Am. J. Hum. Genet. 67:256–258, 2000

Reply to the Letters from Murray et al. and Vianna-Morgante and Costa

To the Editor:

It was with great interest that we read the letter to the editor by Murray et al. (2000 [in this issue]), and we thank the authors for their comments. We agree that it would be very interesting to know whether other investigators observe a parent-of-origin effect in the development of premature ovarian failure (POF) in premutation carriers, as we have stated in our paper (Hundscheid et al. 2000). Murray et al. reevaluated their results and did not observe such an effect. There may be several explanations for this. The etiology of POF is extensive and comprises genetic, nongenetic, and multifactorial components. Therefore, it is not unlikely that there are differences between families with fragile X who are from the United Kingdom and those that are from The Netherlands. More importantly, differences in study design (especially for multicenter studies) will inevitably lead to other results and, subsequently, will lead to other conclusions.

We were very surprised by the mean age at menopause mentioned by Murray et al. In the premutation group, the mean age at menopause was 47.87 years, compared with 52.96 years in the full-mutation and normal groups combined. In our ongoing study, the mean age at menopause in premutation carriers who had experienced spontaneous menopause was 42.0 years (unpublished data); this finding is in line with observations made elsewhere (see Partington et al. 1996). When we performed Kaplan-Meier survival analysis on the entire group of premutation carriers, the mean age at menopause was 45 years. This finding strongly suggests that the study population of Murray et al. differs from ours. To verify our observation of imprinting, the mean age at menopause has to be comparable. A significantly lower mean age may point to a different population, a difference in the occurrence of POF, or the number of premutation carriers that are postmenopausal. If this is the case, then it is obvious that one cannot compare the two studies.

Murray et al. have shown that, of 116 women who were premutation carriers, 51 had a paternally inherited fragile X premutation (PIP), 40 had a maternally inherited fragile X mutation (MIP), and 25 had an unknown parental origin of premutation. Of the 91 women in whom the parental origin of the premutation could be established, only 30 (33%) were of age $\geq$40 years. We have described 148 premutation carriers in whom the parental origin of the premutation could be determined: 106 women with a PIP and 42 women with a MIP. Of the 106 women with a PIP, 82 were of age $\geq$40 years, and, of the 42 women with a MIP, 27 were of age $\geq$40 years. Thus, in our study, 109 (74%) of 148 women were of age $\geq$40 years and did not experience nonspontaneous cessation of menstruation at age <40 years. Therefore, we have to conclude that the proportion of women in whom the occurrence of POF can be established (only in women of age $\geq$40 years) is significantly higher in our study, compared with the study by Murray et al. This may be the result of other methodology, which may also account for the low numbers of observed women with a PIP. This makes it rather impossible to compare the data.

University Hospital Nijmegen has been extensively studying families with fragile X, tracing possible carriers in several generations. We have estimated that the overall frequency of carriers of a PIP is approximately three times higher than that of carriers of a MIP. On the basis of this finding, we cannot reason why Murray et al. identified 51 women with a PIP and 40 women with a MIP. This other PIP:MIP ratio may be an indication that Murray et al. did not study the families to the same extent that we did. Murray et al. might possibly have included a large proportion of first-degree relatives in the younger generation. This will result in a different population with other observations that cannot be compared.

Since the etiology of POF is extensive, we think that it is of paramount importance to check medical histories with attending physicians, to avoid misclassification. Checking the dates with attending physicians may also help to avoid a patient's recall bias; postmenopausal women have a tendency to round off their age at menopause to the nearest age that ends in the numeral 0 or 5 (Partington et al. 1996).

We have to conclude that, in a comparison of our study with that of Murray et al., there are differences in methodology, mean age at menopause, and number of women in whom the occurrence of POF can be established. This probably reflects a different population, and we therefore doubt whether the results can be compared. We agree that it is remarkable that Murray et al. did not observe the same parent-of-origin effect that we observed. Therefore, we would like to invite groups with a population and methodology comparable to ours to verify our observation and to report their findings.

The reply we addressed to the letter to the editor submitted by Murray et al. applies to that submitted by Vianna-Morgante and Costa (2000 [in this issue]) as well. The population in the study by Vianna-Morgante and Costa is very young, compared with that in our study. The population's median age at examination, for women with a MIP (n = 27), was 36.83 years, and, for women with a PIP (n = 32), the median age was 38.875 years. In our study, for women with a MIP (n = 42), the median age at examination was 51.5 years, and, for

those with a PIP (n = 106), the median age was 50.0 years. Besides the fact that the PIP:MIP ratio (32:27) mentioned in the letter by Vianna-Morgante and Costa differs substantially from ours (106:42), we have a much older population. Moreover, this major difference in the study population is again emphasized by the fact that, in the study by Vianna-Morgante and Costa, 15 women with a PIP and 10 women with a MIP were of age ⩾40 years. This number is very low compared with our finding (82 women with a PIP and 27 with a MIP, all of age ⩾40 years). Again, this points out that their population is different than ours—a fact that obviously will lead to other results.

Vianna-Morgante and Costa have compared the occurrence of POF in women with a PIP with that in women with a MIP, and they have concluded that there is no difference between the two groups. However, their analysis incorporated data on women of age <40 years. Since POF is defined as a condition occurring at age <40 years, it can only be established reliably in women of age ⩾40 years. Not only will establishment of the occurrence of POF in women who have not reached the age of 40 years result in a higher risk for misclassification toward POF, but, in the majority of cases, occurrence of the condition cannot even be established. Hence, we think it is remarkable that the authors also included women of age <40 years in their study. Moreover, the mean age of the participants in their study is below the cutoff level for age. Therefore, the numbers they presented probably will not represent the final (as established only in women of age ⩾40 years) occurrence and distribution of POF. In the letter, 14 (24%) of 59 women with either a PIP or a MIP had POF—a finding that is an underestimation of the real (probably even higher) occurrence of POF. Since 16% of women with premutations experience POF (Allingham-Hawkins et al. 1999), we wonder whether the population in the study of Vianna-Morgante and Costa (in which ⩾24% of the women had POF) is randomly selected. In our ongoing study of families with fragile X, we have randomly selected women on the basis of mutation and not on the basis of indication of POF.

Last but not least, for both groups of women, the authors calculated the median and mean age at menopause (see also table 1 in the study by Vianna-Morgante and Costa). The numbers on which these calculations are based are very small (13 women with a PIP and 9 women with a MIP), and, to us, it is not clear which data the authors have included in their calculations. If the authors included women with POF when they calculated mean age at menopause, then their calculation would not result in a reliable mean age at menopause. For instance, for women with a PIP, the mean age at menopause was based on 13 women who have experienced spontaneous menopause. If the authors also used the data on the women with POF, then it can be inferred that only four women who did not experience POF were used in this calculation. The mean age at menopause that is presented is not representative of that in all women with a PIP. Thus, on the basis of these numbers, we do not subscribe to the authors' conclusion that there is no difference between the two groups, as far as age at menopause is concerned.

In conclusion, neither Murray et al. nor Vianna-Morgante and Costa can confirm our observation of a parent-of-origin effect. Both groups have younger populations, other PIP:MIP ratios, and a sample size that is much smaller than ours. We therefore do not follow Vianna-Morgante and Costa's suggestion that a "possible genomic imprinting effect may be peculiar to the Dutch population," since no sufficient convincing evidence of this is provided. Nevertheless, if the parent-of-origin effect that we have observed cannot be demonstrated by other authors, then we have to conclude either that the parent-of-origin effect is unique to the Dutch population with fragile X or that we all are overlooking some other factors (bias or nonbias). Whatever is causing this discrepancy, it will be of major importance with regard to future research (and which methodology is to be used) in this particular field. However, we think that it is too premature to draw final conclusions with regard to the parent-of-origin effect. Our population and methodology differ too much from those described by Murray et al. and by Vianna-Morgante and Costa. Further research is warranted to verify our observation.

Rubin D. L. Hundscheid,[1,2]
Chris M. G. Thomas,[2,3] Didi D. M. Braat,[2]
Ben A. Oostra,[4] and Arie P. T. Smits[1]
Departments of [1]Human Genetics, [2]Obstetrics and Gynecology, and [3]Chemical Endocrinology, University Hospital Nijmegen, Nijmegen, The Netherlands; and [4]Centre for Biomedical Genetics, Department of Clinical Genetics, Erasmus University Rotterdam, Rotterdam

References

Allingham-Hawkins DJ, Babul-Hirji R, Chitayat D, Holden JJ, Yang KT, Lee C, Hudson R, et al (1999) Fragile X premutation is a significant risk factor for premature ovarian failure: the International Collaborative POF in Fragile X study—preliminary data. Am J Med Genet 83:322–325

Hundscheid RDL, Sistermans EA, Thomas CMG, Braat DDM, Straatman H, Kiemeney LALM, Oostra BA, et al (2000) Imprinting effect in premature ovarian failure confined to paternally inherited fragile X premutations. Am J Hum Genet 66:413–418

Partington MW, Moore DY, Turner GM (1996) Confirmation of early menopause in fragile X carriers. Am J Med Genet 64:370–372

Address for correspondence and reprints: Dr. Rubin D. L. Hundscheid, Department of Human Genetics, University Hospital Nijmegen, P. O. Box 9101, Nijmegen, 6500 HB, The Netherlands. E-mail: R.Hundscheid@obgyn.azn.nl

0002-9297/2000/6701-0033$02.00

Am. J. Hum. Genet. 67:258–259, 2000

Inflated False-Positive Rates in Hardy-Weinberg and Linkage-Equilibrium Tests Are Due to Sampling on the Basis of Rare Familial Phenotypes in Finite Populations

To the Editor:

If it is assumed that genotypes of some locus ($\mathbf{G_D}$) are in Hardy-Weinberg equilibrium (HWE) in a population and that these genotypes are correlated with some phenotype (**Ph**), then, among "cases" in the tail of the distribution of **Ph** (equivalently, affected with rare disease), the $\mathbf{G_D}$ will show Hardy-Weinberg disequilibrium (HWD) (Nielsen et al. 1999; Deng et al. 2000; Göring and Terwilliger 2000). However, this does not imply that "generally, in individuals at either end of the quantitative-trait distribution, HWD exists if and only if there exists a whole-population LD [i.e., "linkage disequilibrium"]" (Deng et al. 2000, p. 1030). The "only if" part of this sentence is not correct. Even Deng et al. (2000, p. 1044) point out that "an absence of HWD does not imply that a marker locus and a QTL are not in LD" and that, for completely random marker loci, there will be inflated false-positive rates in tests for HWD (and LD as well), because "cases" of familial disease tend to be more related than "controls," for the following reasons.

Assume that **Ph** is correlated in families, without specifying whether this is due to genetic or shared environmental factors. Let the prevalence, $\phi = P$(individual B is a case), and the familial relative risk $\lambda = P$(individual B is a case|relative A is a case)/ϕ (Weiss et al. 1982; Risch 1990). Then, P(A *and* B are affected|A and B are relatives) = $\lambda\phi^2$, and P(A *and* B are affected) = ϕ^2 if they are randomly ascertained. This implies that P(A and B are relatives|A *and* B are affected) = $\lambda\phi^2 P$(A and B are relatives)/ϕ^2 = λP(A and B are relatives). If $\lambda > 1$, then ascertainment of "cases" ascertains relatives with greater probability than does random ascertainment of "controls," leading to increased false-positive evidence of HWD and LD throughout the genome. This effect will be largest when λ is large, ϕ is small, and the population is small and/or structured (such that P[A and B are relatives] is nontrivial). In a sense, this is related to the problem of population stratification when the phenotype being studied correlates with a familial stratum, regardless of whether the trait is "genetic" (see Chase 1977).

If the "case" phenotype is a good predictor of $\mathbf{G_D}$ (a prerequisite for mapping to be powerful), then a large portion of the "case" sample will share some risk allele IBD from some common ancestor. The coalescent path connecting these chromosomes historically defines the most distant possible relationship among the "cases" carrying this allele, defining an upper bound on how "unrelated" they could possibly be. Again, this implies that ascertainment of affected individuals increases the probability of ascertainment of relatives. And the less frequent the shared risk allele is, the more closely related the "case" individuals will be (see Terwilliger, in press), leading to potential deviations from HWE and LE in unrelated parts of the genome as well.

The more closely related two people are, the larger the proportion of their genomes that they will share, as measured by their kinship coefficient (also see Terwilliger et al. 1997). If cases are "more related" than controls, then they will, with higher probability than will be seen in controls, share alleles IBD at random places in the genome, leading to increased false-positive rates in HWD and LD tests. This anticonservative behavior may be minor in studies of a single marker locus, but, when one considers the effects of testing hundreds of thousands of markers jointly in a genome scan, often making inferences based on the most significant values of the test statistic over the genome, the inflation of the type I error can have significant import. Furthermore, because the effect of small deviations, from HWE and/or LE, that are induced by such sampling is to shift the distribution slightly upward, the anticonservative bias will increase as we look farther out into the tail of the pointwise distribution (data not shown—but similar in shape to what appears in fig. 4 of Göring and Terwilliger 2000), leading to potentially gross inflation of genome-wide false-positive rates. To test for such problems, one can do a Monte Carlo randomization, as was done, in a case-control study of a small genetically homogeneous population isolate, by Hovatta et al. (1999), who kept the genotypes (for the whole genome scan) of all individuals constant and randomized their phenotypes ("case" and "control"). The simulation showed that their sample had approximately twice as many positives as would be expected from the randomization test, consistent with what is expected for reasons described in this note. When the fundamental assumption that "cases" and "controls" are independent and identically distributed with respect to random marker-locus genotype frequencies throughout the genome appears to have been rejected, it is essential to maintain skepticism in the interpretation of the results of such an analysis.

Unfortunately, the conditions in which "cases" are most likely to be relatives (e.g., small populations, rare