HUMAN MUTATION 22:181–182 (2003)

EDITORIAL

Standardizing Mutation Nomenclature: Why Bother?

Johan T. den Dunnen[1]* and Mark H. Paalman[2]

[1]*Department of Human and Clinical Genetics, Leiden University Medical Center, The Netherlands;* [2]*Managing Editor,* Human Mutation, *John Wiley & Sons, Inc.*

Lydon, April 12, 2008—a XBG patient and his parents sued the department of clinical diagnosis in Lydon, the XBG mutation database, and the journal Human Mutation. *The complaint was that serious and culpable mistakes were made during the clinical diagnosis of the pregnancy in the XBG-family, that ultimately led to the birth of an affected child. A paper published in* Human Mutation *listed the sequence variant detected in the family as "nonpathogenic." Careful examination would have revealed that the change was clearly pathogenic (a nonsense mutation). However, the accused parties failed to verify the data of the original report and just copied it.*

Is this imaginary news item pure fiction? Or, might it come true? When a clinical diagnosis is based on the detection of a variant in the DNA sequence (mutation), one wants to be absolutely sure. One of the most reliable decision tools available is to search the literature for confirmative reports. Nowadays, general or gene/disease-specific databases are often available that make this task rather simple. Consequently, when a sequence variant is detected, it is becoming general practice to check these repositories for previous reports of the change and accept the conclusion submitted by the author, "pathogenic" or "not pathogenic." For this process to be reliable, it is critical that mutation reports do not contain errors and that descriptions are unique and unequivocal. For this latter purpose, the HUGO Mutation Database Initiative (MDI) instigated an ad-hoc committee to formulate rules for the description of sequence variants [Beutler, 1993; Beaudet and Tsui, 1993; Beutler et al., 1996]. Based on initial suggestions, the nomenclature committee published several discussion papers describing rules for the description of sequence changes that are currently widely accepted [Antonarakis et al., 1998; den Dunnen and Antonarakis, 2000].

Because of the importance of the issue and the overall consensus on the rules, *Human Mutation* is adopting an editorial policy that requests absolute compliance of these mutation nomenclature rules before manuscripts will be accepted and published.

A quick review across a range of journals that report sequence changes highlights the most offended rules (den Dunnen, in preparation). First, most papers fail to explicitly define which sequence file was used as a reference for numbering residues (nucleotides and amino acids). Consequently, a best guess is made, trying to deduce the numbering used—and errors are introduced. When a cDNA sequence is used as a reference, it is not clearly stated where nucleotide residue 1 is located, i.e., at the start of the sequence file or at the A of the ATG translation initiation codon (the rule). In addition, many papers contain descriptions of intronic sequence changes based on an exon/intron numbering without specifying intron position, intron numbering, and the reference sequence file used. These simple, but basic, omissions make it difficult to correctly deduce the change reported. Intronic changes reported as c.IVS2–1A>G are inconclusive, while a notation like c.123–1A>G is clear.

Second, changes are frequently reported at the protein level without listing the change at the DNA level. DNA description of mutations is absolutely essential, since the amino acid code is degenerate. Errors can occur if one tries to deduce the underlying DNA change simply from the amino acid change. In addition, it is not uncommon that the one-letter amino acid code is used incorrectly: "A" is not only used for Alanine (correct), but also for Arginine, Asparagine, and Aspartic acid (incorrect).

Third, descriptions are often used that are not unequivocal. Examples include:

1. Insertions and deletions are reported in the formats c.123insAAG and c.123delGTG, where some mean starting *at* position 123 and others mean starting *after* position 123. The correct descriptions have the format c.123_124insAAG and c.123_125delGTG, respectively.
2. The use of a "–" (minus) sign to indicate both range and intronic nucleotides 5′ of a splice acceptor site are incorrect. Correct for range is: "_" (underscore), as in c.123_126del.
3. Changes at the DNA level are reported as A786G. That is the format for changes at the amino acid level. Correct notation is: c.786A>G.
4. Describing a variant as "L41L" is wrong. It is uninformative and equivocal (there are five possibilities at the DNA level). So, the description should be given at the DNA level.

Received 11 June 2003; accepted manuscript 11 June 2003.

*Correspondence to: JohanT. Den Dunnen, Mutation Nomenclature Editor, Department of Human and Clinical Genetics, Leiden University Medical Center,The Netherlands. E-mail: ddunnen@LUMC.nl

DOI 10.1002/humu.10262
Published online in Wiley InterScience (www.interscience.wiley.com).

5. Changes in recessive diseases are often reported without specifying whether changes in both alleles were identified and, if yes, in which combination (example: [c.123G>A]+[c.125delG] or [c.123G>A]+[?]).

Another source of confusion occurs when the consequence of a change is discussed and authors do not make clear whether the effect was experimentally verified. "Silent" missense mutations are reported to be benign polymorphisms, but might actually affect a splice enhancer/repressor and, thus, result in a pathologic change. Similarly, changes are often described to cause "errors in splicing," without clearly stating whether RNA was analysed. Consequently, changes end up in summary tables and mutation databases with conclusions attached without the remark "deduced, no experimental proof."

As the official journal of the Human Genome Variation Society (HGVS), *Human Mutation* continues to publish groundbreaking reports in medical and molecular genetics, DNA analysis, mutation detection methods, and valuable reviews on genetic disease. It is a key mission of the Society to support the least ambiguous, most informative mutation nomenclature system possible. To this end, the journal will enforce even more stringently these nomenclature standards. Recommendations for the description of sequence changes (mutation nomenclature rules) can be found linked from the Instructions to Authors and at the HGVS website www.HGVS.org/mutnomen/. The recommendations described at that site are those most recently published (den Dunnen and Antonarakis, 2000), but also include the latest additions and modifications.

REFERENCES

Antonarakis SE, Nomenclature Working Group. 1998. Recommendations for a nomenclature system for human gene mutations. Hum Mutat 11:1–3.

Beaudet AL, Tsui LC. 1993. A suggested nomenclature for designating mutations. Hum Mutat 2:245–248.

Beutler E. 1993. The designation of mutations. Am J Hum Genet 53:783–785.

Beutler E, McKusick VA, Motulsky AG, Scriver CR, Hutchinson F. 1996. Mutation nomenclature: nicknames, systematic names, and unique identifiers. Hum Mutat 8:203–206.

den Dunnen JT, Antonarakis SE. 2000. Mutation nomenclature extensions and suggestions to describe complex mutations: a discussion. Hum Mutat 15:7–12.