Journal of Nematology 26(2):138–143. 1994.

A Genetic Nomenclature for Parasitic Nematodes

DAVID MCK. BIRD[1] AND DONALD L. RIDDLE[2]

Abstract: A uniform system of genetic nomenclature for parasitic nematodes is presented. Conventions for naming strains, genes, alleles, and loci identified as DNA polymorphisms are established, and a standardized system for naming molecular clones is proposed.

Key words: allele, DNA, DNA polymorphism, gene, genetic map, locus, molecular clone, nematode, nomenclature, RAPD, RFLP, strain, STS.

The utility of a genetic approach to elucidate nematode biology has been amply demonstrated by researchers studying the free-living nematode *Caenorhabditis elegans* (1,2). Obviously, a similar approach could reveal considerable information about plant- and animal-parasitic nematodes. Although certain aspects of parasitic nematode biology, such as endoparasitism, limit the identification and maintenance of morphological or behavioral mutants, molecular techniques circumvent this problem. New interest in genetic mapping of parasitic nematodes has been sparked because genomic regions exhibiting DNA polymorphisms between individuals can be identified and mapped as Mendelian characters without the need to define genetic function.

The utility of any genetic map is considerably enhanced if individual researchers are able to use and add to the same map. Therefore, a uniform, unambiguous, and flexible nomenclature is required. Its rigorous application will increase the clarity of publications and improve understanding by biologists in other fields. The nomenclature also will help ensure that important information is not lost or misinterpreted when material is shipped between laboratories.

In this paper we present a system of genetic nomenclature, closely modeled after that developed for *C. elegans* (3). The system includes conventions for naming genetic entities (such as genes and alleles) and also the organisms that harbor them (nematode strains). Code prefixes are assigned to investigators, and an oversight committee to record issuance of gene names and new investigator prefixes is established.

Although this nomenclature is based on the *C. elegans* system, there are important differences. With *C. elegans* the N2 strain (1) is accepted as the wild type, and the nomenclature is designed to describe genetic differences from N2. Because many of the genes of interest in parasitic nematodes will be identified as naturally occurring alleles in wild isolates, no single wild type strain is designated for any species. This lack of a reference wild type requires modification in the use of allele and physical site designations from that practiced for *C. elegans*. Also differing from the *C. elegans* nomenclature is the inclusion of taxonomic prefixes required to identify homologous genes from different nematode species. Similarly, transposon names must also include taxonomic information and thus differ from names in the *C. elegans* system. The naming of transposon insertions and transgenic strains follows unpublished conventions employed by the *C. elegans* community. Although developed for plant- and insect-parasitic species, this nomenclature may be appropriate for other parasitic and free-living nematodes.

Received for publication 18 December 1992.

[1] Department of Nematology and Graduate Genetics Group, University of California–Riverside, Riverside, CA 92521.

[2] Division of Biological Sciences, University of Missouri–Columbia, Columbia, MO 65211.

We thank the many researchers who provided valuable comments and suggestions during the development of this nomenclature. We are particularly grateful to Dr. V. M. Williamson for help with conventions and to the members of the oversight committee (in addition to ourselves): Drs. J. Bakker, P. Castagnone-Sereno, and C. H. Opperman. We also thank Drs. B. C. Hyman and P. A. Roberts for critically reading the manuscript.

Oversight Committee

An essential feature of the nomenclature is that discrete genetic entities are given distinct, universally understood names. The proposed system achieves this by assigning strain and allele descriptor prefixes to individual research groups with the responsibility of assigning accession numbers. To avoid duplication, an oversight committee (OC) will act as a clearinghouse to maintain and distribute the prefix list and add new prefixes when requested. Additionally, the OC will function as a registration point for new gene names, serve as a repository for the accumulating genetic data, and conceivably play a role in the assembly of genetic maps. An OC has been appointed as an ad hoc committee of the Society of Nematologists. Heads of laboratories may request specific prefixes by contacting the chair of that committee, the name and address of whom can be found on the "Information for Contributors" page in the *Journal of Nematology.*

Strain Names

A strain is a line of true-breeding individuals with a particular genotype. All aspects of the genotype may not be known, but the described properties must be genetically stable so that results obtained with a particular strain may be reproduced at any future time. Although a sample of a field population is not a strain, strains may be generated from field isolates by inbreeding. For parthenogenetic species, a line clonally propagated from a single individual could constitute a strain. Strain names consist of two uppercase Roman letters followed by an Arabic accession number. To avoid confusion with numerals, the prefixes should not end in "I" or "O." To ensure that strain names are unique, each two-letter designator is assigned to only one laboratory and is registered with the oversight committee. Strain and allele designators should be different. Currently assigned strain prefixes are listed in Table 1. We anticipate that this list will grow as interest in genetic analysis of parasitic nematodes increases. Examples of strain names are "DB1" and "DB2." If the genetic properties of a strain are changed (e.g., by crossing, mutation, or selection), a new name should be issued. Strain names are never changed.

For each strain, investigators should establish a log of information, such as previ-

Table 1. Laboratory prefixes currently assigned to investigators.

Strain prefix	Allele prefix	Investigator(s)	Location
AN	*ia*	Castagnone-Sereno, Philippe	INRA, Antibes, France
BA	*cu*	Abbott, Albert G.	Clemson University, Clemson, SC, USA
BH	*rc*	Hyman, Bradley C.	University of California, Riverside, CA, USA
CB	*tv*	Vrain, Thierry C.	Agriculture Canada, Vancouver, Canada
CN	*cm*	Novitski, Charles	Central Michigan University, Mt. Pleasant, MI, USA
DB	*rv*	Bird, David McK.	University of California, Riverside, CA, USA
DK	*of*	Kaplan, David T.	USDA Horticultural Research Lab, Orlando, FL, USA
DS	*sp*	Samac, Deborah	University of Minnesota, St. Paul, MN, USA
EC	*dc*	Caswell-Chen, Edward P.	University of California, Davis, CA, USA
GT	*am*	Tylka, Gregory L.	Iowa State University, Ames, IA, USA
IG	*vn*	Glazer, Itamar	The Volcani Center, Bet Dagan, Israel
KS	*ok*	Schubert, Karel	University of Oklahoma, Norman, OK, USA
MN	*mn*	Young, Nevin D.	University of Minnesota, St. Paul, MN, USA
NZ	*jm*	Marshall, John W.	Crop and Food Research, Christchurch, New Zealand
OP	*kr*	Opperman, Charles H.	NC State University, Raleigh, NC, USA
PF	*in*	Ferris, Virginia R. and John M.	Purdue University, Lafayette, IN, USA
RB	*st*	Bolla, Robert I.	St. Louis University, St. Louis, MO, USA
RH	*ug*	Hussey, Richard S.	University of Georgia, Athens, GA, USA
TN	*mu*	Niblack, Terry L.	University of Missouri, Columbia, MO, USA
TP	*ln*	Powers, Thomas O.	University of Nebraska, Lincoln, NE, USA
VW	*da*	Williamson, Valerie M.	University of California, Davis, CA, USA

ously used, unofficial names, the genotype (genes and alleles), the origin (parental strains, method, by whom, and crossing history), a description (genetic and phenotypic), and any references. This information should accompany strains during movement from laboratory to laboratory.

Gene Names

Genes are DNA sequences that function as units of transcription to produce one or more related products. Classically, the existence of a gene has been revealed by observation of the phenotypic effect(s) of a mutation. Independent mutations that confer similar (or sometimes opposite or otherwise functionally related) phenotypes that fail to complement, and that appear to reside at the same location on a genetic map are considered to be allelic (i.e., to define the same gene).

As in the *C. elegans* model (3), gene names refer to general phenotypic or functional categories, and different genes within the same general category are distinguished by numbers. Names are preferably pronounceable and consist of three italicized lowercase letters that reflect the phenotype, followed by an italicized Arabic accession numeral separated from the three-letter phenotypic designation by a hyphen. An example of a gene name is *sec-1* for SECretion. Gene names representing (even partially) acronyms of the word "nematode" or of a particular species should not be used, as this information is redundant in a nematode genetic nomenclature. An example of poor name would be *mic-1* for *Meloidogyne Incognita* Collagen. In cases where a cloned locus is determined to be a pseudogene, a lowercase "ps" may be added as a descriptor (e.g., *sec-1*ps).

New gene names should be registered with the OC to prevent duplication with an existing name. The OC will be responsible for maintaining and distributing a list of gene names with appropriate literature references. If two names have accidentally been assigned to the same gene or locus, then the one published first will take priority. If a gene was initially mapped as a DNA polymorphism but later as a functional gene, then the polymorphic locus will be considered an allele of that gene and renamed accordingly, unless more than one gene is affected.

The taxonomic origin of each gene may be indicated by an italicized two-letter prefix representing the genus and species, separated from the gene name by a hyphen. This prefix is optional but is included when necessary to avoid ambiguity. For example, because the *sec-1* gene was isolated from *Meloidogyne incognita* (Ray and Hussey, pers. comm.), it could be written *Mi-sec-1*. Its homologue from *M. javanica* would be *Mj-sec-1*. Taxonomic prefixes assigned to date are listed in Table 2. We anticipate that many additions will be made to this list and, because nematode taxonomy remains in a state of flux, some names might be removed and permanently retired.

Protein products of genes use the same name as the gene, but in uppercase, nonitalic form. Thus, the product of the *sec-1*

Table 2. Abbreviations used as an optional prefix to a gene, rearrangement, or transposon name to indicate taxonomic origin.

Abbreviation	Taxonomic binomial
As	*Ascaris suum*
Ce	*Caenorhabditis elegans*
Gp	*Globodera pallida*
Gr	*Globodera rostochiensis*
Hb	*Heterorhabditis bacteriophora*
Hc	*Haemonchus contortus*
Hg	*Heterodera glycines*
Hs	*Heterodera schachtii*
Mf	*Meloidodera floridensis*
Ma	*Meloidogyne areneria*
Mt	*Meloidogyne artellia*
Mc	*Meloidogyne chitwoodi*
Mh	*Meloidogyne hapla*
Mi	*Meloidogyne incognita*
Mj	*Meloidogyne javanica*
Na	*Nacobbus aberrans*
Rc	*Radopholus citrophilus*
Rs	*Radopholus similis*
Rv	*Romanomermis culicivorax*
Sc	*Steinernema carpocapsae*
Ss	*Strongyloides stercoralis*
Ts	*Tylenchulus semipenetrans*

gene is SEC-1. Obviously, if the gene product is unambiguously known, then the biochemical name can be used.

Allele Names

Alleles are alternative forms of a gene that may or may not have an obvious effect on a trait. Because most of the strains of parasitic nematodes likely to be used for genetic analysis have been derived from wild isolates, most loci mapped will in fact be "wild-type." Consequently, every allele should be given a number. Allele names consist of two lowercase italicized letters, followed by an italicized Arabic accession number, e.g., *rv1*. This differs slightly from the *C. elegans* system (3), where for historical reasons, one-letter codes also are used. To avoid confusion of letters with numerals, allele names should not end in the letters "i" or "o" or "l," allowing a total of 598 permutations. Each prefix is assigned to only one laboratory and is registered with the OC. Currently assigned allele prefixes (designators) are listed in Table 1.

When gene and allele names are used together, the allele name in parentheses follows the gene name, with no space between the two, such as *ama-1(m118)* in *C. elegans*. Multiple allelic differences (mutations) within a single gene are listed as allele names within a single set of parentheses, e.g., *ama-1(m118m252)*. Genes or alleles on the same linkage group are listed in left-to-right order on the map, separated only by a space, for example, *dpy-13(e184) ama-1(m118)*, whereas genes on different linkage groups are separated by a semicolon. Loci or alleles on different homologues of the same linkage group (i.e., in trans) are separated by a slash. Thus, *dpy-13(e184)/dpy-13(e184)* is homozygous for that allele, whereas *dpy-13(e184)/dpy-13(m399)* is heterozygous, or in this case, heteroallelic. The addition of suffixes designed to convey additional phenotypic information, such as "ts" for "temperature sensitive," is permitted when the author explicitly defines their meaning. An example in *C. elegans* is *unc-54(e1300*ts).

Until tests such as genetic complementation or DNA sequence analysis clearly establish that a gene represents a newly defined locus, an accession number is not assigned; the new gene is distinguished by its allele name, e.g., *sec(ug1)*.

As linkage maps are generated, linkage group numbers (Roman numerals) will be assigned and should be registered with the OC. Where applicable, the sex linkage group should be called "X," and there should be no autosome X. The chromosome name may follow the gene and allele names as an option, e.g., *ama-1(m118)IV*. Cytogenetic analyses of plant-parasitic nematodes have been limited, and their interpretation has been complicated by presumed variability in ploidy levels (4). Consequently, correlation of linkage groups with chromosomes might prove difficult in some species.

Whenever complete genotypes become complex (and possibly confusing to those not familiar with the nomenclature), the interests of clarity may be served by assigning a trivial, easily understood nickname for repeated use in the text of a paper, as long as the trivial name is clearly defined by the complete genotype at an appropriate place in the paper.

Chromosomal Rearrangements

Chromosomal rearrangements are physical differences that may affect more than one gene. They are named with an italicized uppercase letter followed by an italicized lowercase letter between the allele prefix and an accession number (e.g., *rvIn5*). Proposed designations are listed in Table 3. Optionally, affected chromosomes may be designated in parentheses as descriptors of rearrangement names, following the *C. elegans* conventions (3). As an additional option, the taxonomic prefixes (Table 2) may be used to indicate the nematode species. A hypothetical example would be *Hg-krTr1(III;V)* or simply *krTr1*, in which arms of linkage groups III and V are exchanged in a reciprocal translocation. If the two halves of the translocation were separated from each other into sepa-

TABLE 3. Abbreviations used to indicate chromosomal rearrangements.

Abbreviation	Description
Df	Deficiency (deletion)
Dp	Duplication
Ex	Extrachromosomal, transgenic array
In	Inversion
Is	Chromosomally integrated, transgenic array
P	Physical marker†
Tr	Translocation

† Physical sites include RAPDs, RFLPs, and sequence-tagged sites.

rate strains, each half-translocation would then be given a *Dp* (duplication) name.

Free chromosomal duplications are common in *C. elegans* and might also be found in parasitic species following irradiation mutagenesis. Not being attached to any chromosome, such duplications segregate independently and are distinguished by a lowercase, italicized "*f*." The hypothetical duplication *rvDp1(II;f)*, or simply *rvDp1*, would be a rearrangement in which a portion of linkage group II has been duplicated as a free, independently segregating chromosomal fragment. Because karyotypic analysis of most nematode species is difficult, chromosomal aberrations such as translocations will likely be demonstrated by genetic or molecular means rather than by cytological analysis, but free duplications may be detectable as an extra chromosome in species with relatively simple karyotypes.

In *C. elegans*, deficiencies (deletions), *Df*, and duplications, *Dp*, are assigned according to their differences from the wild type, N2 strain. Because wild types are not designated in this nomenclature, determining whether one strain is carrying a duplication or the other is carrying a deficiency will not always be possible unless there is an obvious phenotypic effect. If more than two strains are available for comparison, then the name should correspond to the least frequent form. Investigators are encouraged to test multiple strains before assigning deficiency or duplication names.

Small deficiencies that are completely internal to a gene, such as *unc-54(e675)* in *C. elegans*, retain an allele number and are not assigned a *Df* name. Similarly, a deficiency that includes part or all of a gene and adjacent non-coding sequences should also retain a standard allele designation. However, if at least one neighboring gene is affected by the deficiency, then a *Df* name is appropriate, whether or not the adjacent gene has been identified genetically. Naturally, a *Df* mutant would be viable as a homozygote if no essential genes were deleted.

PHYSICAL SITES

Conceptually, physical sites identified as randomly amplified polymorphic DNAs (RAPDs), restriction fragment length polymorphisms (RFLPs), or sequence-tagged sites (STSs) are equivalent to chromosomal rearrangements and are named in the same manner (Table 3). Because these markers are potentially interchangeable (e.g., a RAPD fragment may be cloned and converted to a RFLP or to an STS), the descriptor "*P*" will be used for all polymorphic or dimorphic physical sites, preceded by an allele prefix and followed by an accession number, e.g., *rvP100*. It is stressed that this name applies to the genetic site of the polymorphism. The various polymorphic forms of the site are named by addition of a second Arabic accession number in parentheses, e.g., *rvP100(1)*. An RFLP will have at least two forms, *rvP100(1)* and *rvP100(2)*; an STS derived from one of the RFLP sequences may be *rvP100(3)*. Optionally, the taxonomic prefixes (Table 2) also may be used.

TRANSPOSONS

Transposons are mobile genetic elements. Names consist of an italicized "*T*" prefix followed by the species abbreviation and an Arabic number (e.g., *TAs1*). Each species may have its own numerical series, but assignments must be coordinated within a species. Numbers will be assigned by the OC to ensure that duplication in

naming does not occur. Before obtaining a *T*-number, an investigator must demonstrate to his or her satisfaction that the element to be named is likely to be a transposable element and is different from all previously named elements in that species (e.g., by hybridization criteria). The relevant information should be conveyed to the OC along with the request for a number.

Transposon Insertions

Mutant alleles generated by transposon insertions may be designated as such by following the allele number with two colons and the name of the inserted element, all in italics, as an optional description, e.g., *gen-1(da55::TAs10)*. When referring generally to such alleles, *gen-1::TAs10* may be used as a shorthand. Revertants demonstrated by molecular methods to be associated with transposon excision (either perfect or imperfect) should receive a new allele number and may carry an "rv" as a descriptor; e.g., a revertant of *da55* might be *gen-1(da62*rv*)*.

When molecular evidence is lacking, as in the case of a spontaneous mutant shown to revert spontaneously, the conventions for complex alleles should be followed. If molecular proof for transposon insertion and excision is later obtained, the allele name should then be changed to conform to the above convention. For example, if *gen-1(da37)* were demonstrated to be *gen-1(da37::TAs2)*, and the *gen-1(da37da42*rv*)* revertant gene was shown to have the transposon excised, its name would properly be changed to *gen-1(da42*rv*)*, or simply *gen-1(da42)*.

Because insertion sites of individual copies of transposons that are defined by molecular analysis but are not inserted in known genes are equivalent to physical sites (typically an RFLP), the insertion sites should be named as described for physical sites. Vacant sites are named as alternative forms of occupied sites.

Transgenic Strains

Cloned genes from parasitic nematodes now may be used to generate transgenic *C. elegans*. In the future, transgenic parasites may also be constructed. In the case of DNA sequences integrated into a chromosome, the genotype of a transgenic strain should be expressed by the laboratory allele prefix followed by "*Is*" and the isolation number (e.g., *rvIs52)*. When transgenes are maintained extrachromosomally, the allele prefix should be followed by "*Ex*" and the isolation number (e.g., *rvEx52)*. The genotype, or partial genotype, of the *Is* or *Ex* may be designated in brackets, e.g., *rvEx52[gen-1(da62)]*. If a gene has been altered by in vitro mutagenesis, it should receive a new allele number.

Molecular Clone Names

A vector containing nematode DNA should be named with a laboratory prefix. To distinguish molecular clones from nematode strains, a "#" is inserted between the prefix and the accession number. Thus, "DB1" is the name of a nematode strain, whereas "DB#1" is the name of a molecular clone constructed in the same laboratory. Because synthetic oligonucleotides are equivalent to molecular clones, they can be named in the same manner. Investigators may find it convenient to reserve a block of accession numbers to be assigned to oligonucleotides.

Literature Cited

1. Brenner, S. 1974. The genetics of *Caenorhabditis elegans*. Genetics 77:71–94.

2. Herman, R. K. 1988. Genetics. Pp. 17–45 *in* W. B. Wood et al., eds. The nematode *Caenorhabditis elegans*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.

3. Horvitz, R. H., S. Brenner, J. Hodgkin, and R. K. Herman. 1979. A uniform genetic nomenclature for the nematode *Caenorhabditis elegans*. Molecular and General Genetics 175:129–133.

4. Triantaphyllou, A. C. 1985. Cytogenetics, cytotaxonomy and phylogeny of root-knot nematodes. Pp. 113–126 *in* J. N. Sasser and C. C. Carter, eds. An advanced treatise on *Meloidogyne*, vol. 1. Raleigh: North Carolina State University Graphics.