

***Phaseolus* Genes and Gene Symbol Nomenclature**

BIC Genetics Committee

The Genetics Committee is a sub-committee of the Bean Improvement Cooperative that organizes and coordinates activities that deal with *Phaseolus* genetics. The committee has served as a clearinghouse for the assignment and use of gene symbols. The committee also maintains the **Guidelines for Gene Nomenclature (last published in the Annual Report of the Bean Improvement Cooperative in 1988, 31:16-19 and supplemented in 1999, 42:vi).**

The Committee strongly recommends that any researcher conducting studies of potentially new, qualitatively inherited traits of common bean submit their manuscript to the committee prior to publication (concurrent submission can be made to the genetics committee and the journal). The committee will evaluate the data to determine

- 1) If sufficient evidence exists to establish the inheritance hypothesis,
- 2) Whether any issue of potential allelism of the trait has been met, and
- 3) Whether the proposed gene symbol has been previously assigned to another gene.

The evidence must include:

- 1) Data from one generation to formulate an hypothesis and
- 2) Data from subsequent generations to test that hypothesis. The population sizes used must be sufficiently large to distinguish (with statistical significance) among potential segregation hypotheses.

Questions or comments should be addressed to the Chair of the Genetics Committee:

Dr. James S. Beaver, Dept. of Agronomy & Soils, University of Puerto Rico, P.O. Box 9030, Mayaguez PR 00681-9030;
ph. (787) 832-4040, ext. 2566; fax. (787) 265-0220; and e-mail;
j_beaver@hotmail.com

Guidelines for Gene Nomenclature (published in the Annual Report of the Bean Improvement Cooperative in 1988, volume 31:16-19). [REPRODUCED BELOW]

A Proposed Revision of Guidelines for Genetic Analysis in *Phaseolus vulgaris* L.

Guidelines for gene symbol nomenclature were published for *Phaseolus vulgaris* in 1977 (BIC 20:1-3). Since that time, the advent of biochemical and recombinant DNA techniques have led to new methods of genetic analysis and to "new" categories of genes. We present here a revised set of rules to be used in dealing with evidence for assigning gene symbols, and nomenclature for gene symbols, linkage groups and chromosomal aberrations.

These guidelines have been taken from rules for soybeans, peas, tomato and corn.

In the past, the uncoordinated use of certain gene symbols has created confusion in the literature. It is our hope that these rules will facilitate the use of future gene symbols.

Evidence Necessary for the Assignment of Gene Symbols

1. A genetic hypothesis is made on the basis of classification of segregating progeny, usually in the F₂ generation, and here called the hypothesis generation.
2. A second generation should be grown to confirm the proposed genetic hypothesis. This second generation may be progeny of the hypothesis generation (usually F₃) or progeny of a testcross (F₁ x recessive homozygote). In the case of Isozyme or RFLP markers, a second cross involving at least one alternate parent may be used in place of a second generation to confirm the proposed hypothesis.
3. Traits strongly influenced by nongenetic factors require verification of the classification scheme by evaluation of the progeny from homozygous plants of the hypothesis generation. Testcross data are not suitable for this purpose.
4. For genes with phenotype similar to that of previously published genes, tests of allelism must be made. This usually requires crossing a homozygous line carrying the newly identified gene with the original sources of previously published genes. For genes conditioning sterility or lethality in the homozygous condition, crosses should be made between heterozygous lines carrying the new gene and the previously known gene.
5. Follow the Rules of Nomenclature to assign a gene symbol.
6. Submit the manuscript to the chair, *Phaseolus* Genetics Committee, for review of the genetic interpretation and approval of the gene symbol (see current BIC for name and address of current Chair).
7. Submit a sample of 100 seeds of the line containing the new gene to the agencies maintaining the genetic type collection [USDA, Pullman & CIAT]. Duplicate collections are to be maintained to minimize loss of genetic stocks.

Rules of Nomenclature

Gene Symbols.

1. A gene symbol shall consist of a base of one to three letters. Previously published gene symbols will have priority. Additional information can be appended as described below.
2. Gene symbols shall be underlined or italicized to differentiate them from chromosome symbols.
3. For cases in which one allele is dominant, the first letter of the base symbol shall be capitalized to designate the dominant allele while all lower case letters will indicate the recessive allele.
4. Mutations at different loci with similar phenotypes or duplicate loci will be given the same base symbol followed by a numeric and italicized (or underlined) locus number. The locus number should be written on the same line as the gene symbol (eg., *msl*, *ms2*,). The first discovered mutant can be written either with or without the numeral "1".
5. Alleles at a locus which differ in phenotype within the same genetic background (e.g., near iso-genic lines) should receive a nonitalicized superscript letter symbol of up to three characters (Example R^{ma} , R^{rho}). The superscript may be written on the same line as the base symbol if preceded by a hyphen.
6. Alleles at the same locus which show identical phenotypes should use identical base symbols followed by a short letter or numeric identifier enclosed in parentheses (eg., *msl* (Prosser 1), *msl* (UI114).). The identifier should not be italicized or underlined.
7. Where a working symbol is needed for a new mutation of a known type, but unknown locus or allele, it shall receive the same base symbol followed by a "?". A hyphen preceding the question mark indicates an unknown allele at a known locus (eg., #-?); the absence of a hyphen indicates an unknown locus (eg., *msl*). An identifier enclosed in parentheses should be appended to identify the "new" mutation. The "?" should be replaced with an appropriate designation when tests of allelism have been performed.
8. The gene product of an enzyme or other protein is designated by an abbreviation consisting of two to four non-italicized capital letters (eg., ADH, GPI, MDH). The accepted names for nearly all enzymes are given in Enzyme Nomenclature. International Union of Biochemistry, Academic Press, 1984. Accepted names for other proteins such as phaseolin and actin are usually available in the recent biochemical literature. Appropriate abbreviations have been established for most enzymes commonly investigated in isozyme laboratories (see Wendel, J. F. and N. F. Weeden, Visualization and interpretation of plant isozymes, in: *Isozymes in Plant Biology* (C. Soltis and P. Soltis, eds), Dioscorides Press, 1988). Loci encoding enzymes or other proteins shall be designated using an appropriate two to four letter abbreviation of the biochemical name for that protein, the first letter of the abbreviation being capitalized (example: *Adh*, *Gpi*, *Mdh*).
9. When more than one isozyme (isozymes being enzymes catalyzing the same biochemical reaction but encoded by different genes) are identified, the isozymes

shall be distinguished by a numeric suffix with the most anodal being given the suffix 1 (e.g., GPI-1, GPI-2). Isozymes identified in later investigations will be given a numerical suffix in the order of discovery. The corresponding gene symbol will have only the first letter of the abbreviation capitalized (eg., *Gpi-l*, *Gpi-2*). In those cases where the isozymes show distinct subcellular localizations, the gene symbol should indicate the subcellular compartmentation of its product by substituting a "c" (for cytosolic), "p" (for plastid), "m" (for mitochondrial), or "mb" (for microbody, peroxisomal, etc.) for the numerical suffix in the isozyme designation (e.g., GPI-1 is coded by the gene *Gpi-p*).

10. Loci identified by Restriction Fragment Length Polymorphism (RFLP) analysis shall be assigned three italic (or underlined) capital letters followed by a number. Any combination of letters may be used except that "C" and "G" in the first letter position are reserved to indicate a cDNA clone or a genomic clone, respectively (eg., *CXX33*, or *GNN34*). The designation should correspond to the laboratory or group that discovered the RFLP. If the RFLP corresponds to a functional gene product then a superscript addition may be used to indicate the gene product. If this gene product corresponds to an existing locus that has been described phenotypically or biochemically, then the symbol for the existing locus should be used in the superscript.

Linkage groups and Chromosome numbers

1. A standard method of chromosome identification has not been accomplished. Therefore, we make two general suggestions. Chromosomes should be numbered according to their length with chromosome 1 being longest and chromosome II being shortest. Arabic numbers should be used.
2. Linkage groups should be designated by Roman numerals until they can be associated with their respective chromosomes.
3. Chromosome aberrations should be designated by non-italicized letters. Translocations are designated by the symbol "T" followed by chromosome or linkage group numbers. A lower case letter will be assigned to differentiate between translocations involving the same chromosomes. (Examples: T (I-2)a, T(I-2)b, T(III-V)). Inversions should be designated by the symbol "In" followed by the chromosome number and a lower case letter to distinguish different inversions of the same chromosome (Example: In(1)a, In(1)b). Deficiencies are distinguished by the symbol "DP and are designated in the same manner as inversions.

Cytoplasmic Factors

1. Symbols for cytoplasmic factors should follow the rules for nuclear factors.
2. Symbols should consist of entirely lower case letters.
3. Mode of inheritance should be listed as cytoplasmic rather than dominant, semi-dominant, or recessive.
4. Cytoplasmic male steriles should be indicated by the italicized symbol *cms* followed by a single letter hyphenated addition to indicate source (eg., *cms-C*, *cms-B*).

Priority and validity of symbols

1. A symbol is valid when it follows the above rules, when published in a refereed scientific journal or when published in the Bean Improvement Cooperative with adequate data support the conclusions. Published information should include description in biological or biochemical terms as appropriate, inheritance data, and tests of allelism.
2. Where different symbols have been assigned to the same factor, the symbol first published shall have priority unless it does not meet the rules for nomenclature or additional evidence shows that a change is necessary. Other symbols shall be regarded as synonyms.
3. A gene symbol is invalid when it does not meet the requirements outlined previously under "Gene Symbols".
4. A gene symbol meeting the above stated requirements are valid. A gene symbol for which a type line no longer exists is extinct. Valid and extinct symbols may not be reused whereas synonyms (see rule 2 above) may be reused provided they meet the requirements of the rules.

Rule Changes

1. These rules may be revised or amended by majority vote of *Phaseolus* Genetics Committee. [END]

Rule Change: (Annual Report of the Bean Improvement Cooperative in 1999, volume 42:vi). [REPRODUCED BELOW]

A tentative gene symbol may be published with the approval of the BIC Genetics Committee when the published rules and requirements have been met except for settling all relevant questions of allelism with related genes under two circumstances:

- 1) the paper supporting the existence of the new gene is already at practical limits of manuscript length and complexity, and the allelism tests can reasonably be expected to require a full length paper for publication or
- 2) the allelism tests cannot be performed successfully by classical techniques because the relevant recombinants do not have phenotypes that can be unambiguously distinguished; thus, requiring some molecular genetic approach, for which the appropriate markers are not currently available. The tentative nature of the gene symbol must be stated in the abstract and the tentative status stated to be due to questions of allelism with related genes not being yet resolved. [END]

GENETICS STOCKS COLLECTION

The Genetics Committee also evaluates materials submitted for inclusion in the Genetics Stocks Collection of the Plant Introduction System (for those rules see the **Annual Report of the Bean Improvement Cooperative in 1995, volume 38:iv-v**). **[REPRODUCED BELOW]**:

Report of the BIC Genetics Committee

J.R. Myers, University of Idaho-Kimberly

A) Coordination of Gene Symbols for Common Bean

The Genetics Committee evaluates genetic data for qualitatively inherited traits, and coordinates use of gene symbols. Rules of nomenclature were last published in the 1988 Ann. Rep. Bean. Impr. Coop. 31:16-19. Researchers who work with common bean, who publish data on simply inherited traits, and who propose gene symbols should be familiar with the Rules. These Rules provide guidelines for gene symbol nomenclature and for minimum data requirements to support the proposed gene designation.

To facilitate coordination of gene symbols, members of the genetics committee review voluntarily submitted manuscripts. The committee determines if the proposed gene symbol uses acceptable nomenclature, does not duplicate previously published gene symbols, and has sufficient data to support the gene designation. Manuscripts are submitted to the Chairman of the BIC Genetics Committee. After review these are returned to the author with anonymous comments from committee members. Current committee members are M. Dron, for Europe; S. Singh for Latin and South America; N. Weeden, M. Bassett, P. Gepts, J. Myers, for North America; and R. Hannan representing the U.S. Plant Introduction Collection. Persons interested in participating on the committee may inquire to the chairman (address below).

B) Submission of Phaseolus genetic stocks to the Plant Introduction Collection

The Western Region Plant Introduction (PI) Station, which houses the *Phaseolus* collection, has begun to accept genetic stocks into the collection. Genetic stocks are well-characterized lines with known genotypes, or may be chromosomal variants. These may include for example, unique, qualitatively inherited genes, the type stock for a widely used gene, translocation tester stocks, and recombinant inbred lines. These accessions are flagged to indicate their special status, and will have information indicating their genetic significance entered into the GRIN database. The collection currently has 25 genetic stocks and 8 translocation stocks.

The PI Station has finite resources, so must limit the number of accessions placed into the collection to those that are most significant. The *Phaseolus* Crop Advisory Committee has asked that the BIC Genetics Committee serve as the clearinghouse for submissions to the genetic stocks collection. If you have genetic stocks that should be preserved in the PI collection, relevant information should be sent to the Chairman of the Genetic Stocks Committee. Documentation may be in the form of a letter, manuscript, or published paper. Sufficient information should be presented for the committee to determine that the stock has been well characterized, and that data support the genetic interpretation of the stock. Upon approval by this committee, seed may then be sent to Rich Hannan, Curator, Western Region Plant Introduction Station, Pullman, WA.

Isogenic lines and well-characterized disease differentials will be accepted. Difficult to maintain material is strongly discouraged. These would include sterile or lethal genes that must be maintained in segregating populations, and pure lines that have low fitness levels. Recombinant inbred populations are generally discouraged, however, a limited number of the most significant populations will be accepted.

[END]

The **procedure for adding material to the genetics stocks collection** is: Submit to the chair of the BIC genetics committee documentation, including published data on inheritance and allelism, and photographs if possible that describes the trait.

Approximately 200 seeds are needed for propagation and preservation. The chair will send this information to members of the committee for review. Lines with traits that are highly deleterious or that require special efforts to propagate are discouraged.

Recombinant inbred populations that are of great significance, and are widely used, may be submitted for preservation. However, the PI collection does not have the space for general preservation of such populations.