

Population Genetic Studies on Nine Aboriginal Ethnic Groups of Taiwan

II. Serum Protein Systems

Isao Yuasa¹, Kazuo Umetsu², Kazutoshi Ago³, Cheih-Shan Sun⁴,
I-Hung Pan⁵, Takafumi Ishida⁶, Naruya Saitou⁷, and Satoshi Horai⁸

¹Department of Legal Medicine, Tottori University School of Medicine, Yonago

²Department of Forensic Medicine, Yamagata University School of Medicine, Yamagata

³Department of Legal Medicine, Faculty of Medicine, Kagoshima University, Kagoshima

⁴Taidong Hospital, Taidong, Taiwan

⁵Department of Bacteriology, College of Medicine, National Taiwan University, Taipei, Taiwan

⁶Department of Biological Sciences, Graduate School of Science, University of Tokyo, Tokyo

⁷Laboratory of Evolutionary Genetics, National Institute of Genetics, Mishima

⁸Department of Biosystems Science, The Graduate University for Advanced Studies, Hayama

(Received May 16, 2001; accepted August 31, 2001)

Abstract As part of our comprehensive genetic study of the aboriginal Taiwanese (Gaoshan), we have reported the distribution of allele frequencies for 13 red cell enzyme and 4 serum protein polymorphisms in all nine aboriginal groups. In this study, 10 additional serum protein polymorphisms for AHSG, C3, C6, C7, GC, HP, ITIH1, PI, TF and ZAG were typed. Common and variant alleles prevailing in East Asian populations were found together with several new variants. The distribution of allele frequencies differed markedly among the nine aboriginal groups and from that in East Asian populations. The large diversity among the aboriginal groups was confirmed by analysis of the genetic affinity on the basis of Nei's standard genetic distance and Cavalli-Sforza's chord distance, which were calculated using the present and previous data from 27 red cell enzyme and serum protein loci. The analysis based on 15 red cell enzyme and serum protein loci revealed that they, in particular the Puyuma and Ami, had strong affinity with Filipino among several East Asian populations.

Keywords: aboriginal Taiwanese, Gaoshan, population study, serum protein polymorphisms

Introduction

The Gaoshan (Mountain Highlander) are one of the 56 recognized ethnic groups in China and are an aboriginal population in Taiwan. They consist of nine linguisti-

Corresponding author: Isao Yuasa

Department of Legal Medicine, Tottori University School of Medicine

86 Nishimachi, Yonago, Tottori 683–8503, Japan

E-mail: yuasai@grape.med.tottori-u.ac.jp

cally and culturally different groups: the Atayal, Saisiat, Bunun, Tsou, Ami, Puyuma, Rukai and Paiwan who live on the main island of Taiwan, and Yami who live on Lanyu Island, located off the southeast coast of the main Island (Du and Yip, 1993; Jin et al., 1999). There have been few comprehensive population genetic studies done on all the nine aboriginal groups in Taiwan at the same time. Our project was undertaken to collect blood samples from all nine groups, to analyze genetic markers including red cell antigens and enzymes, serum proteins, mtDNA and Y chromosome, and to understand the relationships among these groups and the genetic affinity with other populations in East Asia (see Jin et al., 1999). As a part of this project, we preliminarily reported the data on a red cell enzyme, adenosine deaminase (Jin et al., 1995) and on four serum protein polymorphisms including the R subcomponent of the first component of the complement (C1R), complement factor I (IF) and orosomucoid (ORM1 and ORM2) (Umetsu et al., 1994, 1995). Recently we have begun a new series of comprehensive study and have reported the data on 12 red cell enzyme polymorphisms (Jin et al., 1999). Also, another group investigated the HLA system in all nine aboriginal Taiwanese groups (Lin et al., 2000).

Concerning serum protein polymorphisms in the aboriginal Taiwanese, only several polymorphisms including haptoglobin (HP), group-specific component (GC, vitamin D-binding protein), and gamma and kappa chains of immunoglobulin (GM and KM) have been studied in limited numbers of ethnic groups among the aboriginal Taiwanese (Fong, 1974; Nakajima and Ohkura, 1971; Matsumoto, 1988). The Atayal group living in Toroko were studied in detail for six serum protein polymorphisms including the third component of the complement (C3), complement factor B (BF), GC, HP, transcobalamin II (TC2) and Transferrin (TF) as well as 13 other blood groups and red cell enzyme polymorphisms (Chen et al., 1985). Recently, new polymorphisms and microheterogeneity of some serum proteins have been demonstrated by the application of isoelectric focusing. These genetic markers are useful for population studies. Subsequently to the first report of this new series (Jin et al., 1999), this paper presents the typing results of ten additional loci for serum proteins in all nine aboriginal groups of Taiwan.

Materials and Methods

A total of 659 blood samples (Umetsu et al., 1994, 1995; Jin et al., 1995, 1999) were examined for ten genetic markers including alpha-2-HS-glycoprotein (AHSG), C3, the sixth and seventh components of the complement (C6, C7), GC, HP, inter-alpha-trypsin-inhibitor H1 chain (ITIH1), alpha-1-antitrypsin (PI), TF and Zn-alpha-2-glycoprotein (ZAG).

C3 and HP were typed using agarose and polyacrylamide gel electrophoresis, respectively. The others were determined by isoelectric focusing on thin-layer poly-

acrylamide gels. Allele frequencies were calculated by the standard allele counting method. Nei's standard genetic distance (Ds) (Nei, 1987) and Cavalli-Sforza's chord distance (Dc) (Cavalli-Sforza and Edwards, 1967) were calculated by PHYLIP 3.5c (Felsenstein, 1993). Phylogenetic trees were constructed using the unweighted pair-group method with arithmetic mean (UPGMA) and neighbor-joining method (NJ) (Sokal and Sneath, 1967; Saitou and Nei, 1987) with the help of PHYLIP 3.5c (Felsenstein, 1993) and TreeView (Page, 1996).

Results

Distribution of allele frequencies

The distribution of allele frequencies for the ten genetic markers examined is shown in Table 1.

AHSG*1 was predominant over AHSG*2 in the nine groups like all the Asian populations investigated. The allele frequency for AHSG*2 ranged from 34% in the Saisiat to 2% in the Atayal. A combined frequency for AHSG*2 in the nine groups (13%) was low in comparison with other East Asian populations. In the Honshu island and Okinawa of Japan AHSG*2 was observed at the frequency of 27% and 23%, respectively (Yuasa and Umetsu, 1988). The aboriginal Taiwanese had no AHSG*5, which was observed at polymorphic frequencies in Okinawa, a neighboring population located close to the north coast of Taiwan (Yuasa and Umetsu, 1988). The data in the Ami were not in agreement with Hardy-Weinberg's law. The reason for discrepancy is not known.

The C3 system is generally monomorphic due to the occupation of C3*S in East Asian populations studied (Roychoudhury and Nei, 1988; Nishimukai et al., 1985). The Atayal in Toroko was reported to have only C3*S (Chen et al., 1985). Similarly, six of the nine aboriginal groups including the Saisiat, Ami, Puyuma, Rukai, Paiwan and Yami were monomorphic. The three other groups, the Tsuo, Bunun and Atayal, had C3*F07_{tsou} or C3*F06_{atayal}, which subtly differed in mobility from the C3*F07 and C3*F06 found in Europeans, respectively (Nishimukai, personal commun.). It is noteworthy that C3*F07_{tsou} was observed at an allele frequency of as high as 10% in the Tsou. Also, one Bunun individual possessed this variant.

Concerning the C6 system, C6*A, C6*B and C6*B2 were common alleles in East Asian populations including Japanese and Han Chinese, where C6*A was observed at a high frequency of more than 40%, but was somewhat lower in frequency than C6*B (Roychoudhury and Nei, 1988). All aboriginal groups except the Bunun had the three common alleles. The frequencies varied among the nine groups, where C6*A frequency changed from 58% in the Atayal to 18% in the Tsou. A variant allele, C6*A_{taiwanese}, with a slightly lower isoelectric point than C6*A, was observed in the Rukai and Paiwan at polymorphic frequencies of 3.7% and 4.2%, respectively.

Table 1. Allele frequencies of ten loci for serum proteins in nine Aboriginal Taiwanese populations

Alleles	Populations								
	Atayal	Saisiat	Tsou	Bunun	Ami	Puyuma	Rukai	Paiwan	Yami
AHSG*1	0.9800	0.6641	0.9444	0.8678	0.8056	0.8095	0.9167	0.8000	0.9167
AHSG*2	0.0200	0.3359	0.0556	0.1322	0.1944	0.1905	0.0833	0.2000	0.0833
C3*F07 _{tsou}	—	—	0.1000	0.0057	—	—	—	—	—
C3*F06 _{atayal}	0.0050	—	—	—	—	—	—	—	—
C3*S	0.9950	1.0000	0.9000	0.9943	1.0000	1.0000	1.0000	1.0000	1.0000
C6*A	0.5765	0.4766	0.1750	0.2442	0.3889	0.4206	0.5185	0.4250	0.2115
C6*B	0.4184	0.4766	0.8125	0.7558	0.5625	0.5476	0.3981	0.4833	0.6603
C6*B2	0.0051	0.0469	0.0125	—	0.0486	0.0317	0.0463	0.0500	0.1282
C6*A _{taiwanese}	—	—	—	—	—	—	0.0370	0.0417	—
C7*1	0.8776	0.9453	0.7500	0.9244	0.8889	0.8889	0.9352	0.9417	0.8462
C7*2	0.0816	0.0078	0.2438	0.0640	0.0903	0.0873	0.0370	0.0083	0.0833
C7*3 (= C7*5)	0.0102	—	—	—	0.0139	0.0238	0.0093	—	0.0449
C7*4	0.0306	0.0469	0.0063	0.0116	0.0069	—	0.0185	0.0500	0.0256
GC*1F	0.6020	0.6250	0.6420	0.6207	0.5972	0.4444	0.5092	0.6083	0.6987
GC*1S	0.2041	0.2344	0.3025	0.1782	0.2847	0.3413	0.2593	0.2833	0.1667
GC*2	0.1939	0.0938	0.0556	0.1897	0.1111	0.2143	0.2315	0.1083	0.1346
GC*1A8 _{saisiat}	—	0.0469	—	—	—	—	—	—	—
GC*1C _{ami}	—	—	—	—	0.0069	—	—	—	—
GC*2C _{bunun}	—	—	—	0.0115	—	—	—	—	—
HP*1	0.2980	0.2266	0.5000	0.3430	0.4155	0.3810	0.2778	0.3814	0.6667
HP*2	0.7020	0.7734	0.5000	0.6570	0.5845	0.6190	0.7222	0.6186	0.3333
ITIH1*1	0.8900	0.7143	0.5449	0.6524	0.8125	0.7823	0.7593	0.7288	0.7051
ITIH1*2	0.1100	0.2857	0.4551	0.3476	0.1875	0.2177	0.2407	0.2712	0.2949
PI*M1	0.9200	0.8359	0.7716	0.6782	0.8264	0.8016	0.8333	0.7417	0.8846
PI*M2	0.0700	0.0938	0.2160	0.2529	0.1528	0.1349	0.0556	0.1750	0.1154
PI*M3	0.0100	0.0703	0.0123	0.0690	0.0208	0.0635	0.1111	0.0833	—
TF*C1	0.7105	0.8871	0.8264	0.8736	0.8819	0.8730	0.8241	0.8750	0.7500
TF*C2	0.2895	0.1129	0.1736	0.1264	0.1181	0.0952	0.1759	0.1167	0.2500
TF*D _{china}	—	—	—	—	—	0.0317	—	0.0083	—
ZAG*1	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

In the C7 system, the Atayal, Ami, Rukai and Yami had one predominant allele, C7*1 and three minor alleles, C7*2, C7*3 (= C7*5) and C7*4 as seen in other East Asian populations including Japanese and Han Chinese (Roychoudhury and Nei, 1988). However, the Saisiat, Tsou, Bunun and Paiwan were devoid of C7*3 and the Puyuma were lacking in C7*4.

The GC system is determined by three common alleles, GC*1F, GC*1S and GC*2 and many rare variant alleles. The distribution of allele frequencies has been studied in detail all over the world, as reviewed by Kamboh and Ferrell (1986). The Atayal in Toroko had a GC*2 frequency of 17% (Chen et al., 1985). Similarly, the aboriginal Taiwanese were shown to have relatively low GC*2 frequencies (23% in the Rukai to 6% in the Tsou). The frequencies of GC*1F were higher than those in the Japanese, Korean and Han Chinese (Kamboh and Ferrell, 1986). No variants alleles such as GC*1A2, GC*1A3, GC*1A8, GC*1A9 and GC*1C2, characteristic of East Asian and native American populations, were observed in the aboriginal Taiwanese (Constans et al., 1985). A new variant allele, GC*1A8_{saisiat}, located closely but very slightly anodally to GC*1A8, was observed in the Saisiat at a polymorphic frequency of 4.7%. Two other rare variants, GC*1C_{ami} and GC*2C_{bunun}, were found in the Ami and Bunun groups, respectively.

The HP system was controlled by two common alleles, HP*1 and HP*2. The frequency of HP*1 in the nine aboriginal groups varied from 67% in the Yami to 23% in the Saisiat. The frequency in the Atayal group in this study (30%) was similar to that in the Atayal from Toroko (26%) (Chen et al., 1985), which was comparable to the data in most East Asian populations (Roychoudhury and Nei, 1988).

At the ITIH1 locus, three common alleles, ITIH1*1, ITIH1*2 and ITIH1*3 have been observed together with a few variant alleles in several East Asian populations examined so far. ITIH1*1 and ITIH1*2 were observed at frequency of about 50%. However, ITIH1*1 was predominant over the other alleles in Korean, Han Chinese, and Thai. In contrast, Japanese had higher ITIH1*2 frequencies (Vogt et al., 1991; Yuasa et al., 1991; Hou et al., 1993; Harada et al., 1994). In the aboriginal Taiwanese, only ITIH1*1 and ITIH1*2 were observed without variant alleles. The ITIH1*1 was predominant and its frequencies varied from 89% in the Atayal to 55% in the Tsou.

The PI system is determined by three major alleles, PI*M1, PI*M2 and PI*M3 in East Asians, as summarized by Kamboh (1985). The aboriginal Taiwanese had fairly high PI*M1 frequencies, which were comparable to those in Japanese, Korean, Han Chinese and Thai (Kamboh, 1985). PI*M3 was exceptionally predominant over PI*M2 in the Rukai. The Yami were lacking in PI*M3. No variants were detected in any aboriginal groups, while some variant alleles have been observed in East Asian populations: PI*E_{tokyo} was widespread in Japanese and Northern Han Chinese (Yuasa et al., 1984; Ying et al., 1985).

In the TF system, TF*C1 and TF*C2 were found, and TF*C1 was predominant over TF*C2 in East Asian populations, as reviewed by Kamboh and Ferrell (1987). Similarly, TF*C1 was a major allele in the aboriginal Taiwanese. TF*D_{china}, characteristic of East Asian and native American populations, was also observed in the Puyuma and Paiwan. It was reported that the Atayal in Toroko had no TF*D_{china}

Table 2. Genetic distances D_s and D_c among nine Aboriginal Taiwanese populations

Populations	Atayal	Saisiat	Tsou	Bunun	Ami	Puyuma	Rukai	Paiwan	Yami
Atayal	0.2154	0.0198	0.0290	0.0159	0.0136	0.0123	0.0140	0.0143	0.0283
Saisiat	0.0401	0.2255	0.0299	0.0205	0.0131	0.0128	0.0145	0.0134	0.0299
Tsou	0.0531	0.0535	0.2488	0.0141	0.0144	0.0182	0.0234	0.0182	0.0187
Bunun	0.0305	0.0365	0.0282	0.2281	0.0123	0.0088	0.0166	0.0125	0.0213
Ami	0.0348	0.0326	0.0279	0.0263	0.2486	0.0048	0.0108	0.0066	0.0169
Puyuma	0.0375	0.0334	0.0336	0.0210	0.0120	0.2531	0.0059	0.0058	0.0198
Rukai	0.0384	0.0312	0.0392	0.0314	0.0226	0.0157	0.2389	0.0097	0.0226
Paiwan	0.0457	0.0317	0.0356	0.0295	0.0160	0.0167	0.0188	0.2593	0.0235
Yami	0.0436	0.0499	0.0459	0.0429	0.0406	0.0448	0.0441	0.0560	0.2196

Figures above and below the diagonal are D_s and D_c , respectively, and those on the diagonal are average heterozygosity.

Allele frequency data for the 13 red cell enzyme and 14 serum protein loci were used.

(Chen et al., 1985).

The ZAG is known to be almost monomorphic in Japanese, Korean, Han Chinese, Thai and Filipino except for the Papuan. Several rare variants, however, were observed in some East Asian populations (Nakayashiki et al., 1992). No aboriginal groups showed any variation at the ZAG locus.

Genetic distance analysis among nine aboriginal Taiwanese

Average heterozygosities (gene diversities) and genetic distances (D_s and D_c) were calculated using a total of 27 systems including 14 serum protein systems investigated in the present and previous studies (Umetsu et al., 1994, 1995) and 13 red cell enzyme systems reported previously (Jin et al., 1995, 1999). The average heterozygosities are shown on the diagonal in Table 2. The values ranged from 0.215 in the Atayal to 0.259 in the Paiwan. All the nine aboriginal groups except for the Paiwan showed rather small within-population variations in comparison with German (0.254) and Japanese (0.274).

The genetic distances D_s and D_c between each pair of the nine aboriginal Taiwanese groups calculated using the allele frequencies for the 27 loci are given above and below the diagonal in Table 2, respectively. In the D_s distance, the largest distance was obtained between the Saisiat and Tsou or Yami, and relatively high values were observed between the Atayal and Tsou or Yami. The smallest distance was obtained between the Puyuma and Ami, and relatively low values were observed between the Puyuma and Paiwan or Rukai. In the D_c distance, the largest distance was obtained between the Paiwan and Yami, and relatively high values were observed between the Tsou and Atayal or Saisiat. The smallest distance was obtained between the Puyuma

and Ami, and relatively low values were observed between Puyuma and Rukai and between Paiwan and Ami. A distant relationship between Tsou and Saisiat or Atayal, and a close relationship between Puyuma and Ami were confirmed by two genetic distances.

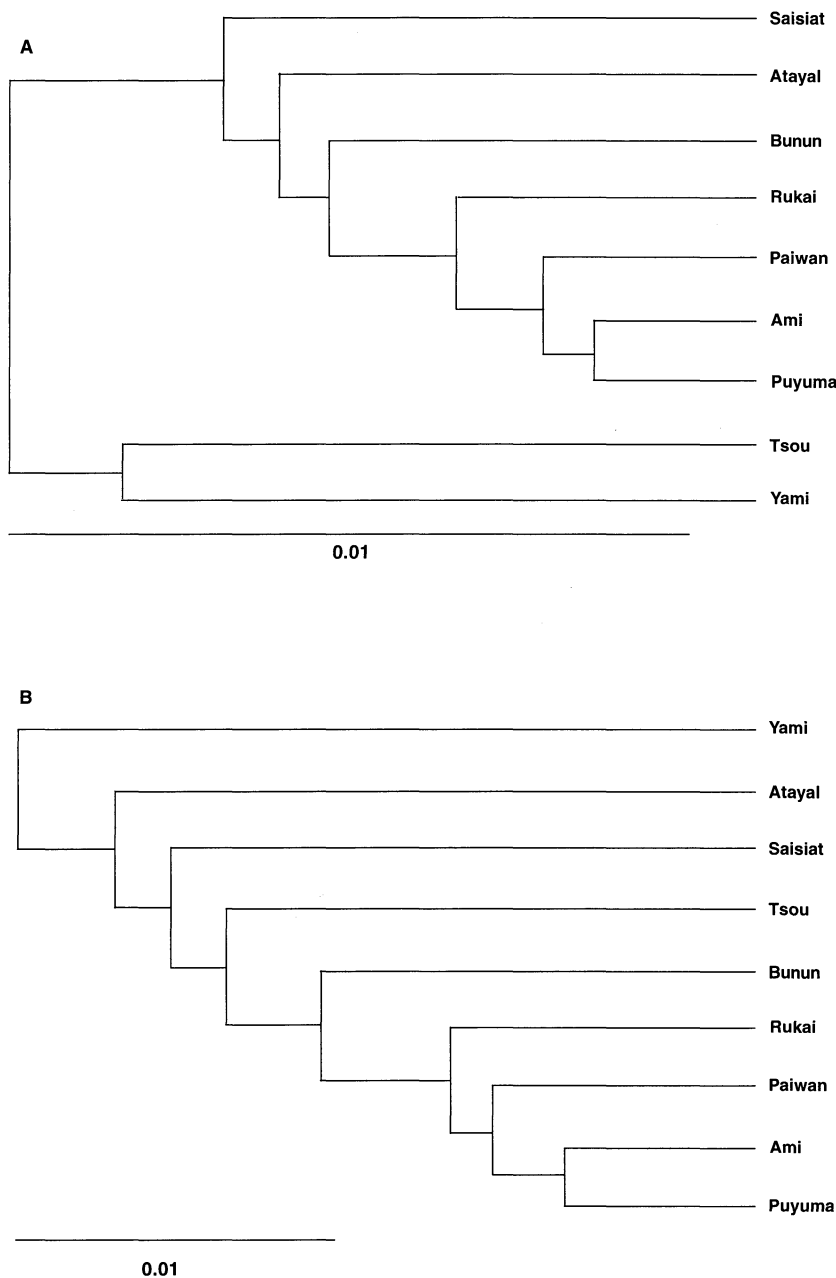


Figure 1. UPGMA trees for the nine aboriginal groups of Taiwan based on Ds (A) and Dc (B) distance matrix from Table 2.

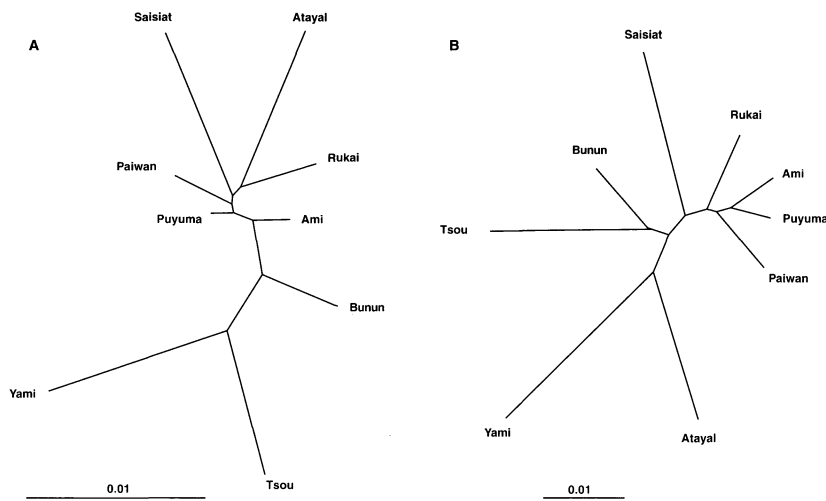


Figure 2. NJ trees for the nine aboriginal groups of Taiwan based on Ds (A) and Dc (B) distance matrix from Table 2.

The UPGMA trees constructed by using the Ds and Dc distances are shown in Fig. 1. The clustering pattern (Fig. 1A) based on Ds distances was somewhat different from that (Fig. 1B) based on the Dc distances. On the Ds-UPGMA tree, the nine aboriginal groups were divided into two major clusters. The Ami and Puyuma were the most closely related to each other, followed by the Paiwan and Rukai. These four groups were closely distributed in the southeastern part of Taiwan. The Bunun, Saisiat and Atayal, living in the central and northern parts of Taiwan, were also included in this cluster. The Yami and Tsou formed another cluster, though the Yami lives on Lanyu Island and the Tsou settles in central region of Taiwan. However, the branch between them was long. The Dc-UPGMA tree confirmed the close relationship among the Puyuma, Ami, Paiwan and Rukai. The Bunun, Saisiat and Atayal were clustered with these four groups. The Tsou and Yami did not form another loose cluster and the Yami was the most remote from the other groups. These two groups were notably different in clustering patterns between the two UPGMA trees.

The NJ trees constructed by using the Ds and Dc distances are shown in Fig. 2. The clustering pattern (Fig. 2A) based on Ds distances was fairly similar to that (Fig. 2B) based on Dc distances. The Puyuma, Ami, Paiwan and Rukai appeared to be in close proximity to each other on the two NJ trees. These results were similar to those observed on the UPGMA trees. The Yami, Tsou and Bunun were loosely clustered and were distant from the Saisiat. The branches of the Atayal, Saisiat, Tsou and Yami were always long. The branching points of the Atayal were remarkably different between the two NJ trees.

Comparison of nine aboriginal Taiwanese populations with eight other populations

As shown in Table 3, average heterozygosities and genetic distances (Ds and Dc) were calculated among 17 human populations including the nine aboriginal Taiwanese groups and eight other human populations using allele frequency data for 15 red cell enzyme and serum protein loci. The average heterozygosities ranged from 0.268 in Atayal to 0.339 in Paiwan in the nine aboriginal Taiwanese groups, while they ranged from 0.324 in German to 0.352 in southern Chinese among the other eight populations. Small within-population variations in the aboriginal groups were confirmed in comparison with those in the other populations, though a limited number of loci were used.

The Ds and Dc distances were also similar to the results obtained using 27 loci. The Puyuma showed a strong affinity with Ami, followed by the Paiwan and Rukai. The largest distances were observed between the Yami and Saisiat in the Ds distance and between the Yami and Paiwan in the Dc distance. In both distances the Yami, Atayal, Saisiat and Tsou apparently showed the fairly large values for some East Asian populations, while the other groups showed the small values for some Asian populations, especially for Filipino. The values were smaller than those between some aboriginal Taiwanese groups. German among the eight populations had the largest values for each aboriginal Taiwanese group.

The UPGMA trees constructed by using the Ds and Dc distances are shown in Fig. 3. The clustering pattern (Fig. 3A) based on Ds distances was fairly different from that (Fig. 3B) based on Dc distances. On the Ds-UPGMA tree, the 17 populations were divided into two major clusters. One consisted of the nine aboriginal groups, but Filipino was included in this cluster with a strong affinity with the Puyuma and Ami. The other consisted of all East Asian populations except for the Filipino. German was also included in this cluster.

On the Dc-UPGMA tree, German formed the most external branch. The Yami was located at the position of the out-group to all the remaining 15 populations, which were divided into two major clusters. One consisted of the eight aboriginal Taiwanese groups, in which the Puyuma and Paiwan showed the strongest affinity and the Saisiat was apart from the others.

The NJ trees constructed by using the Ds and Dc distances are shown in Fig. 4. The nine aboriginal Taiwanese groups formed a large cluster in the both trees. The clustering patterns were somewhat different, but the differences were similar to those observed for 27 loci. The nine aboriginal Taiwanese groups showed a close affinity with Southeast Asian populations including Filipinos, Thai and Southern Chinese, but were distant from Northeast Asian populations. Some aboriginal Taiwanese groups were apparently close to Filipino than other aboriginal Taiwanese groups.

Table 3. Genetic distances D_s and D_c among Aboriginal Taiwanese, German and East Asian populations

Populations	1)	2)	3)	4)	5)	6)	7)	8)	9)	10)	11)	12)	13)	14)	15)	16)	17)
1) Atayal	0.2682	0.0262	0.0255	0.0126	0.0215	0.0171	0.0196	0.0221	0.0338	0.0816	0.0403	0.0334	0.0399	0.0439	0.0279	0.0343	0.0315
2) Saisiat	0.0290	0.2784	0.0352	0.0298	0.0199	0.0222	0.0267	0.0222	0.0468	0.0778	0.0460	0.0425	0.0433	0.0481	0.0424	0.0272	0.0260
3) Tsou	0.0297	0.0348	0.2985	0.0194	0.0126	0.0175	0.0178	0.0176	0.0190	0.0595	0.0454	0.0354	0.0456	0.0484	0.0281	0.0369	0.0167
4) Bunun	0.0196	0.0289	0.0201	0.2980	0.0166	0.0115	0.0209	0.0171	0.0375	0.0613	0.0302	0.0206	0.0333	0.0356	0.0195	0.0274	0.0152
5) Ami	0.0312	0.0277	0.0147	0.0204	0.3191	0.0076	0.0156	0.0113	0.0211	0.0435	0.0236	0.0189	0.0217	0.0251	0.0160	0.0170	0.0079
6) Puyuma	0.0321	0.0281	0.0198	0.0157	0.0108	0.3308	0.0096	0.0089	0.0322	0.0455	0.0251	0.0164	0.0253	0.0287	0.0140	0.0152	0.0099
7) Rukai	0.0297	0.0290	0.0188	0.0228	0.0179	0.0119	0.3092	0.0156	0.0332	0.0688	0.0390	0.0257	0.0394	0.0421	0.0234	0.0242	0.0212
8) Paiwan	0.0363	0.0297	0.0170	0.0214	0.0104	0.0093	0.0168	0.3387	0.0335	0.0723	0.0349	0.0281	0.0374	0.0418	0.0264	0.0248	0.0177
9) Yami	0.0275	0.0356	0.0234	0.0332	0.0303	0.0371	0.0325	0.0405	0.2774	0.0766	0.0463	0.0442	0.0476	0.0501	0.0360	0.0483	0.0283
10) German	0.0840	0.0758	0.0578	0.0546	0.0513	0.0511	0.0680	0.0662	0.0774	0.3241	0.0325	0.0331	0.0249	0.0232	0.0304	0.0325	0.0312
11) Japanese	0.0495	0.0429	0.0362	0.0267	0.0267	0.0235	0.0382	0.0297	0.0448	0.0415	0.3400	0.0042	0.0022	0.0039	0.0089	0.0107	0.0186
12) Ryukyuan	0.0483	0.0452	0.0334	0.0233	0.0254	0.0201	0.0332	0.0290	0.0454	0.0417	0.0043	0.3469	0.0066	0.0078	0.0042	0.0078	0.0126
13) Korean	0.0504	0.0444	0.0390	0.0292	0.0252	0.0240	0.0408	0.0316	0.0486	0.0357	0.0036	0.0066	0.3324	0.0012	0.0098	0.0099	0.0195
14) Mongolian	0.0505	0.0465	0.0383	0.0283	0.0255	0.0249	0.0394	0.0329	0.0488	0.0304	0.0064	0.0095	0.0038	0.3309	0.0105	0.0113	0.0212
15) S-Chinese	0.0426	0.0432	0.0283	0.0205	0.0216	0.0175	0.0304	0.0263	0.0400	0.0355	0.0083	0.0054	0.0071	0.0077	0.3519	0.0071	0.0090
16) Thai	0.0428	0.0305	0.0292	0.0240	0.0199	0.0180	0.0277	0.0232	0.0408	0.0363	0.0107	0.0118	0.0108	0.0105	0.0083	0.3454	0.0099
17) Filipino	0.0400	0.0266	0.0198	0.0166	0.0185	0.0164	0.0265	0.0208	0.0303	0.0342	0.0157	0.0148	0.0172	0.0160	0.0102	0.0071	0.3451

Figures above and below the diagonal are D_s and D_c , respectively, and those on the diagonal are average heterozygosity.

Allele frequency data for 7 red cell enzyme (AK, ACP, ADA, ESD, GPT, PGD, and PGM) and 8 serum protein (AHSG, GC, HP, IF, ORM1, ORM2, PI and TF) loci were used.

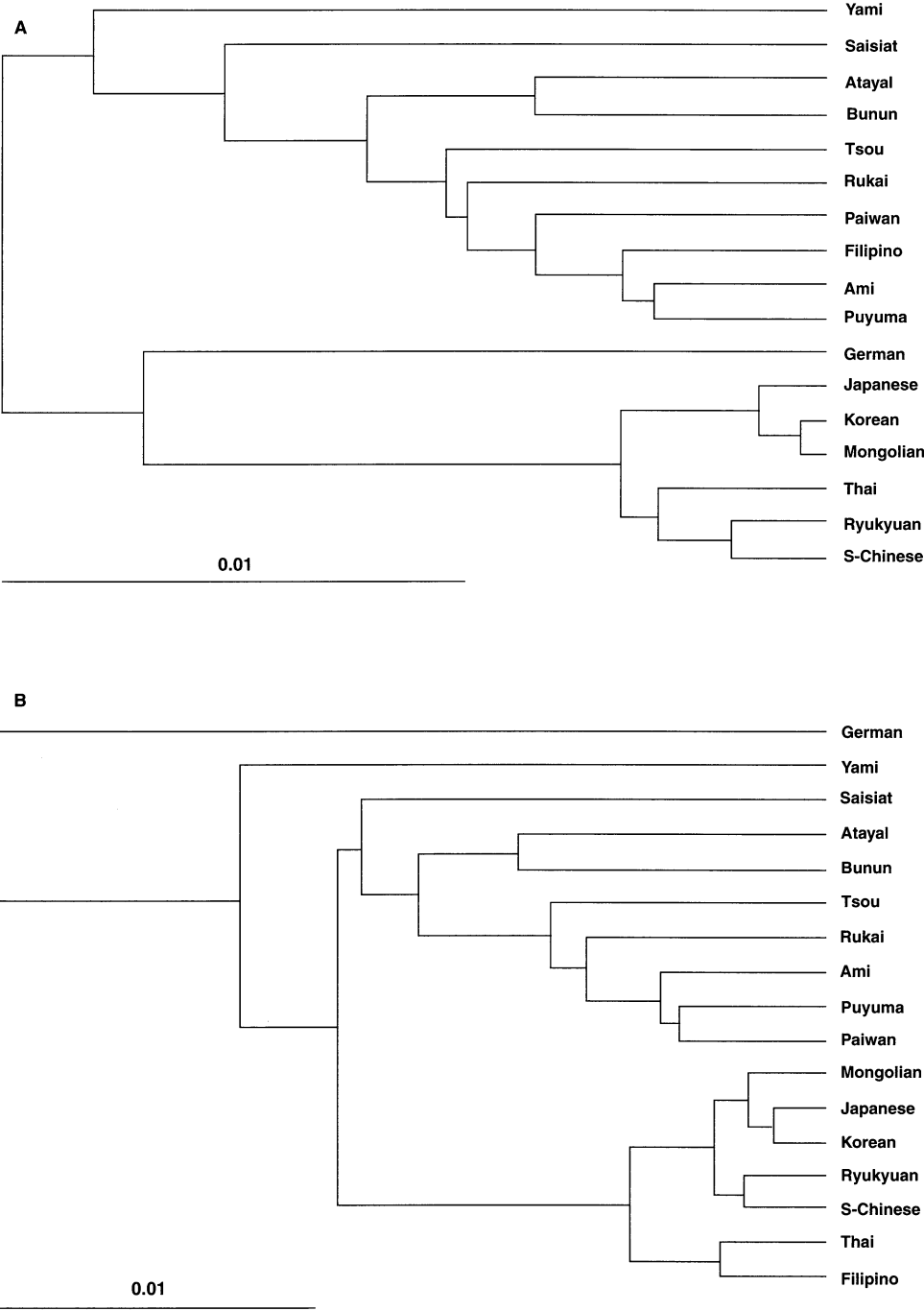


Figure 3. UPGMA trees for the nine aboriginal groups of Taiwan and eight other populations based on Ds (A) and Dc (B) distance matrix from Table 3.

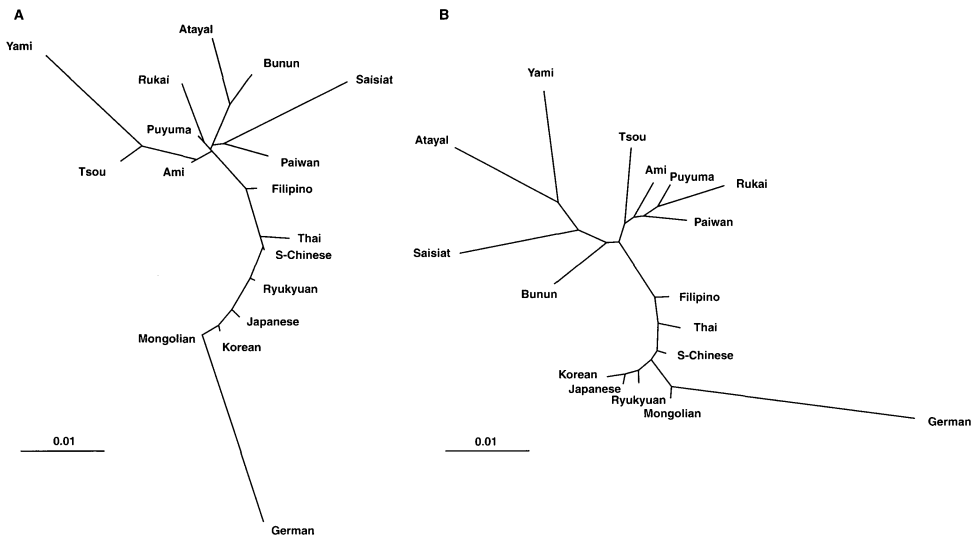


Figure 4. NJ trees for the nine aboriginal groups of Taiwan and eight other populations based on Ds (A) and Dc (B) distance matrix from Table 3.

Discussion

In the early investigations of red cell enzyme and serum protein systems, we observed a significant degree of genetic diversity among the nine groups of the aboriginal Taiwanese (Jin et al., 1995, 1999; Umetsu et al., 1994, 1995). In this study almost all aboriginal Taiwanese groups showed rather small average heterozygosity in comparison with German and Japanese. However, the large diversity among them was confirmed by the present study, in which allele frequencies varied from group to group at as many as nine loci. In particular, the AHSG, C6 and HP systems were conspicuous for their diversity. The HLA study also showed that the nine aboriginal Taiwanese groups were highly homogeneous within each group, but diversified among them (Lin et al., 2000). In the early studies 11 variant alleles including ADA*Taiwan1, ADA*Taiwan2, ESD*7, GPT1*6, PGM1*6, C1R*9, IF*B2, ORM1*B9, ORM1*Q0, ORM2*H2 and ORM2*H19 were observed (Jin et al., 1995, 1999; Umetsu et al., 1994, 1995). In this study seven variant alleles including C3*F07_{tsou}, C3*F06_{atayal}, C6*A_{taiwanese}, GC1A8_{saisiat}, GC*1C_{ami}, GC*2C_{bunun}, and TF*D_{china}, were observed at four of the ten loci. Some of them were new variants specific for the aboriginal Taiwanese and were polymorphic with a frequency of more than 1%. ORM1*Q0 and ORM2*H19 were relatively widespread in the aboriginal Taiwanese, but the other variant alleles were restricted to only a few groups. It is of interest that the Yami group living on Lanyu Island had no such variant alleles. These findings

suggest the low level of gene flow among the nine aboriginal groups. The diversity must also have come from the effect of founder principle and the bottleneck effect.

The tree analysis confirmed the large genetic differentiation among the nine aboriginal Taiwanese groups. In particular, the Atayal, Saisiat, Tsou and Yami always had a long branch and were far apart from each other, though the branching patterns were different from tree to tree. In contrast, the Puyuma, Ami, Paiwan and Rukai were always clustered and related to each other. The Bunun were often included in this group. The clustering patterns were fitted with their geographical distribution because they are neighboring with each other at southeastern parts of Taiwan. Though the Saisiat had a long branch, they were also clustered with a Puyuma-Ami-Paiwan-Rukai cluster on the NJ trees. The results described here, however, were different from those of the HLA study (Lin et al., 2000). It showed that the Rukai and Paiwan were the most closely related to each other, followed by the Tsou and Bunun. These four groups were further clustered with the Atayal and Saisiat. The Puyuma, Ami and Yami became more remote from them.

The languages of the aboriginal Taiwanese are classified into four groups of the Austronesian language: the Atayal and Yami speak Atayalic and Malayo-Polynesian, respectively. Tsou and Rukai use Tsouic. The remaining five groups speak Paiwanic (Ruhlen 1987; see Jin et al., 1999). Thus, genetic relationships based on red cell enzyme and serum protein polymorphisms were fairly identical to the linguistic classification. The long branch of the Atayal, Tsou and Yami on the trees are consistent with the linguistic classification, because each of the three groups has used their own peculiar language diversified by a long isolation. Paiwanic is spoken by the Puyuma, Ami, Paiwan, Bunun and Saisiat, who also formed a large cluster on the NJ trees. However, the Rukai, who use Tsouic, were clustered with a Puyuma-Ami-Paiwan group, but not with Tsou. The Rukai may have genetically been influenced by the Bunun, Puyuma and Paiwan, because they are surrounded with the Bunun, Puyuma and Paiwan, but separated from the Tsou. As described above, the Rukai and Paiwan shared C6*A_{taiwanese} at relatively high frequencies.

According to the 'express-train to Polynesia' model (Bellwood 1991; Diamond 1988, 1997), the homeland of the aboriginal Taiwanese was south China. A study of mtDNA and nuclear DNA variation in the Atayal, Ami, Puyuma, and Bunun suggested that they had temporally deep roots, probably, in central or south China (Melton et al., 1998). The different waves of migration must have brought ancestral populations of aboriginal people to the main island of Taiwan. It is noteworthy that the Atayal and Yami were clustered together with long branches on the Dc-NJ tree. Chen et al. (1985) cited a reference of Hsu and Pan (1945) and described that the Atayal group living in Toroko were believed to be the first of all aboriginal Taiwanese who moved to Taiwan. If it is correct, the Atayal and Yami may be direct descendants of the first people, though they live most distantly from each other

among the nine aboriginal Taiwanese groups.

Chen et al. (1985) studied the Atayal group living in Toroko using 19 red cell antigen, serum protein, and red cell enzyme polymorphisms and concluded that the Atayal group had a strong genetic affinity with populations from the Philippines and Thailand. Melton et al. (1998) described that the aboriginal Taiwanese most resembled populations from the Philippines, but this was the result of migration from Taiwan south to the Philippines. In addition, they deduced that the Ami among the four aboriginal groups investigated was most probable emigrants. Linguistic and archaeological evidence agrees that Taiwan is the most likely Austronesian homeland (Bellwood 1991; Diamond 1988, 1997; Gray and Jordan, 2000). In this study the aboriginal Taiwanese showed genetic affinity with Filipino and to less extent with Thai and southern Chinese. In accordance with the results of Melton et al. (1998), the Ami and Puyuma were most closely related to Filipino. Nevertheless, these results were inconsistent with the linguistic classification. The language of Filipino as well as the Yami belongs to the Malayo-Polynesian, but not to the Paiwanic (Ruhlen, 1987). mtDNA studies have suggested that Taiwanese spread throughout Polynesia by way of Philippines and East Indonesia (Melton et al., 1995; Sykes et al., 1995; Richards et al., 1998). However, the HLA study of all nine aboriginal Taiwanese showed that they were close to Papua New Guinea Highlanders and Javanese in phylogenetic analysis (Lin et al., 2000). A study of the Y chromosome markers suggested that Southeast Asia had provided a genetic source for two independent migrations, one toward Taiwan and the other toward Polynesia through island Southeast (Su et al., 2000).

Thus, there are two major anthropogenetic interests in the aboriginal Taiwanese: their genetic origin and colonization to Polynesia. The aboriginal Taiwanese are a key population to clarify prehistoric Austronesian dispersals. To solve such problems we are now investigating all nine aboriginal Taiwanese groups for the blood group antigens, mtDNA and Y chromosome.

Acknowledgments

We wish to thank Dr. Hiroaki Nishimukai, Department of Legal Medicine, Ehime University School of Medicine, for reference typing of the C3 variants found in this study. We are also grateful to Mr. Minoru Endoh and Ms. Miwa Asahi for their excellent technical clerical assistance. This study was supported by Grants-in-Aid for the International Research Program ("Human Genetic Studies on Native Minority Groups in Taiwan") to S. H. and by Grants-in-Aid for Scientific Research on the Priority Areas ("Prehistoric Mongoloid Dispersals" and "The Origin of the Japanese and Japanese Culture") to I.Y., K.U., N.S. and S.H. from the Ministry of Education, Science, and Culture of Japan.

References

- Bellwood P. (1991) The Austronesian dispersal and the origin of languages. *Scientific American*, 265: 70–75.
- Cavalli-Sforza L. L. and Edwards A.W.F. (1967) Phylogenetic analysis: models and estimation procedures. *American Journal of Physical Anthropology*, 19: 233–257.
- Chen K. H., Cann H., Chen T. C., Van West B., and Cavalli-Sforza L. (1985) Genetic markers of an aboriginal Taiwanese population. *American Journal of Physical Anthropology*, 66: 327–337.
- Constans J., Hazout S., Garruto R.M., Gajdusek D.C., and Spees E.K. (1985) Population distribution of the human vitamin D binding protein: anthropological considerations. *American Journal of Physical Anthropology*, 68: 107–122.
- Diamond J. M. (1988) Express train to Polynesia. *Nature*, 336: 307–308.
- Diamond J. M. (1997) *Guns, Germs, and Steel*. W.W. Norton & Company, New York.
- Du R. and Yip V. F. (1993) *Ethnic groups in China*. Science Press, Beijing.
- Felsenstein J. (1993) PHYLIP: phylogeny inference package. Version 3.5c. University of Washington, Seattle.
- Fong J.M.H. (1974) The distribution of the polymorphic groups of blood, serum protein, and red cell enzyme among three endogamous groups of Takasago. *Bulletin of Osaka Medical School*, 20: 35–45.
- Gray R. D. and Jordan FM (2000) Language trees support the express-train sequence of Austronesian expansion. *Nature*, 405: 1052–1055.
- Harada A., Umetsu K., and Suzuki T. (1994) Inter-alpha-trypsin inhibitor polymorphism: An improved phenotyping procedure and two new alleles. *International Journal of Legal Medicine*, 107: 25–28.
- Hou Y.P., Gou Q., and Wu M.Y. (1993) A study of the genetic polymorphism of human inter-alpha-trypsin-inhibitor (ITI) in the Han population, Chengdu, China. *Human Genetics*, 90: 661–662.
- Hsu T. W. and Pan G. C. (1945) cited by Chen et al. (1985).
- 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 variant detected by isoelectric focusing method. *Anthropological Science*, 103: 227–234.
- Jin F., Saitou N., Ishida T., Sun C.-S., Pan I.-H., Omoto K., and Horai S. (1999) Population genetic studies on nine aboriginal ethnic groups of Taiwan. I. Red cell enzyme systems. *Anthropological Science*, 107: 229–246.
- Kamboh M. I. (1985) Biochemical and genetic aspects of human serum α_1 -proteinase inhibitor protein. *Disease Markers*, 3: 135–154.
- Kamboh M. I. and Ferrell R. E. (1986) Ethnic variation in vitamin D-binding protein (GC): a review of isoelectric focusing studies in human populations. *Human Genetics*, 72: 281–293.
- Kamboh M. I. and Ferrell R. E. (1987) Human transferrin polymorphism. *Human Heredity*, 37: 65–81.
- Lin M., Chu C. C., Lee H. L., Chang S. L., Ohashi J., Tokunaga K., Akaza T., and Juji T. (2000) Heterogeneity of Taiwan's indigenous population: possible relation to prehistoric Mongoloid dispersals. *Tissue Antigens*, 55: 1–9.
- Matsumoto H. (1988) Characterization of Mongoloid and neighboring populations based on the genetic markers of human immunoglobulins. *Human Genetics*, 80: 207–218.
- Melton T., Peterson R., Redd A. J., Saha N., Sofro A.S.M., Martinson J., and Stoneking M. (1995) Polynesian genetic affinities with Southeast Asian populations as identified by mtDNA analysis.

- American Journal of Human Genetics, 57: 403–414.
- Melton T., Clifford S., Martinson J., Batzer M., and Stoneking M. (1998) Genetic evidence for the proto-Austronesian homeland in Asia: mtDNA and nuclear DNA variation in Taiwanese aboriginal Tribes. *American Journal of Human Genetics* 63: 1807–1823.
- Nakajima H. and Ohkura K. (1971) The distribution of several serological and biochemical traits in East Asia. III. The distributions of gamma globulin (Gm[1], Gm[2], Gm[5] and Inv[1]) and Gc groups in Taiwan and Ryukyu. *Human Heredity*, 21: 362–370.
- Nakayashiki N., Umetsu K., Yuasa I., Suenaga K., Omoto K., Ishida T., Misawa S., and Katsura S. (1992) Five new alleles of plasma Zn-alpha 2-glycoprotein variants phenotyped by isoelectric focusing and immunoblotting in twelve populations. *Electrophoresis*, 13: 154–157.
- Nei M. (1987) *Molecular Evolution Genetics*. Columbia University Press, New York.
- Nishimukai H., Kitamura H., Sano Y., and Tamaki Y. (1985) C3 variants in Japanese. *Human Heredity*, 35: 69–72.
- Page R.D.M. (1996) TREEVIEW: An application to display phylogenetic trees on personal computers. *Computer Applications in the Biosciences*, 12: 357–358.
- Richards M., Oppenheimer S., and Sykes B. (1998) mtDNA suggests Polynesian origins in Eastern Indonesia. *American Journal of Human Genetics*, 63: 1234–1236.
- Roychoudhury A. K. and Nei M. (1988) *Human polymorphic genes: World distribution*. Oxford University Press, New York.
- Ruhlen M. (1987) *A Guide to the World's languages*, 1: Classification. Stanford University Press, Stanford.
- Saitou N. and Nei M. (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution*, 41: 406–425.
- Sokal R. R. and Sneath P.H.A. (1963) *Principles of Numerical Taxonomy*. Freeman, San Francisco.
- Su B., Jin L., Underhill P., Martinson J., Saha N., McGarvey S. T., Shriver M. D., Chu J., Oefner P., Chakraborty R., and Deka R. (2000) Polynesian origins: insights from the Y chromosome. *Proceedings of the National Academy of Sciences of the United States of America*, 97: 8225–8228.
- Sykes B., Leibo A., Low-Beer J., Tetzner S., and Richards M. (1995) The origins of the Polynesians: an interpretation from mitochondrial lineage analysis. *American Journal of Human Genetics*, 57: 1463–1475.
- Umetsu K., Yuasa I., Harada A., Suzuki T., Pan I.-H., Ishida T., Saitou N., and Horai S. (1995) Orosomucoid phenotyping with monoclonal antibodies: polymorphic occurrence of ORM1*Q0 in aboriginal Taiwanese populations. *Human Heredity*, 45: 181–185.
- Umetsu K., Yuasa I., Suzuki T., Sun C.-H., Pan I.-H., Ishida T., Saitou N., and Horai S. (1994) Polymorphisms of complement component I and C1R subcomponent of C1 in nine aboriginal Taiwanese populations. *Human Biology*, 66: 339–348.
- Vogt U., Cleve H., Farhud D.D., and Goedde H.W. (1991) The ITI system in South Koreans and Iranians analyzed by an improved classification procedure. Distribution of alleles and description of new phenotypes. *Human Genetics*, 87: 677–679.
- Ying Q., Zhang M., Liang C., Liu X., Huang Y., Wang R., Zhang N., Chen L., Chen L., Yu N., and Muo X. (1985) Geographical variability of alpha-1-antitrypsin alleles in China: a study on six Chinese populations. *Human Genetics*, 69: 184–187.
- Yuasa I., Suenaga K., Gotoh Y., Ito K., Yokoyama N., and Okada K. (1984) PI (α_1 -antitrypsin) polymorphism in the Japanese: confirmation of PI*M4 and description of new PI variants. *Human Genetics*, 67: 209–212.
- Yuasa I., Suenaga K., Saneshige Y., Tamaki N., Ito K., and Okada K. (1991) Inter-alpha-trypsin

inhibitor (ITI): a useful genetic system in paternity testing. Evidence for polymorphic occurrence of ITI*3 and existence of a new allele, ITI*4. *International Journal of Legal Medicine*, 104: 197–199.

Yuasa I. and Umetsu K. (1988) Genetic polymorphism of human α_2 HS-glycoprotein: characterization and application to forensic hemogenetics. *Electrophoresis*, 9: 404–410.

Editor-in-Charge: Yuji Mizoguchi