

### Error rates in prenatal cystic fibrosis diagnosis

SIR—Pembrey and Malcolm<sup>1</sup> state that if a prenatal diagnosis is based on an empirical correlation between an amniotic fluid protein and fetal genotype, an attempt should be made to confirm diagnosis on the abortus after termination of pregnancy. This is certainly standard practice for  $\alpha$ -fetoprotein-based diagnosis of neural tube defects, and has become part of our procedure for microvillar enzyme-based prenatal diagnosis of cystic fibrosis<sup>2</sup>. Apart from providing a continual updating of the sensitivity and specificity of the test, such procedures are useful for monitoring the inevitable errors of sample handling which creep into even the best-regulated laboratories.

What puzzles us is why Pembrey and Malcolm should feel that diagnosis based on linked DNA probes should be excluded from such quality control. The accuracy of such diagnosis is dependent on recombination frequencies. The derivation of recombination frequencies is a form of empirical correlation and only as good as the most recent set of data. In fact, because the total number of informative meioses observable can be quite small, recombination fractions may be subject to dramatic fluctuations. The linkage between the G8 probe and the Huntington's disease gene was initially very tight<sup>3</sup>, but recombinants appeared with the acquisition of new family data<sup>4,5</sup>. The same is true for linkage between the cystic fibrosis gene and the probes, pJ3.11 and met<sup>6,7</sup>. Recombination fractions are usually stated as lying within a given confidence interval, to indicate the uncertainties inherent in their estimation.

There is also the question of genetic heterogeneity. Pembrey and Malcolm suggest that careful linkage studies in a large number of families from different populations will allow exclusion of significant non-allelic heterogeneity. But what do they mean by significant? An average non-allelic heterogeneity level of 5 per cent would be very significant if it meant that a similar proportion of prenatal diagnoses were wrong. It is difficult to see how this figure can go much below 5 per cent when the error rate in diagnosing cystic fibrosis is believed to be of this order<sup>8</sup>.

The point that we were making in our earlier letter<sup>9</sup> is that the whole edifice of quality control is not based on pedantry but on the realization that laboratories do make mistakes. It is possible that, as first-trimester prenatal diagnosis becomes more widely available, the standard procedure of confirmation of diagnosis on abortuses will have to be abandoned. We should, nonetheless, continue to collect and analyse the outcomes of as many of these pregnancies as is possible, since it will improve the reliability of the linkage

data. And let us be clear what we are doing, and not imply — as do Pembrey and Malcolm — that we have acquired new procedures which are now error-free.

DAVID J.H. BROCK

Human Genetics Unit,  
University of Edinburgh,  
Edinburgh, UK

VERONICA VAN HEYNINGEN

MRC Clinical and Population  
Cytogenetics Unit,  
Western General Hospital,  
Edinburgh EH4 2XU, UK

1. Pembrey, M.E. & Malcolm, S. *Nature* **320**, 114 (1986).
2. Brock, D.J.H. *et al. Lancet* **i**, 1175 (1985).
3. Gusella, J.F. *et al. Nature* **306**, 234 (1983).
4. Folstein, S.E. *et al. Science* **229**, 776 (1985).
5. Youngman, S. *et al. J. med. Genet.* **21**, 299 (1984).
6. White, R. *et al. Nature* **318**, 382 (1985).
7. Wainwright, B. *et al. Nature* **318**, 384 (1985).
8. Denning, C. *et al. Pediatrics* **66**, 752 (1980).
9. Brock, D.J.H. & van Heyningen, V. *Nature* **319**, 184 (1986).

### Homology of trichosanthin and ricin A chain

SIR—Trichosanthin is a plant protein extracted from a Chinese herbal medicine, the root tuber of *Trichosanthes kirilowii maxim*, Cucurbitaceae. It has been shown to bind to the trophoblastic syncytial layer killing the cells quite selectively. It can therefore be used clinically to terminate pregnancy and inhibit trophoblastic tumours<sup>1</sup>. This basic single-chain protein consists of 234 amino acid residues. Its sequencing has recently been completed<sup>2,3</sup> and a molecular model has been derived from the X-ray analysis<sup>4</sup>. Here we report a striking degree of sequence homology between trichosanthin and subunit A of ricin-D (ricin A for short). It is the first example of the complete sequence comparison between a ribosome-inactivating protein and cytotoxin.

In recent years, much attention has been paid to cytotoxins and the like for theoretical reason and because of their potential as immunotoxins to treat cancers. Many cytotoxins consist of two kinds of subunits. The toxic action is associated exclusively with subunit A which enzymatically inhibits protein synthesis of eukaryotes, whereas one or more B subunits have the ability to bind to cell surfaces and aid subunit A penetration into

cytosol. Among others, typical protein cytotoxins are diphtheria toxin, cholera toxin (from bacteria), ricin and abrin (from plants). However, it has also been found that many plants contain single chain proteins which are very similar to the toxic subunit A of the above cytotoxins in having the ability to inhibit protein synthesis. Pokeweed antiviral protein (PAP), momorcharin, wheat germ inhibitor and gelonin are good examples (for review see ref 5). *Momordica* and *Trichosanthes* are taxonomically closely related and all belong to the Cucurbitaceae family. It has been reported that momorcharin and trichosanthin have similar biological functions<sup>6</sup>. It led us to think that trichosanthin may also fall into the same group. To test for this possible relatedness, we therefore carried out sequence comparison studies. Sequence comparisons between short pieces of N-terminal segments of some ribosome-inhibiting proteins (RIPs) and cytotoxins have been recently reported<sup>7,8</sup>. Here we report the complete sequence comparison by computer analysis between trichosanthin and ricin A.

Figure 1 demonstrates our alignment of trichosanthin and ricin A, in which 91 residues are identical and 42 conservative. In our analysis, substitutions are taken as conservative only if two conditions are met — a minimum base change of one and that the two side-chain's physico-chemical properties are similar, as with Ser and Thr or Asp and Glu. The identical and conservative residues comprise at least 56 per cent of the trichosanthin sequence.

Further inspection shows that the identical and conservative residues are rather concentrated and that very little deletion needs to be assumed. In particular, there is no deletion at all from residues 97 to 149 in trichosanthin. We therefore believe that the sequence similarity between trichosanthin and ricin A is by no means a random event. It has been reinforced by the probability estimation of the alignment.

It is reportedly very rare for an identical pentapeptide to appear in different chains<sup>9</sup>, the probability is estimated to be about  $1 \times 10^{-6}$ . One may then expect the

