A Thesis

For The Degree of Master of Science

Polymorphisms of the exons 13, 15 and 16 of transferrin gene in Cheju horses

by
Nam-Young Kim

DEPARTMENT OF ANIMAL BIOTECHNOLOGY
GRADUATE SCHOOL
CHEJU NATIONAL UNIVERSITY

2001. 6.

Polymorphisms of the exons 13, 15 and 16 of transferrin gene in Cheju horses

Nam-Young Kim

(Supervised by Professor Young-Hoon Yang)

A thesis submitted in partial fulfillment of the requirements for the degree of Mster of Science

2001. 6.

This thesis has been examined and approved.

Thesis director, Kyu-Il Kim, Prof. of Animal Biotechnology

Hyun-Jong Lee, Prof. of Animal Biotechnology

Young-Hoon Yang , Prof. of Animal Biotechnology

2001. 6.

DEPARTMENT OF ANIMAL BIOTECHNOLOGY

GRADUATE SCHOOL

CHEJU NATIONAL UNIVERSITY

Contents

초 록

본 연구는 제주마집단(Group Ⅰ,제주도 축산진흥원 사육, 137두; Group Ⅱ, 농가사육, 30두)과 더러브렛 품종집단(한국마사회 육성마목장, 43두)을 이용하여 SSCP를 통한 Transferrin exon 13, 15, 16의 다형현상 확인과 각 SSCP 유전자형의 염기서열을 분석하기 위하여 수행하였다.

공시재료에서 SSCP에서 관찰된 band에 의한 분석결과 대립인자는 exon 13, 15 및 16에서 각각 2개(A,B), 3개(A,B,C) 및 3개(A,B,C)가 존재하는 것으로 확인되었다.

Transferrin exon 13의 A와 B의 대립인자는 Group Ⅰ과 더러브렛 품종에서는 빈도가 0.993과 0.007 및 0.977 과 0.023으로 각각 나타나 제주마와 더러브렛 집단 모두 A인자가 매우 높게 분포하고 있음이 확인되었다.

exon 15에서는 관찰된 인자형은 A, B 및 C 의 3종류으로 Group Ⅰ에서는 그 인자형의 빈도는 각각 0.398, 0.463 및 0.139로 조사되었고, GroupⅡ에서는 0.600, 0.283, 0.117로, 더러브렛 품종에서는 각각 0.884, 0.023, 0.093이었다. exon 15에서 높게 출현되고 있는 유전자형은 Group Ⅰ에서 AB(0.445)형, GroupⅡ에서 AA(0.367)형, 더러브렛 품종에서는 AA(0.767) 유전자형이 가장 높은 빈도로 출현되어 제주마 집단간 또는 품종간에 빈도의 차이를 관찰할 수 있었다.

exon 16의 SSCP 분석결과 검출된 인자형은 A, B, C의 3종이 관찰되었다. 대립인자형의 빈도는 Group Ⅰ에서는 A, B, C 각각 0.536, 0.409, 0.055로 나타났으나, GroupⅡ에서는 A 및 B 2종류의 인자형만 0.683 및 0.317로 조사되었고 더러브렛 품종에서는 A(1.000)인자형만 검출되었다. 출현빈도가 높은 유전자형은 Group Ⅰ에서는 AB(0.533)형, GroupⅡ에서는 AA(0.467)형이었고 더러브렛 품종은 AA형(1.000)만 검출되어 제주마와 더러브렛 품종간 유전인자의 빈도차이가 관찰되었다.

exon 15와 exon 16의 SSCP 대립인자형을 친자감별에 이용할 때

parentage exclution확률(친자부정 판정가능확률)은 각각 0.46과 0.374로 추정되었다.

exon 13, 15 및 16의 조합으로 형성된 개체의 유전자형은 전체 13종류가 출현되었고 집단 또는 품종간에 빈도의 차이가 있었다. 이들 전체 유전자형에서 제주마 group I(축산진흥원) 에서는 AA-AB-AB 유전자형, 제주마 group II(농가)와 더러브렛 품종에서는 AA-AA-AA 유전자형이 높은 출현빈도를 보이고 있었다.

SSCP 유전자형에 따른 각 인자들에 대한 염기서열을 분석한 결과 exon 16에서 1개의 새로운 SNP가 발견되었는데 cDNA 2075번에서 Cytosine이 Thymine이로 치환되어 아미노산이 Threonine에서 Methionine으로 치환되고 있음이 확인되었다.

본 연구결과 제주마 transferrin exon 13, 15, 16은 더러브렛 품종에서와 같이 높은 대립인자의 다형성을 보였으며, group II의 일부 제주마는 더러브렛과 교잡된 것으로 검토되었다. 제주마의 유전적 특성을 보호하기 위해서는 등록과 번식에 대한 체계적인 관리가 선행되어야 한다고 사료된다.

I. Introduction

Cheju horses(national treasure No. 347) inhabiting Cheju island have a long history back to 1227 when Mongols brought their horses to the island. They are now used for racing at Cheju Racetrack and for riding.

Many studies have been done on the body types and conformation scores (Lee, 1961; Choung *et al*, 1991; Yang *et al*., 1991; Yang *et al*., 1996), the polymorphisms of blood proteins (Oh *et al*. 1992, Oh *et al*. 1995, Kim *et al*. 1995, and Shin *et al*., 1999) in Cheju horses. Recently, Kim *et al*.(1999) reported the phylogenetic relationships among Cheju and other horse breeds (Mongol, Yunnan, Przewalskii, Swedish, and Thoroughbred) to determine the origin of Cheju horses using the nucleotide sequence polymorphism of mtDNA D-loop region. They suggested that Cheju horses are of mixed origin in their maternal lineage and some extant Cheju horses are descendants of Mongolian matrilines.

Most studies on Cheju horses focused on morphology, blood groups, and blood protein variations. The present study was to investigate the allele frequencies of exons 13, 15 and 16 of the transferrin gene and the relationship of Cheju horses to Thoroughbred breed using the genotypes and their frequencies of this gene.

Ⅱ. Literature Review

Transferrin, which is secreted mainly by the liver, is the major iron-binding glycoprotein found in the serum of all vertebrates. Its major function is to transport iron and also to play important roles in cell proliferation and resistance to infection (Lehninger, 1982; Alberts B. *et al.*, 1989). Equine transferrin occurs in variant forms. Thirteen electrophoretic variants of transferrin were recognized in domestic horses. The segregation of polymorphic bands of plasma transferrin followed autosomal codominant inheritance and these polymorphic bands were used for in routine identifications and pedigree analyses (Bowling, 1999).

The rabbit serum transferrin, which has an $\alpha\beta$ structure of topology similar to human lactoferrin, is composed of two homologus lobes, each of which binds a single ferric ion (Baily *et al.* 1988.; Anderson *et al.* 1989). The ion is bound by two tyrosines, one histidine and one aspartic acid at the iron-binding site, which is located within the interdomain cleft (Baker *et al.* 1987; Baily *et al.* 1988; Anderson *et al.*, 1989). The primary receptor recognition site of human transferrin is on the C-lobe of the protein, and prior binding of this lobe to receptor enables the N-lobe to respond to receptor as well, either directly or by an interaction with the bound C-lobe (Zak *et al.* 1994). The receptor binding is through the C terminal part of the protein, which is in part coded by exon 15 of transferrin gene (Zak *et al.*, 1994). Exon 15 codes for structurally conserved region, including an essential iron binding residue, a region of molecular contact between lobes, and a

domain-linking disulphide bridge (Bailey *et al.*, 1988; Anderson *et al.*, 1989; Baldwin *et al.*, 1993).

The human transferrin, Mouse lactoferrin, and chicken (ovo)transferrin genes have similar intron/exon patterns in which the sizes of the 17 exons are conserved but the intron sizes differ considerably (Schaeffer *et al.* 1987; Shirsat *et al.* 1992; Ghareeb *et al.* 1998). The human transferrin gene which is organized in 17 exons separated by 16 introns has a total size of about 33.5 kb. The chicken ovotransferrin gene comprising also 17 exons and 16 introns has a size of 10.5 kb.

The horse transferrin cDNA sequence is 2305 bp and encodes a protein of 706 residues, including a signal sequence of 19 amino acids (Carpenter *et al.*, 1993). This sequence encodes a protein that had the internally duplicated structure characteristics of the transferrin: the halves had 43% identical amino acids. The horse transferrin had 73.7%, 73.4%, and 72.8% sequence identity with those of the pig, human, and rabbit, respectively (Carpenter *et al.* 1993). Equine transferrin maps to a chromosomal location 16q23 (Lear *et al.* 1999).

Researchers have tried to interpret the polymorphism of plasma transferrin on a DNA sequence basis (Beckman et al, 1998; Baldwin 1993; Carpenter *et al.*, 1994; Brandon *et al.*, 1999). Baily *et al.*(1991) detected a polymorphism in the horse transferrin gene from a Southern analysis with a 700 bp fragment of human transferrin cDNA as a probe. They reported that the digestion with MspI yield a size of 2.4 kb fragment when D, F1, F2, G or H2 alleles of tranferrin protein were present whereas a size of 0.5 kb fragment for O or R protein alleles. Carpenter *et al.* (1994) reported the variant sequences of exon 15, which

had 10 SNPs (Single Nucleotide Polymorphisms). Brandon *et al.* (1999) reported that they identified 23 SNPs in the equine transferrin exons 13 (3 SNPs), 15 (17 SNPs) and 16 (3 SNPs). They observed 16 amino acid changes; 2 in exon 13, 11 in exon 15, and 3 in exon 16.

The single-strand conformation polymorphism (SSCP) analysis is a technique that could detect single base substitutions. It is based on the fact that a single-stranded DNA fragment has a unique folded structure in SSCP process. SSCP analysis is also useful to identify homozygosity or heterozygosity in specific alleles (Orita *et al.* 1989; Kutach *et al.*, 1999).

Ⅲ. Materials and Methods

1. Animals

Three groups of horses were used in this study: 137 Cheju horses from the Cheju Institute (Group Ⅰ), 30 Cheju horses from farms (group Ⅱ), and 43 Thoroughbred horses from Cheju Stud Farm and Training Center established by the Korean Racing Association. The 10 ㎖ of whole blood were collected from each horse for DNA extraction.

2. Genomic DNA preparation

Genomic DNA were extracted from blood samples by a procedure similar to that described by Miller *et al.* (1988). Leukocyte pellets were collected after lysis of erythrocyte with lysis buffer (155 mM NH_4Cl, 10 mM $KHCO_3$, and 10 mM Na_2EDTA) in a 50 ㎖ polypropylene centrifuge tube. Buffy coats of nucleated cells were resuspended with 3 ㎖ nuclei lysis buffer (10 mM Tris-HCl, 400 mM NaCl, and 2 mM Na_2EDTA, pH 8.2). The cell lysates were digested overnight at 50 ℃ after being mixed with 0.2 ㎖ of 10% SDS and 50 ㎕ Proteinase K solution (20 ㎎/㎖). After digestion was complete, 1 ml 6 M NaCl was added to each tube and the tubes were shaken vigorously until the mixture was homogenous, followed by centrifugation at 2,500 × g for 5 minutes. The supernatant containing DNA was transferred to a 15 ㎖ polypropylene tube. About 2 volumes of -20 ℃ absolute ethanol was added and the tube was inverted several times until the DNA

precipitated. The precipitated DNA strands were removed with a pasteurized plastic pipette and transferred to a 1.5 mℓ microcentrifuge tube and then dissolved in 0.5 - 1 mℓ TE buffer (10 mM Tris-HCl, 0.2 mM Na_2EDTA. pH 7.5). The amount of extracted DNA was quantified by a spectrophotometer.

3. Primer design

The following primers were used to amplify the 138 bp, 286 bp and 212 bp fragments of exons 13, 15 and 16 of transferrin gene, respectively.

Exon 13	Forward	5′-TTT CCG TGA AGG CTG TGC CC-3′
	Reverse	5′-CTG AAA GCC CCT GTG TAA CC-3′
Exon 15	Forward	5′-CAG TGA GAG AGC CTT GAC CA-3′
	Reverse	5′-CAC CCG AGA AGA GAA GGT AG-3′
Exon 16	Forward	5′-GTC CTC ATG CAC TTT CTG TC-3′
	Reverse	5′-GAG CAC TGT CTC AGG TTA GC-3′

The transferrin exon 13 primers were the same primers used by Bradon *et al.* (1999). The primers of exons 15 and 16 were designed from a published cDNA sequence (Capenter *et al.* 1993).

4. PCR conditions

PCR was performed in a total of 20 ㎕ reaction containing 50 ng genomic DNA, 5 pM of each primer, 50 mM KCl, 1.9 mM $MgCl_2$, 10 mM Tris-HCl (pH 9.0), 0.1 % Triton X-100, 0.25 mM each dNTP, and 0.75 U Tag polymerase (Promega, Madison, WI). The reaction for exons 13 and 16 included 2 min denaturation step at 94 ℃, followed by 30 cycles of a 92 ℃ denaturation step for 15 sec, a 54 ℃ annealing step for 15 sec, and a 72 ℃ elongation step for 30 sec, and a final 72 ℃ elongation step for 5 min. The reaction for exon 15 was the same as used for exons 13 and 16 except for 58 ℃ for annealing temperature. The reaction was done using a PTC-100 thermocycler (MJ Research, Watertown, MA).

5. SSCP (Single-Strand Conformation Polymorphism)

SSCP gels consisted of 10% w/v acrylamide: Bis-acrylamide (37.5:1), 0.5×TBE (44.5 mM Tris, 44.5 mM boric acid, 1 mM Na_2EDTA pH 8.0), 250 ㎕ of 10% Ammonium persulfate and 45 ㎕ TEMED. The loading mixture was made of 8 ㎕ of PCR product plus 15 ㎕ loading dye (91% N,N-dimethyl formamide, 5% glycerol, 20 mM Na2EDTA, 0.05% bromophenol blue, 0.05% w/v xylene cyanol FF). The mixture was heated to 93 ℃ for 10 min and then placed on ice until loaded. Gels (165㎜×175㎜×1㎜) were run in 0.5×TBE on a V20-set (SCIE-PLAS, Warwickshire, UK) at 100 volts for 12 h (exons 13 and 16) or 16 h (exon 15) at 14 ℃. After running the gels the resulting DNA bands were visualized by silver staining.

6. Sequencing

PCR products of exons 13, 15 and 16 of transferrin gene were seperated on a 1.5% agarose gel and extracted using CONCERT^{TM} Gel Extraction Systems (GIBCO, LIFE TECHNOLOGY). The extracted PCR products were cloned into pCR2.1-TOPO vector using TOPO TA cloning kits (Invitrogen, CA, US). PCR cycle sequencing reactions were accomplished with the SILVER SEQUENCE^{TM} DNA Sequencing System according to the manufacturer's recommendation (Promega, Madison, WI). PCR sequencing reactions were run for the forward and the reverse strands in each sample with M13 forward (17mer) and reverse (17mer) primers. A 6% acrylamide (19:1) gel solution containing 7 M urea, 400 $\mu\ell$ of 10% ammonium persulfate and 50 $\mu\ell$ TEMED was prepared and polymerized in a set of 35 x 45 cm length glass plates. The sequencing PCR products were run on the S3S sequencing apparatus (OWL, Woburn, MA). The PCR sequencing products were fully sequenced in both directions and the sequences were compared for agreement. Ambiguous sequences were checked by re-sequencing.

7. Statistical analysis

To determine whether the observed numbers are consistent with Hardy-Weinberg predictions, a chi-square value was calculated as:

$$\chi^2 = \Sigma [(O - E)^2 / E],$$

O and E are the observed and expected numbers of a particular genotype(Weaver *et al.*, 1991; Gardner *et al.*, 1991).

The probability for parentage exclusion for n alleles was calculated using the following formular (Jamieson *et al.*, 1997);

$$p = 1 + \sum_{i-1}^{n}[p_i^2(2-p_i)]^2-2[\sum_{i-1}^{n} p_i^2(2-p_i)]^2+4(\sum_{i-1}^{n} p_i^3)^2-4\sum_{i-1}^{n} p_i^6$$

n and p_i are the number of alleles and i-th allele frequency.

Ⅳ. Result

1. SSCP analysis of transferrin gene

1) SSCP of exon 13 fragment

SSCP analysis of the PCR products of the transferrin exon 13 revealed three SSCP bands; the fast migrating band (A1), the slow migrating one (A2), and the intermediate band (B) (Fig. 1). The homozygous AA genotype had two bands, A1 and A2. and the heterozygous AB had three distinct bands, A1, A2 and B. No homozygous BB genotype was found in these samples.

The frequencies of alleles A and B were 0.993 and 0.007 in Group Ⅰ, 0.977 and 0.023 in Thoroughbred, respectively. The allele B was not found in group II. The frequency of allele A was much higher than that of the allele B in Cheju horse Group Ⅰ and Thoroughbred horses(Table 1).

The genotype frequencies of exon 13 was 0.985 (AA) and 0.015 (AB) in Group Ⅰ, 1.000 (AA) in GroupⅡ, and 0.953 (AA) and 0.046 (AB) in Thoroughbred horses.

The genotype distribution of exon 13 did not agree with the Hardy-Weinberg genetic equilibrium in Group Ⅰ ($p<0.001$) and Thoroughbred ($p<0.005$).

2) SSCP of exon 15 fragment

Three alleles, A, B, and C were observed in exon 15 and all possible combinations, AA, BB, CC, AB, AC, and AB were found. All the homozygous genotypes showed two SSCP bands, A1 and A2 for AA, B1 and B2 for BB, and C1 and C2 for the CC genotype (Fig. 2). The homozygous AA, BB, and CC genotypes consisted of one fast migrating band and one slow moving band. Because of the similar electorphoretic mobility under the SSCP condition, alleles A1, B1 and C1 were shown the same moving distance in each lane. For this reason heterozygous genotypes AB, AC and BC showed only three bands instead of four.

The allele frequncies of A, B, and C were 0.398, 0.463, and 0.139 in group I, 0.60, 0.283, and 0.117 in group II, and 0.884, 0.023 and 0.093 in Thoroughbred, respectively. The frequency of allele A in Thoroughbred was much higher than those of the other groups (Table 2). Groups I and II showed all six genotypes of transferrin exon 15, but Thoroughbred showed only three genotypes, AA, AB, and AC. The most frequent genotype was AB (0.445) in group I, AA (0.367) in Group II, and AA (0.767) in Thoroubred. The genotype frequencies of Groups I and II agreed well with the Hardy-Weinberg genetic equilibrium($p>0.05$), but not in Thoroughbred ($p<0.025$).

3) SSCP of exon 16 fragment

Three different alleles (A, B and C) and all the genotypes (AA, BB, CC, AB, AC and BC) were identified in the SSCP analysis (Fig 3). The AA genotype had a dark single band which migrated between B1 and B2 bands. The CC genotype had two bands (C1 and C2) which migrated just above the B2 band. The AB and AC genotypes had all three bands; A, B1 and B2 for the genotype AB. and A, C1 and C2 for the genotype AC while the BC genotype showed four bands, B1, B2, C1 and C2. The genotype CC was not detected in the present study groups but detected in some Mongolian horses (data not shown).

Three alleles (0.536 A, 0.409 B and 0.055 C) in group I and two alleles (0.683 A and 0.317 B) in group II and only allele A in Thoroughbred were observed. The genotypes in Cheju horses were AA, BB, AB, AC and BC but only the AA genotype in Thoroughbred.

A total of 13 individual genotypes by the combination of exons 13, 15 and 16 were identified and listed in Table 4. There were differences in the frequencies of the individual genotypes among the three groups. The most frequent genotype was the genotype AA-AB-AB (0.372) for the group I, genotype AA-AA-AA (0.366) for group II, and the genotype AA-AA-AA (0.767) for Thoroughbred.

2. Sequence polymorphism of transferrin gene

The DNA sequences of exon 13 for the genotypes AA and AB are presented in Fig 4. Polymorphism was found at the cDNA position 1582, 1604 and 1626. The genotype AA had one polymorphic site at 1582 (A/G) (Fig. 4). The allele A of the genotypes AA and AB in exon 13 showed the same sequence as AF103832(Genbank accession number). The allele B of the genotype AB showed polymorphism at 1626 (A/G) without amino acd change. G/A subtitution at 1604 reported in Genbank(accesion number AF103826) was not observed in the present study.

Fifteen polymorphic sites were observed in exon 15 (Fig. 5). The polymorphic sites were 1750, 1775, 1783, 1794, 1813, 1827, 1828, 1830, 1837, 1846, 1847, 1897, 1900, 1902, and 1910 in the cDNA. The sequence of allele A showed the same sequence as AF103835. The B allele had 14 SNP sites which were found at cDNA positions 1775(A/G), 1783(A/C), 1794(T/C), 1813(T/G), 1827(C/T), 1828(A/G), 1830(A/G), 1837(T/A), 1846(C/A), 1847(G/A), 1897(T/C), 1900(C/G), 1902(G/A), and 1910(A/G). Translation of the sequence into amino acids revealed that 10 of 15 polymorphic sites caused amino acid substitutions (1750, 1775, 1783, 1813, 1828, 1837, 1847, 1897, 1900, and 1910) (Fig. 5, Fig. 7). The allele B of genotypes BB, AB and BC in exon 15 showed the same sequence as AF103845. The allele C had 15 SNP sites, two of which (position 1750, 1846) differed from those in the B allele of BB type. The 1750 and 1846 sites of SNPs were substitution of A for G and G for C, respectively. The allele C of genotypes CC, AC and BC showed the same sequence as AF103845.

Four polymorphic sites in exon 16 were 1945, 1979, 2011, and 2075 in cDNA (Fig. 6). Allele A of genotypes AA, AB and AC of SSCP was the same sequence as AF103856. The B allele of BB type showed

one SNPs at the position 1979(C/T), which caused amino acid substitution(alanine/valine) at amino acid position 652 (Fig. 7). The B allele of genotypes BB, AB and BC showed the same sequence as AF103864. The C allele of CC type showed three SNP sites which were at the position 1945(C/G), 2011(C/G) and 2075(C/T). The position 2075 was the newly identified SNP site in this study and the substitution at this site resulted in the animo acid change (threonine/methionine) at the amino acid site 684.

V. Discussion

In the SSCP analysis of transferrin gene, the allele A(0.993 in group I, 1.000 in group II, 0.977 in Thoroughbred) of the two alleles(A, B) in exon 13 was predominantly distributed in horse groups. Alleles B and C of the three alleles (A, B, C) in exon 16 were not found in Thoroughbred horses. The differences in the allele and genotype frequencies of the exons 15 and 16 were observed among groups I and II, and Thoroughbred. Shin *et al.* (1999) also reported that TF^{F1} and TF^{F2} showed different frequencies between Cheju horses, Cheju racing horse and Thoroughbreds. However, the intergroup differences in the frequencies of alleles and genotypes of transferrin gene could not be clearly explained as the characteristics of three horse groups. The differences in allelic and genotypic frequencies among groups could be resulted from a small sample size in group II and Thoroughbred and from the different breeding schemes among the groups.

The Mendelian inheritance for these alleles and the multiple allelic frequencies, especially for exons 15 and 16, suggested that SSCP genotyping of this gene could be used to identify the individuals and to test the parentage of offspring. The estimates of the probability for parentage exclusion were 0.46 and 0.374 for the exons 15 and 16 for group I.

Thirteen individual genotypes found by the combination of exons 13, 15 and 16, were different in the observed genotype frequencies among the three groups. The genotype AA-AB-AB (0.372) was the most common in group I, but the genotype AA-AA-AA was common

in II (0.366) and Thoroughbred (0.767). It is not clear that these differences in the genotype frequencies were from the genetic characteristics specific to each group or that the specific type of tranferrin gene was introgressed into group II by crossbreeding between the Cheju horse and Thoroughbred.

A study on the sequences of transferrin gene done by Carpenter *et al.* (1994) reported the 10 SNPs in exon 15 at cDNA positions 1738, 1775, 1783, 1794, 1813, 1827, 1828, 1830, 1837, and 1846. Brandon *et al.*(1999) reported further seven SNPs in exon 15 at cDNA position 1750, 1847, 1897, 1900, 1902, 1909, and 1910. Fifteen SNPs were identified in exon 15 in the present study and the position and base substitution of these SNPs were the same as the previous reports (Carpenter *et al.*, 1994; Brandon *et al.*, 1999) except the base substitutions at positions 1738 and 1999, which were not observed in the present study. Exon 16 had a SNP in the allele C, which substituted cytosine for thymine at the position 2075 and the amino acid also changed from threonine to methionine at 684.

In conclusion, the polymorphism of exons 13, 15 and 16 in the present Cheju horses is as high as in Thoroughbred and other horse groups reported at revious studies. On the consideration of different allele frequencies among the Cheju horse groups I and II and Thoroughbred group, Cheju horse group II may have crossbred with Thoroughbred.

VI. Summary

This study was conducted to determine the polymorphism of transferrin exons 13, 15 and 16 by Single-Strand Conformation Polymorphism(SSCP) analysis and to compare their genotypes of Cheju horse group I (Cheju Institute), Cheju horse group II (farms), and Thoroughbred (KRA).

SSCP of transferrin exons 13, 15, and 16 showed two (A, B), three (A, B, C) and three (A, B, C) codominant alleles, respectively. The frequencies of alleles A and B in the transferrin exon 13 were 0.993, 0.007 in group I, 0.977 and 0.023 in Thoroughbred group, but only allele A in group II. In transferrin exon 15, the frequencies of alleles A, B and C were 0.398, 0.463, and 0.139 in group I and 0.600, 0.283, and 0.117 in group II, and 0.884, 0.023, and 0.093 in Thoroughbred, respectively. The allele frequencies of exon 16 were 0.536, 0.409 and 0.055 for the alleles A, B and C in Group I, 0.683, 0.317 for the alleles A and B in Group II and 1.000 for the allele A in Thoroughbred. The multiple allele frequencies in exons 15 and 16 suggested that the genotyping of this locus could be used to identify an individual and to test the parentage of offspring. The probability for parentage exclusion were 0.46 and 0.374 for exons 15 and 16 for Cheju horse group I.

Among the 13 combined genotypes of exons 13, 15 and 16, the genotype AA-AB-AB (0.372) is the most common in Cheju horse group I, but genotype AA-AA-AA is common in the Cheju horse group II (0.366) and Thoroughbred (0.767), indicating possible crossbreeding between group II and Thoroughbred.

The present study showed one more SNP(not reported before), which was at the cDNA position 2075 (C/T) in C allele of the exon 16 resulting in amino acid change (Threonine → Methionine).

Result of the present study suggests that polymorphism of exons 13, 15 and 16 in Cheju horses is as high as in Thoroughbred and that some Cheju horses in group II may have crossbred with Thoroughbred. Record keeping about the pedigree information and breeding control system are required for the preservation of genetic characteristics of Cheju horses.

Fig. 1 SSCP pattern of equine transferrin exon 13.

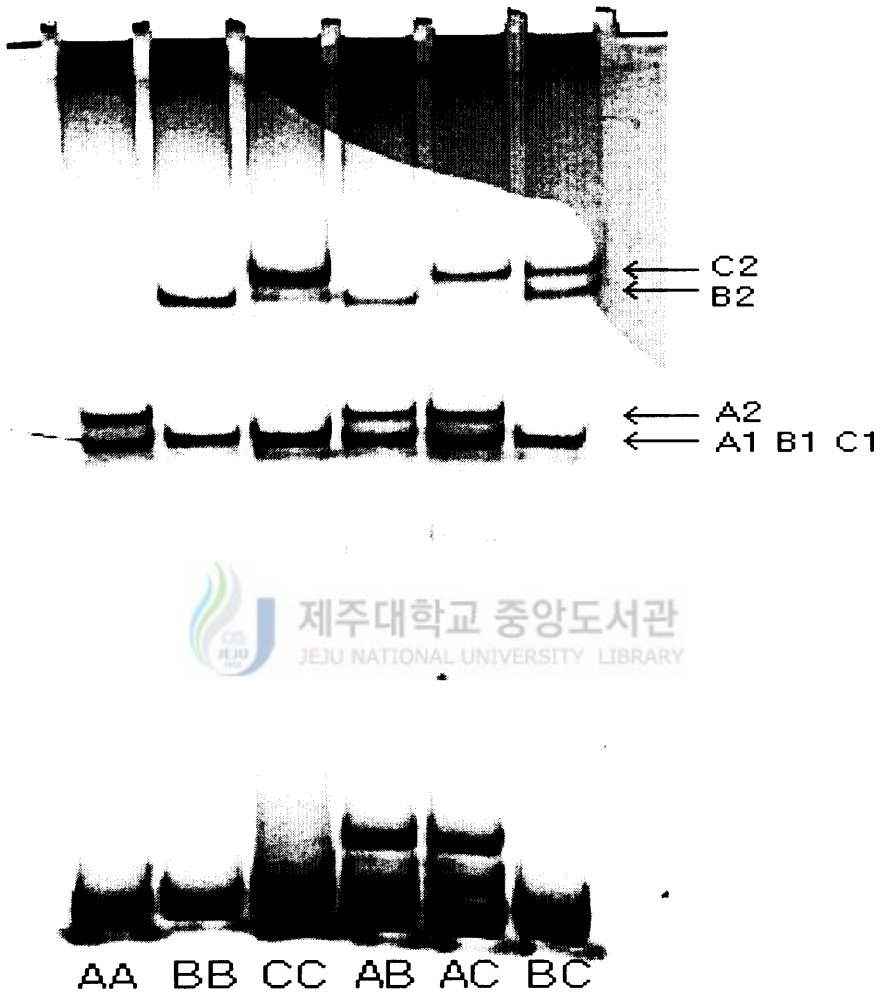

Fig. 2 SSCP pattern of equine transferrin exon 15.

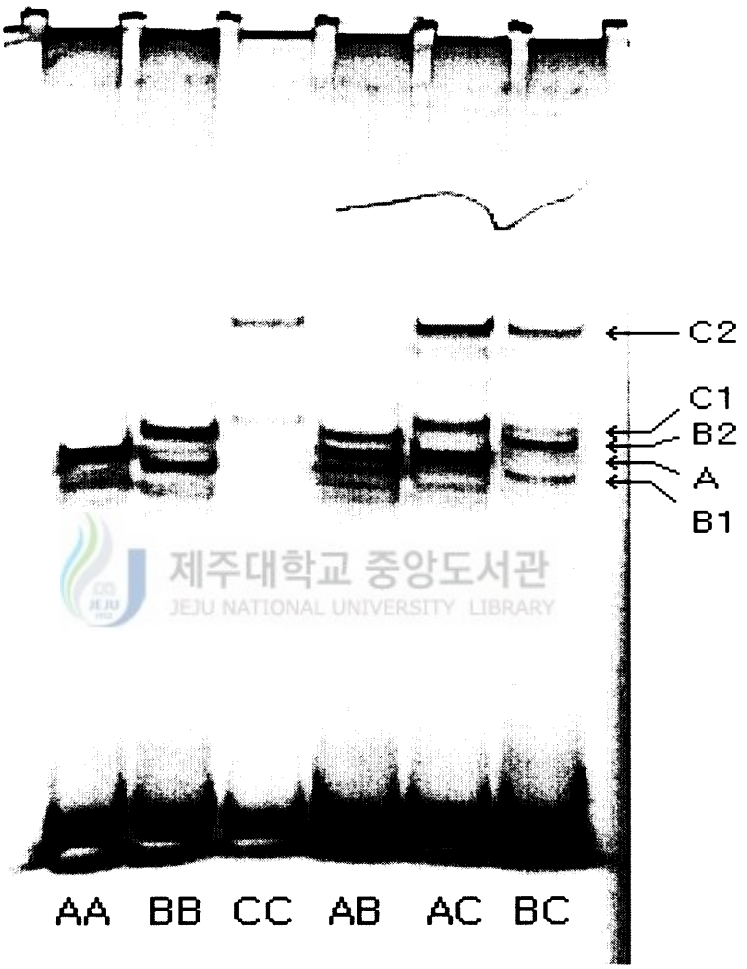

Fig. 3 SSCP pattern of equine transferrin exon 16.

Table 1. Genotype and allele frequencies of transferrin exon 13 in Cheju and Thoroughbred horses.

Breeds		No. of horses	Genotype		Allele		χ^{2*}	p*
			AA	AB	A	B		
Cheju	Group I	137	135 (98.5)	2 (1.5)	272 (99.3)	2 (0.7)	33.37	<0.001
	Group II	30	30 (100)	-	60 (100)	-	-	-
Thoroughbred		43	41 (95.3)	2 (4.6)	84 (97.7)	2 (2.3)	9.88	<0.005

Percentages are in parentheses.

Group I : Horses from Cheju Institute

Group II : Horses from farms

* : Chi-square value and probability for H-W equilibrium test

Table 2. Genotype and allele frequencies of transferrin exon 15 in Cheju and Thoroughbred horses.

Breeds		No. of horses	Genotype						Allele			χ^{2*}	p*
			AA	BB	CC	AB	AC	BC	A	B	C		
Cheju	Group I	137	18 (13.1)	22 (16.1)	2 (1.5)	61 (44.5)	12 (8.7)	22 (16.1)	109 (39.8)	127 (46.3)	38 (13.9)	5.38	>0.1
	Group II	30	11 (36.7)	3 (10.0)	1 (3.3)	10 (33.3)	4 (13.3)	1 (3.3)	36 (60.0)	17 (28.3)	7 (11.7)	0.18	>0.975
Thoroughbred		43	33 (76.7)	-	-	2 (4.7)	8 (18.6)	-	76 (88.4)	2 (2.3)	8 (9.3)	10.41	<0.025

Percentages are in parentheses.

Group I : Horses from Cheju Institute

Group II : Horses from farms

* : Chi-square value and probability for H-W equilibrium test

Table 3. Genotype and allele frequencies of transferrin exon 16 in Cheju and Thoroughbred horses.

Breeds		No. of horses	Genotype no.(%)					Allele freq.			χ^{2*}	p*
			AA	BB	AB	AC	BC	A	B	C		
Cheju	Grpup I	137	32 (23.4)	17 (12.4)	73 (53.3)	10 (7.3)	5 (3.6)	147 (53.6)	112 (40.9)	15 (5.5)	5.39	>0.1
	Group II	30	14 (46.7)	3 (10.0)	13 (43.3)	-	-	41 (68.3)	19 (31.7)	-	0.11	>0.5
Thoroughbred		43	43 (100)	-	-	-	-	86 (100)	-	-	-	-

Percentages are in parentheses.

Group I : Horses from Cheju Institute

Group II : Horses from farms

* : Chi-square value and probability for H-W equilibrium test

Table 4. Genotype frequencies of transferrin gene by combination of exons 13, 15 and 16.

Allele exon13-exon15-exon16	Cheju horse group I		Cheju horse group II		Thoroughbred	
	No. of individual	Frequency	No. of individual	Frequency	No. of individual	Frequency
AA-AA-AA	18	0.131	11	0.366	33	0.767
AA-AB-AB	51	0.372	10	0.333	-	-
AA-AB-AC	9	0.066	-	-	-	-
AA-AC-AA	12	0.088	2	0.066	8	0.186
AA-AC-AB	-	-	2	0.066	-	-
AA-BB-BB	16	0.117	3	0.1	-	-
AA-BB-BC	5	0.036	-	-	-	-
AA-BC-AB	21	0.153	1	0.033	-	-
AA-BC-AC	1	0.007	-	-	-	-
AA-CC-AA	2	0.015	1	0.033	-	-
AB-AB-AA	-	-	-	-	2	0.046
AB-AB-AB	1	0.007	-	-	-	-
AB-BB-BB	1	0.007	-	-	-	-
Total	137	1	30	1	43	1

```
          1558
AF103819  TATAGGCGAA ATTCCACCCT CTGCAATCTG TGTATTGGCT CGGCAAGTGG
AF103826  ---------- ---------- ---------- ---------- ------A---
AF103832  ---------- ---------- ----G----- ---------- ----------
TF13AA    ---------- ---------- ----G----- ---------- ----------
TF13AB-A  ---------- ---------- ----G----- ---------- ----------
       B  ---------- ---------- ----G----- ---------- ----------

          1608
AF103819  TCCAGGAAGG GAGTGTGAAC CCAACAACCA TGAGAGATAC TAT
AF103826  ---------- ---------- ---------- ---------- ---
AF103832  ---------- ---------- ---------- ---------- ---
TF13AA    ---------- ---------- ---------- ---------- ---
TF13AB-A  ---------- ---------- ---------- ---------- ---
       B  ---------- --------G- ---------- ---------- ---
```


Fig. 4. Nucleotide sequences of transferrin exon 13. Reference sequences were from GeneBank(accession number AF103819, AF103826, AF103832). Residues identical with AF103819 sequence are denoted by a dash. The number above sequence is in accordance with cDAN sequence of transferrin gene(GenBank M69020).

```
          1736
AF103835  GACGTAACCC TGACGATTGG GCTAAGGATC TGAAGAGTGA AAACTTTAAG
AF103839  --T------- ---------- ---------- ---------- ----------
AF103845  ---------- ----A----- ---------- ---------G -------C--
AF103848  ---------- ---------- ---------- ---------G -------C--
AF103868  ---------- ---------- ---------- ---------G -------C--
TF15AA    ---------- ---------- ---------- ---------- ----------
TF15BB    ---------- ---------- ---------- ---------G -------C--
TF15CC    ---------- ----A----- ---------- ---------G -------C--
TF15AB-A  ---------- ---------- ---------- ---------- ----------
       B  ---------- ---------- ---------- ---------G -------C--
TF15AC-A  ---------- ---------- ---------- ---------- ----------
       C  ---------- ----A----- ---------- ---------G -------C--
TF15BC-B  ---------- ---------- ---------- ---------G -------C--
       C  ---------- ----A----- ---------- ---------G -------C--
          1786
AF103835  CTGCTATGTC CTGATGGCAC TAGGAAGTCT GTGACTGAAT TCAAAAGCTG
AF103839  ---------- ---------- ---------- ---------- ----------
AF103845  --------C- ---------- -------G-- ---------- -TG-G-----
AF103848  --------C- ---------- -------G-- ---------- -TG-G-----
AF103868  --------C- ---------- -------G-- ---------- -TG-G-----
TF15AA    ---------- ---------- ---------- ---------- ----------
TF15BB    --------C- ---------- -------G-- ---------- -TG-G-----
TF15CC    --------C- ---------- -------G-- ---------- -TG-G-----
TF15AB-A  ---------- ---------- ---------- ---------- ----------
       B  --------C- ---------- -------G-- ---------- -TG-G-----
TF15AC-A  ---------- ---------- ---------- ---------- ----------
       C  --------C- ---------- -------G-- ---------- -TG-G-----
TF15BC-B  --------C- ---------- -------G-- ---------- -TG-G-----
       C  --------C- ---------- -------G-- ---------- -TG-G-----
          1836
AF103835  CTACCTAGCC CGAGCCCCGA ATCATGCTGT AGTCTCACGG AAAGAAAAGG
AF103839  ---------- ---------- ---------- ---------- ----------
AF103845  -A-------- GA-------- ---------- ---------- ----------
AF103848  -A-------- AA-------- ---------- ---------- ----------
AF103868  -A-------- GA-------- ---------- ---------- ----------
TF15AA    ---------- ---------- ---------- ---------- ----------
TF15BB    -A-------- AA-------- ---------- ---------- ----------
TF15CC    -A-------- GA-------- ---------- ---------- ----------
TF15AB-A  ---------- ---------- ---------- ---------- ----------
       B  -A-------- AA-------- ---------- ---------- ----------
TF15AC-A  ---------- ---------- ---------- ---------- ----------
       C  -A-------- GA-------- ---------- ---------- ----------
TF15BC-B  -A-------- AA-------- ---------- ---------- ----------
       C  -A-------- GA-------- ---------- ---------- ----------
          1886
AF103835  CAGCTTGTGT TTGCCAGGAG TTACACAACC AGCAG
AF103839  ---------- ---------- ---------- -----
AF103845  ---------- -C--G-A--- ----G----- -----
AF103848  ---------- -C--G-A--- ----G----- -----
AF103868  ---------- -C--G-A--- ---TG----- -----
TF15AA    ---------- ---------- ---------- -----
TF15BB    ---------- -C--G-A--- ----G----- -----
TF15CC    ---------- -C--G-A--- ----G----- -----
TF15AB-A  ---------- ---------- ---------- -----
       B  ---------- -C--G-A--- ----G----- -----
TF15AC-A  ---------- ---------- ---------- -----
       C  ---------- -C--G-A--- ----G----- -----
TF15BC-B  ---------- -C--G-A--- ----G----- -----
       C  ---------- -C--G-A--- ----G----- -----
```

Fig. 5. Nucleotide sequences of transferrin exon 15. Reference sequences were from GeneBank(accession number AF103835, AF103839, AF103845, AF103848, AF103868). Residues identical with AF103835 sequence are denoted by a dash. The number above sequence is in accordance with cDAN sequence of transferrin gene(GenBank M69020).

```
          1921
AF103856  GCTTCGTATG GAAAAAATGG AAGTCACTGC CCAGACAAGT TTTGTTTGTT
AF103864  ---------- ---------- ---------- ---------- ----------
AF103866  ---------- ---------- ----G----- ---------- ----------
TF16AA    ---------- ---------- ---------- ---------- ----------
TF16BB    ---------- ---------- ---------- ---------- ----------
TF16CC    ---------- ---------- ----G----- ---------- ----------
TF16AB-A  ---------- ---------- ---------- ---------- ----------
       B  ---------- ---------- ---------- ---------- ----------
TF16AC-A  ---------- ---------- ---------- ---------- ----------
       C  ---------- ---------- ----G----- ---------- ----------
TF16BC-B  ---------- ---------- ---------- ---------- ----------
       C  ---------- ---------- ----G----- ---------- ----------

          1971
AF103856  CCAGTCAGCC ACCAAGGACC TTCTGTTCAG GGATGACACA CAATGTTTGG
AF103864  --------T- ---------- ---------- ---------- ----------
AF103866  ---------- ---------- ---------- ---------- G---------
TF16AA    ---------- ---------- ---------- ---------- ----------
TF16BB    --------T- ---------- ---------- ---------- ----------
TF16CC    ---------- ---------- ---------- ---------- G---------
TF16AB-A  ---------- ---------- ---------- ---------- ----------
       B  --------T- ---------- ---------- ---------- ----------
TF16AC-A  ---------- ---------- ---------- ---------- ----------
       C  ---------- ---------- ---------- ---------- G---------
TF16BC-B  --------T- ---------- ---------- ---------- ----------
       C  ---------- ---------- ---------- ---------- G---------

          2021
AF103856  CCAACCTTCA GCCCACAACA ACATATAAAA CCTACTTAGG AGAAAAGTAT
AF103864  ---------- ---------- ---------- ---------- ----------
AF103866  ---------- ---------- ---------- ---------- ----------
TF16AA    ---------- ---------- ---------- ---------- ----------
TF16BB    ---------- ---------- ---------- ---------- ----------
TF16CC    ---------- ---------- ---------- ---------- ----------
TF16AB-A  ---------- ---------- ---------- ---------- ----------
       B  ---------- ---------- ---------- ---------- ----------
TF16AC-A  ---------- ---------- ---------- ---------- ----------
       C  ---------- ---------- ---------- ---------- ----------
TF16BC-B  ---------- ---------- ---------- ---------- ----------
       C  ---------- ---------- ---------- ---------- ----------

          2071
AF103856  CTCACGGC
AF103864  --------
AF103866  --------
TF16AA    --------
TF16BB    --------
TF16CC    ----T---
TF16AB-A  --------
       B  --------
TF16AC-A  --------
       C  ----T---
TF16BC-B  --------
       C  ----T---
```


Fig. 6. Nucleotide sequences of transferrin exon 16. Reference sequences were from GeneBank(accession number AF103856, AF103864, AF103866). Residues identical with AF103856 sequence are denoted by a dash. The number above sequence is in accordance with cDAN sequence of transferrin gene(GenBank M69020).

exon 13

```
          512
AF103819  YRRNSTLCNL CIGSASGPGR ECEPNNHERY Y
AF103826  YRRNSTLCNL CIGSANGPGR ECEPNNHERY Y
AF103832  YRRNSTLCDL CIGSASGPGR ECEPNNHERY Y
TF13-A    YRRNSTLCDL CIGSASGPGR ECEPNNHERY Y
TF13-B    YRRNSTLCDL CIGSASGPGR ECEPNNHERY Y
```

exon 15

```
          572
AF103835  RNPDDWAKDL KSENFKLLCP DGTRKSVTEF KSCYLARAPN HAVVSRKEKA ACVCQELHNQ Q
AF103839  CNPDDWAKDL KSENFKLLCP DGTRKSVTEF KSCYLARAPN HAVVSRKEKA ACVCQELHNQ Q
AF103845  RNPDNWAKDL KSGNFQLLCP DGTRKAVTEF ESCNLAEAPN HAVVSRKEKA ACVRQELRNQ Q
AF103848  RNPDDWAKDL KSGNFQLLCP DGTRKAVTEF ESCNLAKAPN HAVVSRKEKA ACVRQELRNQ Q
AF103868  RNPDDWAKDL KSGNFQLLCP DGTRKAVTEF ESCNLAEAPN HAVVSRKEKA ACVRQELCNQ Q
TF15-A    RNPDDWAKDL KSENFKLLCP DGTRKSVTEF KSCYLARAPN HAVVSRKEKA ACVCQELHNQ Q
TF15-B    RNPDDWAKDL KSGNFQLLCP DGTRKAVTEF ESCNLAKAPN HAVVSRKEKA ACVRQELRNQ Q
TF15-C    RNPDNWAKDL KSGNFQLLCP DGTRKAVTEF ESCNLAEAPN HAVVSRKEKA ACVRQELRNQ Q
```

exon 16

```
          633
AF103856  ASYGKNGSHC PDKFCLFQSA TKDLLFRDDT QCLANLQPTT TYKTYLGEKY LT
AF103864  ASYGKNGSHC PDKFCLFQSV TKDLLFRDDT QCLANLQPTT TYKTYLGEKY LT
AF103866  ASYGKNGSDC PDKFCLFQSA TKDLLFRDDT ECLANLQPTT TYKTYLGEKY LT
TF16-A    ASYGKNGSHC PDKFCLFQSA TKDLLFRDDT QCLANLQPTT TYKTYLGEKY LT
TF16-B    ASYGKNGSHC PDKFCLFQSV TKDLLFRDDT QCLANLQPTT TYKTYLGEKY LT
TF16-C    ASYGKNGSDC PDKFCLFQSA TKDLLFRDDT ECLANLQPTT TYKTYLGEKY LM
```

Fig. 7. Amino acid sequences of the transferrin exon 13, 15 and 16. Reference sequences were from GenBank(accession number AF103819, AF103826, AF103832, AF103835, AF103839, AF103845, AF103848, AF103868, AF103856, AF103864, AF103866).

Reference

Alberts B., D. Bray, J. Lewis, M. Raff, K. Roberts, and J.D. Watson, 1989. Molecular Biology of The Cell. 2nd ed., Garland Publishing, Inc., New York, U.S.A.

Anderson B. F., H. M. Baker., G. E. Norris., D. W Rice and E. N. Baker, 1989. Structure of human lactoferrin: Crystal structure analysis and refinement at 2.8A resolution. Journal of Molecular Biology., 209. 711-734.

Baily. E., T. L. Lear and E. G. Cothran, 1991. Association of MsopI restriction fragment length polymorphisms with transferrin in horse. Animal Genetics., 22. 436.

Baily S., R. W. Evans., R. C. Garratt, B. Gorinsky, S. Hasnain, C. Horsburgh, H. Jhoti, P. F. Lindley, A. Mydin, R. Sarra, and J. L. Watson, 1988. Molecular structure of serum transferrin at 3.3-A resolution. Biochemistry., 27. 5804-5812.

Baker E. N., S. V. Rumball and B. F. Anderson, 1987. Transferrin: insights into structure and function from studies on lactoferrin. Trends in biochemical science., 12. 350-353.

Baldwin G. S., 1993. Comparison of transferrin sequences from different species. Comp. Biochem. Physiol., 106B. 203-218.

Beckman L. E., G. F. Van Landeghem., C. Sikstrom and L. Beckman, 1998. DNA polymorphisms and haplotypes in the human transferrin gene. Human Genetics., 102:141-144

Bowling A. T., 1996. Horse Genetics, P. 84-86. CAB International. Oxford Publications.

Brandon R. B., J. M. Giffard and K. Bell, 1999. Single nucleotide polymorphisms in the equine transferrin gene. Animal Genetics., 30. 439-443.

Carpenter M. A and T. E. Broad, 1993. The cDNA sequence of horse transferrin. Biochimica et Biophysica Acta., 1173. 230-232.

Carpenter M. A and T. E. Broad, 1994. Polymorphismin the coding sequence of the horse transferrin gene. Genome., 37:157-165.

Choung C. C., Y. H. Yang., J. K. Kim and M. S. Kang, 1991. Studies on the classification for the registration of the Cheju Native Horse. I. Body measurements by location, sex, and age. Korean Journal of Animal Science., 33(6). 418-422.

Gardner E. J., M. J. Simmons, and D. P. Smustad, 1991. Principles of Genetics, 8th ed. John Wiley & Sons, Inc.

Ghareeb B.A.A., D. Thepot., C. Puissant., M. Cajero-Juarez and L. M. Houdebine. 1998. Cloning, structural organization and tissue-specific expression of the rabbit transferrin gene. Biochimica et Biophysica Acta., 1398. 387-392,

Jamieson A and St. C. S. Taylor, 1997. Comparisons of three probability formulae for parentage exclusion. Animal Genetics., 28. 397-400.

Kim G. O and M. Y. Oh, 1995. Polymorphism of Plasma Proteins by Two Dimensional Horizontal Electrophoresis in Cheju Native Horse. Korean Journal of Animal Science., 37(6). 603-610.

Kim K. I., Y. H. Yang., S. S. Lee., C. Park., R. Ma., J. L. Bouzat and H. A. Lewin, 1999. Phylogenetic relationships of Cheju horses to other horse breeds as determined by mtDNA D-loop sequence polymorphism. Animal Genetics., 30. 102-108.

Kutach L. S., S. Bolshakov and H. N. Ananthaswamy. 1999. Detection of mutations and polymorphisms in the p53 tumor suppressor gene by single-strand conformation polymorphism analysis. Electrophoresis., 20. 1204-1210.

Lear T. L., R. Brandon., A. Masel., K. Bell and E. Bailey, 1999. Horse alpha-1-antitrypsin, beta-lactoglobulins 1 and 2, and transferrin map to positions 24q15-q16, 28q18-qter, 28q18-qter and 16q23, respectively. Chromosome Research., 7. 667

Lee K. M., 1961. Biostatistic study on the type of the Cheju Horse in Quelport. Korean Journal of Animal Science., 3. 63-73.

Lehninger A. L., 1982. Principles of Biochemistry. Worth Publishers, Inc., New York, U.S.A.

Miller S. A., D. D. Dykes and H. F. Polesky, 1988. A simple salting out procedure for extracting DNA from human nucleated cells. Nucleic Acids Research., 16(3). 1215.

Oh M. Y., M. H. KO., G. O. Kim., S. J. Kim., C. C. Chung and K. I. Kim. 1992. Genetic variations of the blood proteins in Cheju Native Horses. Korean Journal of Genetics., 14(1). 39-50.

Oh Y. S., M. Y. Oh., S. J. Kim., G. O. Kim., M. H. Ko and Y. H. Yang, 1995. Phylogenetic study of Cheju and Tsushima Native Horses. Korean Journal of Animal Science., 38. 324-329.

Orita M., Y. Suzuki., T. Sekiya and K. Hayashi, 1989. Rapid and sensitive detection of point mutations and DNA polymorphisms using the polymerase chain reaction. Genomics., 5. 874-879.

Schaeffer E., M. A. Lucero., J. M. Jeltsch, M.C. Py, M.J. Levin, P. Chambon, G.N. Cohen, and M.M. Zakin, 1987. Complete structure of the human transferrin gene. Comparison with the analogous chicken gene and human pseudogene. Gene., 56. 109-116.

Shin J. A., S. H. Kim., Y. H. Kim., C. I. Chun and K. K. Lee, 1999. Genetic Polymorphism of Serum Protein in Horses on Cheju. Journal of Research Institute for Animal Biotechnology., 14. 149-157.

Shirsat N.V., S. Bittenbender., E. L. Kreider and G. Rovera, 1992. Structure of the murine lactotransferrin gene is similar to the structure of other transferrin-encoding genes and shares a putative regulatory region with the murine myeloperoxidase gene. Gene., 110. 229-233.

Weaver R. F. and P. W. Hedrick, 1991. Basic Genetic. Wm. C. Brown Publishers, IA 52001. U.S.A.

Yang Y. H., C. C. Chang., H. J. Lee and T. S. Kang, 1991. Studies on the classification for the registration of the Cheju Native Horse. Ⅱ. The effect of the Registration Grade on the Body Measurements of Cheju Native Horse. Korean Journal of Animal Science., 33(6). 438-443.

Yang Y. H., J. Kim and D. J. Cho, 1996. Estimation of the Growth Performance of Cheju Native Horse in Cheju Island. Journal of Research Institute for Animal Biotechnology., 11. 9-28.

Zak O., D. Trinder and P. Aisen, 1994. Primary receptor-recognition site of human transferrin is in the C-terminal lobe. Journal of Biological Chemistry., 269. 7110-7114.

Acknowledgement

I am most grateful to Dr. Young-Hoon Yang, major professor, who provided unceasing guidance and profound advice throughout the course of graduate studies, to Dr. Kyu-Il Kim and Dr. Hyun-Jong Lee who contributed actively to my thesis. I wish to thank Dr. Chan-Kyu Park for careful reading of this thesis. I also would like to thank all the professor of Department of Animal Biotechnology, Cheju National University.

I would like to offer my thanks to Young-Hun Kim, Jun Kim, Dae-Cheol Kim and Sung-Soo Lee for their advices and to members of Animal Breeding and Genetics Lab.

Most importantly, I want to express my appreciation for the support and tolerance of my parents, brother and sisters.