

Genetic diversity of the various conserved mtDNA control regions of Cheju Horses

Young-Hoon, Yang

Department of Animal Biotechnology, Cheju National University

제주마의 mtDNA control region에 있어서 부위별 유전적 다양성

양영훈

제주대학교 농업생명과학대학 동물자원과학과

ABSTRACT : The complete mtDNA D-loop sequences of Cheju native horses were used to investigate mtDNA polymorphism in various conserved regions and phylogenetic relationships among other breeds.

Seventy polymorphic sites in the control region, excluding a tandem repeat region, were observed in the seventeen mtDNA D-loop sequences of Cheju horses. These mutation sites comprise 60 transition, 2 transversion and 7 indel sites. The ratios of the base substitution of 5' flank, central domain and 3' flank of the control region were 0.08, 0.07 and 0.04, respectively. The genetic diversity(h) of the control region was 0.91 and the probability of two random individuals having the same genotype was 0.11. Cheju horses showed a broad range of foreign breed affinities in the maximum parsimony tree with the sequences of Thoroughbred, Arabian, Mongolian and Chinese horse breeds. Distribution of mutation sites on the control region was highly unbalanced over various sequence blocks, suggesting that different rates of evolution at different sites should be considered in the analysis models of divergence time and phylogenetic tree among species or populations.

Key Words : polymorphism, D-loop, mtDNA, Cheju horse

INTRODUCTION

Mammalian mitochondrial DNA have been extensively studied in evolutionary and phylogenetic research due to two contrasting features. One is a conservation of size, structure, informational content, and basic gene organization of mtDNA(Attardi, 1985; Clayton, 1991; Wolstenholme, 1992). The other is a high rate of nucleotide sequence evolution(Brown *et al.*, 1979; Brown *et al.*, 1982; Pesole *et al.*, 1999). In addition, mtDNA has other striking features such as maternal inheritance (Hutchinson *et al.*, 1974; Gyllenstein *et al.*, 1991), single copy genome, lack of recombination and clearly defined sequence and gene function. These features make the mtDNA molecule particularly suitable for evolutionary studies (George *et al.*, 1986; Kumazawa and Nishida, 1993; Murray *et al.*, 1998).

The main non-coding region(D-loop) contains the major regulatory elements for replication and expression of the mitochondrial genome and it is the most rapidly evolving part of the mitochondrial genome(Cann *et al.*, 1984). Thus the D-loop has been frequently used for

Corresponding author : Young Hoon Yang, 제주대학교 농업생명과학대학 동물자원과학과, Tel : 064-754-3338.

E-mail : yhyang@cheju.ac.kr

intraspecies phylogenetic analysis to estimate evolutionary relationship and maternal lineage within and between populations. Comparisons of sequence substitutions in mitochondrial control region provides a good tool for the analysis of genetic variation, maternal lineage of a founder population, and the time or rate of molecular evolution(Saccone *et al.*, 1991; Ward *et al.*, 1991; Sbisa *et al.*, 1997; Baasner *et al.*, 1998). The D-loop region(with a length of around 1kb) of mtDNA in mammals is known to be more variable than the sequence in other mtDNA regions and it contains several conserved sequence blocks (Cann *et al.*, 1984; Saccone C. *et al.*, 1987; Ishida *et al.*, 1994; Sbisa *et al.*, 1997; Pessole *et al.*, 1999). The substitution rate of genomic DNA or mtDNA is dependent on the region considered, which suggests that different mutation rates are found in different blocks of the D-loop region and in the calibration of the molecular clock(Saccone *et al.*, 1991; Pesole *et al.*, 1999).

One of the objectives of this study was to investigate genetic diversity in various sequence blocks of mtDNA D-loop of Cheju horses that are being raised in Cheju Institute as protected, heirloom animals. The other objective was to examine the relationships among the Cheju and foreign breeds(Mongolian, Arabian, Chinese Yunnan, and Thoroughbred) by phylogenetic analysis.

MATERIALS AND METHODS

Data of Cheju native horses used in this study were adapted from Yang *et al.*(2002). It

comprises 17 mtDNA D-loop types which were aligned from the complete D-loop sequences of 65 Cheju horses. Genbank accession numbers for the seventeen mtDNA D-loop sequences of Cheju horse type CJ1 through CJ17 are from AF354425 - AF354441.

Genetic diversity of mtDNA sequences (equivalent to heterozygosity) in the sample population was calculated as $h = (1 - \sum x^2) / n$ (where n is the sample size and where x is the frequency of each mtDNA type). The probability of two randomly selected individuals from a population having identical mtDNA types was estimated as $P = \sum x^2$. Phylogenetic analysis of complete mtDNA D-loop sequences was performed using DNAPARS in PHYLIP 3.5(Felsenstein, 1989).

RESULT

Samples of seventeen mtDNA D-loop sequences showed 70 variable positions. The percentage of mutation sites of complete control region was 7.1, and there was a total of 61 transitions, 2 transversions and 7 insertion/deletions. Pyrimidine transitions were a little more frequent than those of purine transitions(Fig.1).

The 5' flank in the control region was the most variable area with 8.2% of the substitution sites among three areas: 5' flank, central domain and 3' flank. But within the small conserved sequence blocks(CSB: Sbisa *et al.*, 1997), CSB1 region in the 3' flank had 12 percent of sites containing a substitution, and ETAS1(Extended Termination Associated Sequence) region in 5' flank had 10 percent of the sites containing substitutions while the CSB3 region had no mutation(Table 1).

were scattered within clusters of various breeds. One Cheju horse D-loop haplotype had a deep-rooted node(CJ13, 100 % of probability) and showed no consistent relationship to each other.

Table 2. Genetic diversity and the probability of maternal lineage identity by various conserved regions within the D-loop

Regions	Number of haplotypes	Genetic diversity(<i>h</i>)	<i>P</i>
Control region	17	0.909	0.105
5' flank	15	0.905	0.109
Central domain	12	0.893	0.120
3' flank	16	0.908	0.106
ETAS1	8	0.752	0.259
ETAS2	6	0.593	0.416
CSB1	6	0.820	0.193
CSB2	2	0.242	0.761
CSB3	0	0.000	1.000

P, probability of two random individuals having the same genotype

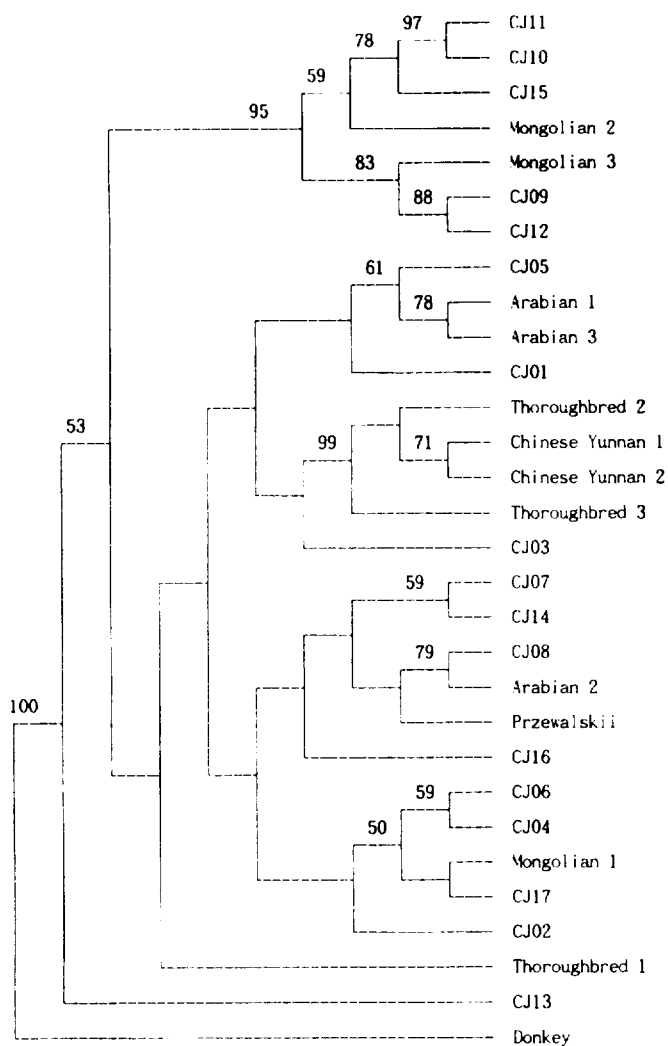


Fig. 2. Phylogenetic tree of Cheju(CJ01-CJ17), Mongolian, Arabian, Chinese Yunnan, Przewalskii, Thoroughbred along with a donkey as an outgroup. The tree was drawn using the maximum parsimony method. Figures on internodes are bootstrap probabilities (in percentage) based on 1000 bootstrapped. Data were adapted from kim et. al.(1999) for Mongolian, Yunnan and Przewalskii, from Flannery(1999) for Arabian and Thoroughbred, and from Xu et. al.(1996) for donkey.

DISCUSSION

Cheju Island has been major horse breeding region of Cheju native horses, crossbreds with foreign breeds and Thoroughbred horses in Korean peninsular. Because of the decreasing value of Cheju native horses for draft or riding in recent times, the number of native horses sharply decreased to about 500 heads in these days.

We consider that Cheju horses comprised various maternal lineages while the origin of Cheju horse has not clearly defined (Kim *et al.*, 1999). Yang *et al.* (2002) reported that Cheju native horse showed 17 maternal lineages in the analysis of mtDNA D-loop. This diversity may be due to a diverse founder population of Cheju horses or due to the results of genetic inflow from other breed or population sources outside Cheju Island in recent times. No sign of bottleneck has been reported under this amount of diversity in history of current population. However, several D-loop types are rare and could be lost without efforts to preserve them.

Phylogenetic analysis showed that Cheju horses did not form a single cluster within the phylogenetic tree among the foreign breeds. Instead they clustered with several non-Cheju horse types included in the analysis. It appears that the population of Cheju horses has a diversity of maternal lineages, indicating that the population of Cheju horses might be a mixed population with origins of a variety of horses brought from outside Cheju Island. This result coincided with the reports of Kim *et al.* (1999) and Yang *et al.* (2002).

Some D-loop types of Cheju horses showed close affinity with Mongolian horse types, suggesting the partial contribution of mongolian ancestor to the present Cheju horse population (Fig. 2).

Previous reports that there was some inflows of Mongolian horses and the dispersion of their genetics characteristics through Cheju native horses was confirmed again over the phylogenetic tree in this study (Lee K. M., 1961; Kang M H., 1965; Oh, Y.S. 1995; Kim *et al.*, 1999). CJ13 has the deepest node within the phylogenetic tree with a 100 percent of probability based on 1000 bootstrapped. This type could be considered to be a deep-rooted maternal lineage of native horses if exist along with those of CJ09 - CJ12 and CJ15.

In general, the degree of genetic diversity of a population in DNA sequence depends on the size of the founding population, chronological duration of the established population, and the degree of isolation of the population from other ones. Populations could undergo a bottleneck event or recombinations between different types of mtDNA but these events reduce genetic diversity and mixed lineages to create a more homogeneous population of mtDNA sequence over time (Ward *et al.*, 1991; Strauss E., 1999). Considering the small local population and geological isolation of Cheju Island from other populations on neighboring continents, a low level of genetic diversity would be predicted. The presence of a substantial level of mtDNA diversity and genetic resemblance to a variety of breeds suggests considerable heterogeneity in maternal lineages of Cheju horses.

The distribution of observed base substitutions showed that most changes are transitions (87% of substitution sites). About 23 percent of transition sites in the D-loop were in the conserved sequence blocks, which were segmented and defined by Sbisà *et al.* (1997) as ETAS1, ETAS2, and CSB1. The Cheju native horse population showed a high amount of genetic diversity at various sequence blocks for such

a small population as now exists as compared to the other horse groups that have been examined(Wang *et al.*, 1994; Kim *et al.*, 1999; Flannery 1999).

As investigated in this study the genetic diversity of the various segmented mtDNA control region varies along with the sequence blocks of control region. This different genetic diversity in various sequenced blocks or conserved regions of mtDNA D-loop suggested that we should consider different rate of evolution at different sites for estimating evolution times and analysing the phylogenetic trees with DNA sequences.

REFERENCES

1. Attardi G., 1985. Animal Mitochondrial DNA: An extreme example of genetic economy. *International Review of Cytology* 93:93-145.
2. Baasner A., Schafer C., Junge A. and Madea B., 1998. Polymorphic sites in human mitochondrial DNA control region sequences: population data and maternal inheritance. *Forensic Science International* 98:169-178
3. Brown W.M., George M. Jr and Wilson A.C., 1979. Rapid evolution of mitochondrial DNA. *Proceedings of the National Academy of Science of USA* 76, 1967-1971.
4. Brown W.M., Prager E.M., Wang A., and Wilson A.C., 1982. Mitochondrial DNA sequences of primates: Tempo and mode of evolution. *J. Mol. Evol* 18:225-239
5. Cann R.L., Brown W.M. and Wilson A.C., 1984. Polymorphic sites and the mechanism of evolution in human mitochondrial DNA. *Genetics* 106:479-499.
6. Clayton D.A., 1991. Replication and transcription of vertebrate mitochondrial DNA. *Ann. Rev. Cell Biol.* 7:453-78.
7. Felsenstein J., 1989. PHYLIP-Phylogeny Inference Package (Version 3.2). *Cladistics* 5:164-166.
8. Flannery A. R., 1999. A study utilizing the DNA sequence variation of the mitochondrial control region to elucidate the domestication pattern of the horse. Thesis of Master of Science at the University of Kentucky
9. George M. and Ryder O.A., 1986. Mitochondrial DNA evolution in the Genus *Equus*. *Mol. Biol. Evol.* 3(6):535-546.
10. Gyllenstein U., Wharton D., Josefsson A. and Wilson A., 1991. Paternal inheritance of mitochondrial DNA in mice. *Nature* 352: 255-257
11. Hutchinson C.A., Newbold J.E., Potter S.S. and Edgell M.H., 1974. Maternal inheritance of mammalian mitochondrial DNA. *Nature* 251:536-538.
12. Ishida N., Hasegawa T., Takeda K., Sakagami M., Onishi A., Inumaru S., Komatsu M., and Mukoyama H., 1994. Polymorphic sequence in the D-loop region of equine mitochondrial DNA. *Animal Genetics* 25:215-221
13. Kang M. H., 1965. Systematic studies on the Korean Native Horse. *Korean J. Anim. Sci.* 7:92-97.
14. Kim K.I., Yang Y.H., Lee, S.S. Park C, Ma R., Bouzat J.L., Lewin H.A., 1999. Phylogenetic relationships of Cheju horses to other horse breeds as determined by mtDNA D-loop sequence polymorphism. *Animal Genetics*, 30, 102-108
15. Kumazawa Y. and Nishida M., 1993. Sequence evolution of mitochondria tRNA genes and deep-branch animal phylogenetics. *J. Mol. Evol.* 37:380-398
16. Lee K. M., 1961. Biostatistic study on the type of the Chuju Horse. *Korean J. Anim.*

- Sci. 3:63-73.
17. Murray-Mcintosh R.P., Scrimshaw B.J., Hatfield P.J., and Penny D., 1998. Testing migration patterns and estimating founding population size in Polynesia by using human mtDNA sequence. *Proc. Nat. Acad. Sci. USA* 95:9047-9052
 18. Oh Y.S., Oh M.Y., Kim S.J., Kim, G.O., Ko M.H., Hiroshi K., and Yang Y.H., 1995. Polymorphisms of blood proteins in Cheju native horses and Tsushima native horses. *Korean J. Zool* 38(3):324-329
 19. Pesole, G., Gissi C., Chirico A. and Saccone, C., 1999. Nucleotide substitution rate of mammalian mitochondrial genomes. *J Mol Evol* 48:427-434
 20. Saccone C., Attimonelli M., and Sbisa E., 1987. Structural elements highly preserved during the evolution of the D-loop-containing region in vertebrate mitochondrial DNA. *J. Mol. Evolution* 26:205-211
 21. Saccone, C., Pesole, G. and Sbisa E., 1991. The main regulatory region of mammalian mitochondrial DNA: Structure-function model and evolutionary pattern. *J. Mol Evol* 33:83-91
 22. Sbisa E, Tanzariello F, Reyes A, Pesole G, Saccone C, 1997. Mammalian mitochondrial D-loop region structural analysis: identification of new conserved sequences and their functional and evolutionary implications. *Gene* 205:125-140.
 23. Strauss E., 1999. Can mitochondrial clocks keep time ? *Science* 283:1435-1438.
 24. Wang W., Liu A.H., Lin S.Y., Lan H., Su B., Xie D.W., Shi L.M., 1994. Multiple genotypes of mitochondrial DNA within a horse population from a small region in Yunnan Province of China. *Biochemical Genetics* 32:371-378
 25. Ward R.H., Frazier B.L., Dew-Jager K., and Paabo S., 1991. Extensive mitochondrial diversity within a single Amerindian tribe. *Proc. Natl. Acad. Sci. USA.* 88: 8720-8724
 26. Ward, R. H., Farazier B. L., Jager K. D. and Paabo S., 1991. Extensive mitochondrial diversity within a single amerindian tribe. *Proc. Nat. Acad. Sci. USA* 88: 8720-8724.
 27. Wolstenholme D.R., 1992. Animal mitochondrial DNA: Structure and evolution. *International Review of Cytology* 141:173-216
 28. Xu X, Arnason U., 1994. The complete mitochondrial DNA sequence of the horse, *Equus caballus*: extensive heteroplasmy of the control region. *Gene* 148:357-362.
 29. Xu X, Arnason U., 1996. The complete mitochondrial DNA(mtDNA) of the donkey and mtDNA comparison among four closely related mammalian species-pairs. *Journal of Molecular Evolution* 43, 438-446.
 30. Yang, Y.H., Kim, K.I., Cothran, E.G., and Flannery, A.R., 2002. Genetic diversity of Cheju hoses(*Equus caballus*) determined by using mitochondrial DNA D-loop polymorphism. *Biochemical Genetics*, vol. 40: 175-186