# Hardy-Weinberg Equilibrium in RFLP Databases

Charles H. Brenner
DNA·VIEW
2486 Hilgard Avenue, Berkeley, California 94709, USA
*email: cbrenner@ccnet.com*

## INTRODUCTION

Population data of RFLP fragment sizes often shows an excess of homozygosity compared to the expectation based on Hardy-Weinberg equilibrium, a fact that has potential repercussions on the use of DNA for identification in criminal work. Three explanations have been suggested:

(i) binning error due to band coalescence (Brenner & Morris 1990),
(ii) bands run off the gel (Chakraborty et al, 1994) or are overlooked, and of course
(iii) disequilibrium in nature (Green & Lander 1991, Giesser & Johnson 1993).

The principal results this paper are these:

A. The mechanism of (i) is explained, and a method of analysis (which eluded us in 1990) is given to get around it.

B. Analysis of a very large collection of data permits a fair perspective on the overall situation. No single explanation is adequate to explain apparent excess homozygosity, but mostly it is explained by (i).

## MATERIALS & METHODS

Databases were accumulated over a six year period from laboratories in the United States and Europe. Most of the data was collected using the DNA·VIEW™ software. In all 228 databases from 20 laboratories were analyzed for the study, covering 13 racial and ethnic groups, 27 probes, and three restriction enzymes. A total of 247,000 people/probe combinations are represented.

Testing whether the rate of homozygosity in a gel -analyzed RFLP system conforms to the Hardy-Weinberg condition of random assortment is confounded by several factors. The first difficulty is the effectively continuous nature of the fragment size spectrum, which precludes a direct computation of $\sum p^2$ . The problem thus arises how to define homozygosity for purposes of a test. One natural way is to call a person homozygous whenever the person's two bands are sufficiently close as to merge (or "coalesce"). This approach has been tried (Brenner & Morris 1990, Lander 1990) but coalescence is so difficult to predict that the method is useless in practice. Finally, in desperation, one considers the distasteful possibility of forcing the round peg of continuous fragment sizes into the square hole of size classes called "bins." Therefore for each database a set of up to 20 bins were determined such that they would have roughly equal numbers of fragments, subject to the proviso that no bin should be too narrow. However a subtle but serious difficulty arises.

## DISCUSSION & RESULTS

### Binning error due to band coalescence

To see the problem, suppose that in some RFLP system a population of four people have the fragment lengths, in bases, as shown in **Table 1**. The initial digit is emphasized to suggest binning by 1000's. From

the point of view of the implied allele categories the four people thus have genotypes *1,1*, *1,2*, *1,2*, and *2,2*. Insofar as only four genotypes can suggest Hardy-Weinberg equilibrium, these do.

**Table 1**

| |
|---|
| *1*331,*1*728 |
| *1*995,*2*001 |
| *2*222,*1*111 |
| *2*468,*2*345 |

However, if the data is measured by gel electrophoresis the genotypic categories will not be correctly inferred. The two bands for the second person will coalesce into a single band (perhaps a bit fat), and will therefore be typed either as *1,1* or as *2,2*, but certainly not as *1,2*. The collection of phenotypes thus may be *1,1*, *1,1*, *1,2*, and *2,2*. In any case the number of pseudo-homozygotes[1] rises from two to three in the transition from nature to observation.

It is worth being clear that the effect is unlike measurement error, which misclassifies as many bands in one direction as in the other. The binning error from coalescence creates a bias in collecting data: when two bands are near each other on opposite sides of a bin boundary one band in effect sucks the other across, whereas there is no compensating mechanism in the opposite direction that would tend to separate two bands that are really in the same bin.

It seems likely that the simple mechanism just explained has not been widely understood. Giesser & Johnson (1993) and Green (1992) guessed that the effect would be insignificant when the bins are large, which is false. Pseudo-homozygosity still can easily be inflated by 15-30%. Lander (1990) said that he had allowed for "the effect of coalescence" but he didn't. Chakraborty et al (1994) tried to explain away all the apparent excess as due to missing bands, which is impossible. Only Devlin et al (1990) correctly, albeit indirectly, circumvent the problem.

In any event the consequence is that Wahlund's (1928) test -- comparing the observed to expected number of people with two bands in the same bin -- is useless for RFLP data.

A Modification of the Wahlund Test

Another way to look at the difficulty with the Wahlund test is depicted in **Figure 1**. Under the classification of fragment sizes into bins **1**, **2**, ... discussed above, each person has a binned phenotype like ***i,j***, represented by a square or, if ***i=j***, a triangle. Band coalescence causes some ***i+1,i*** and ***i,i-1*** genotypes to cross a heavy black boundary line and turn into an ***i,i*** phenotypes. Therefore the binned data is biased toward triangles, i.e. pseudo-homozgotes. Moreover the extent of the misclassification bias is unpredictable, so there is no adequate way to estimate the number of measurements expected to fall into the triangular regions under the hypothesis of Hardy-Weinberg equilibrium.

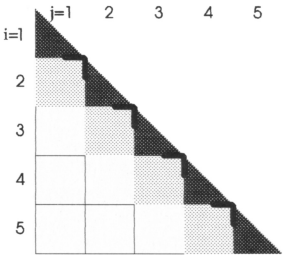


**Figure 1** Modified Wahlund Test

However, the augmented shaded area, obtained by also including the sub-diagonal, *can* be estimated, because there is no systematic misclassification around the new boundary. Therefore, a *Modified Wahlund Test* can be performed using the straightforward procedure that fails for a Wahlund test. In order to compare the results, we'll perform both tests at once.

Let $n_{ij}$ be the number of people counted in bin ***i,j***, $N$ be the total number of people, and $p_i$=band frequency in bin ***i***. Then $O = \sum_{0 \le i-j \le \theta} n_{ij}$ and $E = N \sum_{|i-j| \le \theta} p_i p_j$ represent the observed and expected (assuming HWE and neglecting any bias in binning) numbers to fall in a shaded area, where $\theta$=1 for the modified Wahlund test, and putting $\theta$=0 would give the Wahlund test. **Table 2** compares O vs. E using a binomial model and 1-tailed significance test, for some representative databases.

The new test is successful in showing that in some (but not in all) databases any appearance of excess homozygosity is definitely just factitious -- that is, not reflecting nature (iii) nor even laboratory procedures (ii) but is only an artifact (i) that arises from introducing bins for analysis and using them carelessly.

1 homozygote relative to the binning scheme -- i.e. having two fragments in the same bin

**Table 2** Analysis of homozygote excess for typical databases

| Database | | people N | Wahlund test, θ=0 O=obs homoz | % excess | signif level | modified test, θ=1 O=near homoz | % excess | signif level |
|---|---|---|---|---|---|---|---|---|
| g3 | Hin Caucasian | 300 | 44 | 140% | <0.001 | 71 | 40% | 0.001 |
| pYNH24 | Hin Black | 146 | 23 | 190 | <0.001 | 37 | 59 | 0.001 |
| EFD52 | Hae Black | 16523 | 1045 | 22% | <0.001 | 2484 | 0% | 0.4 |
| EFD52 | Hae Caucasian | 12779 | 1111 | 31 | <0.001 | 2251 | -2 | 0.8 |
| pH30 | Hae Chinese | 276 | 25 | 79 | 0.002 | 47 | 15 | 0.2 |
| g3 | Hin Caucasian | 1220 | 77 | 7% | 0.3 | 191 | 0% | 0.5 |
| pH30 | Hae Black | 2240 | 116 | 4 | 0.4 | 319 | -2 | 0.7 |
| TBQ7 | Hae Chinese | 120 | 6 | -6 | 0.6 | 17 | -9 | 0.7 |
| EFD52 | Hae Hispanic | 466 | 26 | -3 | 0.6 | 62 | -14 | 0.9 |

Significance level = likelihood to observe so much homozygosity by chance if HWE obtains.
% excess = (O-E)/E.

## SUMMARY AND CONCLUSIONS

i. A quarter of the 228 databases, represented by the middle group in Table 1, have an apparently significant (using θ=0) excess of homozygosity that can persuasively be explained away (significance level using θ=1) as binning error. There is no reason to be suspicious of these databases from a forensic point of view.

ii. One sixth, like the first two in Table 1, have an excess beyond binning error. For many of them it is reasonable to suspect running off the gel. For example most of the databases for the g3 probe, which is well-known for missing bands (Niels Morling, personal communication), fall into this category. However, case-by-case followup and careful analysis will be necessary.

iii. The remaining 50+% of the databases both individually and as a group show no statistically significant excess of homozygotes even under the incorrect, Wahlund, test. This is an interesting fact that also requires followup analysis.

Acknowledgments

Thanks to Amanda Soser, Niels Morling, and other members of the DNA·VIEW Users' Group, to Lotte Henke, Theresa Aulinskas, and to the FBI, for friendly collaboration and sharing of data.

## REFERENCES

Brenner CH, Morris JW (1990) Paternity Index Calculations in Single Locus Hypervariable DNA Probes: Validation and Other Studies. In: Promega Corporation (pubs) Proc for the International Symposium on Human Identification 1989, Madison, Wisconsin, pp21-53

Chakraborty R, Zhong Y, Jin L, Budowle B (1994) Nondetectability of Restriction Fragments and Independence of DNA Fragment Sizes Within and Between Loci in RFLP Typing of DNA. Am J Hum Genet 55:391-401

Devlin B, Risch N, Roeder K (1990) No Excess of Homozygosity at Loci Used for DNA Fingerprinting. Science 249:1416-1420

Geisser S, Johnson W (1993) Testing Independence of Fragment Lengths within VNTR Loci. Am J Hum Genet 53:1103-1106

Green P (1992) Population Genetic Issues in DNA Fingerprinting. Am J Hum Genet 50:440-441

Green P, Lander ES (1991) Forensic DNA Tests and HW Equilibrium. Science 253:1038-1039

Lander ES (1991) Lander Reply. Am J Hum Genet 49:899-903

Wahlund S (1928) Zusammensetzung von Populationen usw. Hereditas 11:65-106