ICTVdB

the Universal Virus Database
of the
International Committee on
Taxonomy of Viruses

on the web since 1993

http://phene.cpmc.columbia.edu/

Cornelia Büchen-Osmond
Australian National University
Columbia University
Viruses coming to TaiBNET

How might ICTVdB assist construction of a virus database for TaiBNET?

• briefly introduce viruses, potentially the 8th kingdom of life (Mimivirus)

• describe ICTV, the International Committee on Taxonomy of Viruses that decides on virus nomenclature and classification

• outline distinctive functions of ICTVdB

• show how ICTVdB could be used to add viruses to TaiBNET.
What is a virus?

Viruses are found in all forms of life

- subcellular entities consisting of
  - protein capsids in remarkable diversity
  - may have a lipid envelope
  - nucleoprotein/genome
    - dsDNA, ssDNA, dsDNA-RT,
      dsRNA, ssRNA, ssRNA-RT
  - totally dependent on the host
    - for genome transcription and replication
    - for assembly, maturation and egression
Virus infection is host specific

Viruses usually

- infect specific hosts
  - host from one or more families
  - species specific (Influenza B virus)
  - Influenza A viruses have a wide-spread host range (birds, fish, reptiles, mammals)

- have a high mutation rates
- recombine in the host cell
- can acquire genes from the host
- can transfer genes to another host

Although much reduced forms of life, viruses have been called “master explorers of evolutionary space” and perhaps are a driving force in host evolution and speciation.
Rules of virus nomenclature

The ICTV

- recognizes Orders, Families, Subfamilies, Genera with standardized Latinized endings
- does not accept Linnaean binomial system (genus name followed by species name)
- accepts inclusion of host, symptom, and/or location in species names

species name: *Tobacco mosaic virus* (first virus discovered 1888)
alternative name: *Tobacco mosaic tobamovirus*

species name: *White spot syndrome virus* (found in Taiwan 1992)
species name: *Cercopithecine herpesvirus 1*
synonym: *Herpesvirus simiae*
(early attempt for true binomial nomenclature)
Virus properties used in taxonomy

Since ICTV was founded in 1966 the taxonomic status of a virus has been defined by

- **Virion properties**
  - morphology
  - genome, protein, carbohydrates and lipids

- **Genome organisation and replication**
  - metabolic interaction between virus and host
  - sequence annotations

- **Biological properties**
  - host range and vectors
  - cyto- and histopathology (disease expression)
  - transmission, epidemiology, geographic distribution
ICTV Reports: increase in species

1971 1st Report • 2 virus families
24 floating genera
16 plant virus groups

1990 5th Report • 38 virus families
138 genera/groups

1995 6th Report • 1 order
50 families
164 genera
3600 virus names

2000 7th Report • 3 orders
56 families
233 genera

2005 8th Report • 3 orders
73 families
287 genera
1905 species

2008 ICTVonline • 5 orders
84 families
314 genera
2084 species

but there are about 45,000 recorded virus names
ICTVdB, from the beginning

Anticipated “Three important elements essential for biodiversity information integration” (Ko, 2008)*

• data exchange; linkages to taxonomic and sequence databases, access to images and literature
• global unique identifiers; a taxonomically informative decimal code that gives each virus an “IP” number used as surrogate file name or accession number
• ontologies that record the rapidly evolving taxonomies of viruses

* Burke Chih-Jen Ko; last abstract in this Workshop handbook
Data exchange in ICTVdB

- to integrate information from the molecule (genomics biochemistry and structural biology) to the biosphere (host, geographical distribution to epidemiology)
- to handle semantically challenging nomenclature
- to present a concise and flexible data architecture for interoperability
- to facilitate many activities of ICTV, its interactions with research virologists (on line data entry) and the community at large (a “shop front”)

Many, but not all, of these objectives have been achieved in the face of a conservative discipline
ICTVdB has two principal parts:

- Index of Viruses; the master list of ICTV approved species, detailed species lists for each Family and an alphabetical index of all virus names (with synonyms, alternative names, misnomers, strains and isolates), reference lists and image gallery

- descriptions at all levels of virus taxa, generated from the DELTA formatted database

- interactive key for data retrieval and interrogation

- online data entry system
<table>
<thead>
<tr>
<th>Family Name</th>
<th>Genus</th>
<th>Species</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asfarviridae</td>
<td>Asfarvirus</td>
<td></td>
<td>African swine fever virus</td>
</tr>
<tr>
<td>Iridoviridae</td>
<td>Iridovirus</td>
<td></td>
<td>Invertebrate iridescent virus 6</td>
</tr>
<tr>
<td></td>
<td>Chloridovirus</td>
<td></td>
<td>Invertebrate iridescent virus 3</td>
</tr>
<tr>
<td></td>
<td>Ranavirus</td>
<td></td>
<td>Frog virus 3</td>
</tr>
<tr>
<td></td>
<td>Lymphocystivirus</td>
<td></td>
<td>Lymphocystis disease virus 1</td>
</tr>
<tr>
<td></td>
<td>Megalocytivirus</td>
<td></td>
<td>Infectious spleen and kidney necrosis virus</td>
</tr>
<tr>
<td>Phycodnaviridae</td>
<td>Chiorovirus</td>
<td></td>
<td>Paramaecium bursaria Chlorella virus 1</td>
</tr>
<tr>
<td></td>
<td>Prasinovirus</td>
<td></td>
<td>Micromonas pusilla virus SP1</td>
</tr>
<tr>
<td></td>
<td>Prymnesiovirus</td>
<td></td>
<td>Chrysochromulina brevisflum virus PW1</td>
</tr>
<tr>
<td></td>
<td>Phaeovirus</td>
<td></td>
<td>Ectocarpus siliculosus virus 1</td>
</tr>
<tr>
<td></td>
<td>Raphidovirus</td>
<td></td>
<td>Heterosigma akashiwo virus 01</td>
</tr>
<tr>
<td></td>
<td>Coccolithovirus</td>
<td></td>
<td>Emiliania huxleyi virus 86</td>
</tr>
<tr>
<td>Baculoviridae</td>
<td>Nucleopolyhedrovirus</td>
<td></td>
<td>Autographa californica multiple nucleopolyhedrovirus</td>
</tr>
<tr>
<td></td>
<td>Granulovirus</td>
<td></td>
<td>Cydia pomonella granulovirus</td>
</tr>
<tr>
<td>Nimaviridae</td>
<td>Whispovirus</td>
<td></td>
<td>White spot syndrome virus</td>
</tr>
<tr>
<td>Herpesviridae</td>
<td>Ictalurivirus</td>
<td></td>
<td>Ictalurid herpesvirus 1</td>
</tr>
<tr>
<td>Alphaherpesvirinae</td>
<td>Simplexivirus</td>
<td></td>
<td>Human herpesvirus 1</td>
</tr>
<tr>
<td></td>
<td>Varicellovirus</td>
<td></td>
<td>Human herpesvirus 3</td>
</tr>
<tr>
<td></td>
<td>Mardivirus</td>
<td></td>
<td>Gallid herpesvirus 2</td>
</tr>
<tr>
<td></td>
<td>Iltivirus</td>
<td></td>
<td>Gallid herpesvirus 1</td>
</tr>
<tr>
<td>Betaherpesvirinae</td>
<td>Cytomegalovirus</td>
<td></td>
<td>Human herpesvirus 5</td>
</tr>
<tr>
<td></td>
<td>Muromegalovirus</td>
<td></td>
<td>Murid herpesvirus 1</td>
</tr>
<tr>
<td></td>
<td>Roseolovirus</td>
<td></td>
<td>Human herpesvirus 6</td>
</tr>
</tbody>
</table>
Family 00.103. *Nimaviridae*

Genus 00.103.01. *Whispovirus*

**List of Species Demarcation Criteria in the Genus**

Only a single species within the genus has been identified to date (*White spot syndrome virus 1*). Various isolates with small genetic polymorphisms have been identified. It should be realized, however, that as the family *Nimaviridae* is newly recognized, the species organization may change once existing and new isolates are studied in more detail.

**List of Species in the Genus**

The ICTVdB virus code and the viruses. Official virus species names are in italics. Tentative virus species names, alternative names (%), isolates, strains, serotypes, subspecies, or rejected names are not italicized.

Virus codes, virus names, genome sequence accession numbers [ ] and assigned abbreviations ( ), are:

Species, their serotypes, strains and isolates:

- 00.103.01.001. *White spot syndrome virus 1*
  - [NC_003225] (WSSV-1-Ch)
  - [AF369029] (WSSV-1-Th)
  - [AF332093] (WSSV-1-Tw)
  - (Penaeus monodon SNPV)
  - (Chinese baculo-like virus)
Virus descriptions in ICTVdB

Classification

This is a description of an invertebrate virus at the species level.

ICTVdB Virus Code: 00.103.0.01.001. Virus accession number: 10301001.
NCBI Taxon Identifier NCBI Taxonomy ID: 92652.

Name, Synonyms and Lineage

ICTV approved acronym: WSSV. Virus is the type of the genus 00.103.0.01. Whispovirus; family 00.103. Nimaviridae.

Virion Properties

Morphology

Virions consist of an envelope, a nucleocapsid, and appendages. Virus capsid is enveloped by a trilaminar loose fitting envelope. Virions are ovoid, or bacilliform to ellipsoid, have protrusions that extend through the envelope and have a thread- or tail-like polar extension at one end. Virions measure 120-150 nm in diameter; 270-290 nm in length. A regular capsid structure is present. Capsid elongated and exhibits regular symmetry. The nucleocapsid is rod-shaped; has a tight-fitting capsid with a cross-hatched appearance and a width of 65-70 nm. Nucleocapsid contains DNA-protein core.

Physicochemical and Physical Properties

Virions have a buoyant density in CsCl of 1.22 g cm⁻³ (virions), or 1.31 g cm⁻³ (nucleocapsids). Virions are sensitive to treatment with detergents.

Nucleic Acid

The genome is not segmented and contains a single molecule of circular double-stranded DNA. The complete genome is 300000 nucleotides long. The genome has a guanine + cytosine content of 41 %.

GenBank records for nucleotide sequences; complete genome sequences.
ICTVdB uses a decimal code to uniquely identify each virus

The decimal code

• gives every virus in ICTVdB a unique “IP number”
• indicates its taxonomic status and level
• serves as a link within the whole database (interoperability)
• serves as a surrogate accession number in ICTVdB on the web and as hyperlink from other databases e.g., NCBI and SWISS-PROT or taxonomic databases such as ICTV-online, Species2000, GBIF
• records changing taxonomic decisions by ICTV expert Study Groups, but retains old codes to chart the history of virus taxonomy
The decimal code in ICTVdB

The decimal code for *White spot syndrome virus* indicates its taxonomic context

<table>
<thead>
<tr>
<th>Level</th>
<th>Code</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>Order</td>
<td>00.</td>
<td>...virales</td>
</tr>
<tr>
<td>Family</td>
<td>00.103.</td>
<td>Nimaviridae</td>
</tr>
<tr>
<td>Subfamily</td>
<td>00.103.0.</td>
<td>....virinae</td>
</tr>
<tr>
<td>Genus</td>
<td>00.103.0.01.</td>
<td>Whispovirus</td>
</tr>
<tr>
<td>Species</td>
<td>00.103.0.01.001.</td>
<td><em>White spot syndrome virus</em></td>
</tr>
<tr>
<td>Isolate</td>
<td>00.103.0.01.001.00.003.</td>
<td>WSSV-1-TW (1992)</td>
</tr>
</tbody>
</table>
Interoperability in ICTVdB

- Interoperability is achieved in descriptions
  - via decimal code within ICTVdB
  - from other databases to ICTVdB
- on species level and above via
  - NCBI TaxID to retrieve nucleotide sequences, genomes and PubMed references
- below species level via
  - sequence accession numbers
  - specific accession codes to
    - Databases: CDC, VIPERdB, DPV (CMI/AAB), GBIF
    - Catalogs: ATCC, DSMZ, d’Herelle
    - Publications: ProMed, journals
Since 2007, ICTV-online

- is a new database maintained by ICTV
- it contains the master species list that will be maintained by the ICTV Executive Committee
- updated yearly after ICTV EC meeting
- links to the ICTVdB “Index of Viruses” and virus descriptions
- provides a readily transportable list for integration into other databases
ICTV-online entry for *White spot syndrome virus*
Before viruses are entered in TaiBNET we have to:

- Coordinate a team of specialists to prepare lists of viruses in Taiwan, in
  - clinical medicine
  - agriculture
  - animal husbandry
  - aquaculture
  - natural ecosystems
  - simply, in all forms of life
- obtain taxonomic hierarchy tree data from
  provided by ICTV-online
- prepare short descriptions of isolate data
- customize links to ICTVdB and genomic databases