J Forensic Sci, Nov. 2004, Vol. 49, No. 6
Paper ID JFS2004162
Published 6 Oct. 2004
Available online at: www.astm.org

Commentary on: Budowle B, Shea B, Niezgoda S, Chakraborty R. CODIS STR loci data from 41 sample populations. J Forensic Sci 2001;46:453–489.

Sir:

In an article on allele frequencies for CODIS loci in 41 different populations, Budowle et al. (1) reported no significant deviations from Hardy-Weinberg expectations (HWE) *after making the Bonferroni correction for multiple tests*. However, examination of the data reported in (1) for two Native American populations, Navaho and Salishan, shows significant departures from HWE at three loci for each population. Budowle et al. appear to have relied on an incorrect and unsupportable application of the Bonferroni correction (2,3) to disregard these deviations. These loci should not be used when the product rule is employed to compute the frequency of multi-locus genotypes in these populations unless appropriate corrective actions are taken to account for the observed deviations from HWE.

The Bonferroni correction is a statistical procedure that is invoked when a researcher conducts multiple tests *of the same hypothesis*. The threshold that must be cleared in order for a result to be declared statistically significant is increased to take into account the fact that multiple tests *of the same hypothesis* increase the likelihood that one or more of the tests will, by chance, yield significant results (i.e., the fallacy of multiple tests). We question whether the multiple tests reported by Budowle et al. (1) qualify as tests of the *same* hypothesis. Budowle et al. tested either 12 or 13 loci for HWE within the 41 sample populations in their study. One might argue that each of the loci tested were equivalent, and therefore that Budowle et al. were conducting 12 or 13 tests of the *same hypothesis* for each population (i.e., the hypothesis that the underlying population meets HWE across all equivalent loci). However, Budowle et al. have not demonstrated that the loci tested are equivalent with regard to HWE and discriminating power. Indeed, their study was conducted in part to assess these very issues. Consequently, the use of the Bonferroni correction when testing multiple loci *within populations* is, at the very least, open to debate.

The use of a Bonferroni correction when conducting tests *across multiple populations* as dissimilar as New York Caucasians and Navajo Indians is unquestionably inappropriate. North American Caucasians, African Americans and Native Americans (like Navajo and Salishan Indians) have distinctly different histories, ancestries and allele frequencies (5). (If they were equivalent, there would be no need to determine allele frequencies and perform HWE tests for each individual population in the first place).

Close examination of the pattern of results reported by Budowle et al. (1) further supports the position that the 41 underlying populations are not equivalent. The distribution of departures from HWE within the 41 populations is itself nonrandom. Specifically, of the 524 exact tests (6) performed, 31 yielded statistically significant results ($p < 0.05$). Of those 31, three were in Navajos (FGA:$p = 0.001$; D7S820:$p = 0.030$; THO1:$p = 0.013$) and another three were in Salishans (D3S1358:$p = 0.008$; FGA:$p = 0.049$; D7S820:$p = 0.028$). The chance that 3 of the 31 noted departures from HWE would randomly be found in any one of the 41 populations analyzed by Budowle et al. is itself unlikely ($p = 0.078$). The fact that this unlikely accumulation of departures occurs *twice* (both times in Native American populations) gives very strong evidence ($p = 0.0005$) that these departures are significant and should be taken into consideration when determining genotype frequencies.

Even if the correction for multiple testing could be assumed to be applicable across the tests performed at each of the 12 or 13 different STR loci in a given population (again, a debatable position), two of these loci (FGA in Navajos and D3S1358 in Salishans) would depart significantly from HWE. Budowle et al.'s assertion that "There was little evidence for departures from Hardy-Weinberg expectations in *any* of the populations" (emphasis added) is only true for the Navajo and Salishan populations when these subpopulations are inappropriately considered in the context of literally dozens of very dissimilar subpopulations (apparently raising the threshold for significance from 0.05 to $0.05/524 = 0.000095$). Budowle et al. state that the "majority of the initially found departures from HWE are due to genotypes consisting of rare alleles (e.g., those occurring below five counts in the data)." However, this explanation cannot account for the three Navajo loci that show significant deviations from HWE (FGA, D7S820 and THO1) since not even one of the homozygotes observed for these loci were homozygous for rare alleles. (Genotype information for most of the other populations, including the Salishans was not available).

Significant departures from HWE in Navajo and Salishan populations such as those reported by Budowle et al. as well as other Native American and Inuit tribes are not surprising in light of the distinctive genetic histories of these comparatively small and isolated populations (7–9). They are also consistent with the high estimate of Wright's F_{ST} (10, 11) for Native Americans reported in Table 6 of Budowle et al.'s article (0.0282 for Native Americans relative to 0.0006 for African Americans and −0.0005 for Caucasians) as well as with recent studies of HLA genotypes of other Native American tribes such as the Lakota Sioux (12).

To gain additional insight on this matter, we examined the degree of allele sharing in various populations considered by Budowle et al. for which genotype data were available. Our findings, presented in Table 1, provide yet more evidence of the distinctiveness of Navajos. Table 1 shows the average number of shared alleles observed when making all possible pairwise comparisons of profiles within seven population samples. It reveals that Navajos have significantly higher rates of allele sharing than any of the other populations, which suggests a greater degree of substructure within the Navajo population.

In conclusion, the Budowle et al. study actually provides significant evidence that at least three loci in Navajos (FGA, D7S820, and THO1) as well as Salishans (D3S1358, FGA and D7S820) do *not* adhere to HWE. Given that HWE is a test for independence of alleles within loci and that the logical foundation of the product

TABLE 1—*Pairwise allele sharin within seven different populations completely genotyped at the 13 CODIS loci.*[*]

Population	Population Size	Pairwise Combinations	Average # of Shared Alleles	Standard Deviation
Navajo Indians	182	16,471	11.00	2.19
FBI SW Hispanic	202	20,301	9.33	2.10
FBI Caucasian	194	18,721	9.06	2.16
FBI African Am.	177	15,576	8.92	2.13
Jamaican	157	12,246	8.98	2.20
Bahamian	153	11,628	8.57	2.11
Trinidadian	76	2,850	8.31	2.17

[*]All possible pairwise comparisons between individuals within databases for each population were considered. Given that each genotype contained 13 different loci, the maximum number of shared alleles between any pair of individuals is 26.

rule requires that loci be statistically independent, the product rule should not be used to estimate the rarity of genotypes involving those loci in those populations unless corrective factors are invoked. These departures from HWE also suggest that linkage equilibrium tests (for independence of alleles between loci) should be carefully evaluated prior to using the product rule to estimate multi-locus genotype frequencies for these populations (an issue that is not addressed in the Budowle et al. study).

References

1. Budowle B, Shea B, Niezgoda S, Chakraborty R. CODIS STR loci data from 41 sample populations. J Forensic Sci 2001;46:453–89. [PubMed]
2. Bonferroni CF. Teoria stastica delle classi e calcolo delle probabilita, Pubblicazioni del R Instituto Superiore di Scienze Economiche e Commerciali di Firenze, 1936;8:3–62.
3. Weir BS. Multiple tests. In: Genetic data analysis. Sunderland, MA: Sinauer Associates, Inc., 1990;109–10.
4. Perneger T. What's wrong with the Bonferroni adjustments. British Medical J. 1998;316:1236–8.
5. National Research Council I Report. DNA Technology in Forensic Science. Washington, D.C: National Academy Press, 1992.
6. Guo SW, Thompson EA. Performing the exact test of Hardy-Weinberg proportion for multiple alleles. Biometrics 1992;48:361–72. [PubMed]
7. Smithsonian Institution. Handbook of the American Indians, north of Mexico. Hodge, FW, editor. Bureau of American Ethnology. Bulletin 30, Part I, A – M. Washington, D.C: Government Printing Office, 1907.
8. Anderson GC. Kinsmen of another kind: Dakota – White relations in the upper Mississippi Valley, 1650–1862. Lincoln, NE: University of Nebraska Press, 1997.
9. National Research Council II Report. The evaluation of forensic evidence. Washington, D.C: National Academy Press, 1996.
10. Wright S. Coefficients of inbreeding and relationship. Amer Nat (Chemical Abstracts) 1992;56:330–8.
11. Wright S. The interpretation of population structure by F-statistics with special regard to systems of mating. Evolution 1965;19:395–420.
12. Leffell M, Fallin M, Hildebrand W, Cavett J, Iglehart B, Zachary A. HLA alleles and haplotypes among the Lakota Sioux: Report of the ASHI minority workshops, part III. Human Immunology 2004;65:78–89. [PubMed]

Dan E. Krane, Ph.D., Associate Professor
Biological Sciences
Wright State University
Dayton, OH 45435-0001

Travis E. Doom, Ph.D., Assistant Professor
Department of Computer Science
Wright State University
Dayton, OH 45435-0001

Laurence Mueller, Ph.D., Professor
Department of Ecology & Evolutionary Biology,
University of California, Irvine, CA 92697

Michael L. Raymer, Ph.D., Assistant Professor
Department of Computer Science
Wright State University
Dayton, OH 45435-0001

William M. Shields, Ph.D., Professor of Biology
State University of New York
College of Environmental Science & Forestry
Syracuse, NY 13210

William C. Thompson, Ph.D., JD., Professor
Department of Criminology
Law & Society
University of California
Irvine, CA 92697