Terry Melton,[1] Ph.D.; Mark Wilson,[2] M.S.; Mark Batzer,[3] Ph.D.; and Mark Stoneking,[1] Ph.D.

Extent of Heterogeneity in Mitochondrial DNA of European Populations

REFERENCE: Melton T, Wilson M, Batzer M, Stoneking M. Extent of heterogeneity in mitochondrial DNA of European populations. J Forensic Sci 1997;42(3):437–446.

ABSTRACT: Variation in the mitochondrial DNA (mtDNA) control region as detected by sequence-specific oligonucleotide (SSO) probes is described for 595 individuals from six European or European-derived populations. Estimates of diversity for mtDNA types exceed 0.91 in all populations, while 50% of the 158 types which were observed occur only once. Of 68 shared types, most occur rarely (<3% of the total population); only one type occurs at a frequency greater than 10%, and it is present at comparable frequencies in all six populations (18–29%). An analysis of molecular variance (AMOVA) incorporating genetic distances between types shows that 100% of the variation present in the total sample is attributable to within-population diversity, while there are essentially no between-population differences. Another AMOVA was performed for the first hypervariable region SSO sites only, which included this sample plus an additional 537 SSO types from nine more European populations that were inferred from published mtDNA control region sequence data. Similar results were obtained, with over 99% of the variation overall attributable to within-population differences, and less than 1% of the variation attributable to between-population differences. The Saami were the most different from other populations, which had been observed in an earlier study of nucleotide sequence data. Overall, there is no statistically significant heterogeneity for European populations ($p > 0.001$), and these groups are virtually indistinguishable with respect to mtDNA SSO types. These results demonstrate the utility of mtDNA typing for forensic investigations.

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

Mitochondrial DNA provides a potentially valuable locus for forensic DNA typing. The high number of polymorphisms present in the two hypervariable portions of the noncoding control region readily allow discrimination among individuals and have been used to infer population history and substructure (1–8). The likelihood of recovering mtDNA in small or degraded biological samples is greater than for nuclear DNA because mtDNA molecules are present in high copy number in all cells of the body except red blood cells (9). Therefore, bone, hair, muscle, skin, and blood, even if degraded by environmental insult or time, may provide enough material for typing the mtDNA locus. The use of mtDNA in forensic investigations is increasing (10–14) and procedural guidelines are currently available (15–17).

The use of a DNA locus for forensic typing demands comprehensive understanding of the population genetics of that locus. Undetected subpopulation heterogeneity may in theory result in bias when calculating the probability that a suspect contributed an evidentiary biological sample (18–23). As part of a project to examine variation in the mtDNA control region in worldwide populations, we report here the patterns of variation and extent of heterogeneity for European-derived populations. In a previous paper, we described variation in Asian populations (24), and in the future we will present data on African populations (25). We have relied on sequence-specific oligonucleotide (SSO) typing to detect the sequence variation present at 13 nucleotide positions in the mtDNA control region. This method is useful for rapidly surveying large numbers of individuals and populations. It is also an inexpensive alternative exclusionary technique for forensic investigations (10), compared to labor-intensive sequencing of the mtDNA control region. SSO typing of worldwide populations in a previous study revealed that an enormous amount of variation is detectable by this method, and in fact, there is overall only a 2.6% probability that two unrelated individuals will share the same SSO type (10).

Methods

In this study, 446 individuals of European ancestry were typed for sequence variants at 13 nucleotide positions in eight regions across the mtDNA control region. All samples were purified genomic DNA obtained from maternally unrelated individuals. The four typed populations were from France (Nantes and Brittany; $N = 81$), Switzerland (Bern; $N = 46$), Midwest United States (Michigan; $N = 190$), and Northeast United States (Pennsylvania; $N = 129$). All the above samples were collected for studies of VNTR frequencies or to establish forensic databases. An additional 100 SSO types from British individuals were inferred from published mtDNA control region sequence data (26), and 49 SSO types from Germans (Giessen) were inferred from unpublished sequence data.

SSO types also were inferred from published sequence data for the first hypervariable region of the mtDNA control region (four of the eight variant regions) for an additional 537 individuals from the following nine European populations: Saami ($N = 115$), Finn ($N = 50$), Karelian ($N = 83$), Estonian ($N = 28$), Volga Finnic ($N = 34$), Icelandic ($N = 39$) (all from 7); Swiss ($N = 74$) (27); Sardinian ($N = 69$) (3); and Basque ($N = 45$) (28). These

[1] Postdoctoral fellow and associate professor of anthropology, respectively, The Pennsylvania State University, University Park, PA.
[2] Forensic examiner, Federal Bureau of Investigation, Washington, D.C.
[3] Assistant professor of pathology, Louisiana State University Medical Center, New Orleans, LA.

*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, February 1996, Nashville TN.
Received 20 Feb. 1996; and in revised form 20 Aug. 1996; accepted 9 Sept. 1996.

populations were used for an analysis of molecular variance which also included the first hypervariable region SSO types from the six populations described in the last paragraph. Figure 1 shows the locations of all the populations included in these analyses.

The nonradioactive, sequence-specific oligonucleotide typing method used here and the arrangement into SSO types of individual results of the typing at each variant region are described in detail elsewhere (5,24). The SSO probes used were also previously described (24).

The data set consisting of the control region SSO types from 595 Europeans was examined with respect to type frequency distribution both overall and within the six populations. An unbiased estimate of diversity (h) and its variance were calculated to quantify the amount of mtDNA variation present in each population (24).

An analysis of molecular variance (AMOVA, 29) was applied to the SSO types to measure the apportionment of diversity within and among the six populations for which data were available for both hypervariable regions and for all 15 populations for which data were available for the first hypervariable region. AMOVA as used for these analyses is described in detail elsewhere (24). In general, the method adds information about the genetic distances between pairs of SSO types to a traditional computation of variance components and F-statistics from mtDNA type frequency data to determine whether statistically significant population subdivision exists. A conventional sum of squared deviations was partitioned into variance components attributable to variation among populations (σ_a^2) and variation within populations (σ_b^2). Φ_{ST}, the correlation of random SSO types within populations, relative to that of random pairs of types drawn from the entire data set, was also generated.

Permutational procedures in AMOVA were used to test the significance of Φ-statistics and variance components. To do this, for each AMOVA analysis a null distribution was generated by allocating every individual to a randomly chosen population while holding sample sizes constant over 1000 permutations. Probabilities of observing random variance components and Φ statistics greater than those generated in the analysis were reported. This form of significance testing is useful because concerns about the normality of underlying variance distributions can be ignored. All procedures were carried out with the AMOVA program, provided by L. Excoffier.

Results

Within the data set of 595 complete SSO types from six populations, there were 158 different SSO types (Appendix A). Figure 2 shows the distribution of types within the total sample. Seventy-three percent of all types were rare, occurring once (79 types) or twice (37 types). A single type occurred 131 times, or in 22% of the overall sample. This type is identical to that which would be inferred from the published reference sequence (30), which is known to be a mostly European sequence. Seven other types occurred in from 12 to 38 individuals apiece, or in from 2 to 6% of the total sample.

Table 1 shows the SSO type sharing among the six populations. Ninety types occurred in just one of the six populations. Within this group, there were 79 "unique types," or types which occurred just once, and 11 "population specific types," types found more than once in a particular population. In this latter category, ten types occurred twice in their respective populations (frequencies of <2.5%), and one type occurred three times (frequency 3.7%). The remaining 68 types were "public types," shared by two or more of the six populations, with five types shared by all six populations. Figure 3 shows the breakdown of unique types, population specific types, and public types within each of the six populations (calculated as frequency of kind of type relative to the total number of types). In all populations, there were more public types (always $\geq$ 66%) than unique types (always $\leq$31%), and few population specific types ($\leq$5%), although there were no population specific types in either the German or Swiss. The frequency distribution of unique, population specific, and public types is not significantly different among populations by contingency table analysis ($X^2 = 4.81$, df = 10, $p > 0.05$).

Table 2 shows the number of individuals and number of different SSO types observed in each population, along with estimates of population diversity (h) and their standard errors. Estimates of

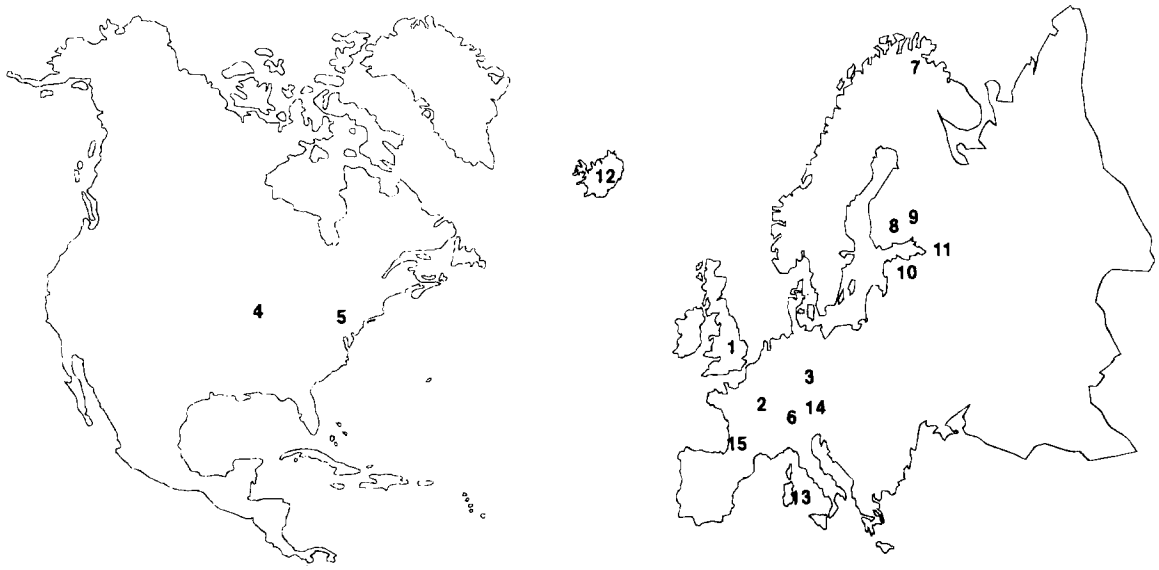

FIG. 1—*Locations of populations in this study: (hypervariable regions 1 and 2) 1-British (26), 2-French, 3-German, 4-Midwest U.S., 5-Northeast U.S., 6-Swiss; (hypervariable region 1 only) 7-Saami, 8-Finn, 9-Karelian, 10-Estonian, 11-Volga Finnic, 12-Icelandic (populations 7–12 are from reference 7), 13-Sardinian (3), 14-Swiss (27), 15-Basque (28).*

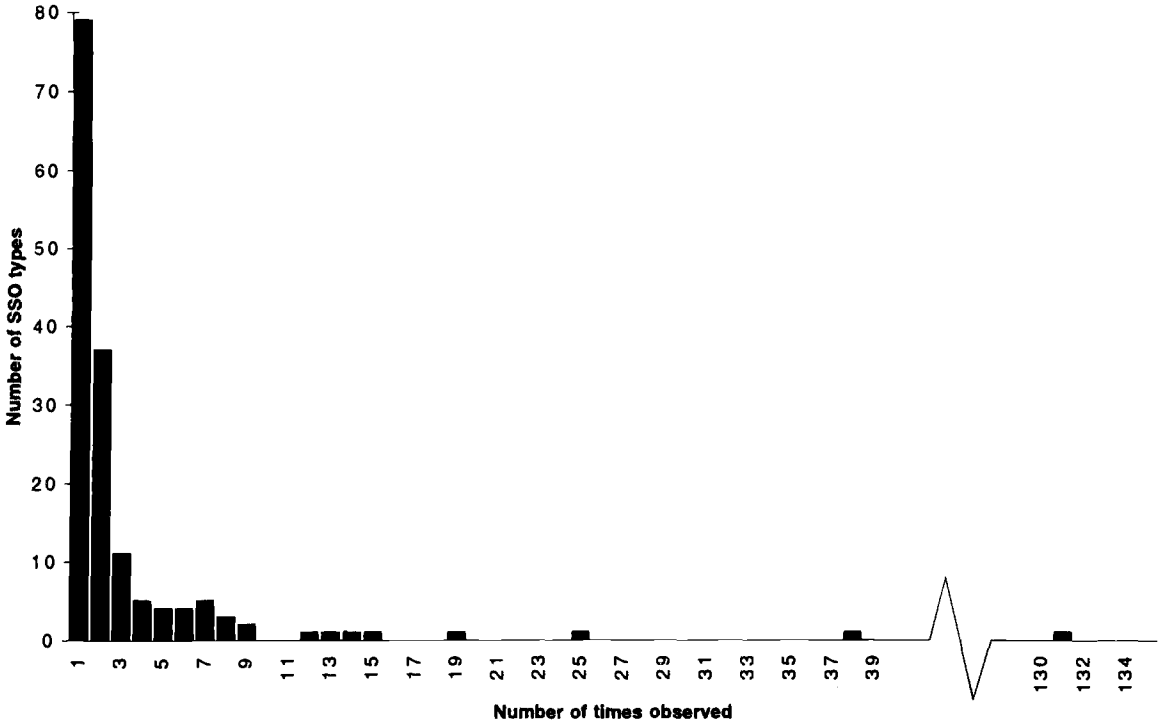

FIG. 2—*Distribution of mtDNA SSO types.*

TABLE 1—*MtDNA SSO type sharing among populations: For example, 33 types are shared by 2 populations.*

Number of mtDNA Types	Number of Populations Sharing
90	1
33	2
14	3
13	4
3	5
5	6

TABLE 2—*MtDNA SSO type diversity* (h) *for 6 European populations.*

Population	N	No. of SSO types	$h \pm$ S.E.
British	100	52	0.957 ± 0.009
French	81	43	0.922 ± 0.016
German	49	29	0.915 ± 0.023
Midwest U.S.	190	80	0.935 ± 0.009
Northeast U.S.	129	63	0.954 ± 0.008
Swiss	46	28	0.945 ± 0.015
TOTAL	595	158	0.941 ± 0.005

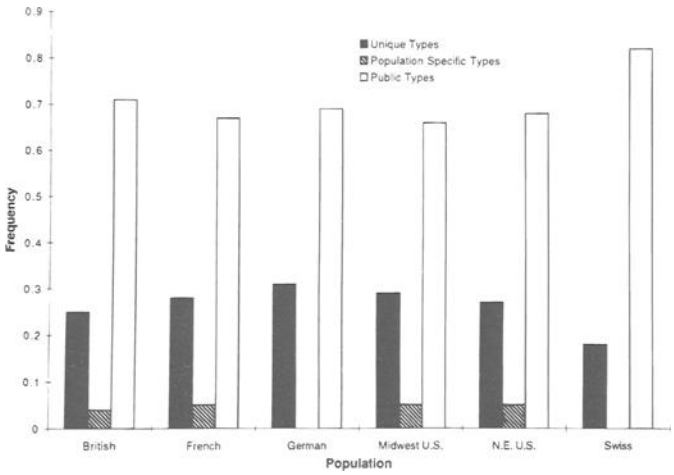

FIG. 3—*Frequency distribution of unique types, population specific types, and shared types among populations.*

diversity were high, exceeding 0.91 for all populations. The mean diversity estimate for all six populations was 0.941 ± 0.005.

The most common SSO types and the SSO types shared by all six populations are described more fully in Fig. 4. Type 40 occurred in from 18% to 26% of individuals in each population, and was seen at highest frequency in Germans, and at lowest frequency in the British. This type has the profile 1-1-1-1-1-1-1-1, and is that which would be inferred from the published reference sequence (30). Types 51, 46, 113, and 48 were observed in 6%, 4%, 3%, and 3% of the total population respectively. Types 51 (1-1-1-1-2-1-1-1) and 46 (1-1-1-1-1-3-1-1) differ from type 40 at one site apiece, indicating that they are closely related. Two additional types, 65 and 122, occur at low frequency (in approximately 2% of the overall sample) but are shared by all populations.

There was generally a low frequency of blank variants (regions for which none of the SSO probes hybridized, indicating unknown sequence variants). However, the frequency of blanks at region IIB ranged from 10 to 22% in the six populations typed at this position (data not shown). Blank variants at this frequency may cause a problem in an AMOVA analysis if they exist solely on the background of a particular SSO type. If different populations have similar frequencies of blanks, but the SSO profiles actually

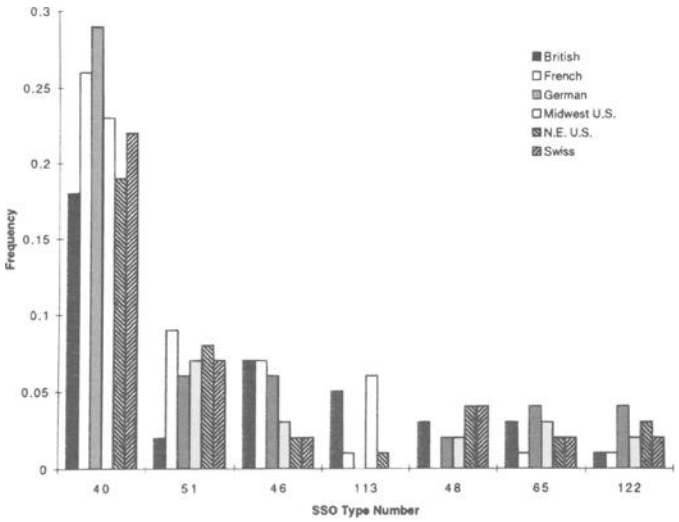

FIG. 4—*Distribution of the most common SSO types in populations.*

represent different sequence variants, populations will appear more similar than they really are. However, for this data set, blanks at region IIB were observed on a multitude of SSO type backgrounds (the profiles of assembled SSO variants from across the entire control region). With 5 to 17 different backgrounds present, depending on the population, the frequency of any individual SSO type with a blank at variant region IIB is low. For example, the British have a frequency of the IIB blank variant of 18%, distributed among 13 different SSO profiles. Examination of the complete nucleotide sequences from which these results were inferred reveals that there are seven different sequence variants at the IIB probe region. If other European populations share the British pattern of variation, there are probably many different substitutions in this region which prevent probe annealing, and no population has a high frequency unique SSO profile which includes a blank for IIB.

AMOVA initially generated interpopulation Φ_{ST} distances (analogous to coancestry coefficients) for the populations in a pairwise distance matrix. The distance matrix for the six populations for which hypervariable regions 1 and 2 data were available is shown in Table 3, as are the significance values based on 1000 permutations, which indicate that there is no significant difference between any two pairs of populations with respect to their SSO types. Pairwise distances are quite small; a negative distance indicates that SSO types in different populations may actually be more closely related to each other than SSO types within the same population (L. Excoffier, personal communication).

TABLE 3—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-British, 2-French, 3-German, 4-Midwest U.S., 5-Northeast U.S., 6-Swiss.*

	1	2	3	4	5	6
1	—	0.318	0.243	0.360	0.642	0.644
2	0.002	—	0.594	0.421	0.535	0.423
3	0.004	−0.003	—	0.114	0.190	0.174
4	0.001	−0.000	0.007	—	0.981	0.769
5	−0.002	−0.001	0.006	−0.005	—	0.670
6	−0.005	−0.000	0.010	−0.006	−0.004	—

Table 4 shows the pairwise distance matrix for the 15 populations for which there were data for the first hypervariable region only, which includes four of the eight possible SSO variant regions, and seven of 13 possible nucleotide positions. Although this sample included 1132 individuals, only 52 different SSO types were observed (Appendix B), out of a possible 192 (for both hypervariable regions, there are a possible 27,648 types, because of the multiplicative effect of more sites), demonstrating the value of using more sites for more informative resolution of individuals. P-values for these distances indicate that, again, there are no significant differences between any two populations, with the exception of comparisons between the Saami and eight other populations, a result which had been reported earlier (7). There are also many negative distances, indicating the overall extreme homogeneity of these populations.

The analysis of molecular variance results are shown in Table 5. For the analysis including 15 populations and hypervariable region 1 only, the variance within populations accounts for 99.27% of the total variation present in the sample, whereas the variance among populations is only 0.73% ($p = 0.005$). When the Saami are removed from the analysis, the variance among populations decreases to 0.29% ($p = 0.158$, $\Phi_{ST} = 0.003$), indicating that this population is responsible for a significant amount of the very minimal substructure that is actually present. When both hypervariable regions are considered, the variance within populations is 100.07% (again a reflection of the extreme similarity between populations), whereas the variance among populations attains a negative value of −0.07% ($p = 0.563$). Figure 5 shows the null distribution of the 1000 variances among populations generated in permutation testing and the observed value. In other words, these populations are virtually indistinguishable with respect to mtDNA SSO types.

Discussion

In comparison with an earlier assessment of Asian mtDNA SSO types (24), this analysis revealed much less complexity for European mtDNA SSO types. Overall, Asian populations were more diverse (0.996 ± 0.001) than European populations (0.941 ± 0.005), while there was statistically significant heterogeneity for the majority of population comparisons which was correlated with geographic relationships. For example, there was less heterogeneity within than between western Asian and eastern Asian populations, and heterogeneity generally increased with distance. European population genetic distances (Tables 3 and 4) do show a wide range of significance values, indicating subtle differences in the degree of relatedness between populations, but lower significance values are not correlated with any consistent geographic pattern, that is, *p*-values are not lower for populations which are most closely geographically affiliated, as they were for Asian populations. Significance values for the population pairwise genetic distances for both hypervariable regions are relatively high, with the lowest values (indicating the most heterogeneity) between the German group and Midwest U.S., Northeast U.S., and Swiss groups.

Although the inclusion of nine additional populations, mostly from Eastern Europe, in an AMOVA analysis on the first hypervariable region indicated the absence of heterogeneity as well, these results should be viewed cautiously. Because there are many fewer SSO types when only seven nucleotide positions are used in the analysis, there will be a higher probability of similarity among populations simply by chance. However, the robustness of the

TABLE 4—*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-Saami, 2-Finn, 3-Karelian, 4-Estonian, 5-Volga Finnic, 6-Swiss, 7-Sardinian, 8-Icelandic, 9-Basque, 10-British, 11-French, 12-German, 13-Midwest U.S., 14-Northeast U.S., 15-Swiss.*

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	—	0.003	0.010	0.028	0.000	0.000	0.014	0.084	0.832	0.006	0.006	0.525	0.000	0.000	0.000
2	0.051	—	0.118	0.299	0.020	0.041	0.064	0.265	0.137	0.141	0.016	0.230	0.031	0.028	0.083
3	0.025	0.012	—	0.918	0.091	0.061	0.350	0.604	0.467	0.153	0.199	0.654	0.499	0.214	0.330
4	0.041	0.008	−0.018	—	0.353	0.298	0.581	0.402	0.486	0.264	0.452	0.408	0.747	0.457	0.693
5	0.104	0.054	0.020	0.000	—	0.064	0.124	0.138	0.027	0.052	0.051	0.022	0.116	0.036	0.585
6	0.036	0.029	0.016	0.004	0.029	—	0.992	0.735	0.226	0.643	0.631	0.117	0.235	0.341	0.322
7	0.027	0.024	0.002	−0.009	0.019	−0.014	—	0.414	0.354	0.376	0.857	0.189	0.571	0.528	0.552
8	0.016	0.005	−0.006	−0.001	0.020	−0.010	−0.003	—	0.676	0.979	0.745	0.954	0.827	0.988	0.609
9	−0.010	0.018	−0.001	−0.003	0.052	0.007	0.000	−0.009	—	0.261	0.377	0.812	0.164	0.222	0.077
10	0.025	0.011	0.007	0.007	0.028	−0.005	0.000	−0.016	0.004	—	0.413	0.432	0.301	0.780	0.388
11	0.029	0.035	0.005	−0.003	0.033	−0.005	−0.009	−0.010	−0.000	−0.001	—	0.256	0.616	0.775	0.298
12	−0.002	0.006	−0.006	0.000	0.049	0.016	0.009	−0.020	−0.012	−0.002	0.005	—	0.170	0.293	0.120
13	0.034	0.022	−0.001	−0.011	0.017	0.003	−0.003	−0.010	0.008	0.001	−0.003	0.006	—	0.819	0.721
14	0.033	0.021	0.004	−0.003	0.031	0.001	−0.003	−0.015	0.006	−0.005	−0.006	0.002	−0.004	—	0.483
15	0.067	0.022	0.003	−0.013	−0.010	0.002	−0.005	−0.008	0.024	−0.000	0.004	0.019	−0.007	−0.002	—

TABLE 5—*AMOVA results for the 6 populations which have SSO type data for both hypervariable regions 1 and 2 and for all 15 populations with SSO type data on hypervariable region 1.*

Control Region	Populations*	Variance Components† σ_a^2	σ_b^2	Φ_{ST}	p (Φ_{ST} and σ_a^2)
HV 1 only	1–15	0.73	99.27	0.007	0.005
HV 1 & 2	1–6	−0.07	100.07	−0.001	0.563

*Populations are: 1-British, 2-French, 3-German, 4-Midwest U.S., 5-Northeast U.S., 6-Swiss, 7-Saami, 8-Finn, 9-Karelian, 10-Estonian, 11-Volga Finnic, 12-Icelandic, 13-Sardinian, 14-Swiss, 15-Basque.

†σ_a^2 = variance among populations, σ_b^2 = variance within populations.

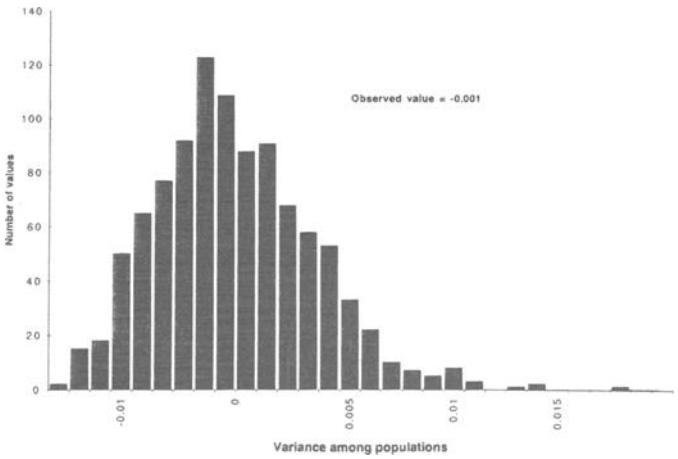

FIG. 5—*Null distribution of the variance component (among populations) obtained through 1000 random permutations of 595 individuals into six populations the sizes of those in this study.*

results (the average high p-value and high frequency of negative pairwise population distances) gives confidence that there is no significant heterogeneity among these populations for these seven sites. In addition, identical results were obtained from the complete nucleotide sequence of HV1 for nine of these populations (7), including the observation of heterogeneity between the Saami and other populations. In fact, the AMOVA distance of −0.009 and significance value of 0.832 in the Saami/Basque comparison supports the curious close relationship between these two populations, which was also previously observed (7). Remarkably, only seven nucleotide sites were necessary to reveal this, indicating that SSO typing has excellent resolving power.

Asian populations had about the same proportion of different SSO types to sample size (0.26) as did European populations (0.27), but unlike Europeans, Asians had no single ubiquitous high frequency type, although there were four types which approached frequencies of 17% in several populations. Almost every European SSO type, with the exception of type 40, could be defined as a low frequency or rare type in Europeans, because frequencies never exceeded 9% in any population, and the vast majority of types occurred only once or twice. This pattern of variation, combined with the high diversity values in each population and lack of subpopulation heterogeneity, indicates that mtDNA is, overall, an excellent locus for forensic SSO typing in Europeans.

However, it should be kept in mind that SSO type identity does not imply complete sequence identity. For example, in the British

sample, SSO type 40 occurs at a frequency of 18% (18 of 100), but in actuality, two of the 18 complete control region sequences which have this type are identical to each other, four other sequences are identical to each other and the remaining 12 are unique (26). Therefore, in a forensic investigation nucleotide sequencing would be necessary to further resolve any additional substitutions if this type were obtained. In fact, in a forensic investigation, an inclusion should always result in sequencing the hypervariable regions for the ultimate resolution in evaluating a match based on SSO typing (17).

Although the Midwest and Northeast U.S. populations are not strictly defined "subpopulations," their lack of significant heterogeneity with other European subpopulations which are more narrowly defined is an indication that most of the mtDNA variation which was originally present in Europe is also present in European Americans. SSO diversity estimates are not different among European and North American populations, nor is the distribution of the most common types described in Fig. 4. The observed pattern is evolutionarily consistent with establishment of European populations with a limited number of mtDNA types, probably most closely related to type 40, with ensuing population expansion from this pool of types throughout Europe and North America. It is expected that further sampling of other North American populations of European ancestry would not reveal substantially different patterns of variation.

The apportionment of variance is not surprising when there are no population pairwise genetic differences. The variance within populations of around 100% indicates that all the variation which is present in the total sample is, on average, present in any single population. The Saami do appear unique compared to all 14 other populations, and removing them from the analysis removed almost all population substructure. Their heterogeneity for SSO types can be attributed to sequence differences at probe variant positions IA, IB, and ID, three of the four regions which were typed. For the purposes of this analysis, we have assumed that the Saami are an unusual relict isolate, much like the aboriginal Taiwanese described elsewhere (24). Of 15 European populations analyzed here, they alone stand out.

The AMOVA Φ_{ST} values of 0.007 (0.003 without the Saami) and -0.001, which are analogous to F_{ST} values, are the lowest values we have yet observed in continental SSO type AMOVA analyses. The lowest value for the Asian continent, obtained with the removal of the aboriginal Taiwanese, was 0.039. Although there is extremely high diversity there, statistically significant substructure is also present. For Europeans, there is also high diversity (somewhat lower than that observed in Asia), but a complete absence of substructure or heterogeneity. This had been observed by others previously (7,27) at the level of DNA sequencing. SSO typing of a larger and geographically more diverse European sample encompassing more of the mtDNA control region confirms these observations and supports the use of mtDNA as a forensic typing locus.

Acknowledgments

We are grateful to Antti Sajantila for providing us with the previously published hypervariable region 1 sequence data. We thank Chris Tomsey and the Pennsylvania State Crime Lab, Greensburg, for contribution of the Northeast U.S. samples, and Hilli Haas of the Institute for Legal Medicine in Giessen, Germany for contribution of the German samples. We also thank Sean Walsh and Roche Molecular Systems for providing SSO probes.

References

1. Cann RL, Stoneking M, Wilson AC. Mitochondrial DNA and human evolution. Nature 1987;325:31–6.
2. Vigilant L, Stoneking M, Harpending H, Hawkes K, Wilson AC. African populations and the evolution of human mitochondrial DNA. Science 1991;253:1503–7.
3. DiRienzo A, Wilson AC. Branching pattern in the evolutionary tree for human mitochondrial DNA. Proc Nat Acad Sci (USA) 1991;88:1597–601.
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. 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 Hu Genet 1995;57:403–14.
6. 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.
7. Sajantila A, Lahermo P, Anttinen T, Lukka M, Sistonen P, Savontaus M-L, et al. Genes and languages in Europe: An analysis of mitochondrial lineages. Genome Research 1995;5:42–52.
8. Soodyall H, Vigilant L, Hill A, Stoneking M, Jenkins T. Mitochondrial DNA control region sequence variation suggests multiple independent origins of an "Asian-specific" 9-bp deletion in sub-Saharan Africans. Am J Hu Gen (in press).
9. Robin E, Wong R. Mitochondrial DNA molecules and virtual number of mitochondria per cell in mammalian cells. J Cell Phys 1988;136:507–13.
10. 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 Hu Gen 1991;48:370–82.
11. Ginther C, Issel-Tarver L, King M-C. Identifying individuals by sequencing mitochondrial DNA from teeth. Nat Genet 1992; 2:135–8.
12. Holland MM, Fisher DL, Mitchell LG, Rodriguez 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.
13. 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.
14. Holland M, Fisher DL, Roby RK, Ruderman J, Bryson C, Weedn V. Mitochondrial DNA sequence analysis of human remains. Crime Lab Digest 1995;22:109–15.
15. 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.
16. Wilson MR, Polanskey D, Butler J, DiZinno JA, Repogle J, Budowle B. Extraction, PCR amplification, and sequencing of mitochondrial DNA from human hair shafts. BioTech 1995; 18:662–9.
17. Stoneking M, Melton T. Forensic applications of mitochondrial DNA analysis. In: Budowle B, editor, Forensic Applications of PCR. In press.
18. Chakraborty R, Kidd KK. The utility of DNA typing in forensic work. Science 1991;254:1735–9.
19. Lewontin RC, Hartl DC. Population genetics in forensic DNA typing. Science 1991;254:1745–50.
20. Devlin B, Risch N, Roeder K. Forensic inference from DNA fingerprints. J Am Stat Assoc 1992;87:337–50.
21. Lander ES, Budowle B. DNA fingerprinting dispute laid to rest. Nature 1994;371:735–8.
22. Weir BS. Population genetics in the forensic DNA debate. Proc Nat Acad Sci (USA) 1992;89:11654–9.
23. Weir BS. The effects of inbreeding on forensic calculations. Ann Rev Genet 1994;28:597–621.
24. Melton T, Stoneking M. Extent of heterogeneity in mitochondrial DNA of ethnic Asian populations. J Forensic Sci 1996;41(4): 591–602.
25. Melton T, Ginther C, Sensabaugh G, Soodyall H, Stoneking M. Extent heterogeneity in Mitochondrial DNA of sub-Saharan African populations. J Forensic Sci (in press).

26. 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.
27. Pult I, Sajantila A, Simanainen J, Georgiev O, Schaffner W, Paabo S. Mitochondrial DNA sequences from Switzerland reveal striking homogeneity of European populations. Biol Chem Hoppe-Seyler 1994;375:837–40.
28. Bertranpetit J, Sala F, Calafell F, Underhill P, Moral P, Comas D. Human mitochondrial DNA variation and the origin of the Basques. Ann Hum Genet 1995;59:63–81.
29. 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.
30. Anderson S, Bankier AT, Barrell BG, de Bruijn MHL, Coulson AR, Drouin J, et al. Sequence and organisation of the human mitochondrial genome. Nature 1981;290:457–65.

Additional information and reprint requests:
Terry Melton, Ph.D.
409 Carpenter Building
Department of Anthropology
The Pennsylvania State University
University Park, PA 16802

APPENDIX A

Data from Complete Control Region

The first column shows the SSO type number and the profile of the SSO type. Numbered columns list the count of SSO types in the following populations: 1-British, 2-French, 3-German, 4-Midwest U.S., 5-Northeast U.S., 6-Swiss. The last column is the total across all populations.

		Population						
SSO type		1	2	3	4	5	6	Grand Total
1	00122311	0	0	0	1	0	0	1
2	01111111	1	0	0	0	0	1	2
3	01112101	0	0	0	1	0	0	1
4	01112311	0	0	1	1	0	0	2
5	01112321	0	0	0	1	1	0	2
6	01122011	0	1	0	0	0	0	1
7	01122201	0	0	0	1	0	0	1
8	01122311	0	0	0	0	1	1	2
9	01312011	0	0	1	1	1	0	3
10	10012111	1	0	0	0	0	0	1
11	10111111	1	2	0	1	2	3	9
12	10112011	0	0	0	0	1	0	1
13	10112101	0	0	1	0	0	0	1
14	10112121	0	1	0	0	0	0	1
15	10112311	0	0	0	0	1	0	1
16	10121101	0	0	0	1	0	0	1
17	10121111	0	0	0	1	1	0	2
18	10302011	1	0	0	0	0	0	1
19	10312011	2	0	0	1	0	0	3
20	10312111	2	0	1	0	1	0	4
21	10312121	0	0	1	0	0	0	1
22	10312211	1	0	0	1	0	0	2
23	10312221	1	0	0	0	1	0	2
24	10312301	1	0	0	0	0	0	1
25	10312311	0	0	0	0	0	1	1
26	10322220	0	0	0	1	0	0	1
27	11011111	0	0	0	0	1	0	1
28	11012011	1	0	0	1	0	0	2
29	11100111	0	0	0	0	1	0	1
30	11101111	1	1	1	0	0	0	3
31	11110111	2	0	0	2	1	0	5
32	11110121	2	0	0	1	1	0	4
33	11110201	0	1	0	0	0	0	1
34	11110311	0	2	0	0	0	0	2
35	11111011	0	0	0	1	0	0	1
36	11111021	1	0	1	0	0	0	2
37	11111100	0	1	0	0	0	0	1
38	11111101	0	1	0	0	0	0	1
39	11111110	0	0	0	1	0	0	1
40	11111111	18	21	14	44	24	10	131
41	11111120	1	0	1	1	1	0	4
42	11111121	1	0	0	4	2	1	8
43	11111211	3	2	0	2	2	0	9
44	11111221	0	0	0	1	0	0	1
45	11111301	0	0	1	0	0	0	1
46	11111311	7	6	3	5	3	1	25
47	11112001	0	0	0	0	0	1	1
48	11112011	3	0	1	4	5	2	15
49	11112021	0	1	0	2	3	0	6
50	11112101	0	0	0	1	1	0	2
51	11112111	2	7	3	13	10	3	38
52	11112121	2	0	0	2	2	1	7
53	11112211	0	0	0	0	1	1	2
54	11112301	0	0	1	0	0	0	1
55	11112311	0	0	0	0	2	0	2
56	11112321	0	1	0	0	0	0	1
57	11121011	0	0	0	0	1	0	1
58	11121111	3	0	1	2	3	0	9
59	11121211	0	0	0	1	1	0	2
60	11121311	1	1	1	2	0	0	5
61	11122011	0	1	0	1	0	0	2
62	11122101	0	1	0	0	0	0	1
63	11122111	0	0	1	1	0	0	2
64	11122311	0	0	0	0	1	0	1
65	11211111	3	1	2	5	2	1	14
66	11211311	0	0	0	1	1	1	3
67	11212011	0	0	0	1	0	0	1
68	11212111	0	0	1	0	0	0	1
69	11221111	0	0	0	1	0	0	1
70	11311011	0	1	0	0	1	0	2
71	11311021	0	0	0	2	0	0	2
72	11311101	1	0	0	1	0	0	2
73	11311111	0	2	0	3	6	2	13
74	11311121	0	0	0	2	0	0	2
75	11311211	0	0	0	0	0	1	1
76	11311311	0	3	0	0	0	0	3
77	11312011	1	2	0	3	2	0	8
78	11312021	0	0	0	1	0	1	2
79	11312110	0	0	0	0	1	0	1
80	11312111	0	1	0	4	1	1	7
81	11312121	0	1	0	1	1	0	3
82	11312201	0	0	0	1	0	0	1
83	11312211	0	0	0	1	0	0	1
84	11312311	0	1	0	4	0	0	5
85	11312331	0	0	0	0	1	0	1
86	11321111	1	0	0	0	3	0	4
87	12012021	0	1	0	0	0	0	1
88	12110021	1	0	0	0	0	0	1
89	12111100	0	0	0	1	0	0	1
90	12112001	1	0	0	1	2	0	4
91	12112021	0	0	1	1	2	1	5
92	12112101	0	1	1	1	3	0	6
93	12112110	0	0	0	1	0	1	2
94	12112111	0	0	0	1	0	0	1
95	12112121	0	0	0	2	0	0	2
96	12112131	0	0	0	1	0	0	1
97	12112311	0	0	1	0	0	0	1
98	12112331	0	0	0	2	0	0	2
99	12122111	0	0	0	0	1	0	1
100	12212131	0	0	0	1	0	0	1

		Population						
SSO type		1	2	3	4	5	6	Grand Total
101	12312131	0	0	0	0	1	0	1
102	20112110	1	0	0	1	1	0	3
103	20112111	0	0	0	0	0	1	1
104	20112210	1	0	0	0	1	0	2
105	20212211	1	0	0	0	0	0	1
106	20322021	0	0	0	1	0	0	1
107	21010111	1	0	0	0	0	0	1
108	21012021	1	0	0	0	0	0	1
109	21102111	1	1	0	0	0	0	2
110	21112011	0	0	1	3	1	2	7
111	21112021	3	0	0	0	1	3	7
112	21112101	2	0	0	2	2	0	6
113	21112111	5	1	0	12	1	0	19
114	21112211	0	1	0	0	1	0	2
115	21112221	0	0	0	1	0	0	1
116	21112301	0	0	0	0	1	0	1
117	21112311	0	1	0	1	1	0	3
118	21112321	1	1	1	3	0	1	7
119	21122111	0	0	0	1	0	1	2
120	21212011	0	1	0	1	0	0	2
121	21212101	0	0	0	1	0	0	1
122	21212111	1	1	2	3	4	1	12
123	21212121	0	1	0	0	0	0	1
124	21212131	0	0	0	0	2	0	2
125	21212211	0	1	0	1	0	0	2
126	21212311	1	0	0	0	1	0	2
127	21222111	0	0	0	1	0	1	2
128	21312111	0	0	0	1	0	0	1
129	21312311	0	1	0	0	0	0	1
130	21312321	0	0	0	0	1	0	1
131	22322302	0	1	0	0	0	0	1
132	30111111	1	0	0	0	0	0	1
133	30312031	0	0	0	0	1	0	1
134	30312221	1	0	0	0	0	0	1
135	31011111	2	0	0	0	0	0	2
136	31111111	3	1	0	2	0	0	6
137	31111301	0	0	0	0	1	0	1
138	31111311	1	0	0	1	0	0	2
139	31112021	0	1	0	0	0	0	1
140	31112111	1	0	0	0	0	0	1
141	31112321	0	0	1	0	0	0	1
142	31121111	0	0	0	0	2	0	2
143	31122121	0	0	1	0	0	0	1
144	31311111	0	1	2	0	0	0	3
145	32112000	1	0	0	0	0	0	1
146	32112021	1	0	0	0	0	0	1
147	32112100	0	0	0	0	1	0	1
148	32112130	0	0	0	0	1	0	1
149	32112300	0	0	1	0	0	0	1
150	32112330	0	1	0	2	0	0	3
151	32122011	0	0	0	1	0	0	1
152	32212130	0	0	0	1	0	0	1
153	32302002	1	0	0	0	0	0	1
154	32311102	0	0	0	1	0	0	1
155	32311302	0	0	0	0	0	1	1
156	32312100	2	0	0	0	0	0	2
157	32312121	0	0	0	0	1	0	1
158	32312130	0	1	0	2	0	0	3
Grand Total		100	81	49	190	129	46	595

APPENDIX B
Data from First Hypervariable Control Region

The first column shows the SSO type number and the profile of the SSO type. Numbered columns list the count of SSO types in the following populations: 1-British, 2-French, 3-German, 4-Midwest U.S., 5-Northeast U.S., 6-Swiss, 7-Saami, 8-Finn, 9-Karelian, 10-Estonian, 11-Volga Finnic, 12-Icelandic, 13-Sardinian, 14-Swiss, 15-Basque. The last column is the total across all populations.

		Population															
SSO type		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	Grand Total
1	0012	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1
2	0111	1	0	1	3	1	1	0	0	2	0	2	0	0	2	0	13
3	0112	0	1	0	1	1	1	0	0	0	0	0	0	0	0	0	4
4	0131	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	3
5	0210	0	0	0	0	0	0	3	0	2	0	0	0	0	0	0	5
6	0211	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1
7	0310	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	3
8	1001	1	0	0	0	0	0	0	0	0	0	0	1	0	0	0	2
9	1011	1	3	1	1	4	3	0	2	1	0	0	0	0	4	1	21
10	1012	0	0	0	2	1	0	0	0	0	0	0	0	0	0	0	3
11	1030	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
12	1031	7	0	2	2	2	1	0	2	1	0	1	1	3	4	0	26
13	1032	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1
14	1101	1	0	0	1	1	0	0	0	1	0	0	3	0	0	0	7
15	1110	1	1	1	0	1	0	0	0	0	0	0	0	0	0	0	4
16	1111	42	43	25	84	58	20	100	23	51	14	16	11	36	37	26	586
17	1112	4	3	3	7	6	0	0	0	2	1	0	2	0	0	1	29
18	1121	3	1	3	7	3	2	1	2	3	2	1	1	3	3	4	39
19	1122	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1
20	1131	2	11	0	23	13	5	1	1	3	2	2	2	5	10	5	85
21	1132	1	0	0	0	3	0	0	0	0	0	0	0	0	0	0	4
22	1201	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1
23	1211	2	1	3	11	7	2	0	7	2	3	0	1	3	2	2	46
24	1212	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	2
25	1221	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1
26	1231	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1
27	1311	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1
28	1321	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1
29	2011	2	0	0	1	2	1	0	0	0	0	1	1	0	0	0	8
30	2012	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	2
31	2021	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
32	2032	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1
33	2101	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
34	2110	1	1	0	0	0	0	0	0	1	0	0	0	0	0	1	4
35	2111	11	4	2	22	8	6	0	4	7	5	6	5	11	7	2	100
36	2112	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	2
37	2121	2	4	2	6	7	1	0	3	2	0	3	1	1	2	0	34
38	2122	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	2
39	2131	0	1	0	1	1	0	0	0	0	0	0	0	1	0	0	4
40	2210	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1
41	2232	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1
42	3011	1	0	0	0	0	0	6	0	0	0	0	0	0	0	1	8
43	3031	1	0	0	0	1	0	0	0	0	0	0	0	1	0	0	3
44	3101	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2
45	3111	5	2	1	3	1	0	0	0	1	1	0	7	3	2	2	28
46	3112	0	0	1	0	2	0	0	0	0	0	0	0	0	0	0	3
47	3131	0	1	2	0	0	0	0	0	0	0	0	0	0	0	0	3
48	3211	2	1	1	2	2	0	0	4	1	0	0	0	0	0	0	13
49	3212	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	2
50	3221	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1
51	3230	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	2
52	3231	2	1	0	3	1	1	0	1	0	0	1	1	1	1	0	13
Grand Total		100	81	49	190	129	46	115	50	83	28	34	39	69	74	45	1132