Terry Melton,[1] *Ph.D. and Mark Stoneking,*[1] *Ph.D.*

Extent of Heterogeneity in Mitochondrial DNA of Ethnic Asian Populations

REFERENCE: Melton, T. and Stoneking, M., **"Extent of Heterogeneity in Mitochondrial DNA of Ethnic Asian Populations,"** *Journal of Forensic Sciences,* JFSCA, Vol. 41, No. 4, July 1996, pp. 591–602.

ABSTRACT: Variation in the mitochondrial DNA (mtDNA) control region as detected by sequence-specific oligonucleotide (SSO) probes is described for 993 individuals in 11 ethnic Asian populations. Estimates of diversity for mtDNA types exceed 0.94 in all populations, while 53% of the 255 types that were observed occur only once. Of 96 shared types, four occur at frequencies of greater than 10% but less than 17% in any one population. There is statistically significant heterogeneity among these 11 populations, however, an analysis of variance incorporating genetic distances between types shows that at least 95% of the variation present in the total sample is attributable to within-population diversity, while only 5% is due to between-population differences. Overall, heterogeneity with respect to mtDNA SSO types is grossly correlated with geographic distance between populations; the most extreme heterogeneity was observed between populations from East Asia and populations from West Asia. With respect to population genetics, the control region of mtDNA exhibits satisfactory qualities as a DNA typing locus.

KEYWORDS: forensic science, mitochondrial DNA, population genetics, Asian populations, sequence-specific oligonucleotide typing, analysis of molecular variance

Polymorphisms observed in the haploid genetic system of mitochondrial DNA provide a valuable source of discrimination between individuals (1–6). MtDNA matching in forensic investigations is increasing (7–10) while guidelines for its use are becoming available to the forensics community (11,12). Although sequencing of the hypervariable control region of mtDNA would provide the highest resolution for matching samples in a forensic investigation, an inexpensive alternative exclusionary technique is sequence-specific oligonucleotide (SSO) typing (7). In this system, oligonucleotide probes can detect an enormous amount of nucleotide variation at 13 sites across the control region. For example, SSO typing of 525 individuals from five ethnic groups revealed that there is only a 2.6% probability that two unrelated individuals will share the same SSO type (7).

The issue of undetected population genetic substructure has been hotly debated for VNTR loci (13–18). When allele frequencies of nuclear genes in unsampled ethnic subpopulations deviate substantially from those measured in major racial groups, match probabilities could, in theory, be biased against a suspect. It is legitimate to raise the question of heterogeneity of mtDNA types within racial groups as well. The maternal mode of inheritance and lack of recombination in mtDNA (19) have justified numerous studies of population history (1,2,5,20–22); individuals within populations share nucleotide substitutions that reflect a common source and that have persisted through population expansions.

SSO typing has detected heterogeneity among African-Americans, Asians, European-Americans, Japanese, and Mexicans based on frequencies of nucleotide sequence variants (7). In addition, logistic regression methods revealed that mtDNA types are somewhat predictive of ethnicity (23). However, within the broad racial classifications which are conventionally used to define U.S. forensics databases (African, European, Asian, Hispanic, and Native American), there has been no examination of heterogeneity of mtDNA types among ethnic subpopulations. As part of a project to examine variation among African, Asian, and European subpopulations, we report here the extent of heterogeneity for mtDNA SSO types within reasonably well-defined ethnic subpopulations of Asians.

Methods

In this study, 993 individuals were typed for sequence variants at 13 nucleotide positions in 8 regions across the mtDNA control region. All samples were purified genomic DNA obtained from maternally unrelated individuals. The 11 populations studied are from the following areas:

- Borneo: 91 samples from the Barito River area (southeast Borneo).
- Bangladesh: 31 samples collected from immigrants (male workers) to Singapore.
- China: 103 samples from southern (Han) Chinese immigrants to Singapore.
- Philippines: 60 samples from immigrants (female workers) to Singapore, 96 samples from the northern island of Ilocano (provided by Roche Laboratories), and 19 samples from Filipino U.S. military personnel.
- Southern India: 73 samples from immigrants to Singapore from southern India and Sri Lanka, primarily of Dravidian origin.
- Northern India: 47 samples from Sikhs who immigrated to Singapore.
- East Indonesia: 93 samples from the Nusa Tenggaras (Alor, 21; Flores, 21; Roti, 26; Timor, 25), and 49 samples from the Moluccas (Hiri, 26; Ternate, 23).

[1]Department of Anthropology, The Pennsylvania State University, University Park, PA.

This project was supported by NIJ grant 92-IJ-CX-K040 to M.S. Portions of this work were presented at the American Academy of Forensic Sciences meeting in Seattle, 14–18 February 1995.

Received for publication 14 Aug. 1995; revised manuscript received 13 Nov. 1995; accepted for publication 22 Dec. 1995.

- Java: 96 samples from rural areas of central Java, near Semarang.
- Malaysia: 81 samples from native Malays.
- Pakistan: 73 samples from Pushtoons of Peshwar in northwest Pakistan.
- Taiwan: 81 samples from four aboriginal groups (Ami, 22; Atayal, 20; Bunun, 19; Paiwan, 20).

Figure 1 shows the locations and sample sizes of these populations and their component subgroups.

The SSO typing method used here was described in detail elsewhere (21). Seventeen of the 21 probes were identical to those used in a previous study (7), while four others used in that study (IA3, IC2, IC3, IIA2) were slightly modified by the addition or removal of nucleotides from the 5′ or 3′ ends to increase specificity (21). Probes IIE1 and IIE2 (7) were not used in this study because of problems with cross-hybridization, which presumably arise because these probes detect variation in the length of a run of consecutive cytosine residues, and not single nucleotide substitutions.

Individual results of the SSO typing at each variant region were arranged into a composite mtDNA SSO type. For example, the mtDNA type 1-1-2-1-2-1-1-0 indicates that probe variant IA1 annealed in the IA region, IB1 annealed in the IB region, etc. A "0" indicates that a blank result for the above example was obtained for the IID variant. A blank result for IID occurs either when a substitution in a nearby site prevents probe annealing, or when a nucleotide other than A or G is present at position 247 (the IID probe specific site). While blanks in different individuals for a probe region could reflect different substitutions, for the purpose of analysis blanks are considered to be the same variant. For forensic purposes, an SSO type, or profile, is equivalent to a single locus which may be compared among individuals; type frequencies are therefore much like allele frequencies.

The data set, consisting of SSO types of 993 Asians, was examined with respect to type frequency distribution both overall and within 11 populations and eight subgroups within three of these populations (Filipino, E. Indonesian, and Taiwanese). An unbiased estimate of diversity (h) was calculated to quantify the amount of mtDNA variation present in each population and subgroup:

$$h = (1 - \Sigma x^2)n/(n - 1) \qquad (1)$$

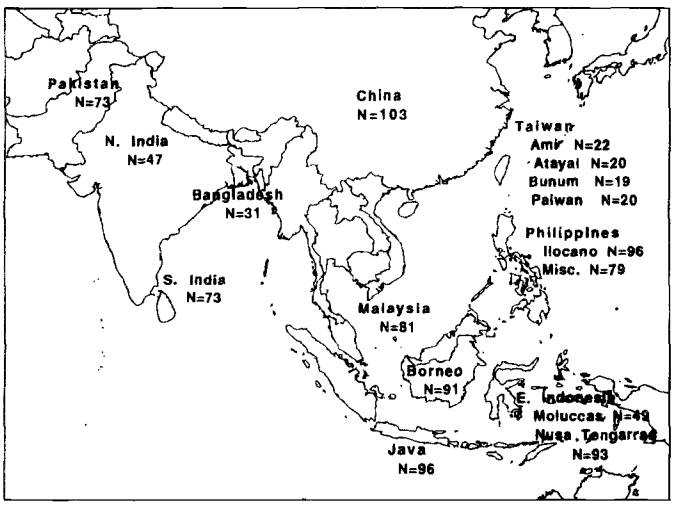

FIG. 1—*Locations and sample sizes of populations in this study.*

where n is the sample size and x is the frequency of each mtDNA type (24). The variance for this estimate of diversity was calculated as

$$V(h) = \frac{2[\Sigma x^3 - (\Sigma x^2)^2]}{n} \qquad (2)$$

(25). Since the probes were chosen because they detect variation, the diversity estimates are biased upwards. However, for the purposes of this analysis comparisons between populations are valid since the same probes were used for each population.

Among-population heterogeneity was examined using several approaches. Either G or chi-square goodness-of-fit tests for heterogeneity of the probe variant frequencies across all 11 populations were used to determine whether populations were significantly different with respect to individual probe variants (26). An analysis of molecular variance (AMOVA, 27) was applied to the SSO types to measure the apportionment of diversity within and among the 11 populations and their component subgroups. AMOVA is especially useful for analysis of mtDNA data since it does not require independence of nucleotide sites. This method incorporates information about genetic distances between pairs of mtDNA types to enhance a more traditional computation of variance components and F-statistics from mtDNA type frequency data, to evaluate population subdivision (28–30). A conventional sum of squared deviations is partitioned into variance components attributable to variation among regions (σ_a^2), among populations within regions (σ_b^2), and within populations (σ_c^2). At present, a maximum of 255 SSO types can be analysed by the available software (L. Excoffier, pers. communication); for this analysis, 993 individuals with 255 mtDNA types were selected from a larger sample of 1007 individuals with 269 SSO types (choosing several different subsets of 255 mtDNA SSO types gave no significant differences in the results, as discussed below).

A genetic distance between each pair of SSO types was calculated for use in the analysis of molecular variance. For each variant region, the total number of nucleotide differences between two SSO types was counted; this count was then summed over all variant regions. In general, this procedure is similar to counting the number of site differences between restriction haplotypes, with changes made for SSO data by adjusting for "blanks" and for probes which detect two polymorphic sites instead of only one. Although a blank result could be the result of one or more nucleotide substitutions, for most populations there is a low frequency of blanks (<5%), usually attributable to a single site change, so the comparison between a blank variant and any other variant is always counted as one difference.

Analogs to Wright's F-statistics (Φ-statistics) were also generated by AMOVA (31). Φ_{ST} is the correlation of random SSO types within populations, relative to that of random pairs of types drawn from the entire data set; Φ_{CT} is the correlation of random SSO types within a regional group of populations, relative to that of random pairs of types drawn from the entire data set, and Φ_{SC} is the correlation between random pairs of SSO types within populations, relative to that of random pairs of SSO types drawn from the region (27). Where the total variance σ^2 is the sum of components σ_a^2, σ_b^2, and σ_c^2, the Φ-statistics can be rewritten as

$$\Phi_{ST} = \frac{\sigma_a^2 + \sigma_b^2}{\sigma^2}, \qquad \Phi_{CT} = \frac{\sigma_a^2}{\sigma^2}, \qquad \Phi_{SC} = \frac{\sigma_b^2}{\sigma_b^2 + \sigma_c^2} \qquad (3)$$

(28,29). Φ_{CT} and Φ_{ST} effectively indicate the extent of population subdivision based on whatever hierarchy of populations is stipulated based on geographic, linguistic, ethnic or historical affinities.

To determine which populations were different or similar with respect to SSO types, Φ_{ST} values for pairs of populations (also analogous to a coefficient of coancestry) were calculated by AMOVA in a matrix containing all populations. Permutational procedures in AMOVA were used to test the significance of Φ-statistics and variance components. For each AMOVA analysis, a null distribution was generated by allocating every individual to a randomly chosen population while holding sample sizes constant over a large number of permutations (generally 500–1000). Probabilities of observing random variance components and Φ-statistics greater than those generated in the analysis were reported. As the null distributions of the Φ-statistics are highly correlated with those of their associated variance components (27), one *P*-value which encompasses both tests was reported. This method of significance testing is useful because concerns about the normality of underlying variance distributions can be disregarded (27).

After determining which pairs of populations were not significantly different with respect to SSO types, trial clusters of populations were designated and analyzed by AMOVA to determine what portion of the total variance was attributable to among region, among population/within region, or within population variation.

Results

As mentioned, 255 mtDNA types (of a possible 269) were included in these analyses (Appendix). The 14 types removed from the original data set to meet AMOVA parameters consisted of the 13 singleton types in the data set of 269 types that had three or more "blank" variants, plus one additional randomly chosen singleton type. These SSO types were removed initially because less information is present in a type with multiple blanks, however, trials of choosing 255 types by removing random singleton types gave nearly identical AMOVA results. Figure 2 shows the distribution of SSO types within the total sample of 993 individuals, while Table 1 shows the distribution of the number of types present in only one population or shared by more than one population. Of 255 types, there were 134 unique types (53%) that occurred once ("unique types"), while there were 25 types (10%) that occurred more than once but occurred in only one population ("population specific types"). Of the latter, all but three SSO types occurred just two or three times. Of these three remaining types, one occurred in five E. Indonesians (distributed among the Moluccas and Nusa Tengarras) while two others occurred five and eight times apiece in the Taiwanese (in the Atayal and Bunun subgroups). Ninety-six types (37%) were observed more than once and were shared among populations, for example, 35 types were shared by two populations, and only one type was shared by nine populations. No mtDNA type was shared by 10 or all 11 populations.

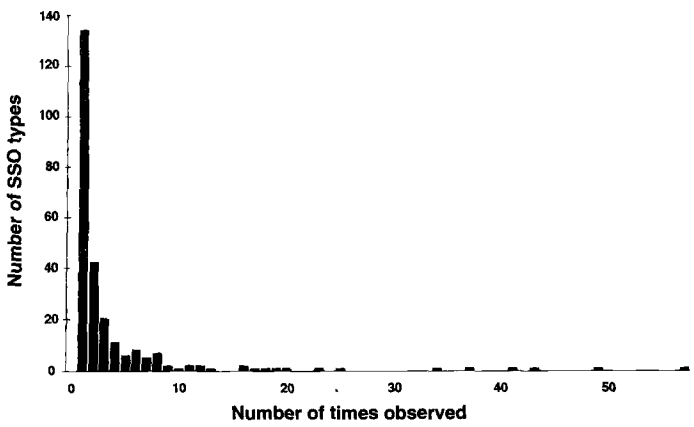

FIG. 2—*Distribution of mtDNA SSO types.*

TABLE 1—*MtDNA SSO type sharing among populations: for example, 35 types are shared by two populations.*

Number of mtDNA types	Number of populations sharing
159	1
35	2
21	3
11	4
9	5
8	6
5	7
6	8
1	9
0	10
0	11

Figure 3 shows the frequency distribution of types within each population based on whether they were unique types (occurred one time in one population), population specific types (occurred more than one time in one population), or public types (shared between two or more populations). The Bangladeshi population possessed 27 unique types (87% of the sample, $N = 31$), and had two shared types which occurred more than once (13% of the sample). By comparison, in the aboriginal Taiwanese population the frequency of population specific types approached 30%. Other populations were intermediate in their distribution of types. In general, in most populations there was a low frequency of population specific types, while the frequency of public types exceeded that of unique types.

Table 2 shows the number of individuals and number of mtDNA types observed in each population along with estimates of population diversity (h) and their standard errors. Also shown are these results for the subgroups within the Filipino, E. Indonesian, and Taiwanese samples. Estimates of diversity were high, exceeding 0.94 for all populations, and were slightly lower in the subgroups.

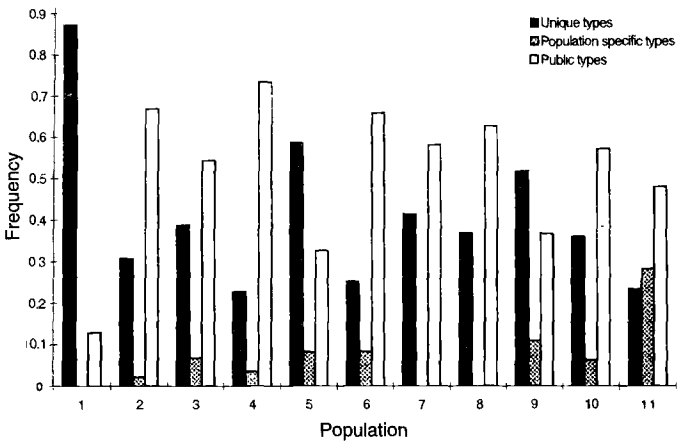

FIG. 3—*Frequency distribution of unique types, population specific types, and public types among populations. 1-Bangladesh, 2-Borneo, 3-China, 4-Philippines, 5-S. India, 6-E. Indonesia, 7-Java, 8-Malaysia, 9-Pakistan, 10-N. India, 11-Taiwan.*

TABLE 2—*MtDNA SSO type diversity* (h) *for 11 Asian populations.*

Population	*N*	No. SSO types	$h \pm$ s.e.
Bangladeshi	31	29	0.996 ± 0.003
Borneo	91	46	0.969 ± 0.005
Chinese	103	59	0.980 ± 0.003
Filipino	175	60	0.949 ± 0.005
Ilocano	96	31	0.927 ± 0.009
Miscellaneous	79	45	0.969 ± 0.006
S. Indian	73	55	0.990 ± 0.002
E. Indonesian	142	60	0.969 ± 0.004
Moluccas	49	26	0.961 ± 0.008
Nusa Tenggaras	93	46	0.966 ± 0.006
Javanese	96	53	0.961 ± 0.008
Malay	81	45	0.975 ± 0.004
Pakistani	73	51	0.987 ± 0.003
N. Indian	47	27	0.968 ± 0.006
Taiwanese	81	36	0.963 ± 0.005
Ami	22	13	0.944 ± 0.015
Atayal	20	9	0.890 ± 0.023
Bunun	19	9	0.866 ± 0.035
Paiwan	20	12	0.937 ± 0.018
Total	993	255	0.996 ± 0.000

For example, the four aboriginal Taiwanese subgroups had diversity estimates ranging from 0.866 to 0.944, all of which were below the estimates for any other population reported here.

The most common mtDNA types are described in Table 3. Any shared mtDNA type which occurred eight or more times (that is, frequency >0.008) in the data set was included. Only four of these 28 types occurred in any population at a frequency greater than 10%; these are further described in Fig. 4. For example, type 86 was observed eight times, distributed in four populations, and was observed in 11% of Northern Indians. Similarly, types 136, 172, and 234 were observed at frequencies of 14% and 13% in Filipinos, and 16% in Javanese, respectively. While type 86 occurs primarily in W. Asian populations, the remaining three types occur primarily in S.E. Asian populations. For all four types occurring in any population at a frequency greater than 10% there is a broad range of frequencies observed within the different populations. The standard errors for frequency of all common types, when calculated across all 11 populations, generally equalled or exceeded the mean, indicating the wide range of population frequencies; that is, a type may have been seen not at all or once in some populations and multiple times in others.

TABLE 3—*Public mtDNA types occurring eight or more times.*

MtDNA type number	Number of times observed	Number of populations	Maximum frequency in single population	Mean frequency over all populations ± s.e.
86*	8	4	0.11	0.01 ± 0.03
168	8	6	0.03	0.01 ± 0.01
169	8	5	0.02	0.01 ± 0.01
204	8	5	0.06	0.01 ± 0.02
224	8	3	0.04	0.01 ± 0.01
231	8	3	0.07	0.01 ± 0.02
176	9	5	0.04	0.01 ± 0.02
197	9	7	0.02	0.01 ± 0.01
110	10	5	0.05	0.01 ± 0.02
115	11	7	0.04	0.01 ± 0.01
252	11	5	0.05	0.01 ± 0.02
109	12	5	0.09	0.02 ± 0.03
245	12	5	0.08	0.01 ± 0.02
47	13	6	0.06	0.02 ± 0.02
69	16	6	0.03	0.01 ± 0.01
139	16	9	0.03	0.02 ± 0.01
126	17	7	0.09	0.02 ± 0.03
133	18	7	0.06	0.02 ± 0.02
111	19	8	0.06	0.02 ± 0.02
105	20	6	0.04	0.02 ± 0.02
241	23	4	0.09	0.01 ± 0.03
59	25	8	0.09	0.03 ± 0.03
174	34	6	0.09	0.03 ± 0.03
52	37	8	0.09	0.03 ± 0.03
136*	41	8	0.14	0.03 ± 0.04
131	43	8	0.09	0.04 ± 0.03
172*	49	8	0.13	0.04 ± 0.04
234*	57	7	0.16	0.05 ± 0.05

*Described further in Fig. 4.

Chi-square and G goodness-of-fit tests applied to probe variant frequencies across all 11 populations showed statistically significant levels of heterogeneity for each set of probes (data not shown). In Asians overall, only region IIA showed low levels of heterogeneity; 97.4% of all individuals carry the variant IIA2. Because all mtDNA sites are tightly linked and there is no recombination, subsequent analyses were limited to those which examine the mtDNA type as a whole, analogous to a restriction haplotype.

Initially, AMOVA was used to generate intrapopulation Φ_{ST} distances (analogous to a coancestry coefficient) for the subgroups represented within the larger groups of Filipinos (Ilocano and Miscellaneous subgroups, the latter made up of the migrant and military samples), E. Indonesians (Moluccas and Nusa Tenggaras subgroups), and Taiwanese (Ami, Atayal, Bunun, and Paiwan subgroups). There was no significant difference with respect to SSO type distribution between the Ilocano and Miscellaneous subgroups within the Filipino population ($P = 0.76$), or between the Moluccas and Nusa Tenggaras subgroups within the E. Indonesian population ($P = 0.05$), although the Indonesian comparison borders statistical significance. With numerous tests being done, to avoid spurious conclusions the significance level was arbitrarily set to be $P < 0.01$. For the Taiwanese subgroups, the Ami and Paiwan clustered together ($P = 0.62$), as did the Atayal and Bunun ($P = 0.49$), with a significant genetic distance between these two larger combinations ($P < 0.002$). For further analyses, therefore, subgroups of the Filipino and E. Indonesian populations were disregarded, while the Taiwanese population was split into the Taiwanese 1 (Ami and Paiwan) and Taiwanese 2 (Atayal and Bunun) populations.

Subsequent analyses included ten original populations plus the two Taiwanese populations derived from subdivision. First, AMOVA Φ_{ST} distances from the pairwise matrix for all 12 populations revealed which pairs of populations were not significantly different (defined here as $P > 0.01$). These pairs are shown in Table 4. Strong associations were indicated by a high probability ($P > 0.10$) of finding similar Φ_{ST} distances with permutation testing (1000 replications). Pairs of populations which were not significantly different with respect to SSO types were Chinese-Malay, Bangladeshi-S. Indian, Javanese-E. Indonesian, Pakistani-N. Indian, and Javanese-Malay. In other words, for example, a random Chinese individual could not be distinguished from a random Malay individual by mtDNA SSO typing. Pairs of populations exhibiting weaker genetic identity, that is, approaching statistically significant difference, are also shown ($0.10 > P > 0.01$). Table 5 shows the pairwise Φ_{ST} values and statistical tests for all twelve populations. All additional comparisons between pairs of populations showed significant differences in SSO variation ($P < 0.01$); hence heterogeneity is apparent for the majority of population comparisons (82%).

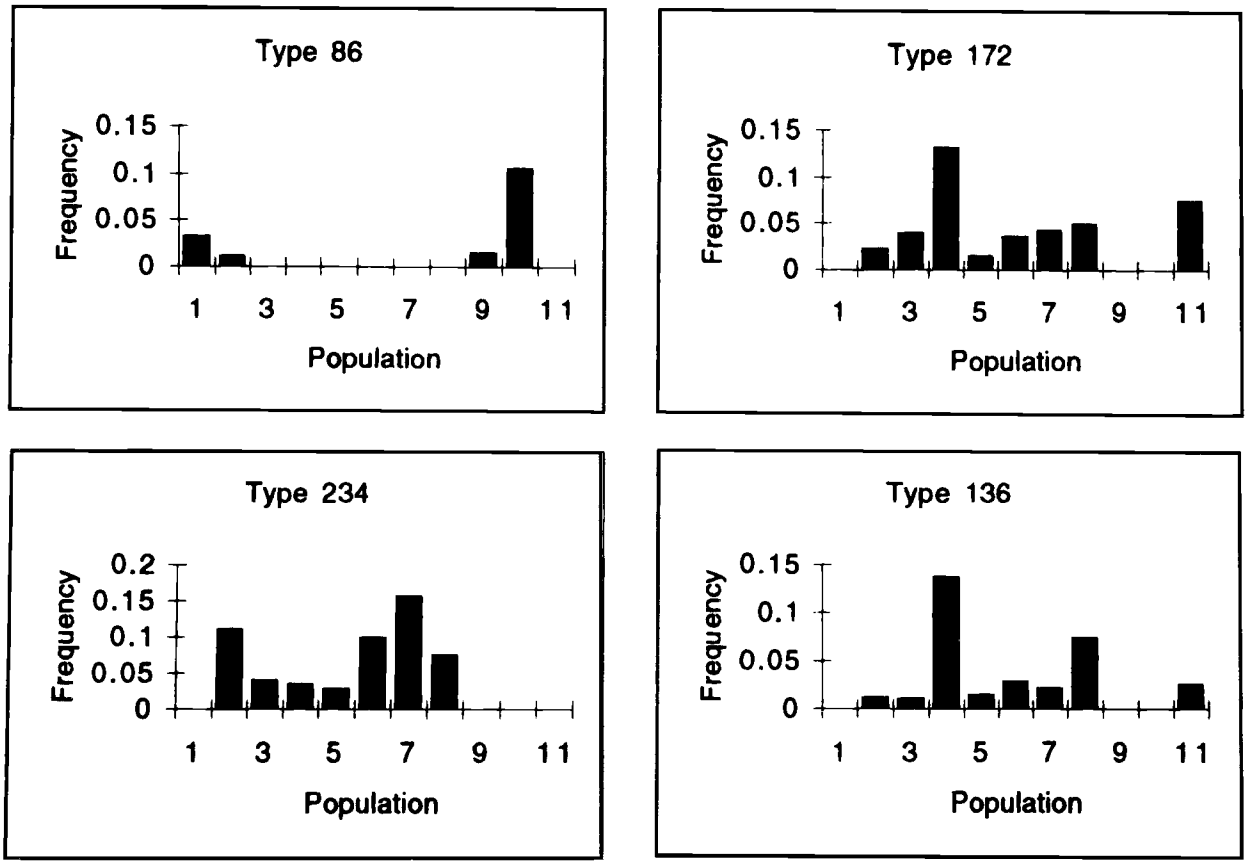

FIG. 4—*Frequencies for types that occur in more than 10% of individuals in a single population. 1-Bangladesh, 2-Borneo, 3-China, 4-Philippines, 5-S. India, 6-E. Indonesia, 7-Java, 8-Malaysia, 9-Pakistan, 10-N. India, 11-Taiwan.*

Selected AMOVA results are shown in Table 6. Since distances from the matrix including all 12 populations had suggested that geographically close populations were similar, populations were assembled into both geographically dissimilar and geographically similar clusters to measure variance components.

In general, variation within populations, regardless of how they were clustered into regional groups, accounted for at least 93% of the SSO type variation observed in the total sample (Φ_{ST} values ranged from 0.39 to 0.45). By extension, therefore, subdivisions of any kind accounted for only up to 7% of the total variation. With the exception of the comparison between the two Taiwanese groups, the most extreme example of regional subdivision was that in which a Pakistani/N. Indian regional cluster was compared to a Javanese/E. Indonesian/Malay regional cluster (Example 6, Table 6). In this case, variation among the regions accounted for 6% of the total variance, while among population/within region variance was less than 1%, and within population variance was 93%. Trials where all populations were randomly and equally divided into two groups never displayed within-population variance components of less than 96%. Therefore, the range of variances accounted for by regional groupings is extremely small (around 4–6%). When all 12 populations are included in one region (Example 1, Table 6), the variance component due to within population differences is 95.47%, while the among population/within region variance component was 4.53%. Heterogeneity among these populations when they are grouped into one large region is significant ($P < 0.001$, 1000 permutations; see Fig. 5).

TABLE 4—*Populations sharing SSO type identities (AMOVA, 1000 permutations).*

Strong: populations not significantly different ($P > 0.1000$)	P
Chinese-Malay	0.671
Bangladeshi-S. Indian	0.423
Javanese-E. Indonesian	0.371
Pakistani-N. Indian	0.232
Javanese-Malay	0.111
Moderate: populations not significantly different ($0.1 > P > 0.025$)	P
Filipino-Taiwanese 1	0.083
E. Indonesian-Malay	0.062
Chinese-Javanese	0.034
Bornean-Javanese	0.033
Weak ($0.025 > P > 0.01$)	P
Bornean-Malay	0.024
Chinese-S. Indian	0.018
Bornean-Chinese	0.011

While Φ_{ST} describes the degree of differentiation present within populations of a specified group, Φ_{CT}, correlated with variance among regions, quantifies the degree of regional subdivision. The highest observed value (0.105, $P < 0.002$) was that for the Taiwanese 1-Taiwanese 2 split (Example 2, Table 6). These two groups ($N = 42$, $N = 39$) have 23 and 16 SSO types respectively, and share only three SSO types which were observed in a total of 15 individuals. The extreme regional clusters of W. Asia versus S.E. Asia (Table 6, Example 6) had the next highest value of 0.061 ($P < 0.002$), although the Φ_{SC} value of 0.005 and P-value of 0.086 (not shown) indicate that the populations within each region are not significantly different from each other. A pair of regional groups that was constructed from all twelve populations randomly

TABLE 5—Φ_{ST} *between pairs of populations are shown in lower left-hand matrix; P-values based on 1000 permutations are shown in upper right-hand matrix. 1-Bangladesh, 2-Borneo, 3-China, 4-Philippines, 5-S. India, 6-Java, 7-E. Indonesia, 8-Malay, 9-Pakistan, 10-N. India, 11-Taiwan 1, 12-Taiwan 2.*

	1	2	3	4	5	6	7	8	9	10	11	12
1	...	0.000	0.006	0.000	0.423	0.000	0.000	0.005	0.001	0.001	0.001	0.000
2	0.081	...	0.011	0.000	0.001	0.033	0.001	0.024	0.000	0.000	0.000	0.000
3	0.035	0.019	...	0.000	0.018	0.034	0.004	0.671	0.000	0.000	0.000	0.000
4	0.071	0.062	0.038	...	0.000	0.000	0.000	0.003	0.000	0.000	0.083	0.000
5	0.001	0.040	0.016	0.061	...	0.000	0.000	0.002	0.003	0.001	0.000	0.000
6	0.062	0.015	0.012	0.037	0.033	...	0.371	0.111	0.000	0.000	0.000	0.000
7	0.066	0.023	0.022	0.026	0.045	0.000	...	0.062	0.000	0.000	0.002	0.000
8	0.038	0.017	−0.003	0.026	0.024	0.006	0.009	...	0.000	0.000	0.001	0.000
9	0.043	0.075	0.060	0.060	0.024	0.045	0.058	0.061	...	0.232	0.000	0.000
10	0.074	0.094	0.082	0.085	0.038	0.069	0.084	0.091	0.005	...	0.000	0.000
11	0.088	0.085	0.058	0.012	0.082	0.075	0.050	0.041	0.100	0.134	...	0.000
12	0.136	0.126	0.087	0.062	0.110	0.062	0.061	0.081	0.097	0.152	0.104	...

TABLE 6—*Representative AMOVA results for 12 populations that describe variance components, Φ-statistics, and their significance values.*

Example	Comparison	Variance components %*			Φ-statistics†		
		σ^2_a	σ^2_b	σ^2_c	Φ_{CT}	Φ_{SC}	Φ_{ST}
1	Populations 1–12§ in one region	...	4.53	95.47	...	...	0.045**
2	Taiwanese 1 v Taiwanese 2 (11 v 12)	10.46	−0.13	89.67	0.105**	−0.001	0.103**
3	NW India v SE India (9,10 v 1,5)	3.45	0.29	96.25	0.035**	0.003	0.037**
4	W Asia v E Asia (1,5,9,10 v 2,3,4,6,7,8,11,12)	3.45	3.09	93.46	0.035**	0.032**	0.065**
5	Mainland v island (1,3,5,9,10 v 2,4,6,7,8,11,12)	1.99	3.55	94.55	0.019**	0.036**	0.054**
6	Extremes (most different) (9,10 v 6,7,8)	6.07	0.44	93.48	0.061**	0.005	0.065**
7	Random (2,4,7,8,10,11 v 1,3,5,6,9,12)	−0.44	4.79	95.66	−0.004	0.048**	0.043**
8	All populations except Taiwanese	...	3.91	96.09	...	...	0.039**

*σ^2_a = variance among regions, σ^2_b = variance among populations/within region, σ^2_c = variance within populations.
†*P*-values <0.002 are indicated with **.
§Populations are: 1-Bangladesh, 2-Borneo, 3-China, 4-Philippines, 5-S. India, 6-E. Indonesia, 7-Java, 8-Malaysia, 9-Pakistan, 10-N. India, 11-Taiwan 1, 12-Taiwan 2.

assorted into two groups (Table 6, Example 7) had a Φ_{CT} value of −0.004, which suggests that SSO types from different populations are, on average, more closely related than those from the same population (AMOVA documentation, L. Excoffier). A comparison of mainland populations with island populations (Table 6, Example 5) gave a Φ_{CT} value of 0.019, indicating that while the regions are significantly different, they are not as different as the group of W. Asian populations compared to the group of E. Asian populations.

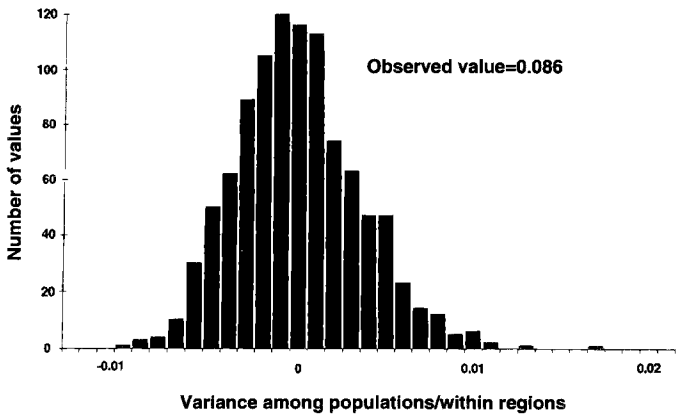

FIG. 5—*Null distribution of the variance component (among populations/within region) obtained through 1000 random permutations of 993 individuals into 12 populations the sizes of those in this study. This trial grouped all populations into one region.*

Discussion

Within this large Asian sample of mtDNAs, SSO typing revealed a substantial amount of variation. In an earlier study of SSO type variation (7), 274 SSO types were detected in a smaller total sample encompassing a more diverse collection of worldwide rather than regional populations (N = 525). Proportionally more different types were found in that study (274/525) than in this one (255/993), which is not surprising since regional variation (in this case, Asia) will be more limited than worldwide variation due to the shallower time depth of regional population expansions. In this study, estimates of SSO type diversity were high, always exceeding 0.94. Even within subgroups of the populations, such as those of the aboriginal Taiwanese, diversity estimates were greater than 0.86.

The number of SSO types observed in the 11 populations ranged between 29 and 60 per population, even though population sample sizes had a much greater range of between 31 and 175 individuals each. It is possible that SSO typing may be approaching the limits of detection of all SSO types present in these particular populations; theoretically, there could be 27,648 possible SSO types for these SSO variants (including a blank variant at any position), but the number of types detected here per population does not increase proportionally with sample size. For example, while the Bangladeshi population contains a unique type for almost every individual,

the Filipino sample has nearly three times more individuals than SSO types.

While many more SSO types are possible than were observed, this was accounted for in part by the relative invariance of certain sites, for example, IA0, IA2, IB0, ID0, IIA0, and IIA1 were often observed at low frequency or were absent in some Asian populations. Site variants such as these, which carry less weight overall due to their infrequency, act to increase homogeneity. In spite of this, diversity values were high for all populations, and no possible site variant was completely absent from the data set.

More than half of the total number of SSO types were unique (53%). Very few types which occurred more than one time were exclusive to one population (10%); 88% of these (22/25) were low-frequency types. Two higher frequency types were observed in the rather isolated aboriginal Taiwanese, who appear to be quite different with respect to SSO types of other Asian populations. Approximately 11% of SSO types appeared eight or more times. No single type occurred at a frequency greater than 6% overall (type 234 was seen in 5.74% of the total sample).

Within populations, only four types occurred at frequencies greater than 10% in any single population. Type 86 appeared primarily in W. Asia, while types 136, 172, and 234 appeared primarily in S.E. Asia; for these common types, substructuring appears to occur along clear geographic lines. The most pronounced regional heterogeneity was observed between Pakistan/N. India and E. Indonesia/Java/Malaysia, with a Φ_{CT} value of 0.61. However, sizeable geographic distance separates these population groups from one other; a similar comparison of Mainland Asia versus Island Asia regions gave a lower Φ_{CT} value of 0.19 ($P < 0.002$), indicating that populations within these two more broadly-defined regions are less easily distinguished.

Inclusion of the aboriginal Taiwanese gave valuable insight into the kind of contribution isolated relict populations would make to a database of mtDNA SSO types. First of all, this population has SSO types which were never observed elsewhere, including several at relatively high frequency. They also lack type 234, the most common type in the entire data set. When the Taiwanese were removed from AMOVA analyses which included all other populations, the coefficient of population subdivision was reduced from 0.045 to 0.039, the largest change which was observed due to the removal of any single population (Example 8, Table 6). Within the Taiwanese themselves, a considerable amount of substructure exists as well. Comparison of the Taiwanese 1 and Taiwanese 2 groups gave a Φ_{CT} value of 0.105, by far the largest coefficient of substructuring observed during all AMOVA analyses. Interestingly, the group described here had no unusually high frequency of any types which otherwise occurred commonly throughout other populations, rather, it had a few private types, some occurring at high frequency, and shared the remaining types at low frequency with other populations. In spite of its uniqueness, the Taiwanese 1 population was not significantly different from the Filipino population ($P = 0.08$). While a population with high-frequency population specific types such as the aboriginal Taiwanese is of particular concern in considering forensic match probabilities, additional surveys of such isolated groups (which tend to be primarily of anthropological and not forensic interest) will help to further define how common this pattern of SSO types is worldwide.

The Φ_{CT} values and variance components that were obtained from numerous comparisons between regional groups comprised of either all populations or just specific subsets always displayed P-values of less than 0.002, indicating that there is statistically significant population heterogeneity. However, heterogeneity appears to be minimal compared to the large amount of variation which is present within populations. Variance attributable to within population variation was greater than 93%, regardless of the groupings of populations which were stipulated. In other words, each population contains on average approximately 93% of the variation which is found in the sample overall. The vast majority of SSO type variation occurs here as both widely shared types which carry the most weight in the analysis of variance, and types unique to populations found in extremely low frequency; few types that occur in high frequency are unique to populations that carry them.

For forensic applications, a desirable genetic locus should combine high diversity with low probability of identity between individuals; by extension, an unsampled subpopulation should not have variants which are present in high frequency relative to the population at large. MtDNA SSO types, as easily screened-for representatives of the variation present in the mtDNA hypervariable regions, conform to these qualifications: 1) diversity estimates exceed 0.94 (except in relict isolate populations), 2) common types are widely distributed in most populations and frequencies of these rarely exceed 10%, and 3) there is a high probability of types which are rare or unique.

Studies of worldwide human mtDNA evolution and expansion predict that African mtDNAs would be most diverse overall, with less diversity present in regions with shorter human evolutionary histories, such as Asia and Europe (2,32). This study is part of a larger effort to measure diversity and heterogeneity in SSO types of African, European, and Asian subpopulations to develop an understanding of the forensic implications of variation present in the mtDNA control region. For the SSO types of Asian populations sampled in this study, diversity is high, and heterogeneity, while present, seems to be in part accounted for by East Asia/West Asia regional substructuring and is minimal in comparison to the large amount of variation present overall. For forensics investigations, the control region of mtDNA exhibits satisfactory qualities as a DNA typing locus, when population genetics parameters of SSO typing are considered.

Acknowledgments

We gratefully acknowledge the contribution of samples from J. Martinson (Institute of Molecular Medicine, John Radcliffe Hospital, Oxford), N. Saha (National University Hospital, Singapore), and A.S.M. Sofro (Gadjah Mada University, Yogyakarta, Indonesia). The collection of East Indonesian samples was supported by a grant from the Wenner-Gren Foundation to Mark Stoneking. We also thank Anne Buchanan and Robert Melton for help with data management and programming, and Sean Walsh and Roche Molecular Systems for providing some probes.

References

(1) Cann RL, Stoneking M, Wilson AC. Mitochondrial DNA and human evolution. Nature 1987;325:31–6.

(2) Vigilant L, Pennington R, Harpending H, Kocher TD, Wilson AC. Mitochondrial DNA sequences in single hairs from a southern African population. Proc Nat Acad Sci (USA) 1989;86:9350–54.

(3) DiRienzo A, Wilson AC. Branching pattern in the evolutionary tree for human mitochondrial DNA. Proc Nat Acad Sci (USA) 1991;88:1597–1601.

(4) Kocher T, Wilson AC. Sequence evolution of mitochondrial DNA in humans and chimpanzees: control region and a protein-coding region. In: Osawa S, Honjo T, editors. Evolution of life: fossils, molecules, and culture. New York: Springer-Verlag, 1991;391–413.

(5) Stoneking M, Jorde LB, Bhatia K, Wilson AC. Geographic variation in human mitochondrial DNA from Papua New Guinea. Genetics 1990;124:717–33.

(6) Ward RH, Frazier BL, Dew-Jager K, Pääbo S. Extensive mitochondrial diversity within a single Amerindian tribe. Proc Nat Acad Sci (USA) 1991;88:8720–24.
(7) Stoneking M, Hedgecock D, Higuchi RG, Vigilant L, Erlich HA. Population variation of human mtDNA control region sequences detected by enzymatic amplification and sequence-specific oligonucleotide probes. Am J Hum Genet 1991;48:370–82.
(8) Ginther C, Issel-Tarver L, King M-C. Identifying individuals by sequencing mitochondrial DNA from teeth. Nat Genet 1992; 2:135–38.
(9) Holland MM, Fisher DL, Mitchell LG, Rodriquez WC, Canik JJ, Merril CR, et al. Mitochondrial DNA sequence analysis of human skeletal remains: identification of remains from the Vietnam War. J Forensic Sci 1993;38:542–53.
(10) Gill P, Kimpton C, Aliston-Greiner R, Sullivan K, Stoneking M, Melton T et al. Establishing the identity of Anna Anderson Manahan. Nat Genet 1995;9:9–10.
(11) Wilson MR, Stoneking M, Holland MM, DiZinno JA, Budowle B. Guidelines for the use of mitochondrial DNA sequencing in forensic science. Crime Lab Dig 1993;20:68–77.
(12) Stoneking M, Melton T. Forensic applications of mitochondrial DNA analysis. In: Budowle B, editor, Forensic Applications of PCR. In press.
(13) Chakraborty R, Kidd KK. The utility of DNA typing in forensic work. Science 1991;254:1735–39.
(14) Lewontin RC, Hartl DC. Population genetics in forensic DNA typing. Science 1991;254:1745–50.
(15) Devlin B, Risch N, Roeder K. Forensic inference from DNA fingerprints. J Am Stat Assoc 1992;87:337–50.
(16) Lander ES, Budowle B. DNA fingerprinting dispute laid to rest. Nature 1994;371:735–38.
(17) Weir BS. Population genetics in the forensic DNA debate. Proc Nat Acad Sci (USA) 1992;89:11654–59.
(18) Weir BS. The effects of inbreeding on forensic calculations. Ann Rev Genet 1994;28:597–621.
(19) Giles RE, Blanc H, Cann RL, Wallace DC. Maternal inheritance of human mitochondrial DNA. Proc Nat Acad Sci (USA) 1980;77: 6715–19.
(20) Merriwether DA, Clark AG, Ballinger SW, Schurr TG, Soodyall H, Jenkins T, et al. The structure of human mitochondrial DNA variation. J Mol Evol 1991;33:543–55.
(21) Melton T, Peterson R, Redd AJ, Saha N, Sofro ASM, Martinson J, et al. Polynesian genetic affinities with southeast Asian populations as identified by mtDNA analysis. Am J Hum Genet 1995;57:403–14.
(22) Redd AJ, Takezaki N, Sherry ST, McGarvey ST, Sofro ASM, Stoneking M. Evolutionary history of the COII/tRNA^{lys} intergenic 9 base pair deletion in human mitochondrial DNAs from the Pacific. Mol Bio Evol 1995;12:604–15.
(23) Connor A, Stoneking M. Assessing ethnicity from human mitochondrial DNA types determined by hybridization with sequence-specific oligonucleotides. J Forensic Sci 1994;39:1360–71.
(24) Tajima F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 1989;123:585–95.
(25) Nei M. Molecular evolutionary genetics. New York: Columbia University Press, 1987.
(26) Sokal RR, Rohlf FJ. Biometry, the principles and practice of statistics in biological research. 2nd ed. San Francisco: W.H. Freeman, 1981.
(27) Excoffier L, Smouse PE, Quattro JM. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics 1992;131: 479–91.
(28) Cockerham CC. Variance of gene frequencies. Evolution 1969;23: 72–84.
(29) Cockerham CC. Analyses of gene frequencies. Genetics 1973;74: 679–700.
(30) Weir BS, Cockerham CC. Estimating *F*-statistics for the analysis of population structure. Evolution 1984;38:1358–70.
(31) Wright S. The genetical structure of populations. Ann Eugenics 1951;15:323–34.
(32) Piercy R, Sullivan KM, Benson N, Gill P. The application of mitochondrial DNA typing to the study of white Caucasian genetic identification. Int J Leg Med 1994;106:85–90.

Address requests for reprints or additional information to
Terry Melton
Department of Anthropology
409 Carpenter Building
University Park, PA 16801
e-mail: TWM107@psuvm.psu.edu

Appendix

MtDNA SSO type numbers are listed in column 1, SSO type profiles are listed in column 2; the count of each type by population in which it was observed is shown in the other columns. 1-Bangladesh, 2-Borneo, 3-China, 4-Philippines, 5-S. India, 6-E. Indonesia, 7-Java, 8-Malaysia, 9-Pakistan, 10-N. India, 11-Taiwan.

SSO type number	SSO type profile	1	2	3	4	5	6	7	8	9	10	11
003	01122331	0	0	0	0	0	0	0	0	0	1	0
004	01212110	0	0	0	0	0	0	1	0	0	0	0
005	01222310	0	0	0	0	0	1	0	0	0	0	0
006	02012111	0	0	0	0	0	1	0	0	0	0	0
008	02112011	0	0	0	1	0	0	0	0	0	0	0
009	02112111	0	0	1	0	0	2	0	0	0	0	0
010	02112201	0	0	0	0	1	0	0	0	0	0	0
011	02112231	0	0	0	0	0	0	0	1	0	0	0
012	02122031	0	1	0	0	0	0	0	0	0	0	0
013	02122201	0	1	0	0	0	0	0	0	0	0	0
014	02312211	0	0	0	0	0	1	0	0	0	0	0
015	02312311	0	0	0	0	0	0	1	0	0	0	0
016	03112111	0	0	0	1	0	0	0	0	0	0	0
017	10022231	0	0	0	0	0	0	1	0	0	0	0
019	10112011	0	0	0	1	1	0	1	0	0	0	0
020	10112021	0	0	0	0	1	0	0	1	0	0	0
021	10112101	0	0	1	0	0	0	0	0	0	0	0
022	10112111	0	0	1	3	1	1	0	0	0	0	0
023	10112211	0	0	0	2	0	0	0	0	0	0	0
025	10122011	0	0	0	0	0	0	1	0	0	0	0
027	10122110	0	0	0	0	0	0	0	1	0	0	0
028	10122210	0	0	0	0	0	0	0	0	0	0	1
029	10212100	0	0	0	0	0	0	0	0	0	0	5
030	10212110	0	0	0	0	0	0	0	0	0	0	8
031	10212200	0	0	0	0	0	0	0	0	0	0	2
032	10212211	0	0	1	0	0	0	0	0	0	0	0
033	10312130	0	0	0	0	0	0	0	0	1	0	0
034	10312211	0	0	0	0	0	0	0	0	1	0	0
035	10312311	0	0	0	0	0	0	0	0	1	0	0
036	11012011	0	0	0	0	1	0	0	0	0	1	0
037	11012110	0	0	0	0	0	0	0	0	1	0	0
038	11012311	1	0	0	0	0	0	0	0	3	2	0
040	11102110	0	0	0	0	0	1	0	0	0	0	0
041	11110311	0	0	0	0	0	0	0	0	1	0	0
042	11111011	0	0	0	1	0	0	0	0	1	0	0
043	11111111	0	0	0	0	1	0	0	0	4	2	0
044	11111121	0	1	0	0	0	0	0	0	0	0	0
045	11111211	0	0	0	0	1	0	0	0	0	0	0
046	11111311	0	0	0	0	0	0	0	0	2	0	0
047	11112011	0	3	0	0	1	2	1	0	3	3	0
048	11112021	0	1	0	0	0	0	1	0	1	0	0
049	11112031	0	0	1	0	0	0	0	0	0	0	0
050	11112101	1	0	0	0	0	0	0	0	0	0	1
051	11112110	0	0	0	0	0	1	0	0	0	0	0
052	11112111	1	7	5	5	0	5	9	4	0	0	1
053	11112121	0	1	1	1	0	0	1	0	0	0	0
054	11112201	0	0	0	0	0	0	1	0	1	0	0
055	11112211	0	2	0	1	1	0	1	1	1	0	0
056	11112221	0	0	0	0	0	1	0	0	1	0	0
057	11112301	0	0	0	0	0	3	0	0	0	0	0
058	11112310	0	0	0	0	0	0	0	0	1	0	0
059	11112311	1	3	0	8	3	0	3	1	0	4	2
039	11112321	0	0	0	0	0	0	0	0	1	0	0
060	11121001	0	0	0	0	0	0	0	0	0	1	0
061	11121111	0	0	0	0	0	0	0	0	2	0	0
062	11122111	1	0	0	1	0	0	0	0	0	0	1
063	11122211	0	0	0	0	0	0	0	1	0	0	0
064	11122310	0	0	0	3	0	0	0	0	0	0	0

SSO type number	SSO type profile	1	2	3	4	5	6	7	8	9	10	11
065	11122311	0	0	0	0	1	0	0	0	1	0	0
066	11212011	0	1	0	0	2	0	0	0	0	0	0
067	11212020	0	0	2	0	0	0	0	0	0	0	0
068	11212101	0	0	0	0	0	0	0	0	1	0	0
069	11212110	0	0	3	6	0	4	1	1	1	0	0
070	11212111	0	1	1	1	0	1	0	0	0	0	0
071	11212120	0	0	1	0	0	0	0	0	0	0	0
072	11212121	0	0	0	0	1	0	0	0	0	0	0
073	11212201	1	0	0	0	0	0	0	0	0	0	0
074	11212211	0	0	0	0	0	0	1	0	0	0	0
075	11212301	0	0	0	0	0	1	0	0	0	0	0
076	11212310	0	0	0	0	0	1	0	0	0	0	0
077	11212311	0	0	1	1	1	0	0	0	0	1	0
078	11222111	0	0	0	1	0	0	1	0	0	0	0
079	11222310	0	0	0	0	0	0	0	0	0	0	1
080	11222311	0	0	1	0	0	0	0	0	0	0	0
081	11311111	0	0	0	0	0	0	0	0	1	0	0
082	11311311	0	0	0	0	0	0	0	0	1	0	0
083	11312101	0	0	0	0	1	1	0	0	1	0	0
084	11312111	0	0	0	2	0	0	0	0	2	0	0
085	11312211	0	1	0	0	0	0	0	0	0	0	0
086	11312311	1	1	0	0	0	0	0	0	1	5	0
087	11312321	0	0	0	0	1	0	0	0	0	0	0
089	11322010	0	0	0	0	0	0	0	1	0	0	1
090	11322211	1	0	0	0	0	0	0	0	0	0	0
091	11322310	0	0	0	0	0	0	1	0	0	0	1
093	12010111	0	0	0	0	0	0	1	0	0	0	0
095	12012101	0	0	0	0	0	0	0	0	1	0	0
096	12012110	1	0	0	0	0	0	0	0	0	0	0
097	12012111	0	1	0	0	0	0	0	0	0	0	0
098	12012310	0	0	0	0	0	0	1	0	0	0	0
100	12102021	0	0	0	0	0	0	0	1	0	0	0
101	12102111	0	0	1	0	0	0	0	0	0	0	0
102	12112001	1	0	1	0	0	0	0	0	1	0	0
103	12110310	0	0	0	0	0	0	1	0	0	0	0
104	12112010	0	0	0	0	0	0	0	1	0	0	0
105	12112011	0	4	4	5	0	0	4	2	0	0	1
106	12112021	0	0	0	1	1	0	0	0	0	0	0
107	12112031	0	0	1	0	0	0	0	0	0	0	0
108	12112100	0	0	0	1	1	0	0	0	1	0	0
109	12112101	1	0	0	1	1	0	0	0	5	4	0
110	12112110	1	0	3	0	1	1	0	4	0	0	0
111	12112111	2	0	2	0	4	4	2	2	1	0	2
112	12112121	0	1	0	0	1	1	0	0	0	0	0
113	12112131	0	0	2	0	4	0	0	0	0	0	0
114	12112201	2	0	0	1	0	0	0	0	0	0	0
115	12112211	0	1	1	2	0	2	0	1	2	2	0
116	12112221	1	1	0	0	0	0	0	0	0	0	0
117	12112231	0	1	1	0	0	0	0	0	0	0	1
118	12112301	1	0	0	1	2	0	0	0	0	1	0
119	12112310	0	0	2	0	0	0	0	0	0	0	0
120	12112311	1	0	1	0	2	1	0	1	0	0	0
121	12112321	0	0	0	0	0	0	1	0	0	0	0
122	12112331	0	0	0	0	0	0	1	0	0	0	0
123	12120231	0	0	0	0	0	1	0	0	0	0	0
124	12122001	0	2	1	1	0	0	0	0	0	0	0
125	12122010	0	0	0	0	0	1	0	0	0	0	0
126	12122011	0	2	9	1	1	0	1	1	0	0	2
127	12122021	0	3	0	0	0	1	0	0	0	0	0
128	12122031	0	2	0	0	0	0	2	1	0	0	0
129	12122101	1	1	1	0	1	0	0	2	0	0	1
130	12122110	0	0	1	0	0	0	0	0	0	0	0
131	12122111	1	0	7	10	1	6	4	7	0	0	7
132	12122120	0	0	0	0	0	0	0	0	0	0	2
133	12122121	1	0	0	4	0	8	1	0	1	1	2
134	12122131	1	0	0	0	0	0	0	0	0	0	0
135	12122211	0	0	0	1	1	0	1	0	1	0	0
137	12122230	0	0	0	1	0	0	0	0	0	0	0
136	12122231	0	1	1	24	1	4	2	6	0	0	2
138	12122301	0	0	2	1	1	0	0	0	0	0	0
139	12122311	1	2	2	1	1	3	2	2	0	0	2
140	12122331	0	0	0	0	0	0	0	1	0	0	0

SSO type number	SSO type profile	1	2	3	4	5	6	7	8	9	10	11
141	12212011	0	0	0	0	1	0	0	0	0	0	0
142	12212101	0	0	0	0	0	0	0	0	0	0	1
143	12212110	0	0	0	0	0	0	0	1	0	0	0
144	12212111	0	0	0	0	0	1	1	0	0	0	0
145	12212131	1	0	0	0	0	0	0	0	0	0	0
146	12212211	0	0	0	0	0	1	0	0	0	0	0
147	12212231	0	0	1	0	0	0	0	0	0	0	0
148	12222131	0	0	0	1	0	0	0	0	0	0	0
149	12312001	0	0	0	0	2	0	0	0	0	0	0
150	12312011	0	1	0	0	0	0	0	1	0	0	0
151	12312021	0	1	1	0	0	0	0	0	0	0	0
152	12312101	0	0	0	0	1	0	0	0	1	1	0
153	12312110	1	0	0	0	0	0	0	0	0	0	0
154	12312111	1	0	2	0	0	0	1	0	1	0	0
155	12312211	0	1	0	0	1	0	0	0	0	0	0
156	12312221	0	0	0	0	0	1	0	0	0	0	0
157	12312311	1	0	0	0	3	0	0	1	0	1	0
158	12322021	0	2	0	0	0	0	0	0	0	0	0
159	12322101	0	0	0	0	0	0	0	0	0	1	0
160	12322111	0	0	0	0	1	0	2	1	0	0	0
161	12322121	0	0	0	0	1	0	1	0	0	0	0
162	12322231	0	0	0	0	0	1	0	0	0	0	0
163	12322311	1	0	1	1	0	1	1	1	0	0	0
164	13012111	0	0	1	1	0	0	0	0	0	0	0
165	13012121	0	0	1	0	0	0	0	0	0	0	0
166	13102111	0	0	0	0	0	0	0	0	0	0	4
167	13112001	0	0	0	0	0	1	0	0	0	0	0
168	13112011	0	1	3	0	1	1	1	0	0	0	1
169	13112021	0	1	0	3	0	2	1	0	0	0	1
171	13112110	0	0	0	1	0	0	0	0	0	0	0
172	13112111	0	2	4	23	1	5	4	4	0	0	6
173	13112201	0	0	0	3	0	0	0	0	0	0	1
174	13112211	0	1	0	13	0	9	1	3	0	0	7
175	13112221	0	0	0	1	0	0	1	0	0	0	0
176	13112311	0	0	1	1	0	2	0	0	3	2	0
177	13112321	0	0	0	0	0	0	0	0	0	1	0
178	13122011	0	0	0	1	0	0	0	0	0	0	0
180	13212211	0	0	0	0	0	0	0	0	0	0	1
181	13312201	0	0	0	0	0	2	0	0	0	0	0
182	13312211	0	0	0	1	0	1	0	0	0	0	4
183	13312221	0	1	0	0	0	0	1	0	0	0	0
184	20112111	0	0	0	0	0	0	0	0	2	0	0
185	20112311	0	0	0	0	0	0	0	0	1	0	0
186	21012110	0	0	0	0	0	1	0	0	0	0	0
187	21112111	0	0	0	0	0	0	0	0	0	1	0
188	21112121	0	0	0	0	0	0	0	0	1	0	0
189	21112211	0	0	1	0	0	0	0	0	0	0	0
190	21112311	0	0	0	0	0	0	0	0	0	3	0
191	21121111	0	0	0	0	0	0	0	0	0	1	0
192	21121311	0	0	0	0	0	0	0	0	2	0	0
193	21122311	0	0	0	0	0	0	0	0	0	1	0
194	21212110	0	0	0	0	1	0	0	0	0	0	0
195	21212111	0	0	0	0	0	0	0	0	1	0	0
196	21312101	0	0	0	1	0	0	0	0	0	0	0
197	21312111	0	0	0	1	1	2	2	1	1	0	1
198	21312211	0	0	0	2	0	0	0	0	0	0	0
200	22012111	0	0	0	0	0	0	0	0	0	1	0
201	22102111	0	0	0	0	0	0	0	0	1	0	0
202	22112001	0	0	0	1	0	0	0	0	0	0	0
203	22112031	0	0	0	0	0	1	0	0	0	0	0
204	22112111	0	0	1	1	0	0	0	1	2	3	0
205	22112121	0	0	0	0	0	0	0	0	1	0	0
206	22112211	0	0	0	0	0	0	0	0	1	0	0
207	22122231	0	0	0	0	0	1	0	0	0	0	0
208	22312031	0	0	0	0	0	0	0	0	0	0	1
209	22312111	0	0	0	0	2	0	0	0	0	0	0
211	30122021	0	0	0	0	1	0	0	0	0	0	0
212	30122321	0	0	0	0	2	0	0	0	0	0	0
213	30212110	0	0	1	0	0	0	0	0	0	0	0
214	30312201	0	0	0	0	0	1	0	0	0	0	0
215	30322311	0	0	0	0	1	0	0	0	0	0	0
217	31012110	0	1	0	0	0	0	0	0	0	0	0

SSO type number	SSO type profile	1	2	3	4	5	6	7	8	9	10	11
218	31012111	0	1	0	0	0	0	0	0	0	0	0
219	31012310	0	1	0	0	0	0	0	0	0	0	0
220	31112001	0	0	1	0	0	0	0	0	0	0	0
222	31112110	0	0	0	0	0	0	0	1	0	0	0
223	31112111	0	0	0	1	0	0	0	0	0	0	0
224	31112131	0	0	0	3	0	2	0	0	0	0	3
225	31112210	0	0	0	0	0	0	0	0	0	1	0
226	31112321	0	0	0	1	0	0	0	0	0	0	0
227	31112331	0	0	0	1	0	0	0	0	0	0	0
228	31122111	0	0	0	1	1	0	0	0	0	1	0
229	31122301	0	0	0	0	0	0	0	0	1	0	0
231	31212010	0	6	1	0	0	0	0	1	0	0	0
232	31212020	0	1	2	0	0	0	0	0	0	0	0
233	31212100	0	0	0	1	0	0	0	0	0	0	0
234	31212110	0	10	4	6	2	14	15	6	0	0	0
235	31212120	0	0	1	0	0	0	1	0	0	0	0
236	31212130	0	0	0	0	0	0	1	0	0	0	0
237	31212220	0	0	1	0	0	0	0	1	0	0	0
238	31212310	0	0	0	0	1	3	0	2	0	0	0
239	31212330	0	0	0	0	0	0	1	0	0	0	0
240	31222300	0	0	0	1	0	0	0	0	0	0	0
241	31222310	0	1	0	8	0	13	1	0	0	0	0
242	32110031	0	0	0	0	0	0	1	0	0	0	0
243	32112001	0	0	0	0	1	0	0	0	0	0	0
244	32112011	0	2	1	0	0	0	1	2	0	0	1
245	32112031	0	7	1	0	0	2	1	1	0	0	0
246	32112101	0	0	0	0	0	0	0	0	0	1	0
247	32112111	1	0	0	0	2	1	0	0	3	0	0
248	32112120	0	0	0	0	0	0	0	0	1	0	0
249	32112121	0	0	0	0	0	1	0	0	0	0	0
250	32112131	0	0	0	0	1	0	0	0	0	0	1
251	32112211	1	0	0	0	0	0	0	0	0	0	0
252	32112310	0	2	1	0	0	0	5	2	1	0	0
253	32122011	0	0	2	0	0	1	0	0	0	0	0
254	32122021	0	0	0	0	1	0	0	0	0	0	0
255	32122031	0	0	1	0	0	0	0	1	0	0	0
001	32122211	0	0	0	1	0	0	0	0	0	0	0
002	32122311	0	0	3	0	0	0	0	0	0	0	0
007	32122331	0	0	0	0	0	0	0	0	0	0	2
018	32312031	0	2	0	0	0	0	0	1	0	0	0
024	32312111	0	0	0	0	0	0	0	1	0	0	0
026	32312121	0	0	0	0	0	1	0	0	0	0	0
092	32312201	0	0	0	0	0	1	0	0	0	0	0
094	32312210	0	0	0	0	0	1	0	0	0	0	0
099	32312211	0	0	0	0	0	5	0	0	0	0	0
170	32312311	0	0	0	0	1	0	1	0	0	0	0
179	32322011	0	0	1	0	0	0	0	0	0	0	0
199	32322130	0	0	0	0	0	0	0	0	1	0	0
210	32322211	0	0	0	0	0	1	0	0	0	0	0
216	32322311	0	0	1	0	0	0	0	0	0	0	0
221	33112021	0	0	1	0	0	0	1	3	0	0	0
230	33112211	0	0	0	0	0	2	0	0	0	0	0
088	33212021	0	0	0	0	0	0	0	1	0	0	0