

References

1. Roy R, Reynold R. Amplitype PM and HLA DQ_a typing from pap smear, semen smear, and postcoital slides. *J Forensic Sci* 1995;40: 266-9.
2. Gall K, Pavicic D, Pavelic J, Audy-Jurkovic S, Pavelic K. PCR amplification of DNA from stained cytological smears. *J Clin Pathol* 1993;46:378-9.
3. Pötsch L, Prager-Eberle M, Penzes L, Rittner Ch. HLA DQ_{alpha}-Typisierung an mikroskopisch spermatischen Ausstrichpräparaten nach PCR. *Z Rechtsmed* 1993;3:54-8.
4. Poljak M, Barlic J, Seme K, Avsic-Zupanc T, Zore A. Isolation of DNA from archival papanicolaou stained cytological smears using a simple salting-out procedure. *J Clin Pathol Mol Pathol* 1995;48: 55-6.
5. Schöch R, Jenisch S, Haferlach T, Müller-Ruchholtz W, Gassmann W, Löffler H. Glass slide smears are a suitable source for RT-PCR-based analysis of chromosomal aberrations in leukaemias. *British J of Haematol* 1996;92:140-42.
6. Li S, Tuck-Müller CM, Yan Q, Wertelecki W, Chen H. A rapid method for PCR amplification of DNA directly from cells fixed in Carnoy's fixative. *Am J Med Genet* 1995;55:116-9.
7. Dimo-Simonin N, Grange F, Brandt-Casadevall C. PCR-based forensic testing of DNA from stained cytological smears. *J Forensic Sci* 1997;42:506-09.
8. De Giorgi C, Finetti Sialer F, Lamberti F. Formalin-induced infidelity in PCR-amplified DNA fragments. *Mol Cell Probes* 1994;8: 459-62.
9. Greer CE, Wheeler CM, Manos MM. Sample preparation and PCR amplification from paraffin-embedded tissues. *PCR Meth Appl* 1994;3:113-22.
10. Inoue T, Nabeshima K, Kataoka H, Kono M. Feasibility of archival non-buffered formalin-fixed and paraffin-embedded tissues for PCR amplification: An analysis of resected gastric carcinoma. *Pathol Int* 1996;46:997-1004.
11. Rish JA, Eisenach KD, Cave MC, Reddy MV, Gangadharan PRJ, Bates JH. Polymerase chain reaction detection of mycobacterium tuberculosis in formalin-fixed tissue. *Am J Respir Crit Care Med* 1996;153:1419-23.
12. Ludes B, Tortel MC, Mangin P. Use of PCR for forensic analysis of DNA from formaldehyde fixed and paraffin embedded human tissues. In: Rittner C, Schneider PM, editors. *Advances in Forensic Haemogenetics* Springer Verlag Berlin 1992:72-4.
13. Fukushima H, Ota M, Tanaka K, Ichinose A, Sato K, Honda K. DNA typing from formalin-fixed paraffin-embedded tissues. In: Rittner C, Schneider PM, editors. *Advances in Forensic Haemogenetics* Springer Verlag Berlin 1992;81-3.
14. Malmgren Ch, Gustavsson I, Hurte I, Saldeen T, Gyllensten U. Reliable typing of DNA amplified from formalin-fixed tissue biopsies. *PCR Meth Appl* 1992;2:175-6.
15. Sajantila A, Lukka M. Proving paternity using formalin fixed paraffin embedded tissue samples from a deceased father. In: Bär W, Fiori A, Rossi U, editors. *Advances in forensic Haemogenetics* Springer Verlag Berlin 1994:232-34.
16. Mannucci A, Casarino L, Bruni G, Lomi A, De Stefano F. Individual identification of flood victims by DNA polymorphisms and autopsy findings. *Int J Leg Med* 1995;107:213-15.

Noëlle Dimo-Simonin
 Fabienne Grange
 Conxita Brandt-Casadevall
 Patrice Mangin
 Institut Universitaire de Médecine Légale
 Rue du Bugnon 21
 CH-1005 Lausanne
 Switzerland

Commentary On: Budowle B, Lindsey JA, Decon JA, Koons BW, Giusti AM, Comey CT. Validation and population studies of the loci LDLR, GYPA, HBGG, D7S8, and Gc (PM loci), and HLA-DQ- α using a multiplex amplification and typing procedure. *J Forensic Sci* 1995 Jan;40(1):45-54

Sir:

A major conclusion of the paper by Budowle et al., was that

"All loci meet Hardy-Weinberg expectations and there is little evidence for associations of alleles between the loci." This conclusion is supported by a number of statistical tests of independence within loci and between pairs of loci. In particular, Table 5 presents tests of the Hardy-Weinberg equilibrium for the five polymarker loci in four U.S. populations maintained by the FBI: African American, Caucasian, Southeastern Hispanic and Southwestern Hispanic. I have now repeated one of the tests used by Budowle et al., the exact test, on all the polymarker loci and all populations. The genotypes of all individuals in the FBI's four data bases were provided to me by the FBI through discovery in a criminal case. In general I have obtained the same results as appear in Table 5 except for one case which I describe in more detail below.

For the HBGG locus in the Caucasian population Budowle et al. report a *p*-value for the exact test of 0.887, suggesting the genotype frequencies observed in this population are well described by the Hardy-Weinberg law. When I repeated this test (using software kindly provided by Dr. Paul O. Lewis, University of New Mexico) I obtained a *p*-value of 0.008, suggesting a highly significant departure from the Hardy-Weinberg equilibrium.

It is unlikely that this difference is due to numerical differences in the software used since the other 23 tests I did gave close agreement to the results in Table 5 of Budowle et al. It is also unlikely that the notation in Table 5 is a typographical error. My results from another test of the Caucasian-HBGG locus, the likelihood ratio test, also differ dramatically from those reported by Budowle et al.: I obtained a *p*-value of 0.001 while they report a *p*-value of 1.000.

It appears that Budowle et al. altered the original data in the Caucasian-HBGG data base when performing the exact test reported in Table 5. The deviation from Hardy-Weinberg that I detected is caused by the presence in the data base of a single individual who is homozygous for the *C* allele. Because the *C* allele is rare, it would be extremely unlikely to find a *CC* homozygote in a sample of 148 people if the population was in Hardy-Weinberg equilibrium, and hence the low *p*-value. I have repeated the exact test under several different alterations of the original Caucasian database and summarize these results in Table 1 below. From the results in Table 1 it would appear that Budowle et al. had altered the genotype of the single *CC* individual by changing it to an *AA* homozygote, since this alteration gives a *p*-value consistent with Budowle et al.'s published value (0.887). I should add that it is unclear if the tests of association between pairs of loci also utilized this altered HBGG genotype.

Unfortunately, there is no description of any alteration of the original data in the paper by Budowle et al. Consequently, there is no justification given for altering the raw data in this fashion. It is now clear that this alteration has a substantial effect on results

TABLE 1—Results from the exact test on the HBGG locus in the Caucasian population under several data base alterations to the single *CC* homozygote.

Change to the FBI Data Base	<i>p</i> -value (based on 10,000 shufflings)
Change <i>CC</i> to <i>BB</i>	1.000
Change <i>CC</i> to <i>AA</i>	0.866
Change <i>CC</i> to <i>AB</i>	1.000
Change <i>CC</i> to <i>BC</i>	1.000
Change <i>CC</i> to <i>AC</i>	0.796
Delete the <i>CC</i> individual from the data base	1.000

of these hypothesis tests. While there may be some reasons one could provide for such an alteration, the acceptability of such a practice may be subject to valid criticism. Only by carefully documenting and justifying how the data were handled can other scientists be expected to repeat and evaluate the study of Budowle et al. Concealing such techniques effectively eliminates informed scientific evaluation.

Laurence D. Mueller, Ph.D.
Department of Ecology & Evolutionary Biology
University of California
Irvine, CA 92697-2525

Authors' Response

Sir:

I welcome the opportunity to respond to the letter to the editor by Mueller. Rather than seeing the issues he raises debated in a courtroom, where the adversary system has a tendency to distort reality, publishing his letter and this response should put these issues to rest.

Dr. Mueller claims that he has discovered errors of a Hardy Weinberg Expectation (HWE) analysis of the HBGG locus in a Caucasian database on the HBGG locus reported by Budowle et al. (1). He contends that "there is no justification given" for pooling the CC homozygote with another class for the exact test analysis and that "concealing such techniques effectively eliminates informed scientific evaluation."

First, the published and raw population data in Budowle et al. (1) have been available to interested parties since 1995. The population data in Budowle et al. (1) clearly display the one individual with the HBGG CC type. Obviously, nothing is concealed because Mueller was able to come to the same statistical conclusion as in Budowle et al. in Table 1 of his letter.

Second, for HWE tests where low frequency, or rare, alleles are observed, genotypes containing these alleles are merged with other genotypes. Budowle et al. (1) merged the CC and AA classes. This approach is very basic in statistics and routine practice, such that I neglected to put a footnote in Budowle et al. (1) describing the practice. There was no attempt to conceal. In fact, I used the Caucasian HBGG database as an instructive example of sensitivity of some HWE tests (e.g., the exact test) to sampling in the basic statistics class taught to over 200 attendees at The Seventh International Symposium on Human Identification at Scottsdale, Arizona in 1996. Regardless, the approach of merging the CC and AA types can be confirmed easily by performing the analysis, as Mueller has done.

Third, Mueller has fallen into the same trap that some novice statistics students encounter. He argues that a HWE test he has performed shows a low p-value and implies there is a problem with the database. Snedecor and Cochran (2) state in their college-level statistics book (page 28):

A test of significance is sometimes thought to be an automatic rule for making a decision either to "accept" or "reject" a null hypothesis. This attitude should be avoided. An investigator rarely rests his decisions wholly on a test of significance. To the evidence of the test he adds knowledge accumulated from his own past work and from the work of others.

The proper approach to analyzing population data is to perform HWE tests. If something shows departures from expectations, then evaluate the data further. This is the practice routinely enjoyed by the forensic science community. In addition, to overcome concerns

when estimating DNA profile frequencies, rare allele frequencies are replaced with a minimum allele threshold frequency (see Budowle et al. (3) or the NRC II Report (4)). Further, the use of the recommended formula with theta by the NRC II Report (4) renders the issue of departures from HWE less meaningful.

Finally, the points raised by Mueller have no impact on the forensic use of these PCR tests. There are a tremendous number of Caucasian databases around the world tested for HWE for the routine PCR-based markers used in forensic science. Any of these databases could be used in lieu of the FBI database, and the end result would not differ.

References

1. Budowle B, Lindsey JA, DeCou JA, Koons BW, Giusti AM, Comey CT. Validation and Population Studies of the Loci LDLR, GYPA, HBGG, D7S8, and Gc (PM loci), and HLA-DQ α Using a Multiplex Amplification and Typing Procedure. *J Forensic Sci* 1995;40(1):45-54.
2. Snedecor GW, Cochran WG. In: *Statistical Methods*, Sixth Edition, Iowa State University Press, Ames, Iowa, 1967.
3. Budowle B, Monson KL, Chakraborty R. Estimating minimum allele frequencies for DNA profile frequency estimates for PCR-based loci. *Int J Leg Med* 196;108:173-176.
4. National Research Council. *The Evaluation of Forensic DNA Evidence*, Washington, D.C., National Academy Press, 1996.

Bruce Budowle, Ph.D.
FSRTC
FBI Academy
Quantico, VA 22135

Additional Commentary On: Budowle B, Lindsey JA, DeCou JA, Koons BW, Giusti AM, Comey CT. Validation and population studies of the loci LDLR, GYPA, HBGG, D7S8, and Gc (PM loci), and HLA-DQ- α using a multiplex amplification and typing procedure. *J Forensic Sci* 1995 Jan;40(1):45-54

Sir:

Information has recently come to light that casts doubt on the accuracy of an important finding reported in this journal. This information also raises disturbing questions about standards of integrity within forensic science.

The underlying scientific issue is whether the FBI's Caucasian data base for the HBGG (polymarker) locus is in Hardy-Weinberg (H-W) equilibrium. Budowle et al. reported that an exact test on this data base produced a p-value of .887, indicating that the locus is in H-W equilibrium and therefore appropriate for use in forensic identification tests. Professor Laurence Mueller (commentary, immediately above, this issue) analyzed the same data base and obtained an exact test p-value of .008, indicating a highly significant deviation from H-W equilibrium. Mueller reported that he can obtain a p-value close to that reported by Budowle et al. only when he alters the original data by transforming a rare CC homozygote to an AA homozygote. This conclusion has been checked and verified by Professor Ranajit Chakraborty (1,2). Although Budowle et al. reported that the data base contains a CC homozygote, they made no mention of changing this genotype to another type when performing statistical tests.

In a recent court hearing, Professor Chakraborty responded to Mueller's assertions by arguing that Budowle et al. had "correctly followed the forensic protocol of dealing with rare alleles" (1). According to Chakraborty, "for HWE tests of loci involving rare alleles, genotypes containing rare alleles are merged with others. . ." (1) Hence, in Chakraborty's view, Budowle et al. were

simply following a standard forensic practice when they changed the CC homozygote to another type before performing the exact test. Chakraborty opined that this type of data manipulation is so routine that it does not warrant mention in a journal article (2).

This matter deserves serious examination by the forensic science community. One issue is whether the "merging" of population data during statistical analysis (such that rare genotypes are treated as more common types) is an acceptable scientific practice. While merging of alleles may be appropriate when using data bases to estimate the frequency of particular DNA profiles, I believe it is inappropriate when testing for H-W equilibrium because it can hide rare homozygotes that signal population structure: it suppresses the very phenomenon that statistical tests of H-W equilibrium are designed to detect.

A second issue is whether data manipulation of this type, when performed, should be reported in journal articles. In my view, a researcher who alters the data in a population data base and fails to mention doing so when reporting statistical tests on the data base has, at a minimum, breached the ethical obligation to report scientific findings accurately. If this breach was committed intentionally, in order to deceive, it is scientific fraud. I believe that Professor Chakraborty's testimony to the contrary is simply unacceptable and should be disavowed by conscientious members of the forensic science community.

References

1. Affidavit of Ranajit Chakraborty, Ph.D. filed in United States v. Stephen Burke, et al., CR 96-50-01/06-M (U.S. Dist. Ct. New Hampshire, August 13, 1997).
2. Transcript of Hearing Before the Honorable Steven J. McAuliffe, in United States v. Stephen Burke, et al., CR 96-50-01/06-M (U.S. Dist. Ct. New Hampshire, August 28, 1997).

William C. Thompson, J.D., Ph.D.
Department of Criminology, Law & Society
University of California
Irvine, Calif. 92697-7080

Sir:

Mr. Thompson has provided one lawyer's viewpoint. Apparently, he is attempting to try, in the scientific literature, a case (United States v. Stephen Burke, et al., CR 96-50-01/06-M), in which upon hearing the arguments, the judge admitted the evidence and a conviction was obtained. I believe it is inappropriate to address the evidence in the specific case; the machinations and deliberations of this case are more appropriate for the particular court. All scientific issues relating to population data on the HBGG locus and Hardy-Weinberg equilibrium have been addressed in my previous letter.

Bruce Budowle, Ph.D.
FSRTC
FBI Academy
Quantico, VA 22135

Further Response to Mueller and Thompson Considerations on the Tests of Independence of Alleles that are Relevant for Forensic Applications

Sir:

The letters by Mueller and Thompson and the responses by Budowle settle two issues in relation to the validation study of the polymarker databases (Budowle et al. (1)). First, the genotype frequencies at the HBGG locus in the Caucasian sample ($n = 148$)

are on conformity with their Hardy-Weinberg expectations (HWE) when the single occurrence of the CC genotype is merged with another genotype (AA, in this instance). Further, when this CC genotype is kept separate, the genotype frequencies show significant departures from HWE ($p = 0.008$). However, Mueller's and Thompson's letters leave two questions open: (i) Did Budowle et al. (1) "alter" their data and/or was there any attempt to "conceal" information to "deceive" the scientific community? and (ii) Is the "merging" procedure employed in their tests scientifically valid as it relates to forensic applications?

Tables 3 and 4 of the validation study of Budowle et al. (1) explicitly record that among the sampled 148 Caucasians, the HBGG-CC genotype was observed in a single individual. This correct representation of the data observation and the distribution of the actual multilocus genotype data from the entire study to the forensic community, as well as to the critics of DNA forensics long before the origin of these letters clearly dispel any notion of intent of data alteration. Thus, the accusations of concealing data and deceiving the scientific community are not only baseless; they lack foundation and are obviously far-fetched. However, the validation study (1) can be criticized only because it did not footnote Table 5, mentioning the merging of single occurrence of the CC homozygote with the AA genotypes.

Turning now to the second question, Thompson opined that with the omission of this information the appropriateness of such a merging procedure for validating forensic databases could not be examined by the scientific community. However, he agrees that for the purposes of DNA profile frequency computations such merging of alleles are appropriate. Therefore, the question becomes, is the assumption of HWE appropriate for the merged allele system, and not the one that Mueller raises, or the one mentioned by Thompson (namely, global adherence to HWE including the rare alleles). Incidentally, "merging" of rare alleles is not unique or new for this validation study only. The forensic community as well as their critics are familiar with allele merging in the context of "re-binning" VNTR fragment sizes in the RFLP analysis of DNA profiles (3), and this concept was approved by both reports of the National Research Council (10,11). Thus, the procedure of merging rare alleles (and consequently rare genotypes) with others in the same database does not equate to data alteration. When such merged data are used for any applications (e.g., for estimating the frequency of a DNA profile), if any assumptions are made, it is sufficient to check the accuracy of the assumption at the level of the merged data. Clearly this is what was done in Budowle et al. (1), and hence, their validation study is not scientifically flawed, in contrast to Mueller's declarations in several courts (e.g. 8,9).

It should also be noted that the concept of merging alleles was conceived not solely to reduce the undue weight on the frequency of a rare genotype (as was the purpose of the re-binning method; see Ref 3); earlier population genetic studies have convincingly shown that, for loci with discrete alleles, the population substructure effects are detectable predominantly for rare alleles, and hence, when they are disregarded or are merged with the common alleles, the population substructure effect on genetic variation is virtually eliminated (see Refs 4 and 6). Thus, merging data on rare alleles for the purpose validating the assumptions of forensic computations has a scientific foundation dating well before the article (1) which Mueller and Thompson now criticize. Also, this discussion clearly spells out the logic that was implied in the testimony given by Chakraborty (5) that Thompson (without giving any rationale) called "simply unacceptable and should be disavowed".

Finally, while correspondence in a scientific forum is the only

appropriate means for resolving the issue, the readers as well as the courts must be informed that the strict assumption of HWE is not a part of the current standard of forensic computations. The recommendations made in NRC (11) clearly allow incorporation of the effect of population substructure in genotype frequency computations at the individual locus level, and thus, even after merging the rare alleles, it is not necessary that the loci obey strict HWE predictions. More recent studies (2,7,12) also suggest that for discrete allele systems, rare allele frequencies may be replaced by a minimum threshold allele frequency in order to adjust for population substructure effects. Since alleles that are adjacent to the rare alleles are somewhat ambiguous for a locus such as the HBGG locus, replacement of rare allele frequencies by an appropriate threshold frequency (depending on the database size and the extent of variation at the locus) also achieves the purposes of eliminating population substructure effect and placing undue weight on frequencies of rare alleles. Therefore, we surmise that Mueller's findings have no impact on forensic applications of the databases described in the validation study (1).

References

1. Budowle B, Lindsay JA, DeCou JA, Koons BW, Giusti AM, Comey CT. Validation and population studies of the loci LDLR, GYPA, HBGG, D7S8, and GC (PM loci), and HLA-DQ α using a multiplex amplification and typing procedure. *J Forensic Sci* 1995; 40(1):45-54.
2. Budowle B, Monson KL, Chakraborty R. Estimating minimum allele frequencies for DNA profile frequency estimates for PCR-based loci. *Int J Leg Med* 1996;108:173-176.
3. Budowle B, Monson KL, Anoe K, Baechtel K, Bergman D. A

preliminary report on binned general population data on six VNTR loci in Caucasians, Blacks and Hispanics from the United States. *Crime Lab. Digest* 1991;18:9-26.

4. Chakraborty R. Genetic profile of cosmopolitan populations. *Anthrop Anz* 1990;48:313-331.
5. Chakraborty R. Affidavit filed in United States vs. Stephen Burke et al., CR 96-50-01/06-M (US District Court, New Hampshire, August, 1997).
6. Chakraborty R, Smouse PE, Neel JV. Population amalgamation and genetic variation: Observations on artificially agglomerated tribal populations of Central and South America. *Am J Hum Genet* 1989;43:709-725.
7. Evett IW, Gill PD, Scranage JK, Weir BS. Establishing the robustness of short-tandem-repeat statistics for forensic applications. *Am J Hum Genet* 1996;58:398-407.
8. Mueller LD. Testimony filed in United States vs. Stephen Burke et al., CR 96-50-01/06-M (US District Court, New Hampshire, July, 1997a).
9. Mueller LD. Affidavit filed in State of Washington vs. Robert Parker, Co. 96-1-07511-2SEA (King County Court at Seattle, Washington, November 1997b).
10. NRC. DNA Technology in Forensic Science. National Research Council, National Academy Press, Washington DC, 1992.
11. NRC. The Evaluation of Forensic DNA Evidence. National Research Council, National Academy Press, Washington DC, 1996.
12. Weir BS. Independence of VNTR alleles defined as fixed bins. *Genetics* 1992;130:873-887.

Ranajit Chakraborty, Ph. D.,
 David N. Stivers, Ph. D.
 Human Genetics Center
 The University of Texas School of Public Health
 P.O. Box 20334, Houston, Texas 77225
 Tel. (713) 500-9820; Fax (713) 500-0900
 e-mail: rc@gc9.sph.uth.tmc.edu