Letter to the Editor

15th November, 2005
Dr Li Jin
The Editor
Human Genomics

Dear Dr Jin

We were pleased to see the important topic of 'Human mutation databases' aired by Dr David Nelson in the pages of your journal (Nelson, D.R. [2005], *Hum. Genom.* Vol. 2, No.1, pp. 70–74), which provides data from an interesting analysis of 'mutation space'. However, the suggestion is made, via the following:

(a) Two questions. 'Do we know everything we need to know about mutants of haemoglobin-beta?' and 'How much mutation data is enough?'.
(b) The quotes: 'An expert is one who knows more and more about less and less' and 'considering cost there must be some point that crosses a practical return on investment'.

The author and the reviewers of the paper have missed a major point, which leads to the flawed conclusion that we do not need to find any more mutations, and that, in any event, they are 'variants of uncertain significance' and are thus of no use. The major point is that the genes are *not* sequenced in the first place purely so that the fault can be studied, and the author seems to ignore this. They are sequenced in the first place for several reasons:

1. Primarily, a clinician has a sick patient in front of them, a disease gene is indicated and a sequence study is ordered to attempt to get a definitive diagnosis, and consequently a prognosis, by the finding of a causative mutation.
2. Once a mutation is found in a gene which is found to cause the disease, the patient can be given a definite diagnosis. Increasingly, such findings can be used prognostically and can guide treatment.
3. This finding then allows the family to be offered the possibility to make reproductive decisions.
4. Increasingly, such findings will be used to define gene therapy tactics.
5. Via up to date mutation databases, clinicians can learn from a case with the same mutation on the other side of the world.

No more needs to be said, but it could be that the title and text treats 'variants of uncertain significance', 'proliferation of human gene disease databases' and 'the cost of analysing each mutation' in isolation, as if it has been done only for the benefit of informaticians and experts of particular genes. This is, of course, a by-product and a benefit. It should also be said that the whole of biology is suffering from Dr Nelson's assertion that leads him to conclude that, 'an expert is one who knows more and more about less and less' and that we 'have more data than we can possibly use'!

All of the above ignores the point that gene-specific databases are knowledge bases, with a host of information for a large number of people (Scriver, C.R. *et al.* [2003], *Hum. Mutat.* Vol. 21, pp. 333–344 and Claustres, M. *et al.* [2002], *Genome Res.* Vol. 12, pp. 680–688).

Yours sincerely

Richard G.H. Cotton
Director, Genomic Disorders Research Centre
President, Human Genome Variation Society

C.R. Scriver
Alva Professor Emeritus of Human Genetics
McGill University
Montreal Children's Hospital Research Institute

Response from Dr David Nelson

Firstly, I think I should clarify, for the readers, who I am and why I wrote the paper. I am the creator and curator of the Cytochrome P450 Homepage (http://drnelson.utmem.edu/CytochromeP450.html), a gene family database in existence since 1995. This database contains nomenclature and sequence information for almost 5,000 cytochrome P450 genes, from hundreds of species. I have been involved in cytochrome P450 research since 1985 and I have been naming these genes since 1987. I am not opposed to gene databases because this is how I spend much of my time. This paper arose from the fact that my wife was diagnosed with breast cancer in 2003. She had her *BRCA1* and *BRCA2* genes sequenced, and a variant of uncertain significance was found.

This sparked my interest in the human gene mutation databases, and led me to write the paper.

The final paragraph, which triggered the reply, raises a point about the cost of a making a set of data, such as mutation data, comprehensive by obtaining by mutagenesis every single point mutant and analysing it. In 1998, Ronald Kaback's laboratory used Cys-scanning mutagenesis and mutated every amino acid in the lacY permease protein and assayed each one for function.[1] This feat was extraordinary, and as far as I know, it has not been repeated on other genes, due to the high cost.

With 30,000 or more human genes, I was trying to ask: should we make comprehensive gene mutation datasets? I am not opposed to diagnostic sequencing, nor to placing the results in mutation databases. These are valuable tools, with considerable benefits. But there is a cost associated with knowledge, and complete knowledge has a very high price.

Reference

1. Frillingos, S., Sahin-Toth, M., Wu, J. and Kaback, H.R. (1998), 'Cys-scanning mutagenesis: A novel approach to structure function relationships in polytopic membrane proteins', *FASEB J.* Vol. 12, pp. 1281–1299.

David R. Nelson Ph.D.

Associate Professor
Department of Molecular Sciences
University of Tennessee Health Sciences Center
Memphis
TN 38163
USA