $f_a$, was 12 for Døle and 7 for Nordland, while the effective number of founder genomes, $f_g$, was calculated as 2.7 for Døle and 1.8 for Nordland (Table 4).

Table 6 shows, as a curiosity, the five most important ancestors in the pedigrees of the reference populations, as represented by their marginal contributions. In sum these five together contributed with more than 50% and 70% to the gene pools of the reference population in the Døle and in the Nordland respectively (Table 6).

## Discussion

In Norway, Norwegian Equine Centre is responsible for registration of the Døle horse and the Nordland horse. From 1990 the rules for registration of horses were changed, so that all horses being born could be registered. Both breeds have developed breeding plans since 1995, where the breeding goals include weighting of conformation, temperament, health and performance (Vangen, 1996). The stallions must be licensed for use in the breeding.

The proportion of founders in the pedigree of the Døle horse was quite high. There were far more female founders, as the data built on those used by Vangen (1983) from the herd books first published for the Døle in 1902 (datatransfer from Norwegian Equine Center in 1989). In these, only approved animals were included for which the pedigree was tabulated for 5 generations. With this practise, the probability of a sire to be recorded as an individual itself was obviously higher than for females, creating far more female founders. So, many of these founders exist only because their ancestral information is missing, not because they founded the population. But still, more than 95% of the ancestors were known for the first six generations in the Døle horse. And even in the 10th generation the completeness was quite good. In the Nordland horse there were few animals and few founders. The first herd book of the Nordland was published as late as 1969, after years of struggling to keep the breed alive. So even if the data of the Nordland only goes back to 1906, the data were very complete in the first six to eight generations of the reference population.

The level of inbreeding was high in both breeds, with an average that is equivalent to half-sib matings. However, relatively the level of inbreeding was higher in the Nordland where the base was only 7.3 generations away, compared with 10.5 generations in the Døle. Already in the 1950's the problem with possible inbreeding depression in the Døle was discussed, especially with respect to reduced fertility. To cope with this growing problem, stallions of the Norwegian cold-blooded trotter were allowed to breed in the Døle horse, for the first time in 1953 (Gaustad, 1953). During the 1970's the breeding policies changed. The cold-blooded trotter was not longer allowed to breed in the Døle. The effect came disturbingly fast, and during 5-10 years the level of inbreeding was higher than ever. The Nordland horse has had difficult periods during the past, where ensuring survival of the breed has been the priority. Back in the late 1930's three men initiated conservation of the Nordland horse (Roaldsøy, 1969). They looked for thoroughbred individuals, and found one stallion in 1939; *Rimfakse*. He was used as the starting point for further breeding, contributing 26% to the reference population (Table 6). The Nordland has had a relatively high level of inbreeding for many years, with no sudden changes during the past 40-50 years. According to Ehiobu et al. (1989) slower rates of inbreeding cause less inbreeding depression, than more rapid rates of inbreeding. There is no recent report of serious health damages in the two breeds, which has to come from practise, as no data is still systematically recorded in these breeds. Most importantly would be to record fertility data to survey a reduction in fitness over time. This requires that Norwegian Equine Center has as first priority to record all matings and all born individuals to the future database, expected established in 2005.

Over the last generation there has not been any significant change in the level of inbreeding in either of the two populations, resulting in a larger effective population size in the Døle (158) than in the Nordland (62). However, in the Døle the effective population size was very dependent on how the inbreeding rate was calculated. Using regression and including the period with a steep increase in level of inbreeding in the 1970's (Figure 1), an effective population size as low as 12 was calculated. This is far below the recommended sizes of approximately 100 (Klemetsdal, 1999).

Seen in a genetic varation setting, the small effective population size indicates a serious loss. This is being supported by the effective number of founders in the reference population of the Døle horse being only 48, against the actual number of 770, which indicates that the contributions from the founders were extremely imbalanced. Even smaller was the effective number of ancestors, and lastly, after having accounted for all losses, the smallest number was found; the effective number of founder genomes, which was only 2.7. In the Nordland horse these parameters also were very small, and showed much imbalance in the contributions. The effective number of founder genomes in Nordland was calculated to be 1.8. These values are comparable with other studies in endangered populations, as Rodrigañez et al. (1998) and Zechner et al. (2002). The $f_e$, $f_a$ and $f_g$ values calculated in the Døle horse and the Nordland horse indicate that the genetic variation could be reduced, and the situation are much more alarming than what was reported in Vangen (1983).

Several calculations in these two breeds indicate a loss of genetic diversity. This means that the breeds should carefully plan the future breeding. The aim must be to recover the genetic variation, and strong selection of any kind (which would further worsen the problems of inbreeding and genetic drift) should not be carried out. There is planned a simulation of these two breeds, to compare the actual situation with current strategies for genetic conservation.

**Acknowledgements**

Thanks to the Norwegian Equine Center for help with completion and correction of the data material, to Morten Svendsen for useful help with programming and to Inger Anne Boman for carefully investigating the calculation of the inbreeding coefficients.

## References

Ballou, J.D. & Lacy, R.C., 1995. Identifying genetically important individuals for management of genetic variation in pedigreed populations. In: Ballou, J.D., Gilpin, M. & Foose, T.J. (eds.) Population management for survival and recovery, Columbia University Press, New York, pp. 76-111.

Boichard, D., Maignel, L. & Verrier, É., 1997. The value of using probabilities of gene origin to measure genetic variability in a population. Genet. Sel. Evol. 29, 5-23.

Caballero, A. & Toro, M.A., 2000. Interrelations between effective population size and other pedigree tools for the management of conserved populations. Genet. Res. Camb. 75, 331-343.

Ehiobu, N.G., Goddard, M.E. & Taylor, J.F., 1989. Effect of rate of inbreeding on inbreeding depression in *Drosophila melanogaster*. Theor. Appl. Genet. 77, 123-127.

Falconer, D.S. & Mackay, T.F.C., 1996. Introduction to quantitative genetics, 4th edn. Longman Group, Essex, England, 464 pp.

Gaustad, M., 1953. Statens utstilling for dølehingster. Dølehestavlen. In: S. Rohdin (Editor). Våre Hester Julenummer, 14-17. (In Norwegian).

Henderson, C.R., 1976. A simple method for computing the inverse of a numerator relationship matrix used in prediction of breeding values. Biometrics 32, 69-83.

Klemetsdal, 1999. Stochastic simulation of sire selection strategies in North-Swedish and Norwegian cold-blooded trotters. Livest. Prod. Sci. 57, 219-229.

Lacy, R.C., 1989. Analysis of founder representation in pedigrees: Founder equivalents and founder genome equivalents. Zoo Biol. 8, 111-123.

Lacy, R.C., 1995. Clarification of terms and their use in the management of captive populations. Zoo Biol. 14, 565-578.

MacCluer, J.W., VandeBerg, J.L., Read, B. & Ryder, O., 1986. Pedigree analysis by computer simulation. Zoo Biol. 5, 147-160.

Quaas, R.L., 1976. Computing the diagonal elements and inverse of a large numerator relationship matrix. Biometrics 32, 949-953.

Roaldsøy, N., 1969. Nordlandshesten. In: Stambok over norske hester, Statens Stambokkontor, Oslo, 309 pp. (In Norwegian).

Rochambeau, H. de, La Fuente, L.F. de, Rouvier, R. & Ouhayoun, J., 1989. Sélection sur la vitesse de croissance post-sevrage chez le lapin. Genet. Sel. Evol. 21, 527-546.

Rodrigañez, J., Toro, M.A., Rodriguez, M.C. & Silió, L., 1998. Effect of allele survival and inbreeding depression on litter size in a closed line of Large White pigs. Anim. Sci. 67, 573-582.

Ruane, J., 2000. A framework for prioritizing domestic animal breeds for conservation purposes at the national level: a Norwegian case study. Conservation Biology 14, 1385-1393.

Vangen, O., 1983. The use of relationship matrices to avoid inbreeding in small horse populations. Z. Tierzûchtg. Zûchtsbiol. 100, 48-54.

Vangen, O., 1996. Breeding goals and testing systems in Norwegian horse breeding. 47th Annual Meeting of the European Association for Animal Production, Lillehammer, Norway, 25-29 August 1996.

Zechner, P., Sölkner, J., Druml, T., Baumung, R., Achmann, R., Bodo, I., Marti, E., Habe, F. & Brem, G., 2002. Analysis of diversity and population structure in the Lipizzan horse breed based on pedigrees information. Livest. Prod. Sci. 77, 137-146.

*Table 1. The number of registered stallions and mares in the reference population; in the Døle and in the Nordland, respectively.*

| | | *Døle* | | | *Nordland* | |
|---|---|---|---|---|---|---|
| Birth-year | Stallions | Mares | Total | Stallions | Mares | Total |
| 1990 | 67 | 91 | 158 | 42 | 60 | 102 |
| 1991 | 64 | 95 | 159 | 56 | 53 | 109 |
| 1992 | 58 | 74 | 132 | 42 | 64 | 106 |
| 1993 | 64 | 87 | 151 | 58 | 51 | 109 |
| 1994 | 78 | 89 | 167 | 39 | 51 | 90 |
| 1995 | 99 | 100 | 199 | 65 | 58 | 123 |
| 1996 | 96 | 88 | 184 | 63 | 68 | 131 |
| 1997 | 115 | 100 | 215 | 83 | 65 | 148 |
| 1998 | 93 | 77 | 170 | 64 | 68 | 132 |
| TOTAL | 734 | 801 | 1,535 | 512 | 538 | 1,050 |

*Table 2. The number of animals in the pedigree and the proportion of founders; in the Døle and in the Nordland, respectively.*

| | | *Døle* | | | *Nordland* | |
|---|---|---|---|---|---|---|
| | Stallions | Mares | Total | Stallions | Mares | Total |
| No. of animals | 3,739 | 27,403 | 31,142 | 727 | 1,260 | 1,987 |
| No. of founders | 1,135 | 10,622 | 11,757 | 19 | 54 | 73 |
| In % | 30.4 | 39.3 | 37.8 | 2.6 | 4.3 | 3.7 |

*Table 3. Calculated parameters describing pedigree completeness of the reference population; in the Døle and Nordland, respectively.*

| | | *Døle* | *Nordland* |
|---|---|---|---|
| Number of animals in reference population | | 1,535 | 1,050 |
| Average number of ancestors per animal | | 22,464 | 524 |
| % of known ancestors in generation | 2 | 99.97 | 99.95 |
| | 4 | 99.51 | 99.94 |
| | 6 | 95.39 | 98.91 |
| | 8 | 84.15 | 71.82 |
| | 10 | 73.58 | 19.77 |
| Maximum number of generations generated ($g_{max}$) | | 26 | 14 |
| Complete generation equivalent (CGE) | | 10.5 | 7.3 |

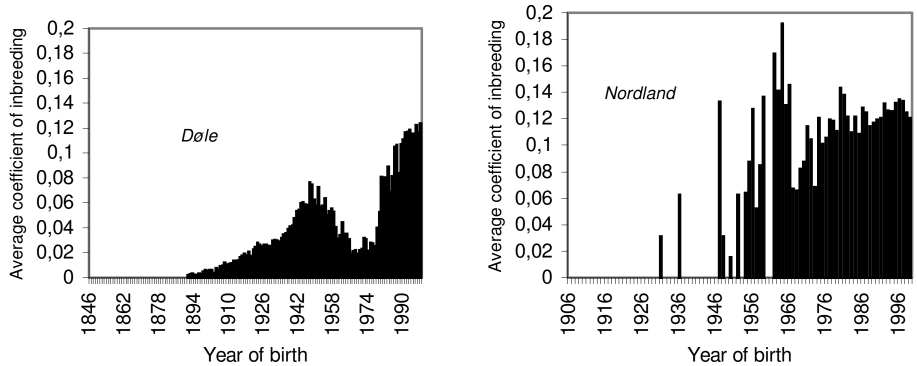


*Figure 1: Average coefficient of inbreeding per year of birth; in the Døle and in the Nordland, respectively.*

*Table 4. Calculated parameters describing genetic diversity of the reference population; in the Døle and in the Nordland, respectively.*

| | *Døle* | *Nordland* |
|---|---|---|
| Average coefficient of inbreeding ($F_t$) | 0.1175 | 0.1277 |
| Increase of average coefficient of inbreeding, last generation ($F_t$-$F_{t-1}$) | 0.0028 | 0.0071 |
| Effective population size ($N_e$) | 158 | 62 |
| Number of founders (f) | 770 | 49 |
| Effective number of founders ($f_e$) | 48 | 14 |
| Effective number of ancestors ($f_a$) | 12 | 7 |
| Effective number of founder genomes ($f_g$) | 2.7 | 1.8 |

*Table 5. Estimated effective population size ($1/N_e = 2\Delta F_y L$) from annual rate of inbreeding ($\Delta F_y$) approximated by regression of individual inbreeding on year of birth of horses ($\beta_1$), generation interval between parents and progeny (L), in three periods; in the Døle and in the Nordland, respectively.*

| | *Døle* | | | | *Nordland* | | | |
|---|---|---|---|---|---|---|---|---|
| | $\beta_1$ | P | L | $N_e$ | $\beta_1$ | P | L | $N_e$ |
| 1990-98 | 0.00096 | 0.0610 | 9.1 | 57 | -0.00008 | 0.8462 | 8.6 | ∞ |
| 1980-98 | 0.00264 | <0.0001 | 9.2 | 21 | 0.00051 | 0.0170 | 8.5 | 116 |
| 1970-98 | 0.00420 | <0.0001 | 9.7 | 12 | 0.00075 | <0.0001 | 8.6 | 78 |

*Table 6. The five most important ancestors in the pedigree of the reference populations of Døle horse and Nordland horse.*

| **Registrationnumber** | **Name** | **Sex** | **Birth-year** | **Marginal contribution** |
|---|---|---|---|---|
| *Døle* | | | | |
| 1185DH | Gjestar | M | 1919 | 0.21673 |
| 0825DH | Brimin | M | 1908 | 0.15894 |
| 1752DH | Tuftar | M | 1946 | 0.06595 |
| 0613DH | Draupner | M | 1898 | 0.04514 |
| 0130DH | Dovre | M | 1914 | 0.03684 |
| SUM | | | | 0.52360 |
| *Nordland* | | | | |
| 0002NH | Rimfakse | M | 1935 | 0.26125 |
| 0010NH | Torgrim | M | 1946 | 0.14834 |
| 00009N | Mona | F | 1935 | 0.13819 |
| 00003N | Bruna | F | 1930 | 0.12669 |
| 0009NH | Bamse | M | 1945 | 0.06071 |
| SUM | | | | 0.73518 |