Mutant Gene Nomenclature in *Medicago truncatula*

Julia Frugoli, Clemson University
jfrugol@clemson.edu

At the 1999 *Medicago truncatula* meeting in Amsterdam, it became obvious to participants that with multiple laboratories generating mutants, the establishment of a policy for gene nomenclature for the species would prevent confusion in the literature. Many people shared experiences from other organisms they hoped rules would solve. Guided by nomenclature policies established for Arabidopsis, a policy was drafted by Julia Frugoli and Kate VandenBosch and circulated to key investigators in the field. After incorporation of comments by these investigators, the policy was published in Molecular Plant-Microbe Interactions. The Medicago steering committee at the time appointed Julia Frugoli to be the curator of gene names and the accompanying website. The publication of the nomenclature rules in 2001 noted that amendments to the rules could be considered at Medicago meetings and selection of the curator would be left to the steering committee, but no amendments have been proposed and the original curator remains.

These rules are for gene identified via mutation. Genes identified by sequence should follow Gene Ontology assignments as much as possible. The original nomenclature publication requested investigators to use the Plant Gene Nomenclature Guidelines for naming genes identified by sequence, but these guidelines and the accompanying Mendel database are no longer freely available on the web.

Before publication, gene names should be registered with the curator. An electronic gene submission process is set up for investigators to submit gene names for approval and posting. Updates or corrections may be submitted in the same manner. Assignment of a gene name does not have to wait until publication of the mutant. In fact, it is advisable to reserve a name for a gene as soon as the information needed to name it is available (see below). If an investigator does not wish the information to be publicly displayed on the website, a simple request keeps the information private until a publication occurs, but reserves the name. As more and more genes are identified via mutation, reserving early becomes increasingly important. Additionally, one does not want to be calling a gene by one name when presenting at meetings, only to have to change the name before publication because it does not meet the guidelines. The curator herself is guilty of this!

The basic rules, extracted from the original journal article and placed on the website are listed below, with a short discussion of how they are applied and common pitfalls.

1. Mutant gene symbols should consist of three lower case letters written in italics or underlined (e.g. *skl*), although symbols longer than 3 letters may be considered if they improve clarity.

Because of the stipulations in rule #5, the use of 4 letters has occurred several times. Thus while three letters is preferred, it is not mandatory.

2. Wild type alleles should consist of the gene symbol in capital letters written in italics or underlined (e.g. *SKL*).

3. Similarly, the full descriptive names of the alleles should be underlined or italicized, and written in lower case for mutants and capitals for wildtype (e.g. *sickle* or *SICKLE*).

These conventions of writing the gene name, as well as that noted in rule 1 were modeled after the system in use in yeast and Arabidopsis.

4. Different genes with the same phenotype, which bear the same three letter symbol, will be distinguished by a number following the three letters (e.g. *dmi1 vs. dmi2*). Different alleles of the same gene will be distinguished by numbers following the symbol, locus number and a hyphen (e.g. *dmi1-1* vs. *dmi1-2*).

It should be noted that "the same phenotype" can depend on the definition of phenotype given by the investigator. Thus, while the *dmi* mutants all do not make infections, they do have distinct phenotypes based on other criteria such as calcium spiking response. Likewise, the *dnf* mutants can be distinguished via molecular phenotypes. Alternatively, genes with similar phenotypes (such as the supernodulators *lss* and *sunn*) can be given different names if the investigator so desires. There are two benefits to naming genes with the same phenotype by the same three letter combination-it aids association of the phenotype (and hopefully pathway) with the name and given the limited number of three letter combinations available that meet all the criteria in rule 5, it can often be easier than finding a new name.

5. Assigning gene symbols to genes identified by mutation:

a. Investigators should make an effort to ascertain, through mapping or allelism tests, that the genes described represent loci that are distinct from previously named loci.

Sometimes this is not possible due to the appropriate stock not being available for crossing. In the initial publication of the rules, investigators were requested to share seed of published mutants for such purposes. Hopefully, with the pending establishment of a *Medicago truncatula* stock center, this step should become easier. Also, occasionally different alleles of the same gene give different phenotypes and allelism is not realized until the actual cloning of the genes. The intent is not to require crosses with every possible mutant before assignment of a gene name, but rather the most likely ones. If genes given different names are later discovered to be allelic, the website annotation will reflect this. While it is hoped that subsequent publications will choose a single name for further work, the original annotations will remain on the website to aid investigators.

Alternatively, a chromosome location identified by mapping that distinguishes the submitted gene from other similar genes is enough for posting on the website.

b. Investigators must avoid redundant use of any three-letter symbol already used for *M. truncatula* genes of distinctly different phenotype (consult the *M. truncatula* gene list).

This is self-explanatory. Two genes can't have the same name.

c. If new mutants are sufficiently similar to previously described *M. truncatula* mutants to imply function in a common pathway or process, it may be appropriate to retain a standard three letter symbol for the phenotype and simply change the locus number.

This is a reiteration of rule 4.

d. Avoid three letter symbols already in use for a family of sequenced genes.

For example, the *CAT* designation is in use for the catalase gene family in plants. Therefore, a mutant in any gene other than a catalase gene cannot be called a *cat* mutant, even though no mutant in Arabidopsis or *M. truncatula* has been published with that name.

e. Avoid symbols used for genes in Arabidopsis, unless the underlying gene is shown to be orthologous (consult the list of Arabidopsis mutant gene names). In the case of orthologous genes, the designation 'Mt' must precede the gene symbol (e.g. *Mtein1*) so that it is not confused with the gene from the original plant.

This is perhaps the rule that causes the most problems. Since the list of Arabidopsis mutants at TAIR is very well curated, freely available, and easily searched, it should be consulted before submission of a name. Use of the three letter symbol for a mutant in Arabidopsis is the most frequent reason for denying use of a gene name in *M. truncatula*.

f. Avoid symbols used to designate symbiotic genes or phenotypes of rhizobia (e.g. *fix*, *nod*, *exo*, and *nif*)

This is to eliminate confusion between mutations in the plants and mutations in their symbionts when discussing the two together.

For the most part, community adherence to the rules has been extremely high, perhaps because of the initial input of many investigators but also because of the nature of the Medicago community itself. The burden of compliance in any nomenclature system is on the investigators themselves and reviewers and editors of journal articles, but to date the investigators have complied without outside prompting. As the Medicago community expands, a continued observance of these simple rules should help keep the literature coherent for years to come.

Links

nomenclature policies established for Arabidopsis
http://www.arabidopsis.org/links/comm_stan.pdf
the policy was published
http://www.apsnet.org/mpmi/freepdfs/2001/1108-01O.pdf
website
http://www.genome.clemson.edu/affiliated_cugi/medicago/index.html
gene submission
http://www.genome.clemson.edu/affiliated_cugi/medicago/submitgenes.html
M. truncatula gene list
http://www.genome.clemson.edu/affiliated_cugi/medicago/genelist.html
Arabidopsis mutant gene names
http://www.arabidopsis.org/servlets/processor?type=genesymbol&update_action=view_symbol&symbol_type=mutant