Genet. Res., Camb. (1973), **21**, *pp.* 291–296
Printed in Great Britain

Gene symbols in *Aspergillus nidulans*

By A. J. CLUTTERBUCK
Department of Genetics, University of Glasgow, Glasgow G11 5JS

(*Received* 31 *January* 1973)

SUMMARY

Proposals are made for symbols consisting of a primary gene symbol of lower case italic letters referring to the mutant phenotype (three letters for new symbols, old symbols unchanged), an italic capital locus specific letter (or a hyphen in mutants not yet tested for allelism) and an italic mutant number. Optional superscripts are recommended to convey additional information. Further proposals relate to symbols for chromosome aberrations and to strain numbers.

1. INTRODUCTION

In the papers of Pontecorvo *et al.* (1953) and Käfer (1958) a convention was established by which mutants of *Aspergillus nidulans* were given a gene symbol consisting of one to five italic letters followed by a mutant isolation number. Loci were given the same symbol as the first mutant to be isolated at that locus. Subsequently, in the papers of Dorn (1965), Apirion (1965) and Foley, Giles & Roberts (1965) the more convenient practice of distinguishing loci by different capital letters after the gene symbol was adopted from bacterial genetics.

Bacterial gene nomenclature has been standardized by Demerec *et al.* (1966) and in recent years attempts have been made to come to a similar agreement for *Aspergillus nidulans* (Sermonti 1968; Clutterbuck 1968, 1969; Roper 1970). In 1969 a questionnaire was distributed with *Aspergillus Newsletter* no. 10, and the results (Roper & De Azevedo 1970) indicated preference for the system proposed by Clutterbuck (1968), although there was also some support for the idea of a unified fungal system. Unfortunately, the current systems in yeast (Von Borstel, 1969) and *Neurospora* (Barratt & Perkins, 1965) are sufficiently different from each other and from the current proposals for Aspergillus to provide little encouragement in this direction.

Demerec *et al.* (1966) proposed that older symbols should be altered if necessary to conform to standard three-letter symbols. The replies to the questionnaire on *Aspergillus* symbols, however, indicate a preference for the retention of existing symbols containing one to five letters. Although typographical uniformity is lost by this procedure, it will avoid problems of recognition of changed symbols and it is hoped that its general acceptability will make for uniformity of usage.

2. PROPOSALS

(i) *Primary gene symbols.* Gene loci and mutants should be designated by three-letter (normally lower-case – see Proposal ii) symbols in italic script. This rule applies to new symbols; symbols in existence, consisting of one to five italic letters, should not be changed.

The gene symbol should refer in some way to the field with which the locus is concerned or its mode of ascertainment, e.g. *pro* = proline requirement. It is convenient that this reference should be to the phenotype of mutants as they are most readily scored. It is a mistake, however, to attempt to be too specific or to incorporate too much information in the symbol, as subsequent research may then render the symbol inappropriate, e.g. *uvs* standing for 'ultra-violet sensitive' is preferable to *uvr* standing for 'ultra-violet repair'. It is similarly unwise to refer to a specific enzyme unless the biochemistry of the situation is beyond all doubt,

Where appropriate, the symbols proposed by Demerec *et al.* (1966) or Sermonti (1969) should be used. It is convenient if a gene symbol is pronounceable.

(ii) *Dominance.* It has been customary to indicate dominant mutants by the use of symbols with initial capital letters, e.g. *Acr* – acriflavine resistance (Roper & Käfer, 1957). It is now clear, however, that dominance is too complex a character to be satisfactorily indicated in this way in most cases. It is therefore proposed that as a general rule dominance should not be indicated in the primary gene symbol. If, however, an author wishes to draw attention to the dominance of a mutant in a particular publication, this may be done either by the use of superscripts (see Proposal viii) or capital letters. It should then be understood, however, that symbols such as *AcrA1* and *acrA1* refer to the same mutation.

(iii) *Locus specific letters.* Different loci which have the same primary symbol should be distinguished by an italic capital letter following the symbol, e.g. *proA*, *proB*. In the symbol of mutants whose locus is unknown, i.e. which have not been tested for allelism, the locus-specific letter should be replaced by a hyphen. Where only a single locus with a particular primary symbol is known it should still carry a locus letter. This will serve firstly to obviate changes if a second locus is discovered, and secondly to allow distinction of mutants which have been tested for allelism from those that have not.

The use of particular letters to denote, for instance, regulatory loci, as suggested by Demerec *et al.* (1966), should only be adopted on the understanding that the letter should not be changed even if the role of the locus is reassessed.

In Table 1 letters are assigned to loci that have previously been distinguished only by numbers.

(iv) *Mutant numbers.* Individual mutants with the same symbol are distinguished by italic serial mutant numbers after the symbol and locus letter, e.g. *proA1*, *proA2*, *proB3*. Where the allelic relationships of a mutant have not yet been determined, the capital letter is replaced by a hyphen, e.g. *pro-99*. No two mutants with the same primary gene symbol should have the same mutant number, even if it is believed that they are at different loci. This avoids difficulties if allelic

Table 1. *Addition of locus-specific letters to gene symbols of* Aspergillus nidulans

(The old locus symbol is given in parentheses after the new version. Gene symbols which have always had a locus-specific letter are not included in the table. For a full list of symbols and description of phenotypes see Dorn (1967), Clutterbuck & Cove (1973) and Clutterbuck (1973).)

abA	(*ab1*)	*galB*	(*gal3*)	*pcnbA*	(*pcnb*)
abaA	(*aba1*)	*galC*	(*gal4*)	*pdhA*	(*pdh*)
acrA	(*Acr1*)	*galD*	(*gal5*)	*phenA*	(*phen2*)
acrB	(*acr2*)	*galE*	(*gal9*)	*phenB*	(*phen6*)
actA	(*Act1*)	*galF*	(*gal2*)	*proA*	(*pro1*)
adC	(*ad1*)	*hisJ*	(*hisEL, his122*)	*proB*	(*pro3*)
adD	(*ad3*)	*iclA*	(*icl*)	*puA*	(*pu*)
adE	(*ad8*)	*ileA*	(*ile1*)	*pyroA*	(*pyro4*)
adF	(*ad9*)	*iodA*	(*Iod1*)	*riboA*	(*ribo1*)
adG	(*ad14*)	*lacA*	(*lac1*)	*riboB*	(*ribo2*)
adH	(*ad23*)	*lacB*	(*lac3*)	*riboC*	(*ribo3*)
adI	(*ad50*)	*luA*	(*lu*)	*riboD*	(*ribo5*)
alpA	(*alp*)	*lysA*	(*lys1*)	*riboE*	(*ribo6*)
anA	(*an1*)	*lysB*	(*lys5*)	*sA*	(*s1*)
anB	(*an2*)	*lysC*	(*lys6*)	*sB*	(*s3*)
aplA	(*allp*)	*lysD*	(*lys7*)	*sC*	(*s0, s12, cys2*)
argA	(*arg1*)	*lysE*	(*lys10*)	*sD*	(*s50*)
argB	(*arg2*)	*lysF*	(*lys51*)	*sbA*	(*sb3*)
argC	(*arg3*)	*malA*	(*mal1*)	*smA*	(*sm*)
biA	(*bi1*)	*masA*	(*mas*)	*sgpA*	(*sgp1*)
blA	(*bl*)	*mauA*	(*mau*)	*sgpB*	(*sgp2*)
brlA	(*brl*)	*meaA*	(*mea*)	*sgpC*	(*sgp3*)
bwA	(*Bw*)	*medA*	(*med*)	*sgpD*	(*sgp4*)
chaA	(*cha*)	*methB*	(*meth3*)	*sgpE*	(*sgp5*)
choA	(*cho*)	*methG*	(*meth1*)	*stuA*	(*stu*)
clA	(*cl4*)	*methH*	(*meth2*)	*suAadE20*	(*su1ad20*)
clB	(*cl6*)	*moA*	(*mo1*)	*suApabaB22*	(*sulpaba22*)
cnxB	(*ni50*)	*moB*	(*mo9*)	*suApro*	(*Su1pro*)
cnxE	(*ni3*)	*moC*	(*mo96*)	*suBpro*	(*Su4pro*)
coA	(*co*)	*niaD*	(*ni7*)	*suCpro*	(*su6pro*)
dilA	(*dil*)	*nicA*	(*nic2*)	*suDpro*	(*su19pro*)
drkA	(*drk*)	*nicB*	(*nic8*)	*sulA*	(*Sul1*)
flA	(*fl*)	*nicC*	(*nic10*)	*teA*	(*tr6*)
fluA	(*flu2*)	*nirA*	(*nir, niiB, ni51, am2*)	*telA*	(*re1*)
fluB	(*flu3*)			*thiA*	(*thi4*)
fluC	(*flu7*)	*ornA*	(*orn4*)	*tsD*	(*ts*)
fluD	(*flu11*)	*ornB*	(*orn7*)	*tsE*	(*ts6*)
fpaA	(*fpA, tyrA*)	*otaA*	(*ota*)	*uapA*	(*uap*)
fpaB	(*fpB*)	*pA*	(*p1*)	*veA*	(*ve*)
fpaD	(*fpC, FpD*)	*pabaA*	(*paba1*)	*wA*	(*w1*)
frA	(*fr1*)	*pabaB*	(*paba22*)	*wetA*	(*wet*)
fwA	(*fw, bge*)	*pantoA*	(*panto1*)	*yA*	(*y*)
galA	(*gal1*)	*pantoB*	(*panto100*)	*ygA*	(*yg*)

relationships are reassessed. It is particularly tempting to repeat mutant numbers where the phenotype is unusual, e.g. for constitutive mutants, but it is in just such cases that the chance of confusion is greatest, since such a mutant may prove to be at the same locus as a loss-mutant with a similar symbol.

(v) *Priority* of gene symbols should be observed wherever possible. In order to

facilitate this, a clearing house has been established by Professor J. A. Roper, Department of Genetics, The University, Sheffield (Roper, 1970). All new gene symbols should be registered with the clearing house before publication, and anyone isolating new mutants with established symbols should consult the clearing house for allocation of available mutant numbers (and if necessary locus letters).

(vi) *Wild-type alleles* should be indicated by a superscript 'plus' sign, e.g. $proA^+$.

(vii) *Multi-marked strains*. Mutant alleles in multimarked strains should be listed from the conventional left to right of each linkage group and with linkage groups in the order I–VIII. Markers on the same linkage group should be separated by a space – those on different linkage groups by a semi-colon. Diploids are best shown as follows, for example:

$$\frac{proA1\ biA1}{+\quad +}\ ;\ \frac{+}{pyroA4}.$$

(viii) *Superscripts*. Optional superscripts may be added to gene symbols in order to indicate special properties of particular mutants. Such superscripts, however, should not be regarded as permanent or essential parts of the symbol.

Superscripts have been found to be of particular value for regulatory mutants, e.g. the constitutive mutant $amdR6^c$ may be contrasted with the loss mutant $amdR11^-$ (Hynes & Pateman, 1970). Superscripts may also be valuable where it is necessary to draw special attention to some phenotypic peculiarity of a particular mutant such as temperature-sensitivity or dominance.

It follows from the optional nature of superscripts that mutants must be distinguishable by unique mutant numbers irrespective of the superscript they may sometimes carry (see Proposal iv). The same mutant may then carry different superscripts on occasions when it is required to draw attention to different properties.

(ix) *A suppressor* should be indicated by the symbol *su* with locus letters and mutant number followed (without a space) by the full gene symbol of the mutant suppressed. The symbol may be simplified if it is discovered that the suppressor is not allele-specific, e.g. *suA1proA1* isolated as a suppressor of *proA1* becomes *suA1proA* and then *suA1pro* when it is found to suppress, firstly, other *proA* alleles, and then also *proB* alleles. Again it is important that all suppressors of one type of mutant should have unique mutant numbers even if they are first isolated as suppressors of different individual mutants.

(x) *Chromosome aberrations* should be indicated by non-italic symbols consisting of a capital initial showing the type of aberration (T = translocation, I = inversion) followed by a serial number (1 should be used where the aberration is unique) and, in parentheses, the relevant chromosome or chromosomes in numerical order separated by a semicolon. Optionally, where a translocation is unidirectional, the semicolon may be replaced by an arrow, but in this case it should be recognized that, for example, T1(III → VIII) and T1(III; VIII) refer to the same translocation. Aberrations, like gene symbols, should be registered with the clearing house.

(xi) *Phenotypic symbols*, if used, should be clearly distinguished from gene

symbols. The unabbreviated word from which the gene symbol is derived is often suitable for indicating the phenotype or, alternatively, a non-italic version of the gene symbol, with the first letter capitalized, can be used. Phenotypic symbols have been more fully discussed by Demerec *et al.* (1966).

(xii) *Strain numbers*. Each strain in an individual collection should carry a letter or letters identifying the collection followed by a serial number. No attempt need be made to unify numbering systems in different collections, but a record of the origin of strains should be kept in each laboratory and this should include the original identification number of strains transferred from other collections. This system will allow distinctions to be made between strains of similar genotype with respect to known markers, but of differing origins.

3. DISCUSSION

A system of genetic symbols should try to balance two opposing aims: firstly, to provide permanent symbols that do not have to be changed in their essentials when new information about the loci they describe becomes available and, secondly, to aid in recognition of the loci referred to by incorporating as much information as possible about the function of the loci.

In this scheme, these aims are reconciled by the use of a primary gene symbol referring to the general area of metabolism with which the locus is concerned, and a unique serial mutant number. These two features should be regarded as the permanent symbol of a mutant. To them is added a locus specific letter which should also be a permanent feature of the symbol although even in its absence identity of the mutant should still be recognizable by the mutant number. Finally, the use of purely optional superscripts allows a writer to convey additional information about a locus or mutant if it is required in a specific instance.

The use of optional features of the gene symbol to some extent reduces the need for phenotypic symbols (cf. Demerec *et al.* 1966). Gene symbols conveying additional information by means of superscripts may be particularly useful, for instance, when describing the phenotypes resulting from the interaction of a number of mutations. Here phenotypic symbols are clearly inappropriate, but superscripts can provide a reminder of the properties of the individual mutants being combined.

I am very grateful to the following who have seen the manuscript and offered suggestions, encouragement, and even, in many cases, agreement: H. N. Arst, B. W. Bainbridge, B. L. Cohen, D. J. Cove, G. L. Dorn, E. Forbes, S. W. Glover, G. J. O. Jansen, E. Käfer-Boothroyd D. W. MacDonald, S. D. Martinelli, J. A. Pateman, E. C. R. Reeve, C. F. Roberts, J. A. Roper, C. Scazzocchio.

REFERENCES

Apirion, D. (1965). The two-way selection of mutants and revertants in respect of acetate utilization and resistance to fluoroacetate in *Aspergillus nidulans*. *Genetical Research* **6**, 317–329.

Barratt, R. W. & Perkins, D. D. (1965). Neurospora genetic nomenclature. *Neurospora Newsletter* **8**, 23–24.

VON BORSTEL, R. C. (ed.) (1969). Proposal for a new system of genetic nomencalture for use in yeast genetics research. Yeast genetics supplement to *Microbial Genetics Bulletin*, no. 31, pp. 1–8.

CLUTTERBUCK, A. J. (1968). Gene symbols and nomenclature: proposals and notes on them. *Aspergillus Newsletter* **9**, 26–29.

CLUTTERBUCK, A. J. (1969). Further comments on gene symbols. *Aspergillus Newsletter* **10**, 26–28.

CLUTTERBUCK, A. J. (1973). *Aspergillus nidulans*. In *Handbook of Genetics*, ed. R. C. King. New York: Van Nostrand Reinhold. (In preparation.)

CLUTTERBUCK, A. J. & COVE, D. J. (1973). The genetic loci of *Aspergillus nidulans*. *Handbook of Microbiology*. The Chemical Rubber Company, Cleveland, Ohio. (In the Press.)

DEMEREC, M., ADELBERG, E. A., CLARK, J. A. & HARTMAN, P. E. (1966). A proposal for a uniform nomenclature in bacterial genetics. *Genetics* **54**, 61–76.

DORN, G. L. (1965). Genetic analysis of the phosphatases in *Aspergillus nidulans*. *Genetical Research* **6**, 13–26.

DORN, G. L. (1967). A revised map of the eight linkage groups of *Aspergillus nidulans*. *Genetics* **56**, 619–631.

FOLEY, J. M., GILES, N. H. & ROBERTS, C. F. (1965). Complementation at the adenylosuccinase locus in *Aspergillus nidulans*. *Genetics* **52**, 1247–1263.

HYNES, M. J. & PATEMAN, J. A. J. (1970). The genetic analysis of regulation of amidase synthesis in *Aspergillus nidulans*. *Molecular and General Genetics* **108**, 97–106.

KÄFER, E. (1958). An eight-chromosome map of *Aspergillus nidulans*. *Advances in Genetics* **9**, 105–145.

PONTECORVO, G., ROPER, J. A., HEMMONS, L. M., MACDONALD, K. D. & BUFTON, A. W. (1953). The genetics of *Aspergillus nidulans*. *Advances in Genetics* **5**, 141–238.

ROPER, J. A. (1970). Clearing house for gene symbols, locus letters and allele numbers. *Aspergillus Newsletter* **11**, 25.

ROPER, J. A. & DE AZEVEDO, J. L. (eds.) (1970). Questionnaire on gene symbols. *Aspergillus Newsletter* **11**, 18–19.

ROPER, J. A. & KÄFER, E. (1957). Acriflavin-resistant mutants in *Aspergillus nidulans*. *Journal of General Microbiology* **16**, 660–667.

SERMONTI, G. (1968). List of proposed symbols. *Aspergillus Newsletter* **9**, 24–26.

SERMONTI, G. (1969). *Genetics of Antibiotic-producing Micro-organisms*. London, New York and Sydney: Wiley.