

## DPVweb: AN OPEN ACCESS INTERNET RESOURCE ON PLANT VIRUSES AND VIRUS DISEASES

Michael J Adams and John F Antoniw, Wheat Pathogenesis Programme, Plant-Pathogens Interactions Division, Rothamsted Research, Harpenden, Herts AL5 2JQ UK describe a major source of information on plant viruses and the diseases they cause

### Keywords

Plant virus, Database, Disease descriptions, Diagnosis, Sequence analysis, Teaching, Symptomatology, Internet

### Introduction

Viruses cause major diseases of humans, animals and plants and have therefore been the subject of extensive study. Of the 1938 virus species officially recognised in 2005, about 900 are plant pathogens and these are responsible for huge losses in crop quality and yield to almost every major crop and in all parts of the world. Some effects of virus infection were known long before the causal agents could be recognised and studied, the earliest record probably being a poetic description of an attractive yellow vein symptom in Eupatorium leaves by a Japanese empress in 752 AD (Saunders *et al.*, 2003). Most viral infections, however, have much less desirable effects. The scientific study of plant viruses themselves began just over a hundred years ago when it was shown that mosaic and stunting on tobacco was caused by an agent smaller than any known bacterium and which we now know as *Tobacco mosaic virus* (TMV). TMV has remained an important model ever since (Scholthof *et al.*, 1999). The studies have expanded as various important technologies (e.g. protein purification, electron microscopy, immunology, molecular biology and genome sequencing) have been developed. The first complete genome sequence of a virus to be determined was that for phage phi X174 (a virus infecting bacteria) in 1977, and the first sequence of a plant virus (*Cauliflower mosaic virus*) followed in 1980. With the ever increasing amount of information available, the Commonwealth Mycological Institute and the Association of Applied Biologists (AAB) launched the first set of (twenty) sheets of Descriptions of Plant Viruses in 1970 (Gibbs *et al.*, 1970). These aimed to provide standardized authoritative descriptions, written by experts and summarising what was known about each virus and the disease(s) it causes. In this article, we describe the subsequent development of this initiative, its incorporation into an electronic form and ultimately into a web-accessible site, DPVweb.

### Descriptions of Plant Viruses

Between 1970 and 1989, a total of 354 descriptions were published. Fifteen of these were descriptions of groups of plant

viruses (corresponding mostly to present day genera) and 28 were revisions of earlier descriptions. The AAB became the sole publishers from set 19 (Description 296, September 1985) onwards. Bryan Harrison and Tony Murant remained editors throughout this period. Each of the 300+ virus (and viroids) descriptions contained sections describing (where known): the Main Diseases caused; Geographical Distribution; Host Range and Symptomatology; Strains; Transmission by Vectors, Seed, Grafting and Dodder; Serology; Relationships with other viruses; Purification and Properties of Particles; Genome Properties; Cytopathology; Ecology and Control. Selected references and photographs of virus symptoms, particles etc., were also provided. After 1989, no more printed Descriptions were published but plans were discussed to provide them in an electronic form and to continue the series. This eventually led to the development of a CD-ROM product that incorporated all the original descriptions with some new ones and much extra information on plant virus classification and genome sequences.

### Descriptions of Plant Viruses on CD-ROM

The CD-ROM for Windows™, launched by AAB in August 1998 (Adams *et al.*, 1998), contained the following elements:

- All the 354 previously published Descriptions originally issued, imported into a computer database and displayed as hypertext. This allowed users to search and display the descriptions more easily than the old paper format. The original format of the descriptions was kept as much as possible, but hypertext links were introduced so that relevant data could be accessed more easily. The text of these descriptions was not changed, except to correct typographical errors and to complete references to papers in press at the time of publication. Usually, the original figures were also provided but in many cases their quality was improved by electronic scanning of the original photographs or by replacing them with colour equivalents where possible.
- An additional 13 Descriptions that had not previously been published.
- An up-to-date taxonomic treatment of plant viruses, satellites and viroids, showing all the recognised families, genera and species and with brief family and genus descriptions.
- Lists of all plant virus, satellite and viroid sequences available from the public databases classified by their



Screen shots from Description 385, *Grapevine fanleaf virus*, in DPVweb. The top left panel shows the heading and menu and the bottom left panel demonstrates part of the text and the display of references in the lower frame. Two of the figures illustrating disease symptoms (which the web site displays in separate Figure windows) are shown in the two right panels.

current species name (information not easily and reliably obtainable directly from the databases themselves).

- For about 400 selected sequences, additional annotation was provided together with separate software (DPV Map) to display interactive feature maps, from which the sequences of particular genes or other features (e.g. coat proteins) could be easily extracted for analysis.

The program therefore provided a comprehensive resource on plant viruses and virus diseases for use in research, quarantine, extension and education. For those who purchased the CD, regular minor updates of the sequence data were provided via the internet and major updates with 10 and 12 new descriptions were issued in Sept 2000 and March 2002, respectively. By then, more than 4000 sequences were provided for interactive display and additional taxonomic and sequence data were provided for viruses of fungi and for animal viruses in the (three) families that also had plant-infecting members.

**DPVweb**

From autumn 2002, we began incorporating most of the information from the CD into a web site (<http://www.dpvweb.net/>). This is an open access web site and now contains most of the features of the CD version and with some more recent additions (including new virus descriptions). The home page provides simple and user-friendly access to all the features of the site. There are five major sections, accessed from a menu bar at the top of the page:

**Home**

This provides an overview of the site and an introduction to plant viruses in general.

**DPV**

This section contains the detailed descriptions of individual viruses. About 410 individual descriptions of plant viruses or virus groups are now included. A menu item allows a search to be made using a text query on the complete set of descriptions, and this can be limited to selected fields (subheadings in the descriptions) if required. Within each virus description, there are hypertext links to access the major sections and to display references in a small, separate, frame at the bottom of the page. Hypertext links are used to load figures and their legends into a separate window. Selected screenshots from the description of *Grapevine fanleaf virus* (DPV385) are shown in the Figure.

**Notes**

The notes are a brief description of each family and genus included within the project, which has now been extended to include all other viruses (animal, fungal, bacterial) except those with a dsDNA genome. This includes (a) a description of particle morphology and genome organisation (where known); (b) a genome map (usually of the genus type member); (c) a representative electron micrograph (for plant viruses only); (d) a list of species with their acronyms and synonyms for each genus with links to the plant virus

Descriptions; (e) a list of accession numbers (and links to the sequences) used by EMBL/Genbank/DDBJ databases for all the sequences in the family or genus together with the description from the sequence header (for viruses, satellites and viroids infecting plants, fungi and protozoa only); (f) a list of “curated sequences” for each virus, providing links to access the sequences of individual genes or other features. A menu item allows a search to be made amongst all the virus (and strain/synonym) names in the database. This returns a list of all matching names, with links to the appropriate genus notes page and, where appropriate, to the specialist description.

### Sequences

This provides lists of the accession numbers used by EMBL/Genbank/DDBJ databases for all the sequences of viruses, satellites and viroids infecting plants, fungi and protozoa. These have been checked to ensure that, as far as possible, they are allocated to their correct species. Each accession number is linked so that the sequence can be fetched from EMBL.

### Analysis

Data from all the curated sequences (currently c. 10,000) have been incorporated into a database to enable rapid access to the features and for comparative analyses. This page provides links to client software that accesses the data from the user's PC and which can be used for various types of molecular analysis (Adams and Antoniw, 2006).

### Management

Individual virus descriptions are written by specialist authors as attributed on each description and are commissioned and edited by a small team of virologists. Editors since 1998 have been Tony Murrant (Scottish Crop Research Institute, SCRI, until 2001), Hugh Barker (SCRI, until 2001), Teifion Jones (SCRI, until 2003), David Robinson (SCRI, until 2004) and Rick Mumford (Central Science Laboratory, CSL from 2003). Additional new editors have recently been appointed. The web site is managed and maintained by John Antoniw (Rothamsted Research) who also designed the databases and wrote all the software provided. Mike Adams (Rothamsted Research) is responsible for the taxonomic information, genus and family notes and for providing sequences with their annotation. The web site is supported by AAB.

### Uses of the CD-ROM and web site

The CD-ROM and web site are designed to be user-friendly and accessible to a wide range of users. We know that they are used in most parts of the world, including developing countries. For those needing to identify and control diseases (e.g. extension pathologists, plant quarantine services), the individual virus descriptions provide detailed information and illustrate the symptoms. Sequence data are useful for designing diagnostic procedures based on the polymerase

chain reaction (PCR) and are also the starting point for many fundamental studies (e.g. of virus gene function or virus evolution). Sequence features and tools available from the web site greatly simplify the process of assembling the required information. For example, we have recently used the data to re-examine species demarcation criteria in the large plant virus family *Potyviridae* and to determine the most appropriate regions of the genome to sequence for diagnostic purposes (Adams *et al.*, 2005). For animal virologists, the site provides an entry into the plant virus literature to enable studies of comparative biology. The site is also a rich educational resource and is used in some university courses to provide a simple introduction to plant virology with many in-depth examples of the intriguing diversity of viruses and virus diseases.

### Conclusion

DPVweb represents a unique resource because of its comprehensive coverage of plant virology. It particularly complements, and interacts with, the major database on virus taxonomy (ICTVdb; Büchen-Osmond, 2003) and the international sequence databases (NCBI/EMBL/DDBJ). For those needing the information in a standalone program, it is intended to release an updated version of the CD-ROM early in 2006.

### Acknowledgements

Rothamsted Research receives grant-aided support from the Biotechnology and Biological Sciences Research Council of the United Kingdom.

### References

- Adams, M. J.; Antoniw, J. F., and Fauquet, C. M. (2005) Molecular criteria for genus and species discrimination within the family *Potyviridae*. *Archives of Virology* 150:459-79.
- Adams, M. J., and Antoniw, J. F. (2006) DPVweb: a comprehensive database of plant and fungal virus genes and genomes. *Nucleic Acids Research* 34:Database issue, in press.
- Adams, M. J.; Antoniw, J. F.; Barker, H.; Jones, A. T.; Murrant, A. F., and Robinson, D. (1998) Descriptions of Plant Viruses on CD-ROM. Association of Applied Biologists, Wellesbourne, Warwick, UK.
- Büchen-Osmond, C. (2003) The Universal Virus Database ICTVdb. *Computing in Science and Engineering* 5:16-25.
- Gibbs, A. J.; Harrison, B. D., and Murrant, A. F. (eds) (1970) CMI/AAB Descriptions of Plant Viruses set 1. CMI, Kew and AAB, Wellesbourne, UK.
- Saunders, K.; Bedford, I. D.; Yahara, T., and Stanley, J. (2003) Aetiology: The earliest recorded plant virus disease. *Nature* 422:831.
- Scholthof, K-B. G.; Shaw, J. G., and Zaitlin, M. (eds) (1999) Tobacco mosaic virus: One hundred years of contributions to virology. APS Press, St Paul, MN, USA.

Michael (Mike) J. Adams is a Principal Investigator at the Plant-Pathogen Interactions Division, Rothamsted Research where he specialises in plant virology.

Similar articles that appeared in *Outlooks on Pest Management* include: 2004 15(3) 125