

M. Lin
C.-C. Chu
S.-L. Chang
H.-L. Lee
J.-H. Loo
T. Akaza
T. Juji
J. Ohashi
K. Tokunaga

The origin of Minnan and Hakka, the so-called “Taiwanese”, inferred by HLA study

Key words:

anthropology; HLA class I and class II genes; HLA haplotypes; Minnan and Hakka; “Taiwanese”

Acknowledgments:

We would like to thank Drs. L.C. Yu and Y.C. Tsai of the Research Department, Mackay Memorial Hospital for Helping us to collect blood samples from Hakka. This work was supported by a grant from the National Health Research Institute of Taiwan (DOH 87-HR-601).

Abstract: The Minnan and Hakka people groups, the so-called “Taiwanese”, are the descendants of early settlers from the southeast coast of China during the last few centuries. Genetically they showed affinities to southern Asian populations, as determined by phylogenetic trees and correspondence analysis calculated from HLA allele frequencies. This corresponds historically with the fact that they are the descendants of the southeast coastal indigenous population (Yueh) of China and should therefore not be considered as descendants of “pure” northern Han Chinese. A33-B58-DRB1*03 (A33-Cw10-B58-DRB1*03-DQB1*02), the most common HLA haplotype among “Taiwanese”, with a haplotype frequency of 6.3%, has also been found to be the most common haplotype among Thai-Chinese and Singapore Chinese, two other populations also originating from the southeast coast of China. These observations suggests that this haplotype is the most well-conserved ancient haplotype of the Yueh.

“Taiwanese”, the major population group in Taiwan, are comprised of the Minnan and Hakka peoples which constitute 73.5% and 17.5%, respectively, of the total population. They are descendants of early settlers from the southeast coast (Fuchien and Kwangton) of China during the past 400 years or more in recent history. Many “Taiwanese” intermarried with the preexisting indigenous tribes after arrival, mainly with the plains tribes but also with the mountain tribes. In our previous study, we found that 13% of “Taiwanese” HLA-A, -B and -C three-locus haplotypes most likely originated from these mountain tribes and also from the Pazeh, who are a disappearing plains tribe. This suggests that only a small proportion of indigenous genes are present in the “Taiwanese” gene pool, although HLA data from the already extinct plains tribes (9 tribes) are not available, and so the degree of contribution of these tribes to the “Taiwanese” gene pool is at present unknown (1). This is in contrast to the traditional account of “Taiwanese” origins handed down through successive generation, either orally or

Authors' affiliations:

M. Lin^{1,2},
C.-C. Chu¹,
S.-L. Chang²,
H.-L. Lee¹,
J.-H. Loo¹,
T. Akaza³,
T. Juji³,
J. Ohashi⁴,
K. Tokunaga⁴

¹Transfusion Medicine
Research Laboratory,
Mackay Memorial Hospital,
Taipei, Taiwan,

²Immunohematology
Reference Laboratory,
Mackay Memorial Hospital,
Taipei, Taiwan,

³Japanese Red Cross
Central Blood Center, Tokyo,
Japan,

⁴Department of Human
Genetics, Graduate School
of Medicine, University of
Tokyo, Tokyo, Japan

Correspondence to:

Marie Lin, M.D.
Transfusion Medicine
Research Laboratory
Mackay Memorial Hospital
45, Ming-Shen Rd.
Tan Sui
Taipei Hsien
Taiwan
Fax: +886 2 25433638
e-mail:
marilin@ms2.mmh.org.tw

Received 26 June 2000, revised,
accepted for publication 1 February 2001

Copyright © Munksgaard 2001
Tissue Antigens. ISSN 0001-2815

Tissue Antigens 2001; 57: 192–199
Printed in Denmark. All rights reserved

through written family pedigrees. In this account, "Taiwanese" have been told that their ancestors originated from the Central Plains of North China but migrated to southeast coastal area sometime after the Han Dynasty, during the invasion of the north by northern pastoral nomads. Hence they are assumed to be descendants of "pure" northern Han Chinese from the Central Plains and thus belong to the great tradition of Han (Hwa-Shia).

In this study, we analyzed the HLA data of both Minnan and Hakka by constructing phylogenetic trees and plotting correspondence analysis, and also by tracing the most common HLA haplotype seen in Minnan and Hakka, and comparing the results with data from other populations. In this way we hope to clarify the truth about the origin of "Taiwanese".

Material and methods

During 1997 and 1998 a total number of 123 families were enrolled for paternity testing at Mackay Memorial Hospital. Among these families, 167 unrelated persons consisting of either the parents of inclusion trios or the mother of exclusion trios, having parents who were either Minnan or Hakka (136 individuals had Minnan parents, 24 had Hakka parents and 7 had one Minnan and one Hakka parent) were studied in this report. Some of these individuals had been included in our previous HLA class I study (1).

Blood samples were tested for HLA class I antigens by a serological method (standard microlymphocytotoxicity test) using Terasaki Chinese HLA-A, -B and -C 72-well Trays (lot 2, 3, 3A, 3B) and the latest Terasaki Special Monoclonal Tray-Asian HLA class I (lot 3). All samples were also tested with a local typing tray for HLA class I antigens. HLA class II DNA typing was performed by using One Lambda Micro SSP Generic HLA Class II DNA Typing Trays (lot 2, 3) of medium-level resolution.

Blood samples were also obtained from 75 Hakka individuals. These individuals were unrelated and had parents who were both Hakka. All samples were submitted as panel cells at the 1998 Japanese Red Cross Central Block Histocompatibility Workshop where 420 antisera collected and specified by the participating labs from Japan, Korea, Thailand, US, South Africa and Taiwan, were used to test for HLA-A, -B and -C antigens. HLA class II DNA typing was also performed. All blood samples were collected in ACD tubes and all testing was performed at the Transfusion Medicine Research Laboratory, Mackay Memorial Hospital.

The HLA-A, B, C, DRB1 and DQB1 gene frequencies of 136 Minnan and 99 Hakka (75+24) individuals were estimated by the maxi-

mum likelihood method developed for the 11th International Histocompatibility Workshop (11th IHW) (2). The A-Cw-B, A-B-DRB1 and A-Cw-B-DRB1-DQB1 haplotype frequencies of these 167 individuals were calculated by direct counting of the haplotypes of parents (inclusion trios) and mothers (exclusion trios) of these families.

The genetic distances D (Nei's standard genetic distance, 1972) (3), and D_A (Nei, 1983) (4) between populations were calculated by using "ODEN" (public domain software) according to the HLA-A, B, C gene frequencies of this report and data reported at the 11th IHW (5). The studied populations are shown in the map of East Asia (Fig. 1). Using genetic distances, phylogenetic trees were constructed by the neighbor-joining (NJ) method (6), the tree using genetic distances D is shown in Fig. 2. The correspondence analysis performed on many populations based on HLA-A, -B and -C gene frequencies (5) by Vista (free software, <http://forrest.psych.unc.edu>) is shown in Fig. 3. These populations included the Minnan and Hakka of this report and many other populations from data reported at the 11th IHW (5).

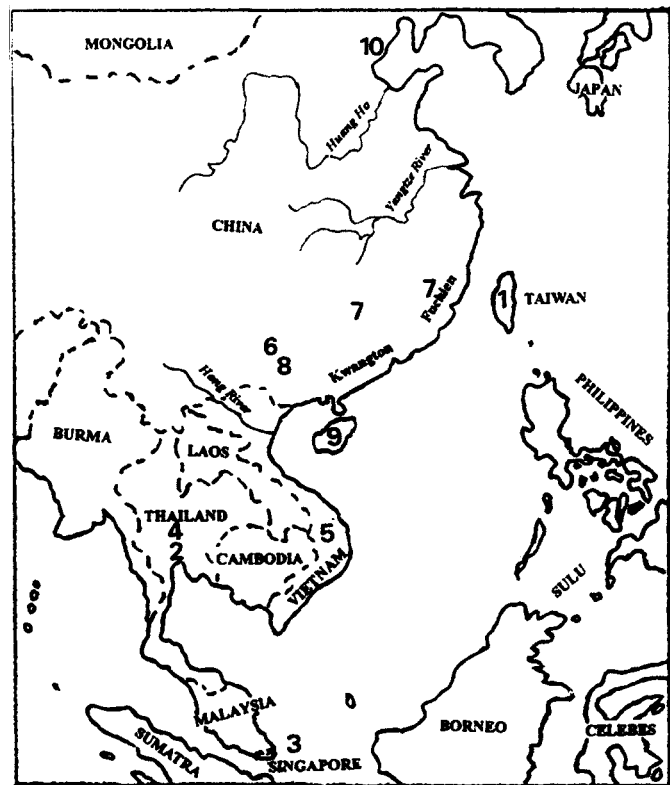


Fig. 1. Map of East Asia. The studied populations: 1, Minnan and Hakka ("Taiwanese"); 2, Thai-Chinese; 3, Singapore Chinese; 4, Thais; 5, Vietnamese; 6, Buyi (Kueichou); 7, Southern Han (Fuchien and Hunan); 8, Miao (Kueichou); 9, Li (Hainan); 10, Northern Han (Beijing and Heilongchian).

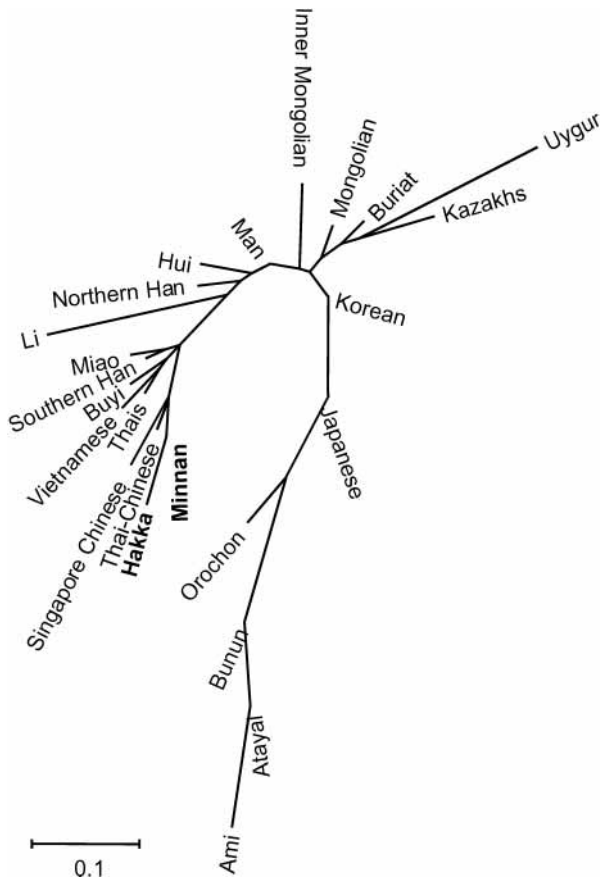


Fig. 2. Neighbor-joining tree of 24 populations in Asia constructed by D (Nei's standard genetic distance, 1972). "Taiwanese" (Minnan and Hakka) are merged together and cluster with 2 other populations (Thai-Chinese and Singapore Chinese) who also originated from the southeast coast of China. Except for Li, all southern Asians are clustered together and are separated from northern Asians including northern Han. Taiwan's indigenous groups (Bunun, Atayal and Ami) are at the far end of the tree.

Results

The HLA-A, B, C, DRB1 and DQB1 gene frequencies of 136 Minnan and 99 Hakka are shown in Table 1. There are significant differences in A2 and B27 gene frequencies ($P < 0.05$) between Minnan and Hakka, most likely due to small samples used here since the frequencies were similar in a large samples study (7). The frequency of Cw8 gene is higher and blank for Cw gene is lower among Hakka, which could be due to 420 antisera were used to type for HLA class I genes for Hakka in HLA workshop in this study. A2, A1101, B60, B46, B58, B13, Cw10, Cw7, Cw1, DRB1*04, DRB1*09, DQB1*06, DQB1*0301 and DQB1*05 showed the high gene frequencies in these two ethnic groups. Analysis of HLA haplotypes for the A-Cw-B, A-B-DR and A-Cw-B-DR-DQ loci of "Taiwanese" (Minnan and

Hakka) demonstrated 126 haplotypes for the three loci A-Cw-B, 212 for the three loci A-B-DR and 247 for five loci A-Cw-B-DR-DQ. Only haplotypes with a haplotype count of six or more for three locus haplotypes (Table 2) and haplotype counts of four or more for five locus haplotypes are shown (Table 3). From Table 2 it can be seen that the five most common haplotypes for the A-Cw-B loci are A33-Cw10-B58 (10.8%), A2-Cw1-B46 (7.8%), A1101-Cw7-B60 (5.7%), A1101-Cw10-B13 (3.9%) and A1101-Cw1-B46 (3.2%). The five most common haplotypes for the A-B-DR loci are A33-B58-DRB1*03 (6.3%), A2-B46-DRB1*09 (3.0%), A33-B58-DRB1*13 (3.0%), A1101-B13-DRB1*15 (1.8%) and A1101-B75-DRB1*12 (1.8%). The three most common haplotypes for the five loci A-Cw-B-DR-DQ, as shown in Table 3, are A33-Cw10-B58-DRB1*03-DQB1*02 (6.3%), A33-Cw10-B58-DRB1*13-DQB1*06 (3%) and A2-Cw1-B46-DRB1*09-DQB1*0303 (2.7%).

The phylogenetic tree constructed from D, as shown in Fig. 2, reveals that Minnan and Hakka merge together and cluster with Thai-Chinese and Singapore Chinese. This group also forms a cluster with the neighboring groups of Thais, Vietnamese and Buyi, and another group of Southern Han and Miao, thus forming a southern Asian cluster. Li is a separate entity, while Northern Han, Hui, Man, Mongolians, Buriat, Uygur, Kazakhs, Korean, Japanese and Orochon form a northern Asian cluster. The phylogenetic tree constructed by D_A revealed that Minnan and Hakka clustered together with Southern Han, Buyi and Miao, then with Thai-Chinese and Singapore Chinese, followed by Thais and Vietnamese, again forming a southern Asian cluster. Northern Han clustered with Li, and with the northern Asian populations, Hui, Buriat, Uygur, Kazakhs, Man and Mongolians.

Correspondence analysis of Minnan and Hakka, and many other populations as shown in Fig. 3 also showed that Minnan and Hakka clustered together with other southern Asian populations including southern Han, Singapore Chinese and Thai-Chinese. Northern Han formed a cluster with Koreans as well as Man and Hui populations. Taiwan's indigenous groups are clustered together and showed an affinity to Highlanders.

Discussion

Because of the high degree of the polymorphism of HLA system, it is a useful genetic marker for the characterization of human populations and analysis of their relationships for anthropological purposes. Differences in the distribution of HLA alleles among various human populations are more marked when compared to other genetic markers. This becomes even more obvious if HLA haplotypes

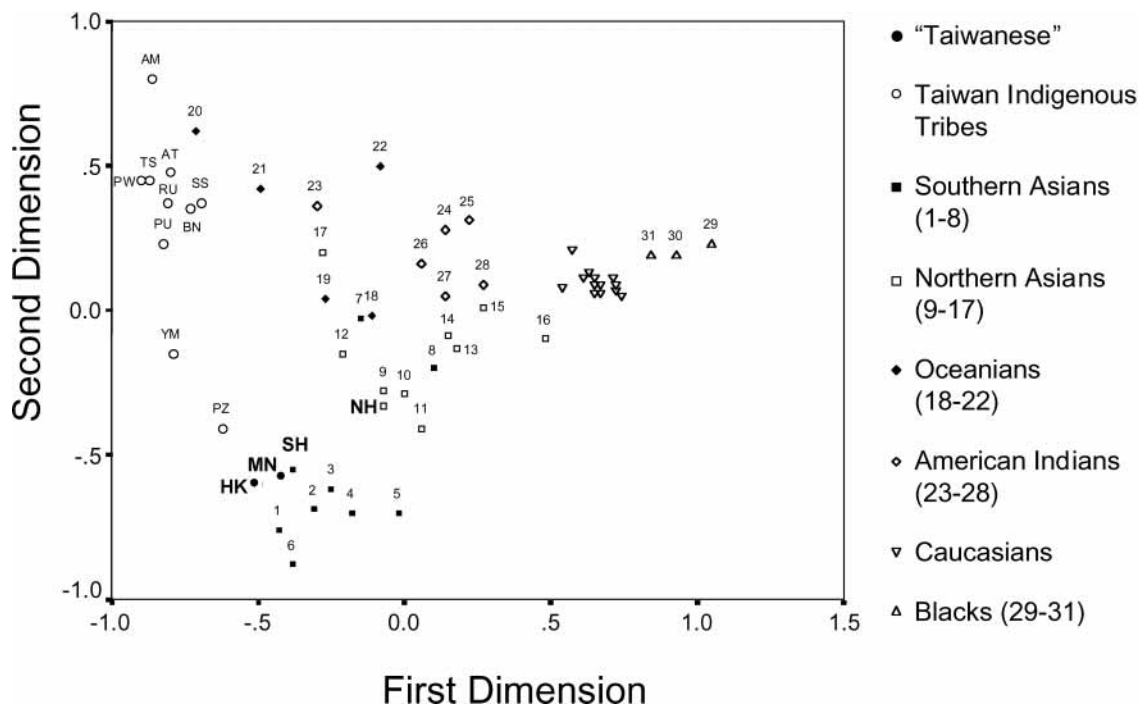


Fig. 3. Correspondence analysis of “Taiwanese” and other ethnic groups. Minnan and Hakka are closely related to the southern Han, and are clustered with other southern Asian populations. Northern Han are separated from the southern Asian cluster, and form a cluster with other northern Asian populations. Populations: MN Minnan; HK Hakka; SH Southern Han; NH Northern Han; AM Ami; AT Atayal; BN Bunun; PU Puyuma; PW Puyuma; RU Rukai; SS Saisiat; TS Tsou; YM Yami; 1. Miao; 2. Singapore Chinese; 3. Thai-Chinese; 4. Thais; 5. Vietnamese; 6. Buyi; 7. Tibetan; 8. Li; 9.

Korean; 10. Man; 11. Hui; 12. Japanese; 13. Buriat; 14. Mongolian; 15. Kazakhs; 16. Uygur; 17. Orochon; 18. Javanese; 19. Timorese; 20. Highlanders; 21. Maori; 22. Australian Aborigine; 23. Inuit; 24. Chilean and Colombian (South Amerinds); 27. Yakut; 28. Brazilian (Amerinds); 29. South African (Black); 30. North American (Blacks); 31. Brazilian (Blacks); Caucasians: British, Dane, French, German, Greek, Italian, Portuguese, Spanish, USA (Whites), Canadian, Mexican, Brazilian (Whites).

(particular combinations of alleles) are used as an index. It is interesting that many haplotypes were found to have a unique organization of HLA genes that have been well-conserved through thousands of years and also each characteristic haplotype shows a limited regional distribution. Therefore the HLA haplotype is a powerful marker and is useful for surveys among closely related ethnic groups (8). Family studies have been found to be the best method for studying multilocus HLA haplotype distribution. Although there are several reports concerning HLA haplotype distribution in Taiwan, this study is the first report based on family study in “Taiwanese” (Minnan and Hakka).

A33-B58-DRB1*03 was found to be the most common and also the best conserved A-B-DR three-locus haplotype among “Taiwanese” (6.3%), and was exclusively related to the most common five-locus haplotype, A33-Cw10-B58-DRB1*03-DQB1*02 (6.3%). Similar to our results, the haplotype A33-B58-DR17 (DRB1*03) was also proven to be the most common haplotype in Taiwan among Minnan ($n=7137$, 5.59%) and Hakka ($n=714$, 5.10%) by the Tzu Chi Taiwan Marrow Donor Registry as estimated by the maximum likelihood

method (7). The three-locus haplotype A33-B58-DRB1*03 has also been found among southeast Asians (Thai-Chinese 7.1%; Singapore Chinese 4.4%) (5) and northern Asians (Han in Urumchi, China 3.5%; Khalha in Mongolia 3.4%; Uygur in Urumchi, China 1.8%; Kazakhs in Urumchi, China 1.8%; Korean 1.2%) (9, 10). The five-locus haplotype has also been found in Thais (Present-day Thais 2.2%; Black Thai 2.8%) (11). It is interesting to note that the HLA-A, B two-locus haplotype A33-B58 among southern Han ($n=138$, samples collected in Hunan and Fuchian) was reported to be only 2.4%, with no significant A-B-DR three-locus haplotype frequency at the 11 IHW (5), while in an earlier report the A33-B17 haplotype frequency among southern Han ($n=844$) was 1.53%, and for southern minorities ($n=621$) was 1.29% (12). In another report, the haplotype A33-Cw10-B58 was found among northern Han ($n=196$, 3.1%), however, the most frequently associated DR gene with this haplotype in northeast Asians is more likely DR13 than DR3 (13), and the distribution of haplotype B58-DR3 was found clearly focused in southeast Asian population (8). Therefore, although three- to five-locus haplotype of A33-Cw10-B58-DRB1*03-DQB1*02 has been

HLA-A, B, C, DRB1 and DQB1 gene frequencies of Minnan and Hakka**Table 1**

HLA-A	Minnan n=136	Hakka n=99	HLA-C	Minnan	Hakka
A2	33.0	18.4	Cw6	1.8	1.5
A24	13.9	16.6	Cw7	20.0	22.1
A26	2.9	5.9	Cw8	1.1	6.5
A1101	29.7	37.0	Cw9	8.0	4.8
A1102	6.3	5.1	Cw10	22.5	30.5
A30	1.1	1.0	Cw14		1.6
A31	2.5	2.0	CBL	25.0	9.2
A33	9.2	11.0	HLA-DR		
ABL	0.4	1.0	DRB1*01	0.7	0.5
A ^a	1.2	3.0	DRB1*15	10.4	13.1
HLA-B			DRB1*16	7.0	4.0
B51	5.1	2.5	DRB1*03	7.4	8.6
B13	8.5	10.1	DRB1*04	14.7	14.6
B62	5.8	5.0	DRB1*11	5.1	6.1
B75	5.8	4.5	DRB1*12	13.8	9.1
B38	5.1	3.0	DRB1*13	3.7	6.6
B39	1.8	3.5	DRB1*14	6.1	9.1
B58	9.6	12.0	DRB1*07	1.2	2.0
B54	3.3	3.0	DRB1*08	9.3	9.6
B55	4.0	2.2	DRB1*09	16.1	15.2
B27	5.1	0.5	DRB1*10	1.1	1.5
B35	2.9	2.5	DR-	3.4	-
B60	22.0	22.5	HLA-DQ		
B61	2.2	2.5	DQB1*05	18.3	18.7
B46	11.6	13.5	DQB1*06	18.3	23.2
B48	1.5	0.5	DQB1*02	8.8	10.6
BBL	1.1	3.0	DQB1*0302 ^c	19.4	16.2
B ^b	4.5	9	DQB1*0302 ^d	7.4	7.6
HLA-C			DQB1*0303 ^e	15.7	15.7
Cw1	17.6	20.9	DQB1*04	8.9	8.1
Cw4	4.1	2.5	DQ-	3.3	-

A^a: A29, A30, A32, A68B^b: B7, B8, B18, B44, B5102, B52, B56, B5603, B67, B71, B76* 0301^c: 0311-012/04/09-10. * 0302^d: 0302/05/07-08. * 0303^e: 03032-033/06

found to be distributed among all east and southeast Asians, however, the highest frequencies have been found in “Taiwanese”, Singapore Chinese and Thai-Chinese, and these three populations are presumed to be the descendants of early settlers from the southeast coast of China. Therefore, this haplotype is probably the most well-conserved haplotype among these three populations.

The second most common three locus haplotype among “Taiwanese” in this study A2-B46-DRB1*09 (3.0%) has also been found

in Singapore Chinese (7.7%); Vietnamese (5.2%); southern Han (5.1%); Thais (4.7%) and Thai-Chinese (2.4%) (5). The corresponding five-locus haplotype A2-Cw1-B46-DRB1*09-DQB1*0303 (2.7%) has also been found in Black Thai (8.3%), Dai Lui (Thais, 5.1%), and present-day Thais (2.1%) (11). This haplotype is thus characteristic for southern Asians (8).

It is well known from the results of HLA studies that there are genetic difference between southern and northern Han Chinese (7,

HLA-A, B, C three-locus haplotypes in "Taiwanese"

Haplotype (n=334)	n	%
A33-Cw10-B58	34	10.2
A2-Cw1-B46	26	7.8
A1101-Cw7-B60	19	5.7
A1101-Cw10-B13	13	3.9
A1101-Cw1-B46	11	3.3
A2-Cw7-B38	9	2.7
A1101-CBL-B75	8	2.4
A24-Cw7-B60	8	2.4
A2-Cw7-B60	7	2.1
A1102-CBL-B27	6	1.8
A24-Cw1-B54	6	1.8
A2-CBL-B51	6	1.8
A2-Cw10-B60	6	1.8

Table 2**HLA-A, B, C, DR, DQ five-locus haplotypes in "Taiwanese"**

Haplotype (n=334)	n	%
A33-Cw10-B58-DRB1*03-DQB1*02	21	6.3
A33-Cw10-B58-DRB1*13-DQB1*06	10	3.0
A2-Cw1-B46-DRB1*09-DQB1*0303	9	2.7 ^A
A1101-Cw10-B13-DRB1*15-DQB1*06	5	1.5
A1101-CBL-B75-DRB1*12-DQB1*0301	4	1.2
A1101-Cw1-B46-DRB1*09-DQB1*0303	4	1.2
A1101-Cw7-B60-DRB1*08-DQB1*06	4	1.2

^A A2-B46-DRB1*09 is the second most common haplotype among "Taiwanese" with HLA-A, B, DR three-locus haplotypes

Table 3

8, 14, 15). This corresponds well with the prehistory and history of China. The Central Plains (Chung Yuan) culture developed in the loess region of the Huang Ho (Yellow River) basin in the north of China, and was dated to sixth or even seventh millennium BC from excavated millets (16). This culture went through the Shia civilization, followed by the Shang and Chou Dynasties and was limited to the north (north of Yangtze River). Only in the Chin Dynasty (221–206 B.C.), there was political control over the central region of the south gained (17). Recent new archeological discoveries showed that in the south of China there was an independent Yueh coastal culture (from the Yangtze delta to the Hong river delta in North Vietnam) in existence almost at the same period of time (18). However, the history of the south only began around 500 B.C. just before the battles between two countries of Yueh and Wu (17). There is a well

known story about the King of the Yueh who slept on firewood for mattress and had a gall hung over his bed to remind him of the bitterness of his defeat by Wu which finally enabled him to win the battle, and resulted in national recovery. The Yueh people lived along southeast coast of China (Chechiang, Fuchien, Kwangtung and Kwangsi), and were named the Hundred (Pai) Yueh before the Han Dynasty (206 B.C.–220 A.D.) due to the great diversity of local cultures. Not very much Yueh history has been available from Chinese history (of the Central Plains) apart from battles between Yueh and Wu during Spring and Autumn Warring States Period, and the dispersion of part of the Yueh peoples to the north during the Han Dynasty, since in Chinese history populations other than those of the Central Plains have been considered as "barbarians" (17, 19). The Minnan (Min) were one of the ethnic groups among the Yueh who lived in Fuchien, and according to Lin (17) and Meacham (19) the present-day Minnan are descendants of indigenous Minnan peoples although probably limited gene flow from the northern Han occurred during the Chin and Han Dynasties and also during the following several centuries, due to invasions of nomads from Mongolia in the north which caused internal migration from north to south (17). As the barbarian status of the Yueh gradually disappeared and they were finally given Han status in history, thus probably resulting in misinterpretation and erroneous self-assertion of present-day Minnan as "pure" descendants of the northern Han. In Chinese history, many ethnic minorities adopted Han culture, and many peoples from these ethnic groups often announced that they were Han, most likely because the Han culture was more dominant at that time and so being a member of a Han ethnic group was both beneficial and a source of pride in the past (17, 19, 20). This also imply to Hakka, as during South Song Dynasty (1127–1279 A.D.) or even earlier there were limited immigrants from Central Plains to south east region of China. These peoples with their dominant Han culture influenced culturally and linguistically the indigenous Yueh peoples especially those who live in Kwangton. Subsequently few initial immigrants and the vast majority of indigenous Yueh peoples forming the Hakka ethnic group.

Two phylogenetic trees constructed by calculation of the genetic distances D and D_A showed a slight difference in the order of clustering among southern Asians. The D tree appears to better fit into the history of "Taiwanese", Thai-Chinese and Singapore Chinese, since they originated from the same area of the southeast coast of China but migrated to different places during the last few centuries. However, the trees constructed both by D and D_A separated the southern Asian populations, including "Taiwanese", from northern Han. Correspondence analysis of "Taiwanese" and many other populations also showed that there was a clear separation between the "Taiwanese" and northern Han. A phylogenetic tree constructed

by HLA gene frequencies on marrow donors also showed that “Taiwanese” (Minnan and Hakka) clustering together with the southern Han are separate from the northern Han (8). Many other studies including HLA as well as other genetic markers such as immunoglobulins (22), blood groups (23), glucose-6-phosphate dehydrogenase (24) and microsatellites (25) also clearly separate the southern Han from the northern Han. A study on Chinese surnames and genetic differences between north and south China also separated southern Han from the northern Han (20). In that study, the southern Han cluster can be further subdivide into three subclusters: the lower Yangtze river centering around Shanghai; most of the Yangtze river basin; and the southeast coastal areas and islands off the coast (including Taiwan). Chinese surnames have a history of probably more than 4,000 years and are transmitted via the male line and can thus be considered as a Y-chromosome gene. The finding of southeast coastal subdivisions of the southern Han corresponds

well with the history and populations of the Yueh. Also in that study, the map of the first two principal coordinates calculated by the gene frequencies for ABO, MN, and Rh(D) of 28 provinces of China plus Taiwan, it was revealed that there was a clear split between the populations of Taiwan and the southeast coastal provinces from all other parts of China, suggesting that “Taiwanese” and its originating southeast coastal populations (Yueh peoples) might be genetically distinct from the others.

These genetic data indicate that the southern Han are basically of southern origin and remain genetically distinct from the northern Han. “Taiwanese” who are descendants of the ancient Yueh peoples preserve the ancestral HLA haplotype A33-Cw10-B58-DRB1*03-DQB1*02 of the Yueh. The genetic distance between “Taiwanese” and southern Han warrant further study because half of the population samples contributing to the HLA data of southern Han used in this study were from Fuchien (5).

References

1. Lin M, Chu LL, Lee HL et al. Heterogeneity of Taiwan's indigenous population: possible relation to prehistoric Mongoloid dispersals. *Tissue Antigens* 2000; **55**: 1–9.
2. Imanishi T, Akaza T, Kimura A et al. Allele and haplotype frequencies for HLA and complement loci in various ethnic groups. In: Tsuji K, Aizawa M, Sasazuki T, eds. *HLA 1991. Proceedings of the 11th International Histocompatibility Workshop and Conference*. Vol 1. Oxford: Oxford University Press, 1992: 1067–74, 1141–9.
3. Nei M. Genetic variation within species. In: *Molecular evolutionary genetics*. New York: Columbia University Press, 1987: 220–1.
4. Nei M, Tajima F, Tatenos Y. Accuracy of estimated phylogenetic trees from molecular data. II. Gene frequency data. *J Mol Evol* 1983; **19**: 153–70.
5. Imanishi T, Akaza T, Kimura A et al. Estimation of allele and haplotype frequencies for HLA and complement loci. In: Tsuji K, Aizawa M, Sasazuki T, eds. *HLA 1991. Proceedings of the 11th International Histocompatibility Workshop and Conference*. Vol 1. Oxford: Oxford University Press, 1992: 76–9.
6. Satiou N, Nei M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* 1987; **4**: 406–25.
7. Shaw CK, Chen LL, Lee A, Lee TD. Distribution of HLA gene and haplotype frequencies in Taiwan: a comparative study among Min-nan, Hakka, aborigines and mainland Chinese. *Tissue Antigens* 1999; **53**: 51–64.
8. Tokunaga K, Imanishi T, Takahashi K, Juji T. On the origin and dispersal of East Asian populations as viewed from HLA haplotypes. In: Akazawa T, Szathmary EJ, eds. *Prehistoric Mongoloid Dispersals*. Oxford: Oxford University Press, 1996: 187–97.
9. Tanaka H, Tokunaga K, Inoko H et al. Distribution of HLA-A, B, and DRB1 alleles and haplotypes in Northeast Asia. In: Charron D, ed. *Proceedings of the 12th International Histocompatibility Workshop and Conference*. Vol 1. Paris: EDK, 1997: 285–91.
10. Park MH, Hwang YS, Park KS et al. HLA haplotypes in Koreans based on 107 families. *Tissue Antigens* 1998; **51**: 347–55.
11. Chandanayingyong D, Bejrachandra S, Kunachiwa W et al. HLA in the Thai population. *Proceedings of 10th Regional Congress of the International Society of Blood Transfusion Western Pacific Region*. Taipei, 1999: 102–9.
12. Lee TD, Zhao TM, Mickey K et al. The polymorphism of HLA antigens in the Chinese. *Tissue Antigens* 1988; **12**: 188–208.
13. Inoue T, Ogawa A, Tokunaga K et al. Diversity of HLA-B17 alleles and haplotypes in East Asians and a novel Cw6 allele (Cw*0604) associated with B*5701. *Tissue Antigens* 1999; **53**: 534–44.
14. Chen RB, Zhao TM, Ye YG et al. Joint report on Mainland Chinese HLA polymorphism. In: Aizawa M, ed. *HLA in Asia-Oceania: Proceedings of the 3rd Asia-Oceania Histocompatibility Workshop and Conference*. 1986: 224–30.
15. Chen RB, Ye GY, Geng ZC et al. HLA polymorphism of the principal minority nationalities in mainland China. In: Tsuji K, Aizawa M, Sasazuki T, eds. *HLA 1991. Proceedings of the 11th International Histocompatibility Workshop conference*. Vol 1. Oxford: Oxford University Press, 1992: 676–9.

16. An Chih-min, Pei-li-kang, Tzu-shan, ho Yang-shao-shih lun Chung-Yuan hsin shih-Chi wen-hua ti Yuan-Yuan Chi fa-chan. *Kao-ku* 1979; **4**: 335–46.
17. Lin Hui-Shiang. *Chuoan Kuo ming Chu shi*. Taipei: Taiwan Commercial Press, 1936.
18. Chang KC. *The archaeology of ancient China*. 3rd edition New Haven: Yale University Press, 1977.
19. Meacham W. Origins and development of the Yueh coastal Neolithic: A microcosm of culture change of the mainland of East Asia. In: Keightly DK, ed. *The origins of Chinese civilization*. Berkeley, CA: University of California Press, 1981.
20. Du R, Yuan Y, Hwang J et al. Chinese surnames and the genetic differences between north and south China. Monograph series 5. *J Chinese Linguistics* 1992.
21. Wu Song-Ti. Hakka Nan Song yuen liou suoh. Soh Hwei Ko Shueh Pang. *Fu Tan Shueh Pau* 1995, No. 5.
22. Matsumoto H. Characteristics of Mongoloid and neighboring populations based on the genetic markers of human immunoglobulins. *Hum Genet* 1988; **80**: 207–18.
23. Lin M, Broadberry RE. Immunohematology in Taiwan. *Transfu Med Rev* 1998; **12**: 56–72.
24. Chu JY. Glucose-6-phosphate dehydrogenase kang Taiwan chu chung te shue yen. *Taiwan Med J* 1999; **42**: 252–6.
25. Chu JY, Huang W, Kuang SQ et al. Genetic relationship of populations in China. *Proc Natl Acad Sci U S A* 1998; **95**: 11763–8.