Population Genetic Studies on the Chinese

Jin Feng, Saitou Naruya, Omoto Keiichi

Interdisciplinary Perspectives
on the Origins of the Japanese
INTERNATIONAL SYMPOSIUM 1996

International Research Center for Japanese Studies

Population Genetic Studies on the Chinese

Jin Feng[1], Saitou Naruya[2], Omoto Keiichi[3]

1. Institute of Genetics, Chinese Academy of Sciences, Beijing, China
2. National Institute of Genetics, Mishima, Shizuoka Prefecture, Japan
3. International Research Center for Japanese Studies, Kyoto, Japan

Introduction

For the study of the origins of the Japanese, genetic data of Chinese populations are of crucial importance. Thus far, however, most published genetic data of the Chinese populations were obtained either from outside the mainland China or from a few big cities like Guangzhou, Shanghai and Beijing, where the populations moved in and out frequently in the history and thus not necessarily useful in the study of regional variation in East Asia (Roychoudhury and Nei, 1988).

The definition of Chinese ethnic minorities is simply based on four main items: common language; common area; common economic condition and common psychological characteristics. The genetic diversity among them, however, has not been studied except for a few instances (Omoto et al., 1993; 1996; Saitou et al., 1994). Even the majority Han Chinese population as we know now may not be geographically homogeneous, judging from phenotypic diversities in physical characters of body as well as languages.

In order to start understanding more precisely the anthropological histories of Chinese populations, of both Han Chinese and ethnic minorities, we are carrying out a series of population genetic studies on blood samples obtained from Han Chinese of different regions and a few ethnic minorities. In the present study, we report part of the results based on the "classical" genetic markers including blood groups and red cell enzyme types including subtypes. Multi-locus genetic distance analyses were carried out emphasizing the relationship with Japanese populations. Although we need much more data for complete elucidation of the origins of the Japanese populations, as well as on genetic regional diversities among the Chinese populations, we report here a few interesting findings.

Materials and Methods

In order to avoid influences of recent migrations from other provinces, we have carefully selected the following areas for sample collection: (1)Taiyuan of Shanxi (山西) Province, Xi'an of Shaanxi (陝西) Province, Yanqing (延慶) County of Beijing and Guan (固安) County of Hebei (河北) Province, as well as Lijiang (麗江) County in Yunnan (雲南) Province for an ethnic minority Naxi (納西) group.

Geographically, Shanxi Province and Shaanxi Province are located on the central to northern part of China, at the latitude of between 33° and 40°. The Yellow River (*Huanghe*) separates the two provinces, and forms the border of the two administration areas. It is said that the ancestor of the Han Chinese was *Huangdi* (黄帝) whose tomb is located in the Shaanxi Province. The sampling place of Shaanxi Province, Xi'an, is also famous as the capital city of *Tang* (唐) Dynasty. Its neighbor, Shanxi Province, takes its name because this province is located on the western side of Taihang (太行) Mountains. In the history, the Yellow River between the two provinces appeared not to be a real barrier for human migrations. People could move between the two regions over the frozen surface of the river in winter.

The third Han Chinese population sampled in this study is in Yanqing County of Beijing, located outside the Great Wall zone, Badaling (八達嶺) Mountain, the entrance of Beijing from Northern areas. Yanqing is known in the history as an ancient war field for 3000 years. Especially, the Mongolians of Yuan (元) Dynasty and the Manchurians of Qing (清) Dynasty conquered this place firstly and entered into Beijing. Therefore, we wondered if the Han Chinese population in this area is genetically related to the northern ethnic minorities.

Another place we chose for Han Chinese population was from Hebei Province, Gu'an County, which is the neighboring county of Beijing. It is about 50 kilometers south of Beijing. The only one ethnic minority investigated in the present study is the Naxis that distributed in Lijiang County of Yunnan Province. Geographically, the Naxis are surrounded by other ethnic minority groups such as Bai (白), Yi (彝) and the Tibetan. The origin of Naxis is so far unknown. The living styles of the Naxis show some cultural influences from the Tibetan of the same area. The facial features of Naxis are suggestive of a northern "Mongoloid" group, in contrast to other ethnic minority groups in Yunnan, such as Tai (傣) and Wa (佤).

In the present study, we collected blood samples from unrelated individuals: 124 for Taiyuans, 126 for Xi'ans, 120 for Guans, 119 for Yanqings and 109 for Naxis. Genetic markers examined include blood groups such as ABO, MNSs, Rh, P, Duffy, Diego, Kidd, Xg, red cell enzyme types including subtypes of 6PGD, AcP, PGM1, ADA, EsD, and GPT. All the results of the observed phenotypes are

in good agreement with that of the expected based on the Hardy-Weinberg's law.

For the comparison of our data with other possibly related populations, we used some unpublished data from different Han Chinese and ethnic minority populations which have already been gathered by us and our Chinese colleagues in Beijing as follows (Fig. 1):

1. Mongolians from Huhhot of Inner Mongolia;
2. Han Chinese from Huhhot of Inner Mongolia;
3. Manchuria from Xiouyan（珲岩）County of Liaoning（遼寧）Province.
4. Hui（回）ethnic minority from the Yinchuan（銀川）, capital city of Ningxia（寧夏）Hui Autonomous Region;
5. Han Chinese from Harbin, the capital city of Heilongjiang（黒竜江）Province;
6. Han Chinese from Guiyang（貴陽）, the capital city of Guizhou（貴州）Province;
7. Han Chinese from Lanzhou（蘭州）, the capital city of Gansu（甘粛）Province;
8. Yi ethnic minority from Liangshan（凉山）Yi Autonomous Prefecture of Sichuan（四川）Province;
9. Hani ethnic minority from Honghe（紅河）of Yunnan Province;
10. Bai ethnic minority from Dali（大理）of Yunnan Province;
11. Yao（瑶）ethnic minority from Bama（巴馬）County of Guangxi（廣西）Zhuang（壮）Autonomous Region.
12. Dong（伺）ethnic minority from Sanjiang（三江）County of Guangxi Zhuang Autonomous Region.
13. Zhuang ethnic minority from Nanning（南寧）City of Guangxi Zhuang Autonomous Region.
14. Han Chinese from Zhengzhou（鄭州）, the capital city of Henan（河南）Province.

Based on the allele frequencies of these genetic marker loci, the modified Cavalli-Sforza genetic distance (DA) have been calculated (Nei et al., 1983). The data for comparison of Japanese populations were those published already and made available for this study by the Japanese counterparts. Because we don't have complete data for every population listed in the present paper, we have to compare only the common data for different groups. Phylogenetic trees (dendrograms) have been constructed based on both UPGMA (Sokal and Sneath, 1963) and the Neighbor-Joining (NJ) method. (Saitou and Nei, 1987).

Results and discussion

In this study a number of phylogenetic trees are constructed based both on Dst and DA genetic distances using different set of genetic loci. The results revealed that if the data of blood group systems of Chinese populations are in-

cluded, the trees sometimes reveal poor correspondence with geographical distribution of these populations. For example, a NJ tree shown in Fig. 2 based on DA distances using data of 12 loci (ABO, MNSs, Rh, Fy, Jk, Xg, Di, P, GPT, ESD, AcP, PGM) is difficult to evaluate, except that ethnic minority groups tend to have extremely diverse positions. The reason is at present not clear, but one explanation would be that some blood group data with relatively rare alleles such as Di and Xg systems have large standard deviations and effects of sampling/typing errors.

On the other hand, it is interesting that trees based on the data of four red cell enzymes including subtypes examined by the present author (JF) himself (PGM, ADA, EsD, GPT and AcP) show good agreement with geographical distribution, anthropological and linguistic findings of the populations examined. The information of subtypes of PGM1 (Bark et al., 1976) and ADA (Jin et al., 1995) which have been made available by means of isoelectric focusing method also improved the usefulness of red cell enzymes as genetic and anthropological marker. In this study, however, the ADA subtype data can not be included, since this

Fig. 1 The map of China indicating the localities of the sample collection in the present study.

subtype was only recently discovered among native ethnic minority groups in Taiwan but not yet examined in the mainland Chinese populations.

Figure 3 shows a NJ dendrogram based on DA distances comparing Japanese (Hondo-Japanese) with nine Han Chinese populations examined. The data of ten genetic loci are used: ABO, MNSs, Rh, Duffy, Kidd, P, GPT, EsD, AcP, and PGM. The result indicates that Japanese population is not well separated genetically from all the Han Chinese populations examined in this study. Among the Han populations there is no clustering of regional groups and it is hard to make interpretations with respect to genetic relationship.

We then tried to exclude all the blood group data and use only four highly polymorphic red cell enzyme loci. The result is shown in Fig. 4, where the same comparison of Japanese and nine Han Chinese populations as used in Fig. 3 was made. This time we used UPGMA method for construction of dendrogram based on DA distances. We consider that this dendrogram is in good accordance with geographic distribution of the populations. Thus, it is apparent that there are two cluster groups: Taiyuan/Xi'an/Beijing/Guan on the one hand, and Guizhou/ Heilongjian/Inner Mongolia/Gansu/Henan on the other. Before this examination, we expected to obtain a separation of northern and southern Han groups, as suggested in the previous studies (Saitou et al., 1994; Omoto et al., 1996). While the first cluster group in Fig. 3 is composed of exclusively northern Han groups, the other cluster includes both northern and southern groups. Thus, the problem of the genetic diversity between northern and southern Han populations is still unsolved at present. More studies including much more populations both from northern and southern China are needed to solve this question.

Figure 5 shows a larger dendrogram comparing Japanese with 17 Chinese populations including ethnic minorities, using data of the same four enzyme loci as in Fig. 4. The result is interesting in several aspects. First, Japanese is clustered together with Mongolian, and then with Manchulia. This finding is consistent with the anthropological hypothesis such as Hanihara's concerning the formation of modern Japanese populations. According to K. Hanihara, the majority of modern Japanese called Hondo-Japanese essentially originated mainly from immigrants from north-eastern Asia since the Yayoi period starting 2,300 BP, while the native Ainu and Okinawa populations were derived from the Jomon people who lived in Japan before the Yayoi period and originated probably in south-east Asia (Hanihara, 1991).

The second interesting point of Fig. 5 is the nature of cluster groups. There are three major population clusters and a few isolated populations. Roughly speaking, each major cluster consists of the populations with similar geographical areas. Thus, in the "northern" cluster there are five populations, Japanese, Mon-

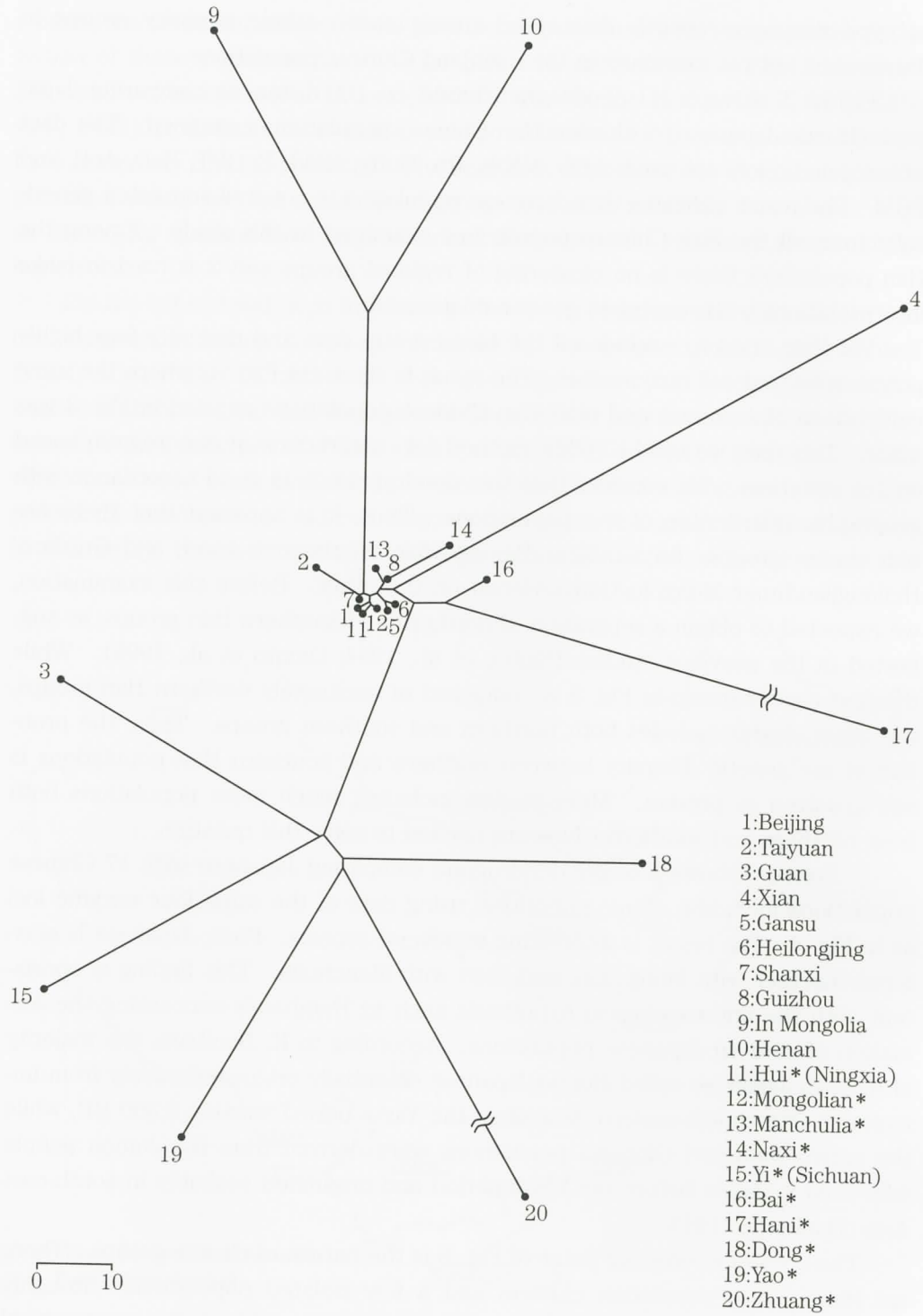

Fig. 2 An example of NJ tree for 20 Chinese populations based on data of 12 genetic loci including those of blood groups

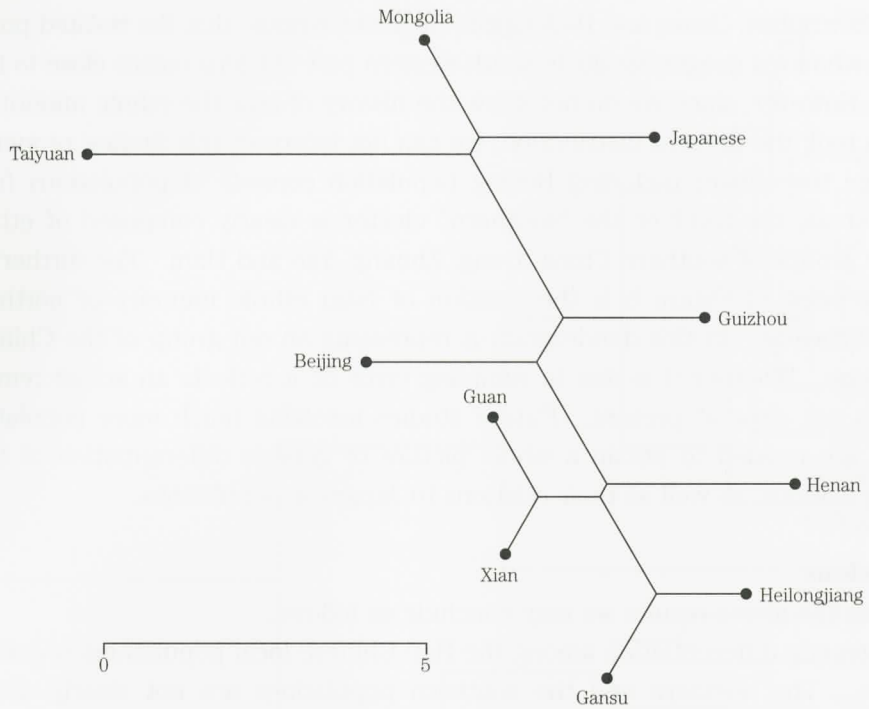

Fig. 3 An example of NJ tree for comparison of Japanese (Hondo-Japanese) population with nine Han Chinese populations on the basis of data of ten genetic loci including those of blood groups.

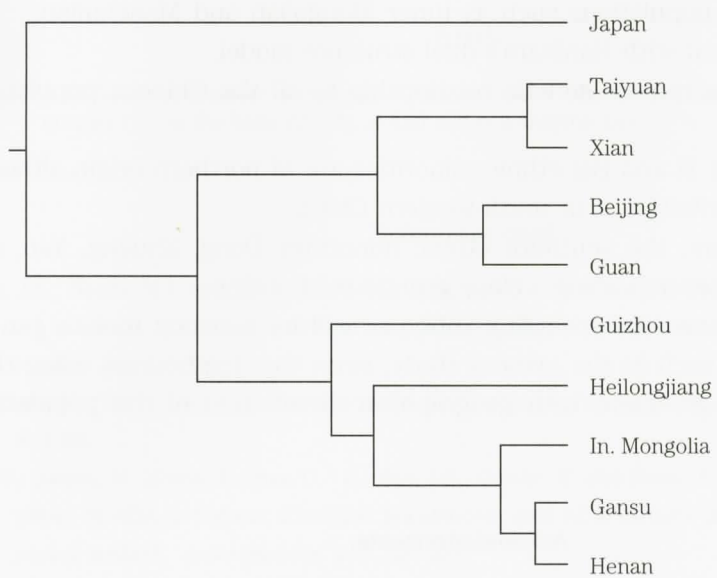

Fig. 4 A UPGMA dendrogram comparing Japanese with nine Han Chinese populations on the basis of data of exclusively four red cell enzyme loci. Han Chinese and Japanese Modified Cavelli-Sforza's tree based on GPT, EsD, Acp, PGM

golian, Manchulian, Gansu and Heilongjian. It seems strange that the isolated population Yi who lives geographically in south-western part of China comes close to this cluster. However, since we do not know the history of how the ethnic minorities of China took the present distribution, we can not interpret this finding properly.

While the cluster including Beijing population consists of populations from various areas, the third or the "southern" cluster is clearly composed of ethnic minority groups of southern China: Dong, Zhuang, Yao and Hani. The further interesting point of Figure 5 is the position of Naxi ethnic minority of northern Yunnan Province. In this dendrogram it represents an out-group of the Chinese populations. Whether it is due to sampling error or it reflects an actual remote origins is not clear at present. Future studies involving much more population samples are needed to obtain a whole picture of genetic differentiation of Chinese populations, as well as their relations to Japanese populations.

Conclusions

From the above results we may conclude as follows:

1) The genetic differentiation among the Han Chinese local populations is considerable. The northern and the southern populations are not clearly distinguished, but the cluster groups in the dendrogram tend to be composed of populations with similar geographical distribution.
2) The Japanese (Hondo-Japanese) population has the highest affinity with north-eastern Chinese populations such as Inner Mongolian and Manchulian. This finding is consistent with Hanihara's dual structure model.
3) The Naxi shows a remote genetic relationship to all the Chinese populations examined.
4) It is probable that Yi and Bai ethnic minorities are of northern origin although their present distribution is in south-western China.
5) In the dendrogram, the southern ethnic minorities Dong, Zhuang, Yao, and Hani share a cluster indicating a close genetic relationship.
6) Red cell enzyme markers including subtypes will be a strong tool to genetic distance analysis such as the present study, since the dendrogram using their data is in good agreement with geographical distribution of the populations examined.

Acknowledgements

We thank The Tokyo Club, Tokyo, for a grant supporting our field trip to Yunnann Province in 1996, and Monbusho, Ministry of Education, Science and Culture, for the grants-in-aid for scientific study. We thank also Ms. Haoluping, Mr. Cao Jinglong for their experimental helps.

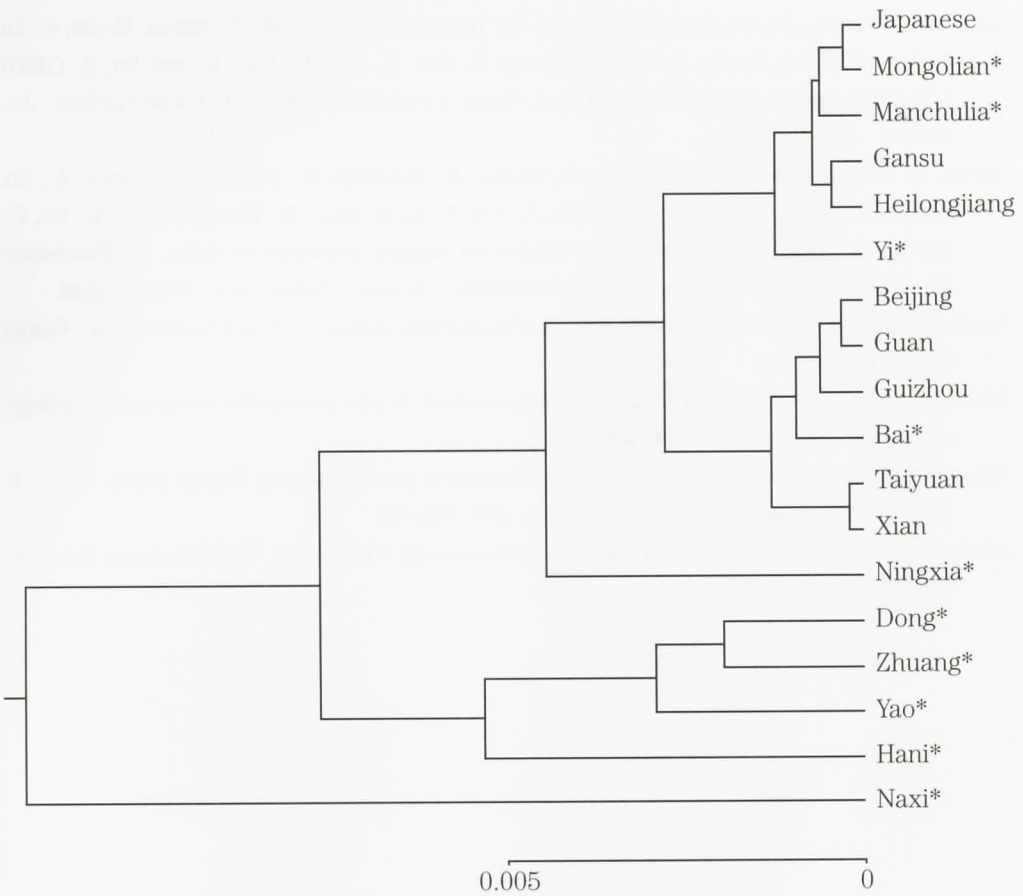

Fig. 5 A UPGMA dendrogram comparing Japanese with 17 Chinese populations including ethnic minorities (*) on the basis of data of four red cell enzyme loci.

References cited

Bark, J.E.; Harris, M.J. and Firth, M. (1976) Typing of the common phosphoglucomutase variants using isoelectric focusing—a new interpretation of the phosphoglucomutase system. *J.Forensic. Sci.Soc. 16*, 115-120.

Hanihara, K. (1991) Dual structure model for the population history of the Japanese. *Japan Review 2*, 1-33.

Jin, F.; Saitou, N.; Ishida, T.; Sun, C. - S.; Pan, I-H., Omoto, K. and Horai, S. (1995) Genetic polymorphism of ADA in Taiwan aboriginal populations: new ADA variants detected by isoelectric focusing method. *Anthropol.Sci. 103*, 227-234.

Nei, M.; Tajima, F. and Tateno, Y. (1983) Accuracy of estimated phylogenetic trees from molecular data.II. Gene frequency data. *J. Mol. Evol. 19*, 153-170.

Omoto, K.; Hirai, M.; Jin, F.; Misawa, S.; Washio, K.; Tokunaga, K.; Yamazaki, K.; Saitou, N.; Du, C.; Lu B.; Luo, Z.; Zeng, Z.; Xu, J.; Zao, H.; Zhang, Z.; Wei, X.; Niu, K., Hao, L.; and Du, R. (1993) Population genetic study in Hainan Island, China. I. Distribution of blood genetic markers. *Anthropol. Sci. 101*, 1-24.

Omoto, K.; Hirai, M.; Harihara, S.; Misawa, S.; Washio, K.; Tokunaga, K.; Saitou, N. Yamazaki, K., Du, R.; Hao, L.; Yuan, Y.; Luo, Z.; Xu, J.; Hu, J.; Wei, X.; Li, S.; Zhao, H.; Zhang, Z.; Niu, K.; Du, C.; and Lu, B. (1996) Population genetic studies on National minorities in China. In *Prehistoric Mongoloid Dispersals* (Akazawa, T. and Szathmary, E. eds.), Oxford Univ. Press, Oxford.

Roychourhury, A.K. and Nei, M. (1988) *Human Polymorphic Genes: A World Distribution*. Oxford Univ. Press, Oxford.

Saitou, N. and Nei, M. (1987) The nighbor-joining method: A new method for constructing phylogenetic trees. *Mol. Biol. Evol.4*, 406-425.

Saitou, N.; Omoto, K.; Du, C. and Du, R. (1994) Population genetic sutdy in Hainan Island, China. II. Genetic affinity analyses. *Anthropol. Sci. 102*, 129-147.

Sokal, R. and Sneath P.A.P. (1963) *Principles of Numerical Taxonomy*. W.H.Freedman, San Francisco.