skewed patterns of X inactivation, familial or not, provide a powerful means of ascertaining mutations that influence cell proliferation.

Barbara R. Migeon and Camille Haisley-Royster
*Department of Pediatrics and*
*Center for Medical Genetics*
*The Johns Hopkins University School of Medicine*
*Baltimore*

## References

Belmont JW (1996) Genetic control of X inactivation and processes leading to X-inactivation skewing. Am J Hum Genet 58:1101–1108

Cremers FPM, van de Pol TJR, Wieringa B, Hofker MH, Pearson PL, Pfeiffer RA, Mikkelsen M, et al (1988) Molecular analysis of male-viable deletions and duplications allows ordering of 52 DNA probes on proximal Xq. Am J Hum Genet 43:452–461

Jacobs PA, Migeon BR (1989) Studies of X-chromosome inactivation in trisomies. Cytogenet Cell Genet 50:75–77

Lee JT, Jaenisch R (1997) Long-range *cis* effects of ectopic X-inactivation centres on a mouse autosome. Nature 386: 275–279

Lee JT, Strauss WM, Dausman JA, Jaenisch R (1996) A 450 kb transgene displays properties of the mammalian X-inactivation center. Cell 86:83–94

Mattei MG, Mattei JF, Vidal I, Giraud F (1981) Structural anomalies of the X chromosome and inactivation center. Hum Genet 56:401–408

Migeon BR (1993) The postulated X-inactivation center at Xq27 is most reasonably explained by ascertainment bias: heterozygous expression of recessive mutations is a powerful means of detecting unbalanced X inactivation. Am J Hum Genet 52:431–432

Migeon BR, Axelman J, Jan de Beur S, Valle D, Mitchell GA, Rosenbaum KN (1989) Selection against lethal alleles in females heterozygous for incontinentia pigmenti. Am J Hum Genet 44:100–106

Migeon BR, Moser HW, Moser AB, Sprenkle JA, Sillence D, Norum RA (1981) Adrenoleukodystrophy: evidence for X-linkage, inactivation and selection favoring the mutant allele in heterozygous cells. Proc Natl Acad Sci USA 78: 5066–5070

Parrish JE, Scheuerle AE, Lewis RA, Levy ML, Nelson DL (1996) Selection against mutant alleles in blood leukocytes is a consistent feature in incontinentia pigmenti type 2. Hum Mol Genet 5:1777–1783

Pegoraro E, Whitaker J, Mowery-Rushton P, Surti U, Lanasa M, Hoffman EP (1997) Familial skewed X inactivation: a molecular trait associated with high spontaneous-abortion rate maps to Xq28. Am J Hum Genet 61:160–170

Plenge RM, Hendrich BD, Schwartz C, Arena F, Naumova A, Sapienza C, Winter R, et al (1997) A promoter mutation in the XIST gene in two unrelated families with skewed X-chromosome inactivation. Nat Genet 17:353–356

Rosenberg T, Niebuhr E, Yang H-M, Parving A, Schwartz M (1987) Choroideremia, congenital deafness and mental retardation in a family with an X chromosomal deletion. Ophthalmic Paediatr Genet 8:139–143

Therman E, Sarto GE, Palmer CG, Kallio H, Denniston C (1979) Position of the human X inactivation center on Xq. Hum Genet 50:59–64

Address for correspondence and reprints: Dr. Barbara R. Migeon, CMSC 10-04, The Johns Hopkins Hospital, 600 North Wolfe Street, Baltimore, MD 21287-3914. E-mail: bmigeon@welchlink.welch.jhu.edu
© 1998 by The American Society of Human Genetics. All rights reserved.
0002-9297/98/6206-0039$02.00

---

*Am. J. Hum. Genet. 62:1557–1558, 1998*

## Reply to Migeon and Haisley-Royster

*To the Editor:*

We thank Drs. Migeon and Haisley-Royster (1998 [in this issue]) for their interest in our research. We are, however, a bit puzzled by their letter to the editor, since they write that they disagree with the interpretation of our results yet then restate what was already written in our previously published article (Pegoraro et al. 1997 ).

The 50-member pedigree that we reported showed an X-linked dominant disorder with male lethality. There is no question of this fact, because we found a deletion mutation of Xq28 associated with skewed X inactivation and recurrent pregnancy loss (LOD = 6.92). The deletion included the factor VIII gene, yet there were no males from 50 females with factor VIII deficiency, again clearly proving that this family had an X-linked dominant disorder with male lethality.

Drs. Migeon and Haisley-Royster appear to wish to address two issues: (1) interpretation of the likely mechanisms that would cause X-inactivation skewing in the females in this family; and (2) transcriptional timing of the deleted gene or gene products in Xq28 and the observed effect on miscarriage detection. There is very little to disagree with in Drs. Migeon and Haisley-Royster's interpretation of our results; they suggest that a growth disadvantage is probably playing a role, which is precisely what we stated in our discussion. We, too, feel that growth disadvantage is the most likely mechanism causing skewing of X-chromosome inactivation. However, in the absence of characterization of the causative genes in Xq28, it seems unreasonable to dismiss the possibility that the gene(s) may actually be involved in the process of X inactivation. This is the least likely mechanism, but it does not seem to warrant exclusion from discussion.

The timing of transcription of the gene products in Xq28 undoubtedly affects when the miscarriage occurs. In fact, the issue of timing is central to the inferred ge-

netic mechanism. A cell-lethal trait expressed very early in embryonic development would be undetectable or perhaps would cause a "biochemical pregnancy." Activation later in embryonic life would still cause male lethality but would be less likely to cause complete skewing of X inactivation in multiple tissues in the heterozygous female. In view of this delicate balance in timing, we feel that the genes in question are most likely to be transcribed early in fetal development and to impart a growth disadvantage rather than being cell lethal. The size of the deletion mutation, however, is less important to when the miscarriage occurs: size is simply being used as a surrogate to the assumed importance of the deletion region and gene(s) contained in that region. In the end, this is all an exercise in mental gymnastics, since the characterization of the causative gene(s) will enlighten us all as to the true mechanism.

ERIC HOFFMAN, ELENA PEGORARO, AND MARK LANASA
*Department of Molecular Genetics and Biochemistry*
*University of Pittsburgh School of Medicine*
*Pittsburgh*

## References

Migeon BR, Haisley-Royster C (1998) Familial skewed X inactivation and X-linked mutations: unbalanced X inactivation is a powerful means to ascertain X-linked genes that affect cell proliferation. Am J Hum Genet 62:1555–1557 (in this issue)

Pegoraro E, Whitaker J, Mowery-Rushton P, Surti U, Lanasa M, Hoffman EP (1997) Familial skewed X inactivation: a molecular trait associated with high spontaneous-abortion rate maps to Xq28. Am J Hum Genet 61:160–170

Address for correspondence and reprints: Dr. Eric Hoffman, Department of Molecular Genetics and Biochemistry, University of Pittsburgh School of Medicine, E1240 Biomedical Science Tower, Pittsburgh, PA 15261. E-mail: eric@hoffman.mgen.pitt.edu

© 1998 by The American Society of Human Genetics. All rights reserved.

0002-9297/98/6206-0040$02.00

*Am. J. Hum. Genet. 62:1558–1560, 1998*

## Difficulties in the Estimation of Ethnic Affiliation

*To the Editor:*

Although I disagree with their results, I am indebted to Shriver et al. (1997) for reawakening my attention to the interesting but tricky subject of the inference of ethnic origin by DNA typing.

They have taken the novel and daunting approach of culling through a vast catalogue of candidate DNA loci to find those which are particularly discriminating. They list a battery of 10 loci, obtained mostly through such a search, which they claim will be effective in determining whether an unknown stain is of African American (AA) or European American (EA) origin. Specifically, they predict that only "0.01% [of individuals will] show log likelihoods <3.0" favoring one origin over the other (Shriver et al. 1997, p. 962). If a prior probability of 50% is assumed for each alternative, this implies the posterior ability to make a correct guess at least 999 times in 1,000. Categorizing Americans as black or white by interviewing them probably does not achieve such a high level of reproducibility, so it seems natural to review with care the basis for such a claim.

I am concerned that the claim rests on serious flaws in statistical methodology. My reanalysis shows that the estimates of efficacy for race determination are significantly overstated because of bias in the algorithm for prediction of likelihood ratios. This is true even for the handful of loci from the literature the authors say that they were able to verify as useful. As for the majority of the recommended loci—those discovered by surveying the catalogue—there is an additional bias that is probably even more serious. I shall discuss a computer simulation that shows that the apparent good performance of the culled loci may be completely illusionary, explainable as mere sampling variation.

These concerns can be conveniently discussed and illustrated in the context of D7S657, the most highly rated of the loci found by the statistical survey. Figure 2 of the Shriver et al. article reveals enough information to allow a check of the calculations for this locus, calculations that assert a typical likelihood ratio of $r = 19$ ($\log_{10} r = 1.276$). I will argue that that number is inflated both by algorithmic errors and by sampling bias. A more realistic likelihood ratio estimation algorithm will reduce the value from 19 to ~8 ($\log_{10} r = 0.9$), and consideration of sampling bias will show that a value of 2.5 ($\log_{10} r = 0.4$) or even less is plausible and consistent with the reported results.

Let $a_1, a_2, \ldots$ and $b_1, b_2, \ldots$ be the allele frequencies at some locus in populations A and B, for alleles 1, 2, ..., respectively. Then for an allele whose true origin is A and for allele frequencies that are known,

$$\log_{10} r_{AB} = \sum a_i \log_{10}(a_i/b_i) \quad (1)$$

is the expected value of the logarithm of the likelihood ratio that the origin is the reference population A rather than the target population B. The formulas in the article by Shriver et al. are equivalent, except that their notation refers to genotypes rather than to single alleles (which explains why their formula has factors of "$\frac{1}{2}$," whereas mine does not), and they formulate a statistic that is