

## *Autosomal STR Variation in Five Austronesian Populations*

E. M. SHEPARD,<sup>1</sup> R. A. CHOW,<sup>1</sup> EPIFANIA SUAFO'A,<sup>2</sup> DAVID ADDISON,<sup>3</sup> A. M. PÉREZ-MIRANDA,<sup>1</sup> R. L. GARCIA-BERTRAND,<sup>4</sup> AND R. J. HERRERA<sup>1</sup>

**Abstract** Human population characteristics at the genetic level are integral to both forensic biology and population genetics. This study evaluates biparental microsatellite markers in five Austronesian-speaking groups to characterize their intra- and interpopulation differences. Genetic diversity was analyzed using 15 short tandem repeat (STR) loci from 338 unrelated individuals from 5 Pacific islands populations, including the aboriginal Ami and Atayal groups from Taiwan, Bali and Java in Indonesia, and the Polynesian islands of Samoa. Allele frequencies from the STR profiles were determined and compared to other geographically targeted worldwide populations procured from recent literature. Hierarchical AMOVA analysis revealed a large number of loci that exhibit significant correspondence to linguistic partitioning among groups of populations. A pronounced divide exists between Samoa and the East (Formosa) and Southeast Asian (Bali and Java) islands. This is clearly illustrated in the topology of the neighbor-joining tree. Phylogenetic analyses also indicate clear distinctions between the Ami and Atayal and between Java and Bali, which belie the respective geographic proximities of the populations in each set. This differentiation is supported by the higher interpopulation variance components of the Austronesian populations compared to other Asian non-Austronesian groups. Our phylogenetic data indicate that, despite their linguistic commonalities, these five groups are genetically distinct. This degree of genetic differentiation justifies the creation of population-specific databases for human identification.

Genetic diversity, characterized at both the intra- and the interpopulation level, forms the basis of forensic biology and population genetics, respectively. Yet these two disciplines are not always fully integrated in studies using human samples. A well-established marker system such as autosomal short tandem repeats

<sup>1</sup>Department of Biological Sciences, Florida International University, University Park, OE 304, Miami, FL 33199.

<sup>2</sup>National Park Service, Pago Pago, American Samoa, 96799.

<sup>3</sup>American Samoa Power Authority, Pago Pago, American Samoa, 96799.

<sup>4</sup>Department of Biological Sciences, Colorado College, Colorado Springs, CO 80903.

*Human Biology*, December 2005, v. 77, no. 6, pp. 825–851.

Copyright © 2005 Wayne State University Press, Detroit, Michigan 48201-1309

KEY WORDS: ISLAND SOUTHEAST ASIA, TAIWAN ABORIGINES, POLYNESIA, TAIWAN, BALI, JAVA, SAMOA, AUSTRONESIAN-SPEAKING GROUPS, AMI GROUP, ATAYAL GROUP, SHORT TANDEM REPEATS, MICROSATELLITE MARKERS, D8S1179, D21S11, D7S820, CSF1PO, D3S1358, TH01, D13S317, D16S539, VWA, TPOX, D18S51, D5S818, FGA, D2S1338, D19S433, PHYLOGENY, FORENSIC BIOLOGY, GENETIC DIVERSITY.

(STRs) represents hypervariable regions that can provide the fine resolution needed to determine relationships among closely related populations in recent evolutionary history (Bowcock et al. 1994; Jorde et al. 1995, 1997; Bosch et al. 2000; Lum et al. 2002; Rowold and Herrera 2003; Pérez-Miranda et al. 2005; Shepard and Herrera 2005) and the discrimination power essential for robust individual probabilities of inclusion (Leibelt et al. 2003; Collins et al. 2004). Also, STRs are used in these studies because of their numerous and relatively even distribution throughout the genome, high levels of polymorphism, a large number of possible alleles per locus, and short amplicon lengths, which facilitate DNA amplification, separation, and detection (Butler 2001; Butler et al. 2003).

The populations composing the Austronesian language family have been the subject of numerous studies from overlapping anthropological disciplines, namely, linguistics, archeology, and molecular biology. Studies in these fields have provided evidence on the complexity of human migration patterns during the Austronesian diaspora (Bellwood 2001; Underhill 2004). The current range of Austronesian-speaking people extends from Taiwan (Formosa) to the north, Easter Island (west of Chile, South America) in the east, New Zealand to the south, and as far as Madagascar (off East Africa) to the west. The distances between these locations cover approximately two-thirds the circumference of the planet. Consequently, because of their wide geographic distribution, the Austronesians are an interesting group from the perspective of population genetics. In addition, because of their relatively recent expansions into the Indian and Pacific Oceans, beginning about 6,000 years B.P. and ending as late as 800 years B.P., Austronesian populations provide an ideal test group to study a major dispersal process from prehistoric time. In terms of forensics, this area is undercharacterized by accepted STR marker sets.

The goal of this study is to investigate the allelic profiles of 15 biparental STR loci common to forensic studies (D8S1179, D21S11, D7S820, CSF1PO, D3S1358, TH01, D13S317, D16S539, VWA, TPOX, D18S51, D5S818, FGA, D2S1338, and D19S433) in five geographically targeted Austronesian populations from the Pacific Ocean. These include two aboriginal Taiwanese populations (the Ami and Atayal), two Indonesian populations from Bali and Java, and a Polynesian population from Samoa. The ultimate aim of our study is to assess the degree of genetic heterogeneity among these five Austronesian populations and to ascertain how they relate phylogenetically to regional and worldwide groups previously studied with the same set of markers. In addition, these well-characterized databases will be of forensic value.

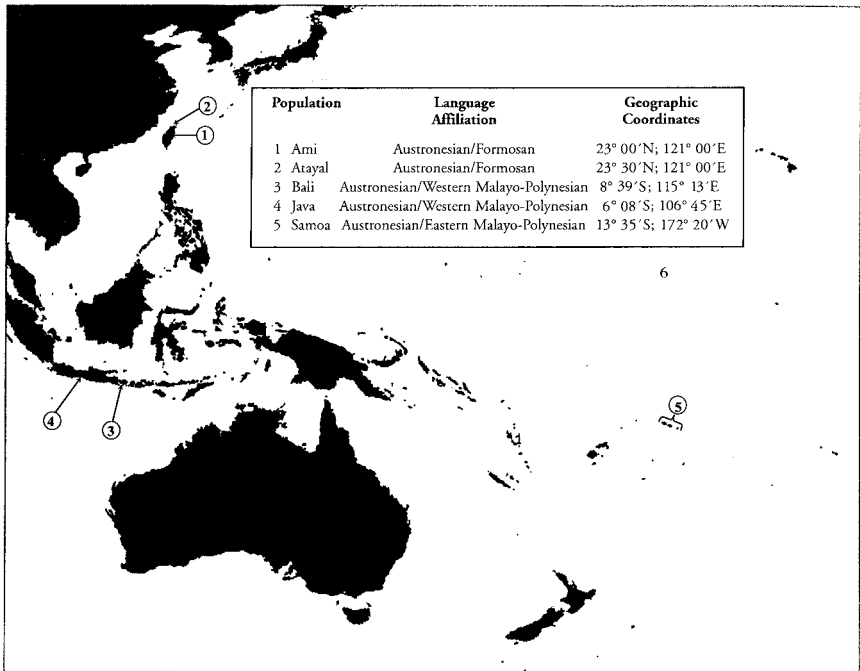
Upon examination of these 15 highly polymorphic loci, we find that when groups of populations are compared, the overall tests of correlation between genetic partitioning with linguistic and geographic differences are statistically significant. Also, most loci exhibit significant correlation at the level of groups of populations along geographic and linguistic lines. Phylogenetic analyses display some thought-provoking results, including an extreme differentiation between

the two aboriginal populations from Taiwan, the Ami and Atayal, despite their overlapping geographic range. Similarly, we detect a clear distinction between the Indonesian populations of Bali and Java, separated by mere miles within the Indo-Malaysian archipelago. Of particular interest is the segregation of Samoa from the other four Austronesian groups into a different clade altogether. Based on this evidence, we conclude that these five populations do not share an obvious genetic link, despite their common language affiliation, the implications of which are discussed further within a framework of autosomal STR analysis.

## **Materials and Methods**

**Populations, Sample Collection, and DNA Isolation.** The five Austronesian populations from the Pacific Ocean investigated in this study include two aboriginal groups from Formosa (Taiwan) (the Ami and Atayal), two populations from islands of the Indonesian chain (Bali and Java), and a fifth population from the Samoan islands in Polynesia (Figure 1). The Samoan samples were collected from both Western and American Samoa as a representative group. Data from 12 worldwide populations (Table 1) were obtained from the literature and used for comparison. Populations were chosen from the literature to be representative of different ethnic groups and biogeographic areas. Individuals were identified by biogeographic information traced back at least two generations. Each collection was arranged through the leaders of each region and supervised by the same. Sample collections were performed according to the ethical guidelines outlined by the Institutional Review Board of Florida International University. All samples were collected as whole blood in Vacutainer tubes containing EDTA. DNA was extracted using the standard phenol-chloroform method (Antunez de Mayolo et al. 2002).

**PCR Amplification and Detection of STRs.** The samples were amplified by PCR using the commercial AmpFISTR Identifiler kit (Applied Biosystems, Foster City, California) at the following loci: D8S1179, D21S11, D7S820, CSF1PO, D3S1358, TH01, D13S317, D16S539, D2S1338, D18S433, VWA, TPOX, D18S51, D5S818, FGA, and amelogenin. Amplifications were performed in a GeneAmp PCR System 9600 thermocycler (Applied Biosystems) using the following cycling parameters: 11 min denaturation at 95°C; 28 cycles of 1 min denaturation at 94°C, 1 min primer annealing at 59°C, and 1 min primer extension at 72°C; and a final soak for 60 min at 60°C. A portion of each amplified sample was mixed with formamide and GS500 LIZ as an internal size standard, as recommended by the manufacturer (Applied Biosystems), and then separated using an ABI Prism 3100 Genetic Analyzer (Applied Biosystems). GeneScan 3.7 was used to determine the fragment sizes, and Genotyper 3.7 NT software was used to designate alleles by comparison with the allelic ladder provided by the manufacturer.



**Figure 1.** Locations of the populations used in this study. Language affiliation classifications were obtained from <http://www.ethnologue.com>. Geographic coordinates for each population were generated according to the geopolitical Mercator projection (Watkins et al. 2003).

**Statistical and Phylogenetic Analyses.** Allele frequencies of the 15 STR loci were calculated using the gene counting method (Li 1976). The Arlequin software package, version 2.000 (Levene 1949; Guo and Thompson 1992; Schneider et al. 2000), was used to assess Hardy-Weinberg equilibrium expectations using Fisher's exact test with the modified Markov chain Monte Carlo method as well as to determine Nei's gene diversity index (GD) (Nei 1987). Hardy-Weinberg equilibrium was evaluated at  $\alpha = 0.05$  and also using the Bonferroni adjustment for the number of loci tested ( $0.05/15 = 0.0033$ ) as a correction for type I errors.

Forensically useful parameters were also examined for all five populations studied, including power of discrimination (PD) and polymorphic information content (PIC), using the PowerStats program, version 1.2 (Tereba 1999; Jones 1972; Brenner and Morris 1990). To determine phylogenetic relationships, the 5 populations studied along with 12 other geographically targeted worldwide reference populations were included in a neighbor-joining tree using Phylip 3.52c software (Felsenstein 2002) based on  $F_{ST}$  distances (Reynolds et al. 1983). Bootstrap consensus scores (1,000 replications) were generated by the SeqBoot and

**Table 1.** Description of and Reference Information for the Studied Populations

<i>Population</i>	<i>n</i>	<i>Description</i>	<i>Reference</i>
African Americans	258	General population of United States	Butler et al. (2003)
Ami	79	Aboriginal tribe, east-central Taiwan	Present study
Angola	110	General population of Cabinda, Angola	Beleza et al. (2004)
Atayal	25	Aboriginal tribe, north Taiwan	Present study
Bali	79	General population of Bali	Present study
Belgium	222	General population, majority from Flanders region of Belgium	Decorte et al. (2004)
Japan	526	General population of Japan	Hashiyada et al. (2003)
Java	60	General population of Java	Present study
Malaysian Malay	210	Malay ethnicity from Malaysia	Seah et al. (2003)
Malaysian Chinese	219	Chinese ethnicity from Malaysia	Seah et al. (2003)
Mozambique	142	General population of Maputo, Mozambique	Alves et al. (2004)
North Poland	145	General population of northern Poland	Szczerkowska et al. (2004)
Samoa	95	General population of American Samoa and Samoa	Present study
Taiwan	597	General population of Taiwan	Wang et al. (2003)
U.S. Caucasians	302	General population of United States	Butler et al. (2003)
U.S. Hispanics	140	General population of United States	Butler et al. (2003)
Venezuela	255	General population of Caracas, Venezuela	Chiurillo et al. (2003)

GenDist options of the Phylip software, and the ConSense programs determined the best-fitting dendrogram.

Multidimensional scaling (MDS) analyses were performed using the Statistical Package for the Social Sciences (SPSS) software program, which is also based on  $F_{ST}$  distances (SPSS Inc. 2001).  $G$  tests were carried out to determine differences in overall genetic variability between populations using Carmody's program (Carmody 1990).

Inter- and intrapopulation genetic variance component values ( $G_{ST}$  and  $H_s$ , respectively) were ascertained for nine Pacific Ocean populations, composed of six Austronesian (Ami, Atayal, Bali, Java, Malaysian Malay, and Samoa) and three non-Austronesian (Japan, Malaysian Chinese, and Taiwan) groups for each locus, according to the DISPAN software program (Ota 1993). Genetic structuring was analyzed for the same nine populations according to both biogeographic lines and linguistic subfamily affiliations through hierarchical analysis of molecular variance (AMOVA) (Excoffier et al. 1992). Linguistic correlations were assessed based on the following subfamily partitioning: two non-Austronesian groups [Japanese (Japan) and Sino-Chinese (Taiwan and Malaysian Chinese)] and three Austronesian groups [Formosan (Ami, Atayal), Western Mayalo-Polynesian (Bali, Java, Malaysian Malay), and Eastern Mayalo-Polynesian (Samoa)]. Geographic correlations were tested on the basis of the following regional groups: northeast Asia (Japan), East Asia (Ami, Atayal, Taiwan and Malaysian Chinese), Southeast Asia (Bali, Java, Malaysian Malay), and Polynesia

(Samoa). Note that the Malaysian Chinese population, while inhabiting Malaysia, was grouped with the East Asian populations because of its Han Chinese ancestry.

## Results

**STR Diversity Within Populations.** The allele-frequency distributions of the Ami, Atayal, Bali, Java, and Samoa populations are listed in Tables 2 through 6. In addition, parameters of importance to population genetics are summarized in Table 7 for each group under study. This table lists the loci in each population that do not meet Hardy-Weinberg equilibrium expectations at  $p < 0.05$  (4 loci out of 75 possible tests). However, after applying the Bonferroni adjustment (see explanation in Materials and Methods section), this is reduced to a single departure from Hardy-Weinberg equilibrium (1 loci out of 75 possible tests).

Although a detailed dissection of allele frequencies is not the purpose of this study, some interesting observations warrant attention. For instance, in the Ami population (Table 2), allele 10 of TPOX, commonly encountered in other Asian databases, is notably absent from the Ami. Conversely, in the Ami, allele 24.2 of FGA (0.0316) is present at a frequency up to 10 times higher than in the published Asian databases included in this study, and is absent altogether from the Atayal, Bali, Java, and Samoa. Observed heterozygosity ( $H_o$ ) for the Ami ranges from 0.6076 in TH01 to 0.8987 in D13S317 and D2S1338. In the Atayal the smallest and largest allele sizes in each locus are often not present (Table 3). This population also lacks four common midsize alleles; most notably absent are alleles 9.3 and 10 of TH01 and alleles 15 and 24.2 of VWA. Observed heterozygosity for the Atayal ranges from 0.5200 in TH01 to 1.0000 in FGA. The Indonesian population from Bali (Table 4) contains a microvariant allele (allele 23.2 of D21S11) not encountered in any of the Asian databases used in our study. There was one deviation from Hardy-Weinberg equilibrium expectations for Bali in the VWA locus ( $p < 0.05$ ) that persists even after application of the Bonferroni adjustment for type I errors. Departures have been reported in other regional populations, specifically the Malay group from Malaysia for this particular locus (Seah et al. 2003). Observed heterozygosity of the Bali population oscillates from 0.6076 in VWA to 0.8608 in D2S1338 and D18S51. In the other Indonesian group from Java (Table 5), allele 13 of TPOX (0.0083) is rarely detected in other Asian groups. In this Javanese population the observed heterozygosity ranges from 0.6333 in CSF1PO to 0.9500 in D21S11. Within the population from Samoa (Table 6), locus D13S317 contains a rare allele with 16 repeats, one of the largest in this locus reported thus far among the Asian data sets.  $H_o$  for Samoa ranges from 0.6105 in TPOX to 0.9158 in D16S539.

**STR Diversity Among Populations.** To investigate genetic affiliations among the five Pacific Ocean populations and their relationships to other worldwide populations, we generated a neighbor-joining tree using  $F_{ST}$  distances. Figure 2 displays a neighbor-joining phylogram based on all 17 populations. There are three major clusters in the dendrogram (bootstrap value of 50%). One consists of those of European/Hispanic descent; another primarily includes African groups or groups of African ancestry, and a third clade represents a cluster made up of Asians and Pacific Ocean populations. Overall, the topology of this tree is robust (only four nodes exhibit bootstrap values under 50% incidence).

The Atayal and Ami segregate into the same group as the Japanese and Chinese from both Malaysia and Taiwan (91%) and then separate from these three populations with a confidence value of 29%. The Ami bifurcate from the Atayal with a bootstrap value of 40%. It is interesting to note that, although there is a genetic relationship between these two Taiwanese aboriginal groups, the extensive branch length of the Atayal, with respect to the Ami, is indicative of their genetic uniqueness. Both Bali and Java bifurcate from the Malaysian Malays (bootstrap value = 71%). However, the genetic distance between Bali and Java is not as pronounced as in the case with the Ami and Atayal. Samoa segregates in an isolated and intermediate position between the African and European/Hispanic groups with a bootstrap value of 100%.

MDS analysis was performed to examine the genetic relationships among the 17 populations based on  $F_{ST}$  distances (Figure 3). Its topology is consistent with that of the grouping in the neighbor-joining dendrogram. As in the neighbor-joining tree, there are three main clusters: (1) Europeans, (2) Africans and groups of African descent, and (3) Asians and Pacific Islanders. The Ami, Atayal, Bali, and Java populations all cluster with the Asian groups on the left side of the plot. The Samoan population lies on the y axis of the crosshairs closer to the African groups than to the European/Hispanic groups. The isolated placement of the Samoans segregating away from all clusters and the extreme outlier position of the Atayal indicate considerable genetic differentiation with respect to the other populations and mirrors the long branch distance observed in the neighbor-joining tree.

The nine Asian and Pacific Ocean populations were split into two groups: Austronesian and non-Austronesian language affiliation. They were then analyzed to determine the allocation of genetic variance at the inter- $G_{ST}$  and intra- $H_s$  population levels (Table 8). The STR markers in the Austronesian groups displayed lower levels of intrapopulation variance than the non-Austronesian groups for 11 of the 15 loci and when calculated across all loci.  $G_{ST}$  values ranged from 2 to 10 times higher in the Austronesians than in the non-Austronesians per locus and more than 8 times higher when assessed across all loci. When pairwise  $G$  tests were performed on our 5 populations and on the 12 geographically targeted worldwide groups taken from the literature, we observed significant genetic differentiation ( $p < 0.05$ ) between all populations.

**Table 2.** STR Allele Frequencies for the Ami (Taiwanese Aborigine),  $n = 79$ 

Allele	D8S1179	D21S11	D7S820	CSF1PO	D3S1358	TH01	D13S317	D16S539	D2S1338	D19S433	VWA	TPOX	D18S51	D5S818	FGA
6						0.1076									
7						0.2342								0.0506	
8			0.1519			0.0570	0.3165					0.4747			
9			0.0570			0.5253	0.0886	0.3101				0.1139		0.0443	
9.3						0.0063									
10	0.1076		0.2215	0.2722		0.0696	0.1582	0.0949						0.3544	
11	0.0759		0.3608	0.2215			0.3101	0.2215				0.3924		0.2089	
12	0.1519		0.1709	0.4114			0.1203	0.2215				0.0190	0.0443	0.2278	
13	0.2595		0.0253	0.0823			0.0063	0.1266		0.2848			0.0253	0.1139	
13.2										0.0190					
14	0.2342		0.0127	0.0127	0.0063			0.0190		0.1519	0.1456		0.1772		
14.2										0.1456					
15	0.1139							0.0063		0.0696	0.0570		0.2658		
15.2										0.2975					
16	0.0380								0.0886	0.0063	0.1266		0.0823		0.0063
16.2										0.0253					
17	0.0190								0.0127		0.2532		0.1709		
18									0.0633		0.2342		0.0759		
19									0.1582		0.1582		0.0886		0.1076
20									0.0443		0.0190		0.0253		0.0696
21									0.0063		0.0063		0.0316		0.1582
22									0.0823				0.0127		0.2405
23									0.2025						0.1709

24		0.2658																		0.1013
24.2																				0.0316
25																				0.0886
26																				0.0253
28																				
29																				
30																				
31																				
31.2																				
32																				
32.2																				
33																				
33.2																				
34.2																				
$H_o$		0.8481	0.7848	0.7595	0.6835	0.6835	0.6076	0.8987	0.7848	0.8987	0.7722	0.7975	0.6582	0.8228	0.7722	0.8734				
$H_e$		0.8279	0.8516	0.7705	0.7052	0.7077	0.6706	0.7611	0.7852	0.8421	0.7852	0.8204	0.6112	0.8493	0.7688	0.8571				
$P$ value		0.33519	0.73645	0.14166	0.92577	0.09143	0.47733	0.21328	0.56048	0.25976	0.59075	0.04676	0.90535	0.06765	0.98115	0.27838				
GD		0.8279	0.8516	0.7705	0.7052	0.6793	0.6536	0.7611	0.7852	0.8421	0.7852	0.8204	0.8204	0.8493	0.7662	0.8571				
PD		0.9354	0.9553	0.9002	0.8566	0.8060	0.8358	0.8678	0.9101	0.9351	0.9146	0.9319	0.7470	0.9438	0.9059	0.9508				
PIC		0.7996	0.8279	0.7297	0.6477	0.6045	0.6060	0.7171	0.7464	0.8175	0.7474	0.7895	0.5278	0.8266	0.7249	0.8349				

$H_o$ , observed heterozygosity.

$H_e$ , expected heterozygosity.

$P$  value: Hardy-Weinberg equilibrium, Fisher's exact test.

GD, gene diversity index.

PD, power of discrimination.

PIC, polymorphic information content.

Statistics calculated using PowerStats, v. 1.2 (Promega).

**Table 3.** STR Allele Frequencies for the Atayal (Taiwan Aborigines), *n* = 25

Allele	D8S1179	D21S11	D7S820	CSFIPO	D3S1358	TH01	D13S317	D16S539	D2S1338	D19S433	VWA	TPOX	D18S51	D5S818	FGA
6						0.2000									
7						0.6400								0.0400	
8			0.1000			0.0200	0.0600					0.3000			
9			0.0200			0.1140	0.0400	0.5200				0.0200			
10	0.3400		0.3600	0.2000		0.3600	0.1000	0.1000				0.1000		0.3800	
11	0.0800		0.3800	0.3400		0.4400	0.1600	0.1600				0.5800		0.4200	
12	0.1400		0.1200	0.3200		0.1000	0.1800						0.0400	0.1200	
13	0.1000		0.0200	0.1200			0.0200		0.2400				0.0200	0.0400	
13.2									0.1800						
14	0.2400			0.0200					0.1200		0.1600		0.4400		
14.2									0.1000						
15	0.0400								0.1800				0.2400		
15.2									0.1600						
16	0.0600										0.0800		0.2000		
16.2									0.0800						
17									0.2000		0.2200		0.0600		
18									0.0600		0.4200				0.0600
19									0.1800		0.1200				0.1600
20									0.0600						0.0800
21															0.1800
22									0.0400						







**Table 5.** STR Allele Frequencies for Java (Indonesia),  $n = 60$

Allele	D8S1179	D21S11	D7S820	CSFIPO	D3S1358	TH01	D13S317	D16S539	D2S1338	D19S433	VWA	TPOX	D18S51	D5S818	FGA
6						0.1250						0.0083			
7			0.0083			0.2667								0.0083	
8			0.2417	0.0083		0.1167	0.2917					0.5083			
9			0.0500	0.0167		0.2833	0.1083	0.1417				0.1333		0.0083	
9,3						0.0583									
10	0.0250		0.1917	0.1750		0.1500	0.2167	0.1583				0.0583		0.4167	
11	0.1083		0.3250	0.4000			0.2417	0.3167				0.2667		0.2000	
12	0.1583		0.1583	0.3250			0.0917	0.2667	0.0583			0.0167	0.0417	0.2083	
13	0.1250		0.0250	0.0750	0.0250		0.0333	0.1167	0.2833		0.0083	0.0083	0.1083	0.1417	
13,2									0.0583						
14	0.2417						0.0167		0.1917		0.2917		0.2500	0.0083	
14,2									0.0417						
15	0.1833								0.1250		0.0417		0.2333		
15,2									0.2250						
16	0.1333								0.0333	0.0083	0.1083		0.1833	0.0083	
16,2									0.0083						
17	0.0250								0.1000		0.2167		0.0833		
18									0.0583		0.1833		0.0083		0.0250
19									0.2333		0.1250		0.0333		0.0583
20									0.1000		0.0167		0.0250		0.0750
21									0.0167		0.0083		0.0083		0.1667
21,2															0.0250
22									0.0417				0.0250		0.2417
22,2															0.0417
23									0.2333						0.1583
23,2															0.0083
24									0.1250						0.0750



**Table 6.** STR Allele Frequencies for Samoa,  $n = 95$ 

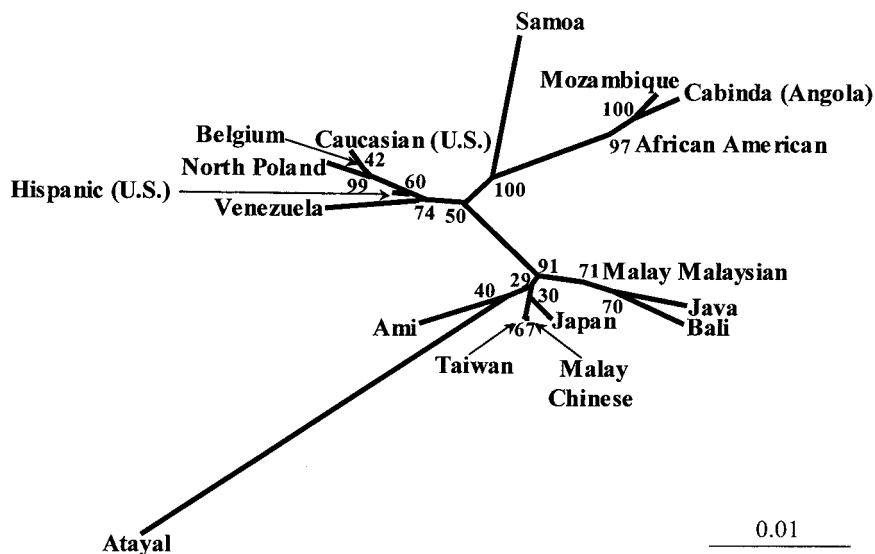
Allele	D8S1179	D21S11	D7S820	CSFIPO	D3S1358	TH01	D13S317	D16S539	D2S1338	D19S433	VWA	TPOX	D18S51	D5S818	FGA
6						0.1263									
7						0.4684									
8			0.1000			0.1368	0.0632					0.3730			
9			0.1105			0.0737	0.1895	0.2000				0.2368			
9.3						0.1526									
10	0.2105		0.1947	0.1789		0.0421	0.0947	0.1684				0.0368		0.2211	
11	0.0474		0.1947	0.3947			0.3368	0.2526				0.3526	0.0053	0.1053	
11.2									0.0316						
12	0.0158		0.2150	0.3474			0.2105	0.1579	0.0211					0.2737	
13	0.2940		0.0842	0.0684			0.0526	0.1053	0.2421				0.0105	0.3105	
13.2									0.1000						
14	0.2316		0.0947	0.0105	0.0105		0.0263	0.1158	0.2053	0.2263			0.0895	0.0684	
14.2									0.1579						
15	0.1000		0.0053		0.3368		0.0158		0.0211	0.2158			0.2263	0.0211	
15.2									0.2211						
16	0.0737				0.3780		0.0105		0.0053		0.1000		0.1000		
17	0.0211				0.2053				0.0789		0.2947		0.3010		
18	0.0053				0.0526				0.1158		0.1158		0.0579		
19					0.0158				0.2263		0.0421		0.1263		0.0632
20									0.0158		0.0053		0.0579		0.0105
21									0.0842				0.0053		0.0158
22									0.2526				0.0053		0.0368
23									0.1211						0.2632



**Table 7.** Statistical Population Genetic Parameters of Five Populations

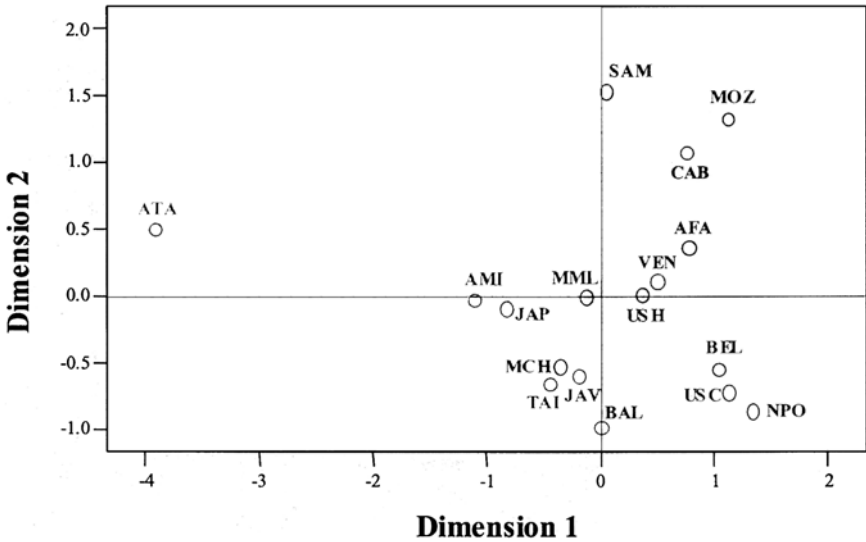
Population	Total Alleles	Combined Power of Discrimination	Average Heterozygosity	Loci with Highest Power of Discrimination	Loci with Lowest Power of Discrimination	Departures from Hardy-Weinberg Equilibrium
Ami	111	0.999999999999999	0.7710	D2S11, D18S51, FGA	TPOX	VWA
Atayal	89	0.999999999999663	0.7269	D2S1338	D3S1358, TPOX	None
Bali	118	0.999999999999999	0.7746	D2S11, D2S1338, D19S433, FGA	TPOX	VWA <sup>a</sup>
Java	129	0.999999999999999	0.7915	D8S1179, D2S1338, FGA	TPOX	None
Samoa	117	0.999999999999999	0.7799	D7S820, D2S1338	CSFIPO, TPOX	D19S433, FGA

a. Persistence of departure from Hardy-Weinberg equilibrium after Bonferroni-like adjustment for number of loci tested (0.05/15 = 0.0033)



**Figure 2.** Neighbor-joining phylogenetic analyses of 17 worldwide populations based on  $F_{ST}$  distances from STR allele frequencies. The GenDist option of the Phylip software created branch distances onto which the corresponding bootstrap values (based on 1,000 replications) were transferred to the corresponding nodes of the neighbor-joining tree.

**Partitioning of Populations Based on Geography and Language.** The distribution of genetic variance was assessed along geographic and linguistic partitioning among the Asian and Pacific Ocean populations using AMOVA. The six Austronesian populations, five from this study and one previously studied (Malaysian Malay), and three non-Austronesian reference populations from Asia (Japan, Taiwan general population, and Malaysian Chinese) were included in this analysis. Table 9 indicates the loci that exhibit statistically significant correlations and their corresponding variance values. Except for marker D19S433, all loci demonstrate no significant correlation ( $p < 0.05$ ) between genetic diversity and linguistic or geographic partitions when populations within groups are compared. In contrast to the lack of significance among populations within groups, the following five loci overlapped, showing significant correlation ( $p < 0.05$ ) between genetic diversity among groups of populations to both linguistics and geography: D8S1179, D7S820, TH01, D16S539, and D2S1338. Genetic diversity in five additional loci exhibited significant correlation ( $p < 0.05$ ) among groups of populations along linguistic lines for 10 out of 15 significant loci. A single additional locus showed significant correlation ( $p < 0.05$ ) between genetic differences and geographic partitioning among groups of populations for 6 of the 15 loci. The overall AMOVA among groups of populations along linguistic ( $p < 0.00001$ )



**Figure 3.** Multidimensional scaling analyses of 17 worldwide populations based on  $F_{ST}$  distances from STR allele frequencies. AFA, African American; AMI, Ami; ATA, Atayal; BAL, Bali; BEL, Belgium; CAB, Cabinda, Angola; JAP, Japan; JAV, Java; MCH, Malaysian Chinese; MML, Malaysian Malay; MOZ, Mozambique; NPO, North Poland; SAM, Samoa; TAI, Taiwan; USC, US Caucasian; USH, US Hispanic; VEN, Venezuela.

and geographic ( $p < 0.01$ ) lines generated significant correspondence to genetic structure. Yet the among-populations within-groups overall AMOVA gave insignificant correlation along linguistic and geographic lines.

## Discussion

These results provide novel databases for 15 autosomal STR loci in four Austronesian populations (Ami, Atayal, Bali, and Samoa). For a fifth Austronesian group (Java) new data was added in the form of loci D2S1338 and D19S433 to a preexisting database (Othman et al. 2004). An inspection of the results reveals that there is a marked lack of commonly encountered microvariant alleles among the Ami, Atayal, and Samoan groups in the highly polymorphic loci D21S11 and FGA. We also detected specific alleles in common among the published Asian and Pacific Ocean population databases used in this study as well as in our Balineses and Javanese groups. These include alleles 28.2, 29.2, and 30.2 in the D21S11 locus and alleles 21.2, 22.2, 23.2, and 25.2 at the FGA locus. Overall, D2S1338 and FGA are the most discriminating loci across all populations, and TPOX is the least discriminating locus; D2S1338 and FGA have the highest number of alleles, and TPOX has the lowest number of alleles within

**Table 8.** Components of Genetic Variance for Nine Austronesian and Asian Non-Austronesian Populations

<i>Locus</i>	<i>Intrapopulation H<sub>s</sub></i>		<i>Interpopulation G<sub>ST</sub></i>	
	<i>Austronesians</i>	<i>Non-Austronesians</i>	<i>Austronesians</i>	<i>Non-Austronesians</i>
8S1179	0.823938	0.841298	0.026409	0.001787
D21S11	0.830586	0.803820	0.023381	0.001264
D7S820	0.772001	0.765179	0.018428	0.002064
CSF1PO	0.706132	0.736505	0.010057	0.002924
D3S1358	0.700171	0.719401	0.011240	0.003285
TH01t	0.704140	0.692647	0.064438	0.008190
D13S317	0.764493	0.798947	0.036658	0.001299
D16S539	0.765375	0.779884	0.037288	0.007956
D2S1338	0.845132	0.866158	0.026127	0.004573
D19S433	0.817751	0.797587	0.015838	0.007098
VWA	0.790375	0.794881	0.026988	0.002716
TPOX	0.602920	0.613072	0.051971	0.009099
D18S51	0.806176	0.862127	0.033837	0.001645
D5S818	0.731385	0.792780	0.034100	0.002064
FGA	0.845817	0.863177	0.033127	0.001464
All loci	0.767093	0.781831	0.029760	0.003691

each population. As expected in most cases, the most polymorphic loci (highest  $H_o$  value) are the most discriminating markers (highest PD value) for each population.

Although four of the five populations under study (Ami, Atayal, Bali, and Java) clustered within the Asian/Pacific Ocean clade of the neighbor-joining tree (see Figure 2), the fact that the Polynesian group from Samoa segregates within the African groups is most notable. It is likely that genetic drift resulting from multiple bottleneck events, founder effects, and/or isolation have contributed to a genetic makeup for this Samoan population that does not reflect its ancestry. This rather unexpected result underscores the need to examine the genetic profiles of individual Pacific islands and not to consider them interchangeable for forensic analysis.

A second observation is the larger branch lengths of the Atayal and Samoan populations. These indicate a large degree of genetic differentiation in these two groups, possibly because of migratory bottlenecks, founder effects, many generations of relative isolation, and/or genetic drift. The relative positions of these two populations in the MDS analysis (see Figure 3), most notably that of the Atayal, corroborates this notion. As mentioned previously, it is interesting that the Atayal and Samoa are two of the three groups that lack microvariant alleles.

Another important observation is that the three Indo-Malaysian island groups in this study (Bali, Java, and Malay Malaysia) segregate in a different branch of the clade distant from the two Taiwanese aboriginal groups, the Ami

**Table 9.** Significant AMOVA Values for Nine Austronesian and Asian Non-Austronesian Populations<sup>a</sup>

<i>Linguistic Partitioning</i>		<i>Geographic Partitioning</i>	
<i>Among Groups of Populations</i>	<i>Among Populations Within Groups</i>	<i>Among Groups of Populations</i>	<i>Among Populations Within Groups</i>
D8S1179 (0.37)	D19S433 (3.17)	D8S1179 (0.38)	D19S433 (2.88)
D21S11 (0.89)		D7S820 (1.01)	
D7S820 (1.21)		CSF1PO (0.99)	
TH01 (2.12)		TH01 (2.61)	
D16S539 (1.70)		D16S539 (1.64)	
D2S1338 (1.10)		D2S1338 (0.98)	
VWA (0.50)			
TPOX (1.69)			
D18S51 (1.05)			
D5S818 (1.39)			

a. Numbers in parentheses refer to percentage of variance, considered significant when  $p < 0.05$ .

and Atayal. From these results it appears that the Austronesian language affiliation of these five groups is not reflected in their present genetic relationship. Analysis of variance components (see Table 8) revealed two related points: (1) The level of interpopulation differentiation among our Austronesian populations is higher than in the Asian non-Austronesian groups, and (2) conversely, the intrapopulation variance is lower in the Austronesians than in the Asian non-Austronesians for most loci. It would be expected that during geographic and cultural isolation, forces such as founder effects, inbreeding, and limited gene flow would mitigate within-population variance while acting to augment interpopulation differences.

An examination of the AMOVA results (see Table 9) for nine Austronesian and Asian non-Austronesian populations indicates that most of the loci exhibit significant genetic partitioning along linguistic and geographic lines among groups of populations. On the other hand, only one locus, D19S433, showed significant correlations with both language and geography among populations within groups. It is likely that high levels of polymorphism and differences in allele frequencies particular to this locus among the nine populations provide the fine resolution necessary to detect genetic variability at the among-populations within-groups level.

Similarly, overall statistically significant correlation along linguistic and geographic partitioning was detected only among groups of populations. It is expected that a greater number of loci will generate significant correlations with linguistic and geographic partitioning among groups of populations rather than among populations within groups because genetic differences are generally

greater at the among-populations level. It is worth mentioning that a greater number of loci display significant correspondence between genetics and linguistics compared to genetic and geographic partitioning at the among groups of populations level. These results indicate that division based on language groups is in better agreement with the genetic structure of these nine Austronesian and Asian non-Austronesian populations.

Focusing on a smaller geographic scale, the aboriginal populations of Formosa provide a unique opportunity to dissect the effects of social and cultural relationships on the genetic makeup of neighboring populations. The Ami and Atayal groups represent two of the nine extant indigenous tribes of Taiwan. The Ami inhabit the narrow eastern seacoast plains of the island and represent the largest tribal group, approximately 130,000 in number. The Atayal are the second largest tribe (about 90,000) and reside adjacent to the Ami in the mountainous terrain of northern Taiwan. Historical accounts cite continuous waves of migrants from the Asian mainland who displaced the indigenous tribes and forced them into the less accessible areas of the island, hence leading to their current distribution (Knapp 1980).

A number of previous studies using mitochondrial DNA (Lum et al. 1994; Redd et al. 1995; Melton et al. 1995, 1998) found that the Ami and Atayal aboriginal groups have a large amount of mtDNA sequence homology, suggesting a common ancestral source in central or southern China. However, the investigators also reported evidence of a prolonged isolation from mainland Chinese and other Asian influences in the recent past. Studies based on autosomal (Sewerin et al. 2002; Chow et al. 2005) and mtDNA (Horai et al. 1995) markers have also demonstrated genetic uniqueness among the indigenous groups, implying varying temporal and/or spatial sources for the initial colonization of Formosa. One paternal lineage study found that the Ami stand out from the other aboriginal groups because of their closer genetic association with both South China and the Philippines. This is illustrated by the Ami's high frequency of haplogroup L in contrast to the other four aboriginal groups, which lack this haplogroup altogether.

The distinctness of aboriginal groups from each other is especially evident in the extremely homogeneous Atayal, whose Y chromosomes are almost entirely of a single haplogroup (haplogroup H) (Capelli et al. 2001). This theme is echoed in older studies using classical markers (Cavalli-Sforza et al. 1994). In the present study both the Ami and the Atayal separate from the other populations from Asia and the Pacific Ocean in both the dendrogram and the MDS plot. In turn, the extreme branch distance of the Ami and, in particular, the Atayal is consistent with the intertribal group differentiation described in the literature. Although the ranges of these two tribes share a border on the northeast of the island, it is likely that regional geographic, cultural, and linguistic barriers played a role in the genetic differentiation of these two Taiwanese aboriginal groups. These findings support previous genetic studies with regard to differences between tribes.

Similar to the Taiwanese tribal groups, the populations of the islands of Bali and Java lie in close proximity within the Indo-Malaysian archipelago, separated by mere miles. In a recent study of the general Javanese population (Othman et al. 2004) that reported a battery of autosomal STRs that coincide with 13 of the loci reported in the present work, we found no marked differences in allele frequencies compared to our data. Both Bali and Java belong to the same Western Malayo-Polynesian subgroup of the Austronesian language family and, not surprisingly, segregate together within our neighbor-joining tree and MDS analysis. Also, as could be anticipated, the Malaysian Malay population clusters together in the same subclade with the Bali and the Java groups. Contrary to expectations based on geography alone, the two populations are clearly more distinct from each other than even the population of Han Chinese is from the Japanese. In the MDS analysis the Javanese group plots closer to the Han Chinese from Malaysia and Taiwan than to its nearest neighbors from Bali. This may be a reflection of admixture of the Javanese population with an influx of Muslim, Indian, and mainland Chinese in the recent past. In historical times Java has been subject to waves of Buddhist, Hindu, and Muslim migrations, generating cultural diversity and possibly genetic heterogeneity. In addition, Java is more than 23 times larger in area than Bali, and it is possible that this sheer difference in area allows a larger effective population size. On the other hand, Bali has remained relatively isolated both culturally and genetically.

Samoa, the most distant population from the Asian mainland examined in the present study, segregates as an outlier in both the dendrogram and MDS analyses. These islands lie near the boundary dividing Polynesia from both Melanesia to the west and Micronesia to the northwest and therefore represent a population at the crossroads of the Pacific Ocean. Our phylogenetic analyses of this Polynesian group indicate no apparent genetic relationship with the other Austronesian populations. In fact, the Samoans segregate at an intermediate position within the African cluster in the neighbor-joining tree and plot closer to the African groups in the MDS analysis. The relatively long branch length in the phylogram and isolation in the MDS analysis is indicative of the genetic uniqueness of this particular group compared to the other populations examined in this study.

In a previous work using five Y-chromosome STRs, similar results were noted in a Western Samoan population and were attributed to unique allele distributions compared to other Asian and Pacific Ocean groups (Parra et al. 1999). Parra and colleagues (1999) suggested founder effects stemming from the initial Austronesian settlement of Oceania as a likely cause. It is possible that the pointed genetic distinctiveness of the Samoans that provides for their segregation within the African cluster is the result of extreme genetic drift.

It is well documented that early Polynesians had reached as far as Samoa by means of the northern islands of Tonga by about 3,000 years B.P. Here, the Polynesian language evolved into two different subgroups: the Tongic and Nuclear Polynesian subgroups. The Nuclear Polynesian subgroup contains the Eastern Polynesian and Samoan Outlier languages. Archeological and linguistic data

suggest that further colonization of the eastern islands of Polynesia occurred via Samoa within the last 2,000 to 800 years before present (Bellwood 1978; Kirch 1997). This recent 1,000-year layover in Samoa may have led to the severe genetic distinctness that is observed in both the dendrogram and the MDS plot and paralleled by the documented linguistic subdifferentiation.

Our results also could be interpreted as indicating a considerable genetic contribution to the Samoan gene pool from another source, such as neighboring Melanesia. In this scenario, depending on the proportion of a Papuan genetic component, the Samoans would be expected to appear as genetically distinct from the other groups examined in this study. This possibility is supported by the findings of a previous study in which biparental STRs displayed a pattern consistent with an initial Austronesian expansion into Remote Oceania from Southeast Asia followed by a significant amount of gene flow from Near Oceania (Lum et al. 2002). However, although our phylogenetic data on the Samoan population may be thought-provoking, inclusion of additional relevant Pacific Ocean populations, including the Papuans, needs to be studied in order to examine these issues.

*Acknowledgments* We gratefully acknowledge Laisel Martinez, Diane J. Rowold, and Maria Christina Terreros for their constructive criticism of the manuscript.

*Received 13 May 2005; revision received 5 September 2005.*

## Literature Cited

- Alves, C., L. Gusmao, A. Damasceno et al. 2004. Contribution for an African autosomal STR database (AmpFISTR Identifiler and Powerplex 16 system) and a report on genotypic variations. *Forensic Sci. Int.* 139:201–205.
- Antunez de Mayolo, G., A. Antunez de Mayolo, P. Antunez de Mayolo et al. 2002. Phylogenetics of worldwide human populations as determined by polymorphic *Alu* insertions. *Electrophoresis* 23:3346–3356.
- Beleza, S., C. Alves, F. Reis et al. 2004. 17 STR (AmpFISTR Identifiler and Powerplex 16 system) from Cabinda (Angola). *Forensic Sci. Int.* 141:193–196.
- Bellwood, P. 1978. *The Polynesians*. London: Thames and Hudson.
- Bellwood, P. 2001. Early agriculturalist population diasporas? Farming, languages and genes. *Annu. Rev. Anthropol.* 30:181–207.
- Bosch, E., F. Calafell, A. Perez-Lezaun et al. 2000. Genetic structure of northwest Africa revealed by STR analysis. *Eur. J. Hum. Genet.* 8:360–366.
- Bowcock, A., A. Ruiz-Linares, J. Tomfohrde et al. 1994. High resolution of human evolutionary trees with polymorphic microsatellites. *Nature* 368:455–457.
- Brenner, C., and J. Morris. 1990. Paternity index calculations in single locus hypervariable DNA probes: Validation and other studies. In *Proceedings for the International Symposium on Human Identification 1989*. Madison, WI: Promega, 21–53.
- Butler, J. M. 2001. *Forensic DNA Typing: Biology and Technology Behind STR Markers*. London: Academic Press.

- Butler, J. M., R. Schoske, P. M. Vallone et al. 2003. Allele frequencies for 15 autosomal STR loci on U.S. Caucasian, African American, and Hispanic populations. *J. Forensic Sci.* 48(4):908–911.
- Capelli, C., J. F. Wilson, M. Richards 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.
- Carmody, G. 1990. *G-Test*. Ottawa, Canada: Carleton University.
- Cavalli-Sforza, L. L., P. Menozzi, and A. Piazza. 1994. *The History and Geography of Human Genes*. Princeton, NJ: Princeton University Press.
- Chiurillo, M. A., A. Morales, A. M. Mendes et al. 2003. Genetic profiling of a central Venezuelan population using 15 STR markers that may be of forensic importance. *Forensic Sci. Int.* 136:99–101.
- Chow, R. A., J. L. Caeiro, S. J. Chen et al. 2005. Genetic characterization of four Austronesian-speaking populations. *J. Hum. Genet.* (in press).
- Collins, P. J., L. K. Hennessy, C. S. Leibelt et al. 2004. Developmental validation of a single-tube amplification of the 13 CODIS STR loci, D2S1338, D19S433 and amelogenin: The AmpFISTR Identifiler PCR Amplification kit. *J. Forensic Sci.* 49(6):1265–1277.
- Decorte, R., M. Engelen, L. Larno et al. 2004. Belgian population data for 15 STR loci (AmpFISTR SGM Plus and AmpFISTR Profiler PCR amplification kit). *Forensic Sci. Int.* 139:211–213.
- Excoffier, L., P. E. Smouse, and J. M. Quattro. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. *Genetics* 131:479–491.
- Felsenstein, J. 2002. *Phylogeny Inference Package (PHYLIP), Version 3.6a3*. Distributed by author. Seattle: Department of Genetics, University of Washington.
- Guo, S., and E. Thompson. 1992. Performing the exact test of Hardy-Weinberg proportion for multiple alleles. *Biometrics* 48:361–372.
- Hashiyada, M., Y. Itakura, T. Nagashima et al. 2003. Polymorphism of 17 STRs by multiplex analysis in Japanese population. *Forensic Sci. Int.* 133:250–253.
- Horai, S., K. Hayasaka, R. Kondo et al. 1995. Recent African origin of modern humans revealed by complete sequences of hominoid mitochondrial DNAs. *Proc. Natl. Acad. Sci. USA* 92:532–536.
- Jones, D. A. 1972. Blood samples: Probability of discrimination. *J. Forensic Sci. Soc.* 12:355–359.
- Jorde, L. B., M. J. Bamshad, W. S. Watkins et al. 1995. Origins and affinities of modern humans: A comparison of mitochondrial and nuclear genetic data. *Am. J. Hum. Genet.* 57:523–538.
- Jorde, L. B., A. R. Rogers, M. Bamshad et al. 1997. Microsatellite diversity and the demographic history of modern humans. *Proc. Natl. Acad. Sci. USA* 94:3100–3103.
- Kirch, P. V. 1997. *The Lapita Peoples: Ancestors of the Oceanic World*. Cambridge, MA: Blackwell.
- Knapp, R. G. 1980. *China's Final Frontier: Studies in the Historical Geography of Taiwan*. Taipei: SMC Publishing.
- Leibelt, C., B. Budowle, P. Collins et al. 2003. Identification of a D8S1179 primer binding site mutation and the validation of a primer designed to recover null alleles. *Forensic Sci. Int.* 133(3):220–227.
- Levene, H. 1949. On a matching problem arising in genetics. *Ann. Math. Stat.* 20:91–94.
- Li, C. C. 1976. *First Course in Population Genetics*. Pacific Grove, CA: Boxwood Press.
- Lum, J. K., L. B. Jorde, and W. Schiefenovel. 2002. Affinities among Melanesians, Micronesians, and Polynesians: A neutral biparental genetic perspective. *Hum. Biol.* 74(3):413–430.
- Lum, J. K., O. Rickards, C. Ching et al. 1994. Polynesian mitochondrial DNAs reveal three deep maternal lineage clusters. *Hum. Biol.* 66:567–590.
- Melton, T., S. Clifford, J. J. Martinson et al. 1998. Genetic evidence for the Proto-Austronesian homeland in Asia: mtDNA and nuclear DNA variation in Taiwanese aboriginal tribes. *Am. J. Hum. Genet.* 63:1807–1823.
- Melton, T., R. Peterson, A. J. Redd et al. 1995. Polynesian genetic affinities with Southeast Asian populations as identified by mtDNA analysis. *Am. J. Hum. Genet.* 57:403–414.

- Nei, M. 1987. *Molecular Evolutionary Genetics*. New York: Columbia University Press.
- Ota, T. 1993. *DISPAN: Genetic Distance and Phylogenetic Analysis*. University Park, PA: Institute of Molecular Evolutionary Genetics, Pennsylvania State University.
- Othman, M. I., L. H. Seah, S. Panneerchelvan et al. 2004. Allele frequencies for the PowerPlex 16 STR loci in Javanese population from Malaysia. *J. Forensic Sci.* 149(1):190–191.
- Parra, E., M. D. Shriver, A. Soemantri et al. 1999. Analysis of five Y-specific microsatellite loci in the Asian and Pacific populations. *Am. J. Phys. Anthropol.* 110(1):1–16.
- Pérez-Miranda, A. M., M. A. Alfonso-Sánchez, A. Kalantar et al. 2005. Allelic frequencies of 13 STR loci in autochthonous Basques from the province of Vizcaya (Spain). *Forensic Sci. Int.* 152(2–3):259–262.
- Redd, A. J., N. Takezaki, S. T. Sherry et al. 1995. Evolutionary history of the COII/tRNA<sup>Lys</sup> intergenic 9 base pair deletion in human mitochondrial DNAs from the Pacific. *Mol. Biol. Evol.* 12:604–615.
- Reynolds, J., B. S. Weir, and C. C. Cockerham. 1983. Estimation of the coancestry coefficient: Basis for a short term genetic distance. *Genetics* 105:767–779.
- Rowold, D. J., and R. J. Herrera. 2003. Inferring recent human phylogenies using forensic STR technology. *Forensic Sci. Int.* 133:260–265.
- Schneider, S., J.-M. Kueffer, D. Roessli et al. 2000. *Arlequin v. 2000: A Software for Population Genetics Data Analysis*. Geneva: Genetics and Biometry Laboratory, University of Geneva.
- Seah, L. H., N. H. Jeevan, M. I. Othman et al. 2003. STR data for the AmpFISTR Identifier loci in three ethnic groups (Malay, Chinese, Indian) of the Malaysian population. *Forensic Sci. Int.* 138:134–137.
- Sewerin, B., F. J. Cuza, M. N. Szmulewicz et al. 2002. On the genetic uniqueness of the Ami aborigines of Formosa. *Am. J. Phys. Anthropol.* 119:240–248.
- Shepard, E. M., and R. J. Herrera. 2005. Iranian STR variation at the fringes of biogeographical demarcation. *Forensic Sci. Int.* (in press).
- SPSS Inc. 2001. *SPSS for Windows, Release 11.0.1*. Chicago: SPSS Inc.
- Szczerkowska, Z., E. Kapinska, J. Wysocka et al. 2004. Northern Polish population data and forensic usefulness of 15 autosomal STR loci. *Forensic Sci. Int.* 144:69–71.
- Tereba, A. 1999. Tools for analysis of population statistics. In *Profiles in DNA*, I. MacIver, ed. Madison, WI: Promega Corporation, v. 2, pp. 14–16.
- Underhill, P. A. 2004. A synopsis of extant Y chromosome diversity in East Asia and Oceania. In *The Peopling of East Asia: Putting Together Archeology, Linguistics, and Genetics*, L. Sagart, R. Blench, and A. Sanchez-Mazas, eds. London: Routledge Curzon, ch. 17, pp. 301–319.
- Wang, C.-W., D.-P. Chen, C.-Y. Chen et al. 2003. STR data for the AmpFISTR SGM Plus and Profiler loci from Taiwan. *Forensic Sci. Int.* 138:119–122.
- Watkins, W. S., A. R. Rogers, C. T. Ostler et al. 2003. Genetic variation among world populations: Inferences from 100 *Alu* insertion polymorphisms. *Genome Res.* 13:1607–1618.