

# Y-chromosomal STRs haplotypes in the Taiwanese Paiwan population

Fang-Chin Wu · Chin-Wen Ho · Chang-En Pu ·  
Kuang-Yu Hu · Sascha Willuweit · Lutz Roewer ·  
David Hwang Liu

Received: 20 November 2008 / Accepted: 22 December 2009 / Published online: 28 January 2010  
© The Author(s) 2010. This article is published with open access at Springerlink.com

**Abstract** The distribution of Y-chromosomal short tandem repeat (Y-STR) haplotypes was determined in a population of Taiwanese Paiwan aboriginals. Using 17 Y-STR markers, a total of 135 haplotypes were observed, 102 of which were unique. The overall haplotype diversity for the 17 Y-STR loci tested was 0.9922 and the discrimination capacity was 0.6490. In addition, three novel intermediate alleles at the DYS448 locus were also found.

**Electronic supplementary material** The online version of this article (doi:10.1007/s00414-009-0416-x) contains supplementary material, which is available to authorized users.

F.-C. Wu · C.-E. Pu · D. H. Liu  
Department of Forensic Science, Investigation Bureau,  
Ministry of Justice,  
Taipei County, Taiwan, Republic of China

F.-C. Wu · C.-W. Ho  
Department of Bioengineering, Tatung University,  
Taipei City, Taiwan, Republic of China

K.-Y. Hu  
Department of Biochemistry, National Defense Medical Center,  
Taipei City, Taiwan, Republic of China

S. Willuweit · L. Roewer  
Department of Forensic Genetics, Institute of Legal Medicine and  
Forensic Sciences, Charité—Universitätsmedizin Berlin,  
Berlin, Germany

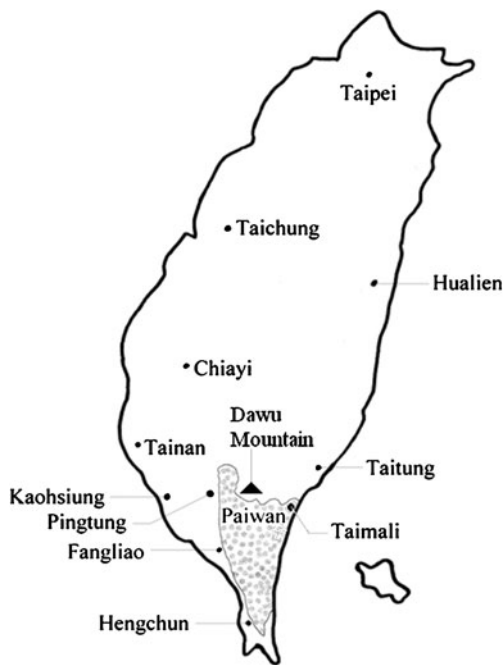
F.-C. Wu (✉)  
No. 74, Chung-Hua Road, Xindian City,  
Taipei County, 23149 Taiwan, Republic of China  
e-mail: wfcank@ms35.hinet.net

C.-W. Ho (✉)  
No. 40, Chungshan N. Rd., Sec. 3,  
Taipei 104 Taiwan, Republic of China  
e-mail: cwwho@ttu.edu.tw

**Keywords** Y chromosome · Paiwan · Y-STR · DYS19 ·  
Yfiler · Haplotype

## Introduction

Taiwan's population of about 23 million is heterogeneous and is made up of 97.9% Han Chinese and 2.1% indigenous people. Han Chinese are composed of Minnan (68.9%), Hakka (15%) and mainland Chinese (14%) who arrived from China after World War II. Minnan and Hakka are the so-called “Taiwanese” who have immigrated to Taiwan from the southeast coast of China since the seventeenth century [1]. Physical and cultural anthropology has led to the classification of some 14 different indigenous Taiwanese people [2]: Ami, Paiwan, Atayal, Bunun, Truku, Rukai, Puyuma, Tsou, Saisiyat, Yami (Tau), Kavalan, Thao, Sakizaya and Sediq. The Paiwans, with a population size of about 84,500, reside around in the southern mountainous regions of Taiwan (Fig. 1) and belong to the Formosan sub-branch of the Austronesian language group [3, 4]. Linguistic and archaeological evidence suggests that Austronesian language speakers expanded early from Southern China into Taiwan [5, 6]. Due to the recent theory of Taiwan being the stepping-stone of the Austronesian demographic expansion, many anthropological studies have investigated the distribution of Y-STR haplotypes amongst Taiwanese indigenous people. Major limitations to these studies however, have been the relatively small sample size. None of the previous publications on male genetics of the Paiwan [6–14] comprise more than 50 chromosomes, which sometimes provide less information for studying. Here, we present the largest Y-STR study on the Paiwans to date.



**Fig. 1** Map of the region of residence of the Paiwan population in Taiwan (shaded). (Modified from <http://www.apc.gov.tw/>)

## Materials and methods

After obtaining consent, buccal swab samples were collected from 208 unrelated male Paiwan volunteers residing in the southern mountainous regions of Taiwan. Genomic DNA was extracted and purified using the Genomic DNA Extraction Kit (Yeastern Company, Taipei County, Taiwan, ROC.), and DNA concentrations were determined with the 7300 real-time PCR system using the Quantifiler™ Y Human Male DNA Quantification Kit (Applied Biosystems, Foster City, CA, USA).

Target DNA (0.1 ng/μL) was amplified by PCR, using the AmpFISTR® Yfiler™ PCR Amplification Kit (Applied Biosystems), at the following loci: DYS19, DYS385ab, DYS389I, DYS389II, DYS390, DYS391, DYS392, DYS393, DYS437, DYS438, DYS439, DYS448, DYS456, DYS458, DYS635 and GATA H4. The efficacy of the kit was determined by the identification of a 9-marker European minimal haplotype (minHt, the haplotype set of DYS19, DYS385ab, DYS389I, DYS389II, DYS390, DYS391, DYS392 and DYS393) and an 11-marker SWGDAM (Scientific Working Group on DNA Analysis Methods) core set (minHt plus DYS438 and DYS439). Amplification was performed in 25 μL reactions containing 10 μL target DNA, 5 μL primer set, 9.2 μL PCR reaction mix and 0.8 μL DNA polymerase using a GeneAmp PCR System 9700 (Applied Biosystems) under the following conditions: 95°C for 11 min, followed by 30 cycles at 94°C for 1 min, 61°C for 1 min and 72°C for 1 min, with a final

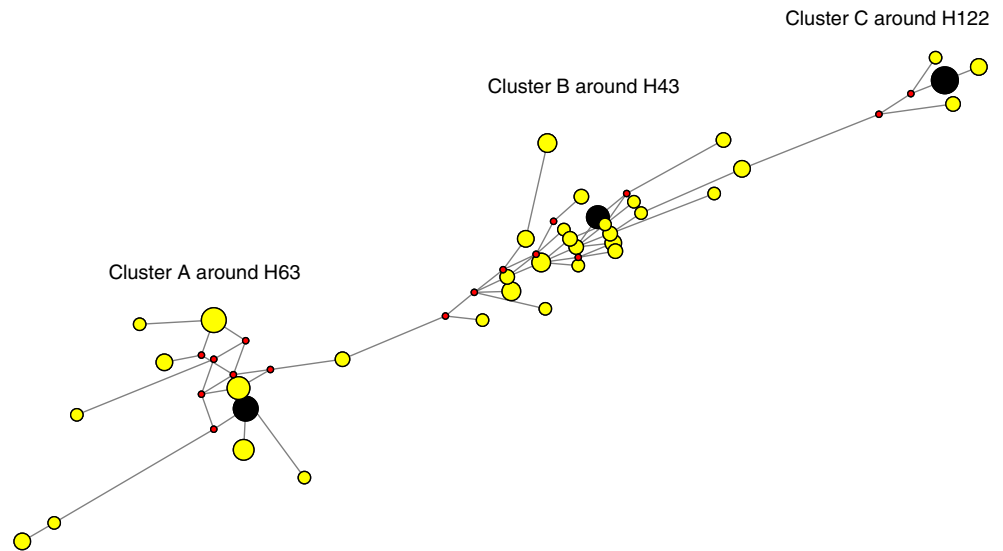
extension step at 60°C for 80 min. Electrophoresis of the PCR products, together with GeneScan-500 Internal Lane Size Standard (LIZ-500) to determine the base-pair sizes, was performed on the ABI Prism® 3100 Avant Genetic Analyzer (Applied Biosystems). GeneMapper ID Software Version 3.1 (Applied Biosystems) was used in order to determine the allelic repeats by applying Yfiler Allelic Ladder. Alleles were named as suggested by the DNA Commission of the International Society of Forensic Genetics [15].

Allelic frequencies were estimated by direct counting. Gene diversity and haplotype diversity were calculated according to Nei [16]. Comparisons among populations were computed by an analysis of molecular variance (AMOVA) test. Pairwise values of  $\Phi_{st}$ , an analogue of  $F_{st}$  that takes the evolutionary distance between individual haplotypes into account [17, 18], were calculated to measure genetic distances between minimal 9-locus haplotypes of the Paiwan and the published data from ten other populations [19–27] from the region ( $n=2112$ ) with the statistical significance determined by a permutation test (10,000 replicates). We used an implementation of AMOVA provided at the YHRD website [28]. The DYS389I allele length was obtained by subtracting the shorter allele from the longer allele at DYS389I/II. To illustrate the relationship between populations based on pairwise  $\Phi_{st}$ , an MDS plot was created by using the “Population analysis” tools of the YHRD [28]. A median-joining network [29] based on 12-locus Y-STR haplotypes (omitting the complex repeats) was calculated using the NETWORK 2.0b software (available online, [www.fluxus-engineering.com/sharenet.htm](http://www.fluxus-engineering.com/sharenet.htm)).

## Results and discussion

The Y-STR allele frequencies of the Paiwan population are shown in Table S1. Forensic indices such as the locus diversity and haplotype diversity values were calculated from the allelic frequency for each locus. Of the 17 markers analyzed, DYS385ab showed the greatest degree of diversity, 0.8354, and DYS438 showed the lowest, 0.1014. A total of 135 haplotypes were identified in the 208 individuals studied, of which 102 were unique (Table S2). The overall observed haplotype diversity reached  $0.9922 \pm 0.0010$ , and the discrimination capacity was 0.6490. The most frequent 17-locus haplotype was H122 that was identified in nine individuals (Table S2). The diversity values were markedly reduced when compared to the major population in Taiwan [30]. In addition, three intermediate or duplicated alleles at locus DYS448, not yet observed in the YHRD 3.0, (Release 29 from June 15th, 2009; 15.2, 16.2 and 15.2, 18), were identified in this study. A further search of the YHRD release

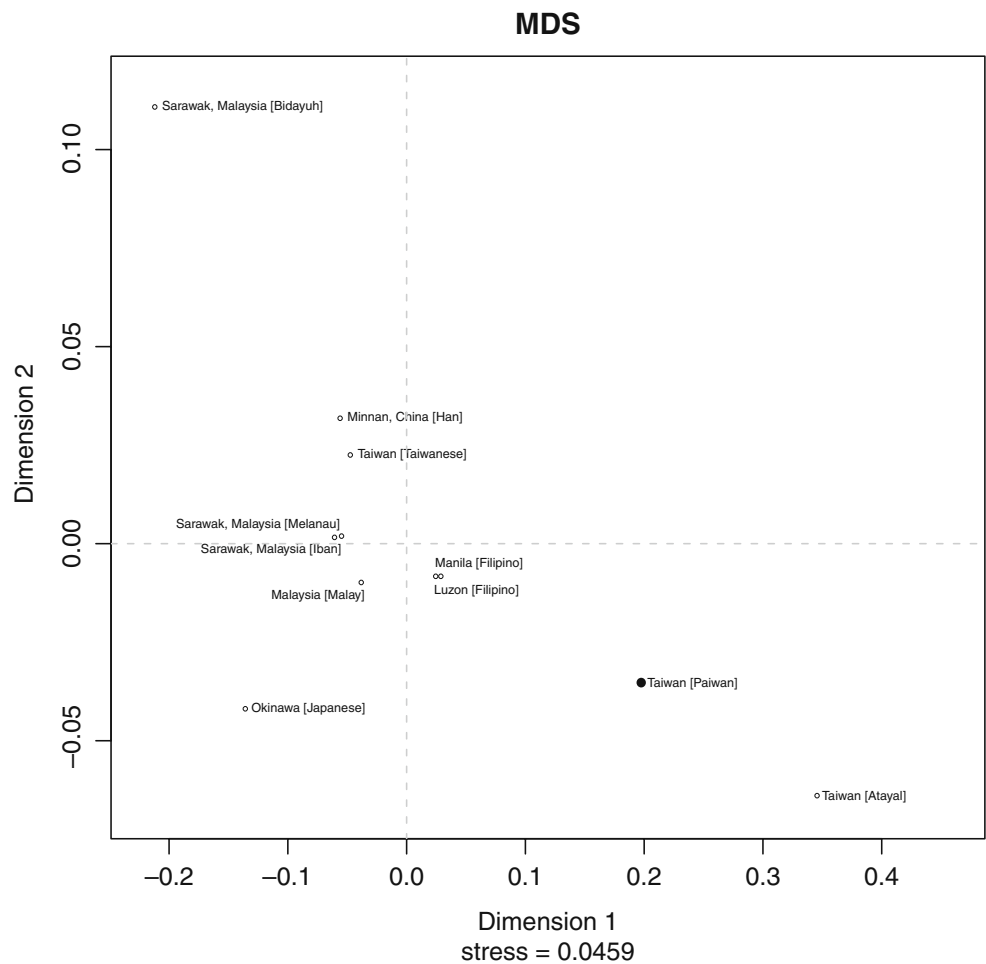
**Fig. 2** Median-joining network for 36 different 12-locus haplotypes ( $n > 2$ ) from the Paiwan population. Three clusters surrounding dominant haplotypes (black circles) are clearly discernible



reveals no match among 17,384 haplotypes. Only one match (H102) was observed with the Atayal population [21]. A median-joining network analysis for the Paiwan sample based on the YFiler haplotypes (omitting the loci with mutations at different repeat blocks DYS389II, DYS390,

DYS385ab and DYS448) shows 36 different haplotypes occurring more than once. Most of these haplotypes fall in three clearly discernible clusters, A–C, which surround the predominant H43, H63 and H122 haplotypes (Table S2, Fig. 2). The network illustrates the reduced haplotype

**Fig. 3** MDS plot based on pairwise  $\Phi_{st}$  values calculated for the Paiwan (filled circle) and ten reference populations (open circles)



variability and the dominance of certain patrilineal clans within the Paiwan population. Taken together, this data indicate a patrilineal type of residence of the Paiwan which live in relative isolation in the southern mountainous regions of Taiwan.

In order to analyze the relationship of the Paiwan tribal people to other Austronesian-speaking (Atayal, Malay and Filipino) and non-Austronesian-speaking neighbouring populations from Taiwan, South Eastern China and Japan, we compared the most widely studied minimal 9-locus haplotypes (minHt) of the Paiwan via AMOVA with published data from several reference populations. In total, 2112 reference samples from the following populations (all published and fully referenced in the YHRD) we used: 805 Taiwanese from Taiwan, YHRD Accession Numbers 003340 and YA003193 [19, 20], 170 Atayal from Taiwan, YA003524 [21], 109 Han Chinese from Minnan, YA003308 [22], 76 Filipinos from Manila, YA003202 [23], 211 Filipinos from Luzon, YA003206 [24], 334 Malays from Malaysia, YA003278 [25], 113 Bidayuh from Malaysia, YA003415 [26], 103 Iban from Malaysia, YA003416 [26], 104 Melanau from Malaysia, YA003417 [26] and 87 Japanese from Okinawa, YA003205 [27]. All pairwise comparisons between the indigenous Paiwan and Atayal and the other populations show large  $F_{st}$  values and are significant at the 5% level (see Fig. 3, Table S3). As shown by AMOVA and illustrated in the MDS plot (Fig. 3), the Paiwan and Atayal indigenous tribes show little genetic relationship to other Taiwanese, Chinese, Malaysian and Japanese populations tested. A close association between the Paiwan and Tagalog- and Cebuano-speaking populations of the Philippines (which also belong to the Austronesian language group) was also not detectable with Y-STR markers. Surprisingly, while intermarriage between the early settlers from the mainland with pre-existing mountain tribes occurred [19], any admixture of the male gene pool of Chinese settlers on Taiwan and indigenous people is not supported by our Y chromosome analysis. Accordingly, the studied native Paiwan population from Taiwan is seen as a population isolate with a male gene pool formed by caste endogamy and a patrilineal residence type resulting in strong genetic drift [31].

While the national Y-STR database for the Taiwanese populations is growing with currently 1183 haplotypes (refer to [www.yhrd.org](http://www.yhrd.org), National databases, Taiwan) the heterogeneous populations structure must strictly be observed by construction of separate database branches for major and minor subpopulations.

**Acknowledgements** We would like to thank the National Science Council of R.O.C. for supporting the study and appreciate the Pingtung County Government, the Pingtung County Public Health Bureau, and the Majja and Manjhou Township Health Centers for collecting samples.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

## References

1. Lin M, Chu CC, Chang SL et al (2001) The origin of Minnan and Hakka, the so-called “Taiwanese”, inferred by HLA study. *Tissue Antigens* 57:192–199
2. Government Information Office of ROC (2009) The Republic of China yearbook 2008. Government Information Office Press, Taipei
3. Bellwood P (1991) The Austronesian dispersal and the origin of languages. *Sci Am* 265:70–75
4. Chu JY, Huang W, Kuang SQ et al (1998) Genetic relationship of populations in China. *Proc Natl Acad Sci USA* 95:11763–11768
5. Gray RD, Jordan FM (2000) Language trees support the express-train sequence of Austronesian expansion. *Nature* 405:1052–1055
6. Li H, Wen B, Chen SJ et al (2008) Paternal genetic affinity between western Austronesians and Daic populations. *BMC Evol Biol* 8:146–157
7. Forster P, Kayser M, Meyer E, Roewer L, Pfeiffer H, Benkmann H, Brinkmann B (1998) Phylogenetic resolution of complex mutational features at Y-STR DYS390 in aboriginal Australians and Papuans. *Mol Biol Evol* 15:1108–1114
8. Hagerberg E, Kayser M, Nagy M et al (1999) Molecular genetic evidence for the human settlement of the Pacific: analysis of mitochondrial DNA, Y chromosome and HLA markers. *Phil Trans R Soc Lond B* 354:141–152
9. Hurler ME, Nicholson J, Bosch E, Renfrew C, Sykes BC, Jobling MA (2002) Y chromosomal evidence for the origins of oceanic-speaking peoples. *Genetics* 160:289–303
10. Kayser M, Brauer S, Weiss G, Underhill PA, Roewer L, Schiefelhövel W, Stoneking M (2000) Melanesian origin of Polynesian Y chromosomes. *Curr Biol* 10:1237–1246
11. Kayser M, Krawczak M, Excoffier L et al (2001) An extensive analysis of Y-chromosomal microsatellite haplotypes in globally dispersed human populations. *Am J Hum Genet* 68:990–1018
12. Hsieh HM, Shen RK, Yuen TY et al (2003) Screening of Y-chromosome STR loci in Taiwanese populations. *Forensic Sci J* 2:69–76
13. Su B, Jin L, Underhill P et al (2000) Polynesian origins: insights from the Y chromosome. *Proc Natl Acad Sci USA* 97:8225–8228
14. Capelli C, Wilson JF, Richards M et al (2001) A predominantly indigenous paternal heritage for the Austronesian-speaking peoples of insular Southeast Asia and Oceania. *Am J Hum Genet* 68:432–443
15. Gusmao L, Butler JM, Carracedo A et al (2006) DNA commission of the international society of forensic genetics (ISFG): an update of the recommendations on the use of Y-STRs in forensic analysis. *Forensic Sci Int* 157:187–197
16. Nei M (1987) *Molecular evolutionary genetics*. Columbia University Press, New York
17. Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* 131:479–491
18. Roewer L, Kayser M, Dieltjes P, Nagy M, Bakker E, Krawczak M, de Knijff P (1996) Analysis of molecular variance (AMOVA) of Y-chromosome-specific microsatellites in two closely related human populations. *Hum Mol Genet* 5:1029–1033

19. Wu FC, Pu CE (2001) Multiplex DNA typing of short tandem repeat loci on Y chromosome of Chinese population in Taiwan. *Forensic Sci Int* 120:213–222
20. Liu HM, Chen PS, Chen YJ, Lyou JY, Hu HY, Lin JS, Tzeng CH (2007) Y-chromosome short tandem repeats analysis to complement paternal lineage study: a single institutional experience in Taiwan. *Transfusion* 47:918–926
21. Wu FC, Ho CW, Pu CE, Hu KY, Liu DH (2009) Genetic polymorphisms of 17 Y-chromosomal short tandem repeat loci in Atayal population of Taiwan. *Croat Med J* 50:313–320
22. Hu S (2006) Genetic polymorphism of 12 Y-chromosomal STR loci in the Minnan Han Chinese in Southeast China. *Forensic Sci Int* 159:77–82
23. Kwak KD, Jin HJ, Shin DJ et al (2005) Y-chromosomal STR haplotypes and their applications to forensic and population studies in East Asia. *Int J Legal Med* 119:195–201
24. Lessig R, Willuweit S, Krawczak M et al (2003) Asian online Y-STR Haplotype Reference Database. *Leg Med (Tokyo)* 5:160–163
25. Chang YM, Perumal R, Keat PY, Kuehn DLC (2007) Haplotype diversity of 16 Y-chromosomal STRs in three main ethnic populations (Malays, Chinese and Indians) in Malaysia. *Forensic Sci Int* 167:70–76
26. Chang YM, Swaran Y, Phoon YK, Sothirasan K, Sim HT, Lim KB, Kuehn D (2008) Haplotype diversity of 17 Y-chromosomal STRs in three native Sarawak populations (Iban, Bidayuh and Melanau) in East Malaysia. *Forensic Sci Int Genetics* 3:e77–e80
27. Uchihi R, Yamamoto T, Usuda K et al (2003) Haplotype analysis with 14 Y-STR loci using 2 multiplex amplification and typing systems in 2 regional populations in Japan. *Int J Legal Med* 117:34–38
28. Willuweit S, Roewer L, the International Forensic Y chromosome User Group et al (2007) Y Chromosome Haplotype Reference Database (YHRD): update. *Forensic Sci Int Genet* 1:83–87
29. Bandelt H-J, Forster P, Röhl A (1999) Median-joining networks for inferring intraspecific phylogenies. *Mol Biol Evol* 16:37–48
30. Tsai LC, Yuen TY, Hsieh HM et al (2002) Haplotype frequencies of nine Y-chromosome STR loci in the Taiwanese Han population. *Int J Legal Med* 116:179–183
31. Chai CK (1967) Taiwan aborigines: A genetic study of tribal variations. Harvard University Press, Cambridge