PLANT MOLECULAR BIOLOGY REPORTER
Volume 1, Number 4, Fall 1983 pages 38-43

GENETIC RESOURCES

Proposals for the Naming of Chloroplast Genes

Richard B. Hallick

Department of Chemistry, Campus Box 215, University of Colorado, Boulder, CO 80309 USA

Warwick Bottomley

CSIRO Division of Plant Industry, P.O. Box 1600, Canberra City, ACT, Australia

At the conference on "Structure and Function of Plant Genomes" held in Portese, Italy, in September 1982, a meeting was held to discuss the possibility of making a standard nomenclature for naming genes coding for chloroplast proteins and RNAs. Those in attendance agreed that such a standard nomenclature would be valuable and nominated a committee of two, W. Bottomley from Canberra and R. B. Hallick from Boulder, to draft a set of guidelines. These were formulated and circulated at the meeting and the response was, in general, favorable. Subsequently, the draft proposal was circulated to various interested parties between September 1982 and April 1983, with an invitation to suggest changes or modifications.

At the conference on "Biosynthesis of the Photosynthetic Apparatus: Molecular Biology, Development, and Regulation" held in Keystone, Colorado in April, 1983 a second meeting was held to discuss all suggested amendments to the nomenclature, and to formalize a draft proposal for publication. The present set of guidelines was agreed to by all who attended this special workshop on chloroplast gene nomenclature. Those in attendance were Warwick Bottomley, Don Bourque, Rose Ann Cattolico, Marvin Edelman, A. A. Gatenby, John Gray, R. B. Hallick, Lee McIntosh, Laurens Mets, Jeff Palmer,

Offprint requests to: Richard B. Hallick

Jean-David Rochaix, William F Thompson, Elaine Tobin, and John C Watson

General Principles

In setting up the proposal for chloroplast gene nomenclature, some general principles were followed

A The gene name should contain the maximum information about the gene product to allow ease of identification.

B The name for a gene product should not be identical to the name for the corresponding gene

C In general, the nomenclature for bacterial genes should act as a guideline

D The gene name should consist of two parts The first part, generally a 3-letter code in lower case letters (preferably italicized) is used to designate the group to which the product belongs (ribosomal subunit, enzyme, membrane complex, etc) The second part, generally of one capitalized letter, but up to 2 or 3 letters or numbers, is used to designate the specific gene For groups of polypeptide genes, the capitalized letters or numbers used to designate specific genes within a group do not necessarily carry any connotation as to gene product relationships or polypeptide MW hierarchy

Recommended Nomenclature for Ribosomal RNA Operons and Ribosomal RNA Genes

Ribosomal RNA Operons

RNA operons are designated *rrn*A, B, C, etc., depending on the number of gene sets This is the standard *Escherichia coli* nomenclature For chloroplast genomes with two rRNA operons located in inverted repeat DNA segments (i e , maize, spinach, etc), when the larger single copy region is oriented at the top of the map and *rbc*L is on the left side of the map, the operon on the right side of the map is designated *rrn*A The operon on the left side of the map is *rrn*B. For tandemly arrayed operons (i e., *Euglena gracilis*), the 5'-proximal operon is *rrn*A The remaining operons are designated successively *rrn*B, C, etc. Partial operons do not receive *rrn* designation.

Ribosomal RNA Genes

Gene	*Gene Product*
16 S rDNA	16 S rRNA
23 S rDNA	23 S rRNA

5 S rDNA	5 S rRNA
4.5 S rDNA	4.5 S rRNA
7 S rDNA	7 S rRNA
3 S rDNA	3 S rRNA

Recommended Nomenclature for Transfer RNA Genes

General Principles

Transfer RNA genes are designated "*trn*" to designate the group of genes, and with the single-letter amino acid code to specify the esterified amino acid. Where identified, isoaccepting species can be designated with the anticodon following the amino acid code. Where the anticodon is unknown or as an alternative to anticodons sequential numbers can be used to distinguish isoaccepting species.

Specific Examples

Gene	*Gene Product*
*trn*F	$tRNA^{Phe}$
*trn*W	$tRNA^{Trp}$
*trn*C	$tRNA^{Cys}$
*trn*L1 or *trn*L-UAA	$tRNA^{Leu}_{1}$
*trn*L2 or *trn*L-CAA	$tRNA^{Leu}_{2}$
*trn*L3 or *trn*L-UAG	$tRNA^{Leu}_{3}$

Recommended Nomenclature for Chloroplast Protein Genes

The nomenclature is primarily designed for chloroplast DNA coded chloroplast protein genes. In some cases there are obvious extrapolations to related nuclear genes, i.e., for ATP synthase subunit genes, chloroplast ribosomal protein genes, etc. Some nuclear gene names are therefore proposed. However since the existence of multigene families for nuclear coded chloroplast proteins complicates this issue, we feel that the question of gene nomenclature for plant nuclear genes, i.e., chl A/B binding proteins, should be treated by experts in this field.

Chloroplast Ribosomal Protein Genes

Genes are designated in two groups. These are "*rps*" for ribosomal proteins for the small, or 30 S subunits, and "*rpl*" for ribosomal proteins of the large, or 50 S subunit. Where a clear analogy exists to an *E. coli* ribosomal protein gene, the gene is designated with the same number as the *E. coli* counterpart. Where there is no analogy to *E. coli* proteins, then letters can be used in place of numbers.

Gene	*Gene Product*
*rps*1, 2, 3, etc	30 S ribosomal protein analogous to *E. coli* S1, S2, S3, etc
*rpl*1, 2, 3, etc	50 S ribosomal protein analogous to *E. coli* L1, L2, L3, etc

Subunits of the ATP Synthase

Gene	*Gene Product*
*atp*A	CF_1 α subunit
*atp*B	CF_1 β subunit
*atp*C	CF_1 γ subunit
*atp*D	CF_1 δ subunit
*atp*E	CF_1 ε subunit
*atp*F	CF_0 subunit I
*atp*G	CF_0 subunit II
*atp*H	CF_0 subunit III (DCCD binding)

Stromal Polypeptides

Gene	*Gene Product*
*rbc*L	RuBP carboxylase large subunit
*rbc*S	RuBP carboxylase small subunit
*tuf*A	elongation factor Tu

Thylakoid Membrane Polypeptides

It is recommended that thylakoid membrane protein genes be grouped into three categories. These are "*psa*" for components of photosystem I, "*psb*" for components of photosystem II, and "*pet*", i.e., "photosynthetic electron transport," for proteins involved in the photosynthetic electron transport pathway. A number of specific recommendations are offered.

Gene	*Gene Product*
*psa*A	p700 apoprotein of PSI
*psb*A	"32 kd polypeptide"
*psb*B	46 kd reaction center polypeptide of PSII

*psb*C	43 kd reaction center polypeptide of PSII
*psb*D	*"D-2 Protein"*
*pet*A	cytochrome f
*pet*B	cytochrome b_6
*pet*C	Rieske Fe-S protein
*pet*D	polypeptide IV of cytochrome complex
*pet*E	plastocyanin
*pet*F	ferredoxin

Central Register

In addition it is recommended that a central clearing house register should be established. People wishing to publish new gene names would be encouraged to first check with the clearing house to avoid duplication and maintain consistency. Suggested revisions should be addressed to either Richard B. Hallick, Department of Chemistry, Campus Box 215, University of Colorado, Boulder, CO 80309, USA or to Warwick Bottomley, CSIRO, Division of Plant Industry, P.O. Box 1600, Canberra City, ACT, Australia.

Distinguishing Between Nuclear and Chloroplast Coded Genes

In many circumstances it is not necessary to explicitly distinguish between chloroplast and nuclear genes (i.e., as part of a gene name), since the coding locus is evident from the context of the work. However, Nicholas Gillham and John E. Boynton have noted that in genetic crosses involving both nuclear and cytoplasmic markers, it is often necessary to make such a distinction. They suggest using brackets or parentheses to separate one class of genes from another. This is comparable to bacterial crosses involving chromosomal markers, episomes, and viral genes. An example of a hypothetical genetic cross in *Chlamydomonas* involving nuclear and chloroplast genes might be described as follows:

$$\text{nic7}^{+}\ \text{arg2}\ \text{atpE}\ \text{rps3}\ \text{mt}^{+}\ (\text{atpA}\ \text{rpsl})$$

$$\times$$

$$\text{nic7}^{+}\ \text{arg2}^{+}\ \text{atpE}^{+}\ \text{rps3}^{+}\ \text{mt}^{-}\ (\text{atpA}^{+}\ \text{rpsl}^{+})$$

Endorsements

In addition to those individuals named above, the following people have reviewed this proposal and endorsed the chloroplast gene nomenclature suggestions. Hans Bohnert, John E. Boynton, Dennis Buetow, Nam-Hai Chua, Stephanie Curtis, Tristan Dyer, John Ellis, Nicholas Gillham, Wilhelm

Gruissem, Maureen Hanson, Elizabeth Harris, Martin Hartley, Robert B. Helling, Hans Kossel, Regis Mache, Eugene Nester, Michael Reith, Barbara Sears, Diter von Wettstein, and Herbert Weissbach. Revisions of this nomenclature proposal will be published in future issues of *The Plant Molecular Biology Reporter.*

Received August 8, 1983