

---

# Letter to the Editor

Dr Li Jin  
The Editor  
Human Genomics

We read with interest the article by Dr David Nelson in the September 2005 issue of *Human Genomics* dealing with the nomenclature of the rat *Cyp* genes and the problems of gene nomenclature in general [Nelson, D.R. (2005), 'Gene nomenclature by default, or BLASTing to Babel', *Hum. Genomics* Vol. 2, pp. 196–201].

After the decision was made to coordinate nomenclature for human, mouse and rat genes wherever possible, the Rat Genome Nomenclature Committee assigned the Rat Genome Database (RGD) the task of periodically reviewing and updating rat gene nomenclature. The RGD serves as a repository of genomic, genetic and physiological information about the rat as a model organism for research, and of information on comparative genomics between the rat and other organisms. As such, the RGD is a community resource and is both responsible to and reliant upon the research community to present correct and up to date information. This includes the assignment of both homologies and nomenclature. Every effort is made to determine correct orthologies/homologies using informatic means, manual review and the homology resources of the Mouse Genome Informatics, Homologene and Ensembl. Nowhere is this more of a challenge than in dealing with families of closely related genes, such as the cytochrome P450 gene family. In cases such as these, we are grateful to researchers such as Dr Nelson who are able to advise us on the correct nomenclature for genes and/or gene families with which they have worked and are knowledgeable.

Dr Nelson highlights one of the reasons why community input into the various scientific databases is so vital. Databases rely on the expertise of the wider research community not only to supply them with data, but to review the records and correct the information when problems are discovered. Databases in general, and the RGD in particular, both need and encourage user input.

Dr Nelson's review of cytochrome P450 nomenclature was clear and concise. We would like to thank him for presenting both the challenges of gene family nomenclature

and a solution for the confusion which the current *Cyp* gene nomenclature may engender.

Yours sincerely  
*Professor Göran Levan*  
*Chair, Rat Genome Nomenclature Committee*  
*Lundberg Laboratory for Cancer Research*  
*Department of Pathology*  
*Göteborg University*  
*Göteborg, Sweden*

*Howard J. Jacob, PhD*  
*Principal investigator,*  
*Rat Genome Database*  
*Medical College of Wisconsin*  
*Milwaukee, Wisconsin, USA*

---

## Response from Dr David Nelson

Nomenclature is at the heart of communication and understanding. For gene nomenclature to be valuable, it must provide a useful (ie short) name that is widely used and recognised. If possible, the name should convey relationship information to other genes in the same family. The *CYP* nomenclature for cytochrome P450 attempts to do this by naming genes based on their sequence relatedness by using families and subfamilies. In a single species such as human, each gene is given a name and there is a grouping of 57 genes into 18 families and 43 subfamilies. When a second species (such as the rat) is added, again, each gene is given a name, but now there should be cross-referencing to human. Many of the genes are orthologues and, ideally, they should receive the same name. A number of the genes are paralogues, however, and this is where we often get into trouble. The automated naming systems employed by genome annotators based on best BLAST score can misassign names. These systems subscribe to the Star Wars system of nomenclature, 'There's no substitute for a good BLASTER at your side kid'. Once this happens, inaccurate names are perpetuated and distributed from the source to other databases and into the literature. As a human curator of a nomenclature system, I find this frustrating. As a counter to this, I have worked with members of the human

and mouse gene nomenclature committees, and the *Arabidopsis*, *Drosophila*, *Anopheles*, *Populus* and *Caenorhabditis elegans* nomenclature committees or genome annotators to get the names right at the 'official' source for names. This may have a curative effect, as newer database compilations may refer to these master name compendiums. It is nearly impossible to go to places such as Genbank and try to get things fixed there, since only the submitter has the right to change a submission.

There is a strong need for editorial lines in Genbank records, commenting on the errors in the entry, whether these are nomenclature errors or other errors. This would go a long way to fixing inaccuracies. As always, I would be happy to

work with the rat gene nomenclature committee, or any other committee, to correct any *CYP* names at the source.

David R. Nelson PhD  
Associate Professor,  
Department of Molecular Sciences,  
University of Tennessee Health Sciences Center,  
Memphis,  
TN 38163,  
USA