Forensic Science International: Genetics 5 (2011) 541–542

Contents lists available at ScienceDirect

Forensic Science International: Genetics

journal homepage: www.elsevier.com/locate/fsig

Corrigendum

Corrigendum to "Linkage disequilibrium analysis of D12S391 and vWA in U.S. population and paternity samples" [Forensic Sci. Int.: Genet. (in press), doi:10.1016/j.fsigen.2010.09.003]

Kristen Lewis O'Connor *, Carolyn R. Hill, Peter M. Vallone, John M. Butler

Biochemical Science Division, National Institute of Standards and Technology, 100 Bureau Drive, Gaithersburg, MD 20899-8312, United States

The above paper, which is published in this issue of the Journal, requires the following Corrigendum to be made.

Following the availability of this paper on the Journal's Articles in Press website, we (the authors) were contacted by a number of experts in the field of population genetics directly and via the editors of the Journal to point out an error in the original analysis of independence between D12S391 and vWA using father/son paternity samples. This correspondence has led to the reanalysis of the father/son paternity data and the writing of this Corrigendum.

In the original paper, father/son pairs were used to determine the gametic phase of alleles at D12S391 and vWA by determining which alleles were transmitted from father to son as a haplotype, assuming no recombination during meiosis. After phasing there were four D12S391/vWA haplotypes for each father/son pair—two haplotypes for the father and two haplotypes for the son. A father and son share one of their two D12S391/vWA haplotypes that was transmitted from the father to the son during meiosis. The phased father/son dataset was used for linkage disequilibrium analysis in the original paper. Also, the frequencies of the D12S391/vWA haplotypes were calculated using the four phased genotypes of the father/son pairs and were provided in the original Supplementary Data.

We became aware that the phased father/son dataset mistakenly included both pairs of haplotypes so that the paternally transmitted haplotype was counted twice (i.e., once for the father and once for the son). A transmitted haplotype is not independent between a father/son pair and should not have been counted separately for linkage disequilibrium analysis and haplotype frequency calculations. Also, it was determined that the phase of some father/son genotypes could not be definitively determined without the genotype of the mother, which was unavailable in this dataset.

Two parallel methods were used to revise the D12S391/vWA haplotype data used for linkage disequilibrium analysis:

Dataset 1: The father/son dataset was reduced by removing one count of the paternally transmitted haplotype in addition to removing sample pairs if the gametic phase was ambiguous. The revised dataset included three haplotypes per father/son pair (one haplotype for the father and two haplotypes for the son) for all unambiguously phased father/son pairs. Linkage disequilibrium analysis was performed as described in "Materials and Methods" section of the original paper.

Dataset 2: The expectation-maximization (EM) algorithm was used to estimate the four parental haplotypes (two paternal and two maternal) for each son, assuming a recombination fraction of 0.108 [1]. From these estimated haplotypes, a maximum likelihood approach estimated the parental haplotype frequencies [2]. Note that the non-transmitted maternal haplotype provided little information for calculating haplotype frequency estimates. The haplotype frequency estimates were used as the revised dataset for linkage disequilibrium analysis. For linkage disequilibrium analysis, a likelihood ratio test was performed that compared the likelihood of the data based on the estimated haplotype frequencies (i.e., assuming linkage disequilibrium) to the likelihood of the data assuming no association (i.e., linkage equilibrium) [3]. Since the likelihood ratio test does not perform well if there are many haplotypes with low frequencies, alleles with frequencies below 5% were pooled for each locus prior to linkage disequilibrium analysis.

Upon revised analysis of D12S391/vWA haplotypes, no significant evidence of linkage disequilibrium was observed between the D12S391 and vWA loci in the U.S. population groups (Table 1). These findings are consistent with the linkage disequilibrium analysis results using unrelated U.S. population samples as described in the original paper. Furthermore, the conclusion of independence between D12S391 and vWA is consistent with findings from two recent articles [1,4]. Thus, the single-locus genotype probabilities for the D12S391 and vWA loci may be multiplied to determine the profile match probability when unrelated individuals are involved.

After acceptance of the original paper, the authors became aware of a study by Budowle et al. that reported no significant linkage disequilibrium between D12S391 and vWA using a set of unrelated U.S. population samples [1]. Using multi-generation family samples, this study estimated a recombination fraction of 0.108 between D12S391 and vWA, thus, indicating "close" linkage between the loci [1]. The effect of linkage is an increased tendency for alleles at physically close loci to be transmitted together during meiosis. Since there is no evidence for linkage disequilibrium between the two loci, it does not seem necessary to use haplotype

DOI of original article: 10.1016/j.fsigen.2010.09.003

* Corresponding author. Tel.: +1 301 9755205; fax: +1 301 9758505.
E-mail addresses: m.gooley@elsevier.com, kristen.oconnor@nist.gov (K.L. O'Connor).

doi:10.1016/j.fsigen.2011.05.004

Table 1
P-value results from analysis of linkage disequilibrium (LD) of the D12S391 and vWA loci using U.S. father/son paternity samples. The original data included four haplotypes for each father/son pair, including some pairs with ambiguous phase. Dataset 1 consisted of three definitively phased haplotypes for each father/son pair. Dataset 2 consisted of estimated parental haplotype frequencies. Significance level, $p < 0.05$ (in bold). *N* = number of haplotypes.

Population	Original data		Dataset 1		Dataset 2	
	N	LD	*N*	LD	*N*	LD
African American	356	0.0275	214	0.4888	356	0.8756
Caucasian	396	0.0001	250	0.2195	396	0.1685
Hispanic	380	0.0915	228	0.7105	380	0.8857
Asian	396	0.0031	217	0.1317	396	0.6541

frequencies. Instead, an independent occurrence of alleles at both loci for the calculation of (two-locus) genotype frequencies can be assumed. Nevertheless, completely ignoring linkage between the two loci can lead to incorrect inference in certain kinship scenarios, at least if transmission over more than two meioses is observed [5,6]. This erratum provides the maximum likelihood estimates of the D12S391/vWA haplotype frequencies (alleles have not been pooled) and should be used in lieu of the haplotype frequency table provided as Supplementary Data in the original paper.

We thank Rolf Fimmers and Andreas Tillmar for their assistance with the different approaches for data reduction and phasing.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.fsigen.2011.05.004.

References

[1] B. Budowle, J. Ge, R. Chakraborty, A.J. Eisenberg, R. Green, J. Mulero, R. Lagace, L. Hennessy, Population genetic analyses of the NGM STR loci, Int. J. Legal Med. 125 (2011) 101–109.
[2] R. Ceppellini, M. Siniscalco, C.A. Smith, The estimation of gene frequencies in a random-mating population, Ann. Hum. Genet. 20 (1955) 97–115.
[3] K. Lange, M. Boehnke, Extensions to pedigree analysis. V. Optimal calculation of Mendelian likelihoods, Hum. Hered. 33 (1983) 291–301.
[4] C. Phillips, L. Fernandez-Formoso, M. Garcia-Magariños, L. Porras, T. Tvedebrink, J. Amigo, M. Fondevila, A. Gomez-Tato, J. Alvarez-Dios, A. Freire-Aradas, A. Gomez-Carballa, A. Mosquera-Miguel, Á. Carracedo, M.V. Lareu, Analysis of global variability in 15 established and 5 new European Standard Set (ESS) STRs using the CEPH human genome diversity panel, Forensic Sci. Int. Genet. 5 (3) (2011) 155–169.
[5] M. Krawczak, Kinship testing with X-chromosomal markers: mathematical and statistical issues, Forensic Sci. Int. Genet. 1 (2) (2007) 111–114.
[6] M. Nothnagel, J. Schmidtke, M. Krawczak, Potentials and limits of pairwise kinship analysis using autosomal short tandem repeat loci, Int. J. Legal Med. 124 (3) (2010) 205–215.