

## Rules for Nomenclature and Gene Symbolization in Barley

Jerome D. Franckowiak<sup>1</sup> and Udda Lundqvist<sup>2</sup>

<sup>1</sup>Department of Plant Sciences

North Dakota State University

Fargo, ND 58105, U.S.A.

<sup>2</sup>SvalöfWeibull AB

SE-268 81 Svalöv, Sweden

In this volume of the Barley Genetics Newsletter the recommended rules for nomenclature and gene symbolization in barley as reported in BGN 2:11-14, BGN 11:1-16, BGN 21:11-14, BGN 26:4-8, and BGN 31:76-79 are again reprinted. Also, the current lists of new and revised BGS descriptions are presented by BGS number order (Table 1) and by locus symbol in alphabetic order (Table 2) in this issue.

1. In naming hereditary factors, the use of languages of higher internationality should be given preference.
2. Symbols of hereditary factors, derived from their original names, should be written in Roman letters of distinctive type, preferably in italics, and be as short as possible.

AMENDMENT: The original name should be as descriptive as possible of the phenotype. All gene symbols should consist of three letters.

COMMENTS: All new gene symbols should consist of three letters. Existing gene symbols of less than three letters should be converted to the three-letter system whenever symbols are revised. When appropriate, one or two letters should be added to existing symbols.

For example, add the letters "*ap*" to "*K*" to produce the symbol "*Kap*" to replace "*K*" as the symbol for Kapuze (hooded). As another example, add the letters "*ud*" to "*n*" to produce the symbol "*nud*" to replace "*n*" as the symbol for naked seed. Similarly the letter "*g*" can be added to "*ms*" to produce the symbol "*msg*" for genetic male sterility and the letter "*e*" can be added to "*ds*" to produce the symbol "*des*" for desynapsis. When inappropriate or when conflicts arise, questions should be referred to the Committee on Genetic Marker Stocks, Nomenclature, and Symbolization of the International Barley Genetics Symposium for resolution.

3. Whenever unambiguous, the name and symbol of a dominant begin with a capital letter and those of a recessive with a small letter.

AMENDMENT: When ambiguous (co-dominance, incomplete dominance, etc.) all symbols should consist of a capital letter followed by two small letters that designate the character, a number that represents a particular locus, and a letter or letters that represents a particular allele or mutational event at that particular locus.

COMMENTS: As an example, the letters "*Mdh*" can be used to designate the character malate dehydrogenase, "*MdhI*" would represent a particular locus for malate dehydrogenase and "*Mdh1a*", "*Mdh1b*", "*Mdh1c*", etc. would represent particular alleles or mutational events at the "*MdhI*" locus. Row number can be used as an example of symbolizing factors showing incomplete dominance. At the present time, the symbol "*v*" is used to represent the row number in *Hordeum vulgare*, "*V*" is used to represent the row number in *Hordeum distichum*, and "*V'*" is used to represent the row number in *Hordeum deficiens*. According to the amendment to Rule 3, if row number were to be designated by the letters "*Vul*", the designation of the locus on chromosome 2 would then become "*VulI*" and the alleles "*v*", "*V*", and "*V'*" would be designated "*Vul1a*", "*Vul1b*", and "*Vul1c*".

SUPPLEMENTARY AMENDMENT: A period should be placed before the allele symbol in the complete gene symbol.

COMMENTS: Since DNA sequences similar to those of the original locus may occur at several positions in the *Hordeum vulgare* genome, a three-letter symbol plus a number is inadequate to represent all potential loci. Also, both numbers and letters have been assigned to specific mutants and isozymes in *Hordeum vulgare*. The six-rowed spike locus is used as an example although the symbol *VulI* for row number in *Hordeum vulgare* is not recommended because the botanical classification of *Hordeum* spp has changed. The locus symbol *vrsI* and the name six-rowed spike 1 are recommended for the *v* locus. Gene symbols recommended for common alleles at the *vrsI* locus are *vrsI.a*, *vrsI.b*, *vrsI.c*, and *vrsI.t* for the "*v*", "*V*", "*v<sup>h</sup>*", and "*V'*" genes, respectively.

4. Literal or numeral superscripts are used to represent the different members of an allelic series.

AMENDMENT: All letters and numbers used in symbolization should be written on one line; no superscripts or subscripts should be used.

5. Standard or wild type alleles are designated by the gene symbols with a + as a superscript or by a + with the gene symbol as a superscript. In formulae, the + alone may be used.

AMENDMENT: This rule will not be used in barley symbolization.

6. Two or more genes having phenotypically similar effects are designated by a common basic symbol. Non-allelic loci (mimics, polymeric genes, etc.) are distinguished by an additional letter or Arabic numeral either on the same line after a hyphen or as a subscript. Alleles of independent mutational origin may be indicated by a superscript.

AMENDMENT: Barley gene symbols should consist of three letters that designate the character, a number that represents a particular locus, and a letter or letters that represents a particular allele or mutational event at that particular locus. All letters and numbers should be written on the same line without hyphens or spaces. Alleles or mutational

events that have not been assigned to a locus should be symbolized by three letters that designate the character followed by two commas used to reserve space for the locus number when determined, followed by a letter or letters representing the particular allele or mutational event. After appropriate allele testing, the correct locus number will be substituted for the commas. Where appropriate (when assigning new symbols or when revising existing symbols) letters representing alleles or mutational events should be assigned consecutively without regard to locus number or priority in discovery or publication.

COMMENTS: The use of the proposed system of symbolization can be illustrated by the desynaptic mutants. Two loci are known: *lc* on chromosome 1 (7H) and *ds* on chromosome 3 (3H). These will be resymbolized as *des1a* and *des2b*. A large number of desynaptic mutants have been collected. They will be designated *des,,c*, *des,,d*, *des,,e*, etc. If allele tests show that *des,,c* is at a different locus than *des1* and *des2*, *des,,c* will become *des3c*. If allele tests show that *des,,d* is at the same locus as *des2*, *des,,d* will become *des2d*. In practical use, the symbol *des* will be used when speaking of desynapsis in general or if only one locus was known for the character. The symbol *des2* will be used when speaking of that particular locus, and the symbol *des2b* will be used only when speaking of that particular allele or mutational event. If additional designation is needed in particular symbolization, it can be obtained by adding numbers behind the allele letters, and, if still further designation is needed, letters can be added to the symbol behind the last number. Symbolization consisting of alternation of letters and numbers written on the same line without hyphens or spaces will allow for the expansion of the symbol as future needs arise. In any work with large numbers of polymeric gene mutants, every mutant has to be given a designation not shared by any other mutant of this polymeric group and this designation should become a part of the permanent symbol representing that particular allele or mutational event. This requirement can be met by assigning allele designations in consecutive order without regard to locus number.

SUPPLEMENTARY AMENDMENT: A period should be used instead of two commas in gene symbols for mutants within a polymeric group that can not be assigned to a specific locus.

COMMENTS: The *des* symbol should be used when referring to desynapsis in general; *des1* and *des2*, for specific loci; *des1.a* and *des2.b* for specific genes or alleles at their respective loci; and *des.c*, *des.d*, *des.e* etc., for desynaptic mutants not assigned to a specific locus.

SUPPLEMENTARY AMENDMENT:

Even if the locus in question is the only one known that affects a given phenotype, the three-letter basic symbol is followed by a serial number.

7. Inhibitors, suppressors, and enhancers are designated by the symbols *I*, *Su*, and *En*, or by *i*, *su*, and *en* if they are recessive, followed by a hyphen and the symbol of the allele affected.

#### AMENDMENT

This rule is no longer applicable and will not be used in barley symbolization.

8. Whenever convenient, lethals should be designated by the letter *l* or *L* and sterility and incompatibility genes by *s* or *S*.

AMENDMENT: This rule will not be used in barley symbolization.

COMMENTS: J.G. Moseman (BGN 2:145-147) proposed that the first of the three letters for designating genes for reaction to pests should be *R*. The second and third letters will be the genus and species names of the pest.

SUPPLEMENTARY COMMENT: A motion was passed during the workshop on "Linkage Groups and Genetic Stock Collections" at the Fifth International Barley Genetics Symposium in 1986 (Barley Genetics V:1056-1058, BGN 17:1-4), that the International Committee for Nomenclature and Symbolization of Barley Genes should "recommend use of *Ml* as the designation of genes for resistance to powdery mildew."

9. Linkage groups and corresponding chromosomes are preferably designated by Arabic numerals.

SUPPLEMENTARY AMENDMENT: The current wheat homoeologous group numbering scheme (the Triticeae system) is recommended for *Hordeum vulgare* chromosomes. Arabic numerals followed by an H will indicate specific barley chromosomes. The *H. vulgare* chromosomes should be 7H, 2H, 3H, 4H, 1H, 6H, and 5H instead of 1, 2, 3, 4, 5, 6, and 7, respectively.

10. The letter *X* and *Y* are recommended to designate sex chromosomes.

AMENDMENT: This rule will not be used in barley symbolization.

11. Genic formulae are written as fractions with the maternal alleles given first or above. Each fraction corresponds to a single linkage group. Different linkage groups written in numerical sequence are separated by semicolons. Symbols of unlocated genes are placed within parenthesis at the end of the formula. In euploids and aneuploids, the gene symbols are repeated as many times as there are homologous loci.
12. Chromosomal aberrations should be indicated by abbreviations: *Df* for deficiency, *Dp* for duplication, *In* for inversion, *T* for translocation, *Tp* for transposition.
13. The zygotic number of chromosomes is indicated by 2n, the gametic number by n, and basic number by x.
14. Symbols of extra-chromosomal factors should be enclosed within brackets and precede the genic formula.

The following recommendations made by the International Committee for Nomenclature and Symbolization of Barley Genes at the Fourth International Barley Genetics Symposium in 1981 (Barley Genetics IV:959-961) on gene and mutation designations were as follows.

AMENDMENT:

- A. Present designations for genes and mutations. - Most of the present designations should be maintained. However, new designations may be given, when additional information indicates that new designations would aid in the identification of genes and mutations.
- B. New designations for genes and mutations. - New genes or mutations will be designated by characteristic, locus, allele, and then the order of identification or mutational event. Three letters will be used to identify new characteristics. Consecutive numbers will be used to identify the order of identification or mutational event. Loci will be designated by numbers and alleles by letters when they are identified. For example, *des-6* indicates that this is the sixth gene or mutation identified for the characteristic *des* (desynaptic). *des 1-6* and *des 2-7* indicate that gene or mutational events 6 and 7 for the desynaptic characteristic have been shown to be at different loci and those loci are then designated 1 and 2, respectively. *des 1a6* and *des 1b8*, indicate that the gene or mutational events 6 and 8 for the characteristic desynaptic have been shown to be at different alleles at locus 1 and those alleles are then designated a and b.

SUPPLEMENTARY COMMENT:

A motion was passed during the workshop of the "Nomenclature and Gene Symbolization Committee" at the Fifth International Barley Genetics Symposium in 1986 (Barley Genetics V:1056-1058) that "the recommended systems for Nomenclature and Gene Symbolization of the International Committee be published annually in the Barley Genetics Newsletter."

SUPPLEMENTARY COMMENT 2:

At the workshop for "Recommendations of Barley Nomenclature" held at Saskatoon, July 31, 1996 and adopted at the General Meeting of the Seventh International Barley Genetics Symposium, it was recommended that a period instead of a dash be used to designate the allele portion of the gene symbol. Consequently, the first gene symbol for the characteristic *des* (desynapsis) should be expressed as *des1.a*. The code *des1* identifies a specific locus. The period indicates that the symbol *a* identifies a specific allele or mutational event that produces a desynaptic phenotype. (The allele symbol *a* will be always associated with this specific desynaptic mutant even if the locus symbol is changed based on subsequent research results.)