

Geminivirus nomenclature: the need to set taxonomic standards

D. E. Rochester¹, R. N. Beachy¹, and C. M. Fauquet²

¹ ILTAB/TSRI, ² ILTAB/ORSTOM, Department of Cell Biology, The Scripps Research Institute MRC-7,
La Jolla, California, U.S.A.

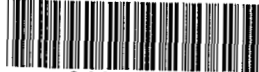
Summary. The geminiviruses are plant viruses of significant economic impact. Different nomenclature has been used in various studies in descriptions of their genomes. We propose in this communication a uniform nomenclature to be used in future studies.

*

Geminiviruses comprise a group of plant viruses with genomes of ssDNA. Several biological and molecular characteristics of these viruses have been used to subdivide them into three subgroups (I-III). Until recently, placement within a subgroup was on the basis of the number of genome components, host range, and naturally transmitting insect vector [1]. Subgroup I contains at least 10 geminiviruses with monocotyledonous hosts and are transmitted by leafhoppers in a persistent manner. Subgroup II has only one well-characterized virus, beet curly top virus (BCTV) which infects dicotyledonous hosts and, like subgroup I, is transmitted persistently by a leafhopper. Members of subgroups I and II have only one genomic DNA and are difficult to transmit mechanically. Subgroup III is the largest subgroup with thirty members identified to date. These geminiviruses usually have a bipartite genome, have a dicotyledonous host range, and are transmitted in a persistent manner by the whitefly *Bemisia tabaci*; some members have been experimentally transferred in a mechanical fashion. Unfortunately, the rapid influx of information from sequence analysis of cloned geminiviruses has led to a partitioning of the nomenclature used in describing sequence elements of their genomes. In this communication we are proposing simple rules to be used in future geminivirus descriptions.

Nomenclature of viruses within a subgroup

To date, all studied members of subgroup I and II have genomes comprised of a single ssDNA molecule (2.7-3.0 kb) encapsidated in twinned (i.e., geminate) particles. Subgroup III geminiviruses were considered bipartite in nature with two genomic components designated A- and B-DNA (2.4-2.9 kb). These viruses have been further subdivided on the basis of the geographical regions where the virus was first isolated [2]. Three research groups have now described different geminiviruses of the Old World that infect tomato and are most likely to be associated with subgroup III. These viruses include three isolates of tomato yellow leaf curl virus from Thailand (TYLCV-Th), Israel (TYLCV-I) and Sardinia (TYLCV-S) [3-5]. TYLCV-Th is apparently bipartite in composition with A- and B-DNAs analogous to previously described bipartite members of subgroup III. In the case of TYLC-



I and TYLCV-S, exhaustive and well-designed experiments have shown that these viruses contain sufficient genetic information for natural infection on a single DNA molecule analogous to the A-DNA of the bipartite geminiviruses. Although these three geminiviruses share the same vernacular name, sequence comparisons of their genomes and predicted open reading frames (ORFs) indicate a degree of diversity [5]. Unfortunately, similar and even identical names have been used for geminiviruses that are not identical at the molecular level.

We are currently analyzing a geminivirus isolated from India designated tomato leaf curl virus. Preliminary findings indicate that this virus is substantially different at the molecular level from a recently described TLCV isolated from Australia [6], and we have tentatively designated this virus as Indian geminivirus (TLCV-In).

Proposal #1

We propose that plant virologists continue to name geminiviruses according to the symptoms produced during infection (e.g., tomato yellow leaf curl).

Proposal #2

We propose to distinguish geminiviruses with identical vernacular names, by adding the country of isolation after the name. For example, tomato yellow leaf curl virus from Thailand was named TYLCV-Th.

Proposal #3

We propose to use at least two letters to name the country of isolation (e.g. Th for Thailand). The first letter of the country designation should be capitalized.

Proposal #4

We propose to use alphabetic characters in cases where the genome is multipartite (e.g., A- and B-DNA). In this case, A-DNA is the component which contains the replicase gene.

Nomenclature for intergenic regions

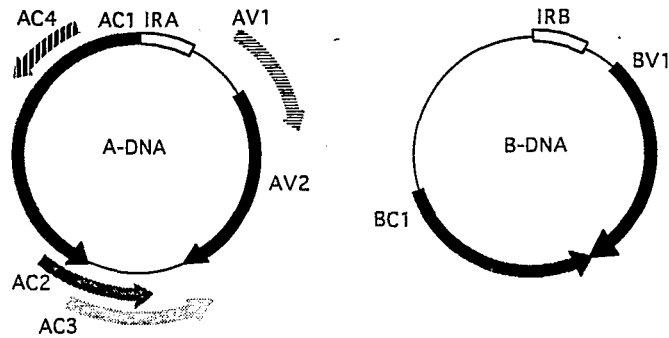
Within the bipartite geminiviruses, a 200 base region with highly conserved nucleotide sequence identity is found on both components. This sequence is referred to as the "Common Region or CR" and contains various conserved sequence motifs with many potential regulatory functions [7]. Monopartite geminiviruses also contain a similar region with the conserved sequence features; it has generally been referred to as "Intergenic Region or IR".

Proposal #5

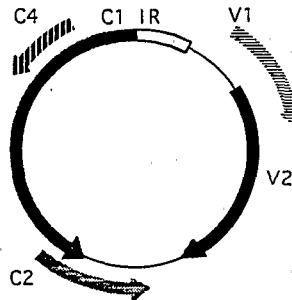
We propose that the noncoding regions which contain known conserved sequence elements be referred to as intergenic regions (IRs). IRs from different components should be further designated with alphabetic characters (e.g., IRA, IRB) (see Fig. 1).

Proposal #6

We propose that all published sequences of geminivirus genomes begin (i.e. nucleotide 1) at the first base 3' of the ATG of the replicase gene.



BIPARTITE GEMINIVIRUSES



MONOPARTITE GEMINIVIRUSES

Fig. 1. Schematic representation of generic bipartite and monopartite geminivirus genomic maps. *IR* Intergenic region, *V* open reading frame encoded on virion strand, *C* open reading frame encoded on complementary strand

Nomenclature for open reading frames (ORFs)

Sequence analysis of the genomes of at least 20 geminiviruses has shown that there are analogous ORFs based on sequence conservation and genomic position. Unfortunately, two different nomenclature systems are currently in use, depending on the number of genomic DNAs and/or personal preference. For clarity and uniformity, we propose that studies involving bipartite geminiviruses use nomenclature that is generally used for monopartite geminiviruses (see Fig. 1).

Proposal #7

We propose to label the open reading frames (ORFs) of geminiviruses "A" or "B" depending on which component the ORF is found. The letter is followed by "V" or "C" depending on whether the ORF is encoded by the *Virion* or *Complementary* strand. Furthermore, each designation is followed by a number given in the order of description (e.g., AV1). These numbers do not necessarily give relative distance from the intergenic regions. In the case of geminiviruses with a single component, the first letter would be deleted (e.g., V1) (see Fig. 1).

With development and application of modern molecular biological techniques, many gem-

iniviruses are being rapidly cloned and sequenced. We hope that researchers will continue to carefully use a nomenclature that is both useful and consistent.

References

1. Hull R, Stanley J, Briddon R (1991) In: Francki RIB, Fauquet CM, Knudson DL, Brown F (eds) Classification and nomenclature of viruses. Arch Virol [Suppl] 2, pp 173-177
2. Howarth AJ, Vandemark GJ (1989) Phylogeny of geminiviruses. J Gen Virol 70: 2717-2727
3. Rochester DE, Kosiratana W, Beachy RN (1990) Systemic movement and symptom production following agroinoculation (with a single DNA of tomato yellow leaf curl geminivirus Thailand). Virology 178: 520-526
4. Navot N, Picherski E, Zeidan M, Zamir D, Czosnek H (1991) Tomato yellow leaf curl virus: a whitefly-transmitted geminivirus with a single genomic molecule. Virology 185: 151-161
5. Kheyr-Pour A, Bendahmane M, Matzeit V, Accotto GP, Crispi S, Gronenborn B (1991) Tomato yellow leaf curl virus from Sardinia is a whitefly-transmitted monopartite geminivirus. Nucleic Acids Res 19: 6763-6769
6. Dry IB, Rigden JE, Krake LR, Mullineaux PM, Rezaian MA (1993) Nucleotide sequence and genome organization of tomato leaf curl geminivirus. J Gen Virol 74: 147-151
7. Lazarowitz S (1987) Molecular characterization of geminiviruses. Plant Mol Bio Rep 4: 177-192

Authors' address: ILTAB/ORSTOM, Department of Cell Biology, The Scripps Research Institute MRC-7, La Jolla, CA 92037, U.S.A.