UNIVERSITY OF CALIFORNIA, BERKELEY

BERKELEY · DAVIS · IRVINE · LOS ANGELES · RIVERSIDE · SAN DIEGO · SAN FRANCISCO   SANTA BARBARA · SANTA CRUZ

SCHOOL OF PUBLIC HEALTH
DEPARTMENT OF BIOMEDICAL AND
ENVIRONMENTAL HEALTH SCIENCES

EARL WARREN HALL
BERKELEY, CALIFORNIA   94720

FAX TO: VICTOR McKUSICK / OSKAR ZABORSKY
      Fax: 301-955-4999
      Tel: 301-955-6641

FROM: GEORGE SENSABAUGH
      Fax: 510-643-5163
      Tel; 510-642-1271   6 + 11 pages

Oct. 28, 1991

Dear Victor,

I regret to inform you that I can not sign off on the report as
it stands. My main concern is chapter 3. When I signed off last
May, I indicated that I had reservations about this chapter but
couldn't add anything to the comments I had already made.
Frankly, I had hoped that the reviewers comments might induce
some of the changes I had lobbied for. I did not see their
comments so I don't know what they said. However the 10/15/91
draft is not that different from the draft I was concerned about
in May.

My major reservations about chapter 3 are outlined below.

1. The chapter takes a very strong position that the US
    population is highly substructured and that accordingly any
    population frequency estimates based on existing data are
    meaningless. Although some of the phrasing of this draft is
    softened relative to previous drafts, the position is not
    significantly softened. This position is stronger than much
    of the testimony we heard from other population geneticists,
    some of whom questioned the existence and/or relevance of
    substructuring. It is clear that there is not unanimity of
    opinion on this point and I think the report should reflect
    this. At least, there should be some statement that the
    impact of substructuring on frequency estimation is being
    actively debated by population geneticists and that there
    may be legitimate differences of opinion.

    I appreciate the core concern that extreme frequency
    estimates have been quoted in court and I know we are all
    concerned that this might lead to greater weight being
    placed on the DNA evidence than is warranted. The
    population substructure argument posits that the general

1

population contains pockets of individuals who possess certain combinations of genes at high frequency and that this possibility is obscured by the citation of statistics for the general population. In other words, the use of general population statistics may imply absolute identification when, due to these pockets, there is a realistic possibility that other individuals of the same genetic type may exist. I think this concern can be adequately met by providing guidelines for the generation of very conservative frequency estimates. The ceiling approach accomplishes this. We could also recommend a ceiling on the cumulative frequency, e.g., no figures be cited less than one in a million. The population substructure argument, however, I think lays a trap for the future; I comment on this below.

2. The central issue in the statistical interpretation area is the legitimacy of combining individual locus genotype frequencies by multiplication. By multiplying enough single locus frequencies together, one can get infinitely low combination frequencies. The legitimacy of multiplication depends on the statistical independence of allelic combinations of the loci in question, i.e., allelic combinations at different loci cannot be correlated.

We propose the ceiling approach to get at the problem of making safe allele frequency estimates. We then state that this conservative approach solves the multiplication problem. However, one element of the population substructure argument is the existence of allelic correlations across loci (the tall, blond, blue-eyed nordic argument); we can't wave this away simply by taking a conservative approach to allele frequency estimation. Moreover, testing small (n=100) homogeneous populations is not likely to reveal any but the most extreme correlations. I suspect our position could be shown to be in error if alleles occur at different loci in a reasonably correlated way although I have no idea what level of correlation it would take. It would be better to address the correlation issue directly: What tests exist for detecting allelic correlations and what are their limits? At what level does allelic correlation begin to significantly erode frequency estimation? Are there correction factors that can be put on frequency estimates at various levels of correlation? What, if anything, can be done with existing databases?

The allele correlation question is likely to become more of an issue as PCR based typing systems come on line. It will become possible to test a large number of moderately polymorphic loci (heterozygosities 60-90%) and the question of multiplication will inevitably come up.

2

3. The summary recommendation (p. 3-40) that population
    frequencies be based on counting until appropriate
    population studies (which we define) be done is undesirable
    on two counts. First, and most important in my mind, it
    questions genetics; it implies that multiplication of
    frequencies is mere theory and is not applicable in the real
    world. To retreat so completely from genetics signals to
    lay people and lawyers that something is wrong about its
    principles. (In some places, we sound almost as dismissive
    of genetics "theory" as creationists are of evolutionary
    "theory".) Second, the recommendation leaves us in a policy
    limbo. The N values for the existing databases are on the
    order of a thousand or two. These sorts of 1/N numbers can
    be obtained with conventional blood group and protein
    typing; should we retreat from DNA until the "appropriate"
    population studies are done? If DNA typing of an evidence
    sample at one locus matches none in my database, why type
    additional loci since I can already make my 1/1000
    statement. (Studies on the Home Office database show single
    locus matches are uncommon and two locus matches are very
    rare.)

    This recommendation should be amended to state that this is
    an absolute upper limit frequency statement and that genetic
    principles can be used to generate lower frequency
    estimates. This brings the recommendation more into accord
    with the text (p. 3-37).

4. What happens if the NCFDT never gets off the ground? It
    appears to be the final arbiter for the choice of the
    representative homogeneous population groups from which the
    allele frequency estimates are to be derived. If there is
    no NCFTD, then again we end up in a policy limbo where
    nothing that has been done previously would be considered
    acceptable. This is not an idle concern; in the past,
    courts have considered apparent unresolved issues to be
    grounds for exclusion. This concern extends as well to the
    other recommendations that depend on the NCFTD.

    I am not opposed to additional population studies; indeed I
    welcome them. Rather, I would like to see more
    consideration given to the studies that have been done as at
    least providing first approximations.

5. The 10% gene frequency ceiling is unnecessarily high. We
    know from classical genetics that there are genes that are
    uncommon (< 1%) in all populations and there is no need to
    set an artificially high platform for them. In any case, it
    should be clearly stated that the ceiling principle yields
    an estimate for a worst case scenario - that the frequency
    estimates used are higher than would be expected in any real
    population.

6. The chapter focuses primarily on the problems associated
    with the VNTR loci with a very large number of "continuous"
    alleles. We should be providing guidance, however, on the
    upcoming PCR based systems that detect discrete alleles for
    here some of the issues are different. For loci with
    smaller numbers of discrete alleles (5-15), the assessment
    of allele frequencies becomes more straightforward and the
    issue of correlation becomes more tractable; we should
    provide guidance.

7. With all due respect to Lewontin, the main message of his
    1972 apportionment of diversity paper was that the greatest
    proportion of diversity was at the individual level. Within
    race population variation and between race variation
    accounted for only about 15% of the total variation, 8% and
    7% respectively, as I recall. We should make this point
    clear since as the draft now implies that the major
    proportion of variation is between populations.

    It should also be stated up front rather than as a coda that
    the example illustrating variation between Poles and
    Italians represents the extreme values for these two groups.
    Also, and unfortunately from a blood grouping perspective,
    these are not terribly good examples since both the Rh and
    Kell allele frequencies are inferred from results obtained
    by testing for the dominant antigen; both are
    dominant/recessive systems and the allele frequency
    estimates may not be very good.

8. It is suggested that laboratory error rates be included as
    part of the overall population frequency statement.
    Although I am a strong advocate for open access to
    proficiency test results, it must be appreciated that these
    results have limited relevance to particular cases. First,
    when labs make errors, it is incumbent upon them to try to
    identify the source of the error and to take corrective
    actions. Thus proficiency test error rates should not
    reflect a steady state rate. Second, and more important,
    each case is different. In some cases, internal consistency
    checks can be incorporated in the testing scheme that
    virtually guarantees the accuracy of the result; in other
    cases, this may not be possible. None of this can be
    reduced to a simple number; the quality of the work product
    must be evaluated individually in each case.

9. Finally, we have skirted the ultimate question: at what
    point does a population frequency become a statement of
    identification. Is it at 1/million, 1/billion, 1/trillion?
    Or do we want to state that identification requires a
    subjective leap? We make a number of statements that
    absolute identification is on the horizon; how will we know
    when we are there.

4

Some other general comments, by page and line:

ES-9, 23 and 2-38, 20:   delete too; if the sample is too small,
     nothing can be done with it.

ES-20, 7-10 and corresponding section in Ch 4:   I object to the
     presumption that any non-accredited laboratory is not
     meeting standards. It would be better to say that any non-
     accredited laboratory should be expected to demonstrate that
     it is operating at the same level of standards as accredited
     laboratories. This makes it harder but not impossible for
     non-accredited labs to do work, for example, as in the case
     of a non-forensic lab repeating the analysis of a forensic
     lab. Our objective is to ensure high standards, not to
     discriminate on the basis of accreditation.

ES-22, 3-4:   "list of genetic types" instead of "list of fragment
     sizes".

1-16, 18:   "numeric value" -> "digital code"
1-16, 22-23:   It is premature to predict that this system might
     lead to absolute identification; it needs to undergo the
     same validation as other DNA typing systems. Moreover,
     since the work is not yet published in a peer reviewed
     journal (although I have no doubt it will be soon) we should
     not wax too enthusiastic.

2-1, 11-16:   "patterns" -> "types". The patterns are the basis
     for assigning type designations but are not the types
     themselves.

2-4, 19-20:   Lab results have both qualitative and quantitative
     elements, e.g., bands on an autoradiogram are characterized
     both by position and intensity.

2-10, 11-12:   Where did the recommendation for 10 or fewer easily
     distinguished alleles come from; this is new to me. This
     takes out all the currently used VNTR systems. Do we really
     want to say this?

2-22, 11-12:   The yield gel is used to assess both the quantity
     and quality of DNA in a sample; properly done, that is, with
     appropriate quantitative standards, it will give an adequate
     estimate (within 3X) of the quantity of high molecular
     weight DNA. Fluorescence assays can't distinguish degraded
     from undegraded DNA. There is no real gain in information
     with a pure quantitative assay.

2-40, refs: If there is to be a reference in this chapter to my
     writing on the analysis and interpretation of blood group
     and protein markers, the better reference is Sensabaugh, GF.

5

Biochemical markers of individuality, in: <u>Forensic Science Handbook</u> (R. Saferstein, ed.) Prentice Hall, 1982, pp. 338-415.

3-generally: The term "odds" is used throughout the chapter where "estimated frequency of occurrence" is intended. "Odds" has a defined meaning in probability theory, specifically a ratio of probabilities as in log odds scores.

3-6, 5ff: Combined frequency estimates for the conventional markers often are less than 1/100. These frequencies have not been verified by simple counting but by multiplication. The justification for multiplication is the lack of correlation of alleles at different loci as indicated by pairwise comparisons. With the conventional markers, no evidence of significant allele correlations has been noted.

As a final note, shouldn't we have seen all the reviewers comments? I got 7 pages of comments faced by a page labeled "general".

For your information, I include the Sept. 1991 report of the DNA Commission of the Society for Forensic Haemogenetics.