# Genetic Structure and Composition of Genetic Diversity in the Kouchi Sub-breed of the Japanese Brown Cattle Population

**Takeshi Honda**\*, **Toshihide Fujii**[1] **and Fumio Mukai**[2]

Food Resources Education and Research Center, Kobe University, Uzurano, Kasai, Hyogo 675-2103, Japan

**ABSTRACT :** Japanese Brown cattle, one of the four domestic beef breeds in Japan, are suffering from numerical reduction due to economic pressure from profitable breeds. In this study, all the reproductive cows in the Kouchi sub-breed of the Japanese Brown cattle that were alive in July 2005 were investigated by pedigree analysis to clarify genetic structure and composition of genetic variability. In addition, genetically important individuals for the maintenance of genetic variability of the sub-breed were also identified through the core set method. The number of cows analyzed was 1,349. Their pedigrees were traced back to ancestors born around 1940, and pedigree records of 13,157 animals were used for the analysis. Principal component analysis was performed on the relationship matrix of the cows, and their factor loadings were plotted on a three-dimensional diagram. According to their spatial positions in the diagram, all the cows were subdivided into five genetically distinctive subpopulations of 131 to 437 animals. Genetic diversity of the whole sub-breed, which is estimated to be 0.901, was decomposed into 0.856 and 0.045 of within-subpopulation and between-subpopulation components. Recalculation of genetic diversity after removal of one or several subpopulations from the five subpopulations suggested that three of them were genetically important for the maintenance of genetic variability of the sub-breed. Applying the core set method to all the cows, maximum attainable genetic diversity was estimated to be 0.949, and optimal genetic contributions assigned to each cow supported the previous results indicating relative importance of the three subpopulations as useful genetic materials. (**Key Words :** Genetic Diversity, Genetic Structure, Japanese Brown, Pedigree Analysis, Principal Component Analysis)

## INTRODUCTION

Because of economic pressure from a limited number of highly profitable breeds, a large number of minor breeds are currently at high risk of extinction in many livestock species worldwide (Barker, 2001). Likewise, difference of importance among four domestic beef breeds in Japan (i.e., Japanese Black, Japanese Brown, Japanese Shorthorn, and Japanese Poll) has been recently widening. As a result of the liberalization of beef import restrictions in 1991, domestic production of high-quality beef has been emphasized in Japan to compete with the economical imported beef. Domestic beef production has greatly depended on the Japanese Black cattle because of its superior meat quality. In contrast, the declining population

of the other domestic breeds, which have relatively inferior meat quality, has been an increasing problem.

The Kouchi sub-breed of Japanese Brown cattle is one of the domestic beef breeds suffering from numerical reduction. Although its meat quality has depreciated compared to that of the Japanese Black, the Kouchi sub-breed is characterized by some superior production traits, such as growth rate, feed utilization efficiency, and grazing aptitude. In our previous work (Honda et al., 2006), we suggested that recovering from the numerical reduction by the improvement of profitability with these unique characteristics is crucial for the persistence of the sub-breed and preservation of potential future breeding materials.

Because of the numerical reduction and the increased relationships among the selected sires, the Kouchi sub-breed has been also suffering from serious reduction of genetic diversity. Nomura et al. (2005) recently estimated that the effective population size of the Japanese Black cattle was 20.3 and suggested that this was lower than the standard estimates of 30 to 70 for various foreign cattle breeds. However, the effective size of the recent Kouchi sub-breed was estimated as only 6 (Honda et al., 2006),

\* Corresponding Author: Takeshi Honda. Tel: +81-790-49-3122, Fax: +81-790-49-3122, E-mail: takhonda@diamond.kobe-u.ac.jp
[1] Graduate School of Science and Technology, Kobe University, Nada, Kobe 657-8501, Japan.
[2] Faculty of Agriculture, Kobe University, Nada, Kobe 657-8501, Japan.

implying a remarkably rapid decreasing rate of genetic diversity. Therefore, recovery from numerical reduction should be accompanied by the recovery of genetic variability to establish a self-sustained population thereafter. Investigation of genetic resources distributed in the sub-breed will provide useful information for this purpose.

In this study, pedigree analysis was conducted to examine genetic structure and composition of genetic diversity of the Kouchi sub-breed. As a complementary study, genetically important individuals for the maintenance of genetic diversity of the sub-breed were also identified.

## MATERIALS AND METHODS

### Pedigree data

This study investigated 1,349 reproductive cows that were alive in July 2005. The Wagyu Registry Association has kept pedigree records of the cows that trace ancestry as far as possible to the base population around 1940. The total number of animals in the pedigree file was 13,157.

### Principal component analysis

The genetic relationship between two reproductive cows, $i$ and $j$, was estimated as relationship coefficient ($r_{ij}$), and a relationship matrix including all the pairs of animals ($\mathbf{R} = [r_{ij}]$) was constructed. Principal component analysis (PCA) was applied to the $\mathbf{R}$ matrix, regarding each cow as an operational taxonomic unit (OTU) and $\mathbf{R}$ matrix as the correlation matrix among them. To visualize the genetic configuration of the cows, the first three principal components (PCs) were extracted, and factor loadings (i.e. correlations between OTUs and principal components) of the cows were plotted on a three-dimensional diagram of the three PCs. Based on their spatial positions in the diagram, all the cows were classified into several subpopulations, and the constitution of genetic diversity of the entire sub-breed was investigated.

### Composition of average coancestry and genetic diversity

To examine the constitution of total genetic diversity of the sub-breed, average coancestry ($\bar{f}$) and expected heterozygosity ($1 - \bar{f}$; Nei, 1973) of all the individuals were partitioned into within-subpopulation and between-subpopulation components. Suppose a population of $N_T$ individuals subdivided into $n$ subpopulations, each of which consists of $N_i$ individuals ($\sum_{i=1}^{n} N_i = N_T$), and let $\bar{f}_{ij}$ be the average coancestry between individuals of subpopulations $i$ and $j$, including all $N_i \times N_j$ pairs. Caballero and Toro (2002) showed that average coancestry of all the individuals ($\bar{f}$) is obtained as

$$\bar{f} = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} \bar{f}_{ij} N_i N_j}{N_T^2}$$

and can be decomposed into

$$\bar{f} = \frac{\sum_{i=1}^{n} \bar{f}_{ii} N_i}{N_T} - \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} \mathbf{D}_{ij} N_i N_j}{N_T^2} = \tilde{f} - \overline{\mathbf{D}} \tag{1}$$

where $\mathbf{D}_{ij}$ is Nei's minimum distance (Nei, 1987) between subpopulations $i$ and $j$:

$$\mathbf{D}_{ij} = \frac{\bar{f}_{ii} + \bar{f}_{jj}}{2} - \bar{f}_{ij}$$

Equation (1) indicates that the average coancestry of the whole population ($\bar{f}$) depends on within-subpopulation component ($\tilde{f}$) and averaged distance among the subpopulations ($\overline{\mathbf{D}}$). Thus, total genetic diversity of the whole population ($GD_T$) expressed as expected heterozygosity is also decomposed as

$$GD_T = 1 - \bar{f} = (1 - \tilde{f}) + \overline{\mathbf{D}}$$

(Caballero and Toro, 2002), implying that the total genetic diversity of the whole population is partially ascribed to within-diversity components of each subpopulation ($1 - \tilde{f}$) and to between-diversity components among all the pairs of subpopulations ($\overline{\mathbf{D}}$).

### Core set method

As a complementary study for the population-based investigation described above, optimal genetic contributions of individuals for maximizing the expected heterozygosity of a hypothetical gamete pool were calculated. This calculation was conducted by the core set method of Eding et al. (2002), which was originally proposed for the ranking of animal breeds regarding conservation of genetic diversity within a livestock species. Suppose a situation where $g_i$ gametes are sampled from animal $i$ ($i = 1,\ldots,1349$) to obtain a hypothetical gamete pool with $g_T$ total gametes, so that the genetic contribution of individual $i$ is $c_i = g_i/g_T$ and

$$\sum_{i=1}^{1349} c_i = 1 \tag{2}$$

Letting $f_{ij}$ be coancestry between two individuals, $i$ and $j$, and $\mathbf{F} = [f_{ij}]$ be the coancestry matrix of all the animals, the

**Table 1.** Genetic relationships within and among the five subpopulations measured by average coancestries (diagonal and below diagonal) and Nei's minimum distances (above diagonal)

| Subpopulation (No. of animals) | A | B | C | D | E |
|---|---|---|---|---|---|
| A (437) | 0.145 | 0.075 | 0.055 | 0.046 | 0.049 |
| B (307) | 0.093 | 0.191 | 0.073 | 0.042 | 0.066 |
| C (286) | 0.088 | 0.092 | 0.141 | 0.048 | 0.044 |
| D (131) | 0.103 | 0.130 | 0.099 | 0.153 | 0.043 |
| E (188) | 0.057 | 0.063 | 0.059 | 0.067 | 0.067 |
| $F$ [a] | 0.057 | 0.088 | 0.074 | 0.104 | 0.061 |
| $\alpha$ [b] | -0.103 | -0.128 | -0.078 | -0.058 | -0.006 |

[a] Average inbreeding coefficient within subpopulation. [b] Degree of deviation from Hardy-Weinberg proportion.



**Figure 1.** Plots of factor loadings of 1,349 reproductive cows on a three-dimensional diagram of the first three principal components (PC1, PC2, and PC3).

probability of two alleles randomly chosen from the gamete pool being identical by descent is

$$\bar{f}_c = \mathbf{c'Fc}$$

where **c** is a column vector of $c_i$. Thus, maximization of expected heterozygosity of the hypothetical gamete pool $(1 - \bar{f}_c)$ is equivalent to the problem of finding optimal **c** ($\mathbf{c}_{opt}$) such that $\bar{f}_c$ is minimized under the restriction of equation (2). This problem can be solved by applying the Lagrange multiplier method, and the optimal contributions of each animal is obtained as

$$\mathbf{c}_{opt} = \frac{\mathbf{F}^{-1}\mathbf{1}}{\mathbf{1'F}^{-1}\mathbf{1}}$$

where **1** is a unit vector (for detailed derivation, see Eding et al. (2002)).

**RESULTS AND DISCUSSION**

Applying PCA to the **R** matrix, the first three PCs with a cumulative contribution of 26.5% (19.2% for PC1, 4.2% for PC2, and 3.1% for PC3) were extracted. Figure 1 plots factor loadings of reproductive cows on a three-dimensional diagram of the three PCs. According to the spatial positions in the diagram, all the cows were genetically classified into five subpopulations.

The number of animals in each subpopulation, average coancestry between two subpopulations ($\bar{f}_{ij}$) and within each subpopulation ($\bar{f}_{ii}$), and Nei's minimum distance ($\mathbf{D}_{ij}$; Nei, 1987) between two subpopulations are presented in Table 1. The numbers of animals in each subpopulation range from 131 (subpopulation D) to 437 (subpopulation A). The coefficient $F$ is the average inbreeding coefficient of each subpopulation, and $\alpha = \left( (F_i - \bar{f}_{ii})/(1 - \bar{f}_{ii}) \right)$ indicates the degree of deviation from Hardy-Weinberg proportion within each subpopulation, which gives an estimate of $F_{IS}$ of Wright's $F$-statistics (Wright, 1951; 1969). Negative values of $\alpha$ in all the subpopulations suggest that no genetically subdivided structure exists within each of them.

Among the first four subpopulations, A to D, average coancestry within subpopulation B was prominent, while those within the other subpopulations ranged from 0.14 to 0.15. On the other hand, average coancestries among the four subpopulations were approximately equal to or less than the average coancestry of all the animals ($\bar{f} = 0.099$), except for that between B and D, indicating that each subpopulation has distinct genetic characteristics. In contrast to this apparent subdivided structure among the four subpopulations, the relatively small average coancestries both within and between subpopulation E implies that this subpopulation consists of animals that have low relationships to any other animals in the whole sub-breed.

Contributions of each subpopulation to the average coancestry of the whole sub-breed, decomposed into the within-subpopulation component and average distance to
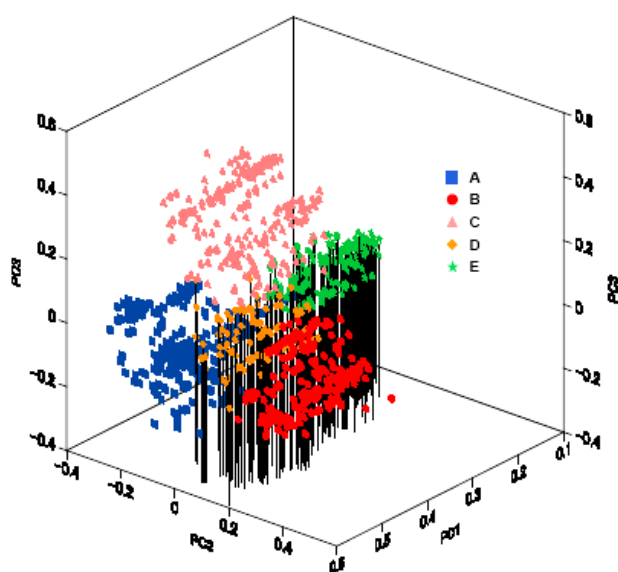
**Table 2.** Contributions of subpopulations to the average coancestry of the whole sub-breed ($\bar{f}$), decomposed into within-subpopulation component ($\tilde{f}$) and averaged minimum distance to the other subpopulations ($\overline{\mathbf{D}}$)

| Subpopulation | $\tilde{f}$ | $\overline{\mathbf{D}}$ | $\bar{f}$ | Equal subpopulation sizes [a] | | |
| --- | --- | --- | --- | --- | --- | --- |
| | | | | $\tilde{f}$ | $\overline{\mathbf{D}}$ | $\bar{f}$ |
| A | 0.047 | 0.013 | 0.034 | 0.029 | 0.009 | 0.020 |
| B | 0.043 | 0.012 | 0.031 | 0.038 | 0.010 | 0.028 |
| C | 0.030 | 0.010 | 0.020 | 0.028 | 0.009 | 0.019 |
| D | 0.015 | 0.004 | 0.011 | 0.031 | 0.007 | 0.023 |
| E | 0.009 | 0.006 | 0.003 | 0.013 | 0.008 | 0.005 |
| Total | 0.144 | 0.045 | 0.099 | - | - | - |

[a] Estimates under the assumption of equal subpopulation sizes.

**Table 3.** Changes of total genetic diversity of the sub-breed ($GD_T$), within-subpopulation component ($GD_{WS}$), and between-subpopulation component ($GD_{BS}$), according to the removal of subpopulations and their proportional changes in percentages

| Removed subpopulation | $GD_{WS}$ | $GD_{BS}$ | $GD_T$ | Changes (%) | | |
| --- | --- | --- | --- | --- | --- | --- |
| | | | | $GD_{WS}$ | $GD_{BS}$ | $GD_T$ |
| None [a] | 0.856 | 0.045 | 0.901 | - | - | - |
| A | 0.856 | 0.041 | 0.897 | 0.0 | -0.4 | -0.4 |
| B | 0.870 | 0.034 | 0.904 | 1.5 | -1.1 | 0.4 |
| C | 0.855 | 0.041 | 0.896 | -0.1 | -0.4 | -0.5 |
| D | 0.857 | 0.045 | 0.902 | 0.1 | 0.1 | 0.2 |
| E | 0.843 | 0.044 | 0.887 | -1.4 | -0.1 | -1.5 |
| B and D | 0.873 | 0.032 | 0.905 | 1.9 | -1.4 | 0.5 |
| A and C | 0.855 | 0.034 | 0.888 | -0.1 | -1.3 | -1.4 |
| A and E | 0.836 | 0.038 | 0.874 | -2.2 | -0.8 | -3.0 |
| C and E | 0.838 | 0.038 | 0.875 | -2.0 | -0.8 | -2.8 |
| A and C and E | 0.821 | 0.018 | 0.838 | -3.9 | -3.0 | -6.9 |

[a] Genetic diversity of the whole sub-breed without any removal of subpopulations.

the other subpopulations, are presented in Table 2. The average coancestry of the whole sub-breed is composed as 0.144 of within-subpopulation component minus 0.045 of minimum distance averaged over all the pairs of subpopulations. The largest proportion of $\bar{f}$ is ascribed to subpopulation A because of its highest contributions to the within-subpopulation component. In contrast, subpopulations D and E account for a relatively small proportion of $\bar{f}$. The right-hand side of the table shows the results obtained under the assumption of equal subpopulation sizes, which represent relative amount of contributions of randomly chosen alleles from each subpopulation to $\bar{f}$ and its components. This assumption alters the order of the relative amount of contributions of subpopulations. For example, the contribution of subpopulation A, which was the largest contributor, becomes moderate; and subpopulation D, which was the second smallest contributor, becomes the second largest one.

The effects of each subpopulation or several subpopulations on the genetic diversity of the whole sub-breed can be examined by removing them from the whole population and recalculating the genetic diversity of the remaining population. This is one of the methods to prioritize animal breeds in the preservation of genetic diversity of livestock species (Petit et al., 1998; Caballero and Toro, 2002). Genetic diversity and its proportional changes according to the removal of subpopulations are provided in Table 3. Genetic diversity of the whole sub-breed ($GD_T$: $1-\bar{f} = 0.901$) is decomposed into 0.856 ($GD_{WS}$) of within-diversity component and 0.045 ($GD_{BS}$) of between-diversity component. Since subpopulation B has an extremely high average coancestry and relatively large distances from the other subpopulations (see also Table 1), removal of this subpopulation causes the largest within-diversity ($GD_{WS}$) increase and the largest between-diversity ($GD_{BS}$) decrease, enhancing overall genetic diversity as a consequence. Enhancement of overall genetic diversity also results from the removal of subpopulation D and that of both B and D (see row of B and D), indicating that these two subpopulations are not genetically important in terms of genetic management of the sub-breed. In contrast, overall diversity is decayed when subpopulations A, C, and E are individually removed; in particular, the amount of reduction caused by the removal of E is substantially large, due to its low average coancestry (see also Table 1). In addition, simultaneous removals of these subpopulations (see rows of A and C, A and E, C and E, and A and C and E) have larger effects than the sum of individual effects, indicating that all three of these subpopulations are necessary for the genetic conservation of the population.

**Table 4.** Numbers of animals that were assigned non-zero contributions in the core set method ($N_s$), and sums of the contributions within each subpopulation ($c_s$)

| Subpopulation | $N_s$ | $c_s$ |
|---|---|---|
| A | 19 | 0.13 |
| B | 13 | 0.04 |
| C | 24 | 0.10 |
| D | 2 | 0.01 |
| E | 76 | 0.71 |
| Total | 134 | 1.00 |

Maximum genetic diversity of the hypothetical gamete pool ($1 - \bar{f_c}$) obtained by applying the core set method to the reproductive cows was 0.949. The distribution of the numbers of animals with non-zero optimal contributions and the sum of the contributions within each subpopulation are given in Table 4. Almost all the genetic contributions, which summed up to 0.95, were assigned to the animals belonging to subpopulations A, C, and E. This method determines the most effective way to make use of the genetic materials distributed over the population. However, it should be difficult to obtain a sufficient number of progenies according to the suggested optimal contributions because the number of progenies produced by each animal would be restricted by such practical problems as low economic merit of productive traits and limitation of reproductive ability. Thus, individual-based analysis of the core set method and the previous population-based analysis, which can recognize alternatives to the selected animals in the core set method, may complement one another. The optimal genetic contributions assigned to each cow, which is summarized in Table 4, support the results indicating the importance of subpopulations A, C, and E as useful genetic materials.

Recovering from numerical reduction by improving profitability is the key issue for the persistence of the sub-breed, and it is also recommended from the viewpoint of avoidance of deleterious inbreeding effects. Reduction of population size inevitably accumulates the relationships among breeding animals and advances inbreeding (Frankham et al., 2002). If productive or reproductive traits consequentially exhibit definite inbreeding depression, the sub-breed will increasingly depreciate and may fall into a vicious circle of reduction of population size and deleterious inbreeding effects. Recent simulation study by Theodorou and Couvet (2006) suggested that deleterious effect of inbreeding on the fitness becomes larger when the population growth rate after the numerical reduction becomes lower. Smooth expansion of population size with the effective use of genetic variability should be an urgent task also for avoiding these potential side effects. In addition, a strict mating system should be planned, based on the actual population structure, for a compromise between genetic improvement and maintenance of genetic variability, as proposed by Oyama et al. (2007) for a highly inbred line of beef cattle in Japan.

## REFERENCES

Barker, J. S. F. 2001. Conservation and management of genetic diversity: a domestic animal perspective. Can. J. For. Res. 31:588-595.

Caballero, A. and M. A. Toro. 2002. Analysis of genetic diversity for the management of conserved subdivided populations. Conserv. Genet. 3:289-299.

Eding, H., R. P. M. A. Crooijmans, M. A. M. Groenen and T. H. E. Meuwissen. 2002. Assessing the contribution of breeds to genetic diversity in conservation schemes. Genet. Sel. Evol. 34:613-633.

Frankham, R., J. D. Ballou and D. A. Briscoe. 2002. Introduction to Conservation Genetics. Cambridge University Press, Cambridge.

Honda, T., T. Fujii, T. Nomura and F. Mukai. 2006. Evaluation of genetic diversity in Japanese Brown cattle population by pedigree analysis. J. Anim. Breed. Genet. 123:172-179.

Nei, M. 1973. Analysis of gene diversity in subdivided populations. Proc. Natl. Acad. Sci. USA. 70:3321-3323.

Nei, M. 1987. Molecular Evolutionary Genetics. Columbia University Press, New York.

Nomura, T., T. Honda and F. Mukai. 2005. Monitoring and preservation of genetic diversity in livestock breeds: a case study of the Japanese Black cattle population. Curr. Top. Genet. 1:59-71.

Oyama, K., M. Nojima, M. Shojo, M. Fukushima, K. Anada and F. Mukai. 2007. Effect of sire mating patterns on future genetic merit and inbreeding in a closed beef cattle population. J. Anim. Breed. Genet. 124:73-80.

Petit, R. J., A. El Mousadic and O. Pons. 1998. Identifying populations for conservation on the basis of genetic markers. Conserv. Biol. 12:844-845.

Theodorou, K. and D. Couvet. 2006. On the expected relationship between inbreeding, fitness, and extinction. Genet. Sel. Evol. 38:371-388.

Wright, S. 1951. The genetical structure of populations. Ann. Eugen. 15:323-354.

Wright, S. 1969. Evolution and the Genetics of Populations. vol. 2. The Theory of Gene Frequencies. University of Chicago Press, Chicago.