Q&A for webinar: The ICTV and how to navigate our web site (03/21/2024)

Question	Asker Name	Answer	Answer Name
In plant viruses what determines strain identities within an existing virus?	Antony Kigaru	*live answered*	Murilo Zerbini
The originally discovered/sequenced virus isn't always the Exemplar, is it?	Gene Godbold	The ICTV does not classify viruses below the level of species. Some Study Groups have established strain demarcation criteria. For example, the Geminiviridae SG established a 94% identity threshold for the demarcation of strains in the genus Begomovirus. But this is a voluntary SG initiative, and this particular example must not be extrapolated to other genera.	Murilo Zerbini
The originally discovered/sequenced virus isn't always the Exemplar, is it?	Gene Godbold	I'm sorry - the above answer refers to the question by Antony Kigaru.	Murilo Zerbini
The originally discovered/sequenced virus isn't always the Exemplar, is it?	Gene Godbold	Answering Gene's question, no, the exemplar is not necessarily the one originally described.	Murilo Zerbini

Question	Asker Name	Answer	Answer Name
There are tons of virus names that are the same as the species name still.	Detlef Leipe	They will all be replaced by binomial names as soon as the ratification vote is finished and the website is updated, probably by mid-April. In the meantime, you can see the proposed new names by accessing the Pending Proposals in the following menu path: Taxonomy -> Proposals -> Pending Proposals -> Select a Subcommittee -> Proposals for final ratification vote (2023)	Murilo Zerbini
What is the % similarity in nucleotide sequence that determines a new plant virus species?	Celia Chalam	This will vary from one family (or genus) to another.	Elliot Lefkowitz
What is the % similarity in nucleotide sequence that determines a new plant virus species?	Celia Chalam	Species demarcation criteria (including % identities) are established by Study Groups. Thus, they vary from family to family and even within families in some cases.	Murilo Zerbini

Question	Asker Name	Answer	Answer Name
With respect to the genome type: linear, circular. Is that information available on the VMR?	Walter Betancourt	Not at this time. We are preparing a new resource on virus properties to display information such as this.	Elliot Lefkowitz
Why is the coat protein gene a point of focus for most plant RNA viruses for example?	Antony Kigaru	We will have to get our plant virus taxonomy experts to answer that.	Elliot Lefkowitz
Why is the coat protein gene a point of focus for most plant RNA viruses for example?	Antony Kigaru	The past emphasis on the CP was a throwback from the days when serology was a major taxonomy criterion, but it no longer is. Most (maybe all ?) plant virus SGs require whole genome sequence comparisons these days.	Murilo Zerbini

Question	Asker Name	Answer	Answer Name
Hello everyone. Greetings from São Paulo, Brazil. First of all, thank you for this excellent opportunity to learn about ICTV and its work. My group works in the development of bioinformatics tools for viral species demarcation and detection. Our main approach is based on the development of taxonomically specific profile HMMs. VMR, as a reference list of approved viral taxa, is a major advance. I intend to produce taxon-specific models for viral detection and discrimination, using VMR as a golden standard of viral names and taxonomy. One limitation that I faced is the fact that there is still some inconsistency in the data. A given viral group, may have different nomenclatures for its genome segments. For instance, Bunyavirales may have segments L, M and S for some families and 1, 2, 3, 4, etc. for other families. This makes any effort to establish reliable sequence references for model design very challenging. Is there any data curation on the way to standardize this segment naming scheme? Thank you.	Arthur Gruber	We try and make sure that a standard nomenclature is used within a family, but it has been much harder to do this between families, and there are often good historical reasons for not changing how segments are named. The ICTV controls taxonomic names, but not virus names etc. so it's not something that could be enforced. Does that clarify things - although I guess it doesn't help you much? Donald Smith	Donald Smith
At least in the transition period from the old nomenclature to the binomial one, a tool to correlate the old names with the new ones would be useful to have.	Elisavet Chatzivassiliou	We provide a separate tab on the MSL spreadsheet that maps old names to new names.	Elliot Lefkowitz
Do the virus reports have PubMed ID numbers? Those look really useful.	Gene Godbold	The Taxonomy Profiles that are published in the Journal of General Virology for each Report Chapter are indexed in PubMed and have a PMID.	Elliot Lefkowitz

Question	**Asker Name**	**Answer**	**Answer Name**
What about phage Taxonomy?	Taiwo Faloni	Phage are handled like any other virus. Everything we show is available for bacterial viruses.	Elliot Lefkowitz
What about phage Taxonomy?	Taiwo Faloni	Indeed ! There is a Bacterial Viruses Subcommittee and also an Archaeal Viruses Subcommittee.	Murilo Zerbini
Thank you, Donald. It did clarify and I understand the point you raised. The challenge occurs when we want for instance to use all reference RDRP sequences of the order, but we don't know which numbered segments code it (it's not standardized like L segment, for instance). If at least we could know that segment 1 codes for RDRP, segment 2 glycoprotein, etc. this would be very helpful. Perhaps an additional column specifying this information would be very useful, without the need of changing the classical naming conventions.	Arthur Gruber	The problem might be that for another higher taxon it might be a different segment that was important and I can see this being tricky to curate. In most cases you might be able to screen for the segment coding for the RdRP on size alone	Donald Smith

Question	Asker Name	Answer	Answer Name
It is always a great experience to listen to ICTV members!! I am Manish Kumar from University of Georgia, USA. My question is, Tomato spotted wilt virus new nomenclature is Orthotospovirus tomatomaculae. While writing an article, for adding an abbreviation what should we use TSWV or the binomial nomenclature? Thank You!	Manish Kumar	If you are referring to the species, then the full binomial name should be used. But as you transition to talking about the virus, then, abbreviations are appropriate.	Elliot Lefkowitz
It is always a great experience to listen to ICTV members!! I am Manish Kumar from University of Georgia, USA. My question is, Tomato spotted wilt virus new nomenclature is Orthotospovirus tomatomaculae. While writing an article, for adding an abbreviation what should we use TSWV or the binomial nomenclature? Thank You!	Manish Kumar	Exactly. Acronyms such as TSWV are very useful but have no taxonomic value (and are not regulated by the ICTV) because they refer to the virus name (which remains "tomato spotted wilt virus" in the English language), not to the species name. So when you mention the species name, the acronym should not be used. But if you refer to the virus name, then it's ok to use it.	Murilo Zerbini
If the most recent taxbrowser release is 2022, then it is not up-to-date with the VMR?	Detlef Leipe	The current VMR corresponds to the current, 2022 MSL#38 release. We will make a new VMR available in late April, after the new MSL (#39, 2023) is ratified.	Elliot Lefkowitz

Question	**Asker Name**	**Answer**	**Answer Name**
I didn't get the reason for why the name polio disappeared after 2005.	Anthony Ike	I'll show that when we look at the history pages.	Curtis Hendrickson
I didn't get the reason for why the name polio disappeared after 2005.	Anthony Ike	(*live answered*) Did that explain sufficiently that poliovirus was merged and renamed, and how to find the proposal documents that give the reasons for that?	Curtis Hendrickson
Did you folks write/code this yourself or do you have contractors? (I'm just curious, I don't code but I've helped build a number of databases.)	Gene Godbold	With the exception of the Visual Taxonomy Browser, everything has been generated in-house (by Curtis, Elliot, Logan, and myself).	Don Dempsey
Inspired by Murilo's presentation at the ICTV/EVBC Workshop on Automating Virus Taxonomy (August 2023), who described the use of the Sequence Demarcation Tool for some viral groups, we felt stimulated to develop a new tool that can perform all-against-all genome comparisons, generating heatmap and frequency distribution plots. Protein sequences are used as input to generate plots of protein identity %, similarity %, and maximum-likelihood distance (using phylogenetic model). Also, the program can use 3D structure (PDB) files and perform all-against-all comparison with TM-align and generate TM-score distance plot, as well as use Foldseek and 3Di characters to generate 3Di similarity plots. Clustered data e distance trees are generated, helping to establish taxon demarcation criteria, even for remote viruses due to the structural analysis, which is more sensitive than pairwise sequence comparisons. Anyone interested to test/validate the tool is welcome to contacting me in this regard (argruber@usp.br)	Arthur Gruber	Thanks for letting us know about what you're working on, Arthur. Sounds very interesting!	Don Dempsey

Question	Asker Name	Answer	Answer Name
Is it possible to have access to a phylogenetic tree for a family?	Cica Urbino	If there is a Report chapter for the family then there will usually be a phylogenetic tree there along with the alignment used and the treefile. These trees are produced by the Study Group for that family, so the methods used and regions studied vary between families.	Donald Smith
I am curious to know, are there any plans to incorporate machine learning and AI algorithms to the site to automate specific tasks?	Suneesh Narayanan Sulabha	Potentially. One possibility is to extract demarcation criteria from proposals and Report chapters for virus taxa.	Elliot Lefkowitz
Acronyms are not unique!	Detlef Leipe	That's a good point, Detlef. Thanks for pointing that out. We will very likely return multiple results, along with a "score" indicating how close the match is.	Don Dempsey

Question	**Asker Name**	**Answer**	**Answer Name**
Do you have any connections with SwissProt/UniProt? I think they collaborate with ViralZone folks, but I've been urging the few people I know there to better annotate viral genes. In most cases, the polyprotein gets labeled with ALL the functions of the individual sequences, and they don't have complete protein lists for viruses (just one or a couple of polyprotein suffice, they think). Sorry this is long.	Gene Godbold	This is a bit beyond the remit of the ICTV - the VMR relates virus species to virus isolates, but annotating the genomes is a different kind of task …	Donald Smith
Let us know how we can help in consuming the NCBI Taxonomy appropriately - there are now ways to access our information using the datasets API (still being developed) for instance	Conrad Schoch	Thanks, Conrad. I'd like to hear more about the datasets API. For now, we are downloading the database export and importing into a local database.	Don Dempsey
It would be nice to have an updated page for demarcations and criteria for novel virus discovery. At least the most important criteria like similarity or profile homology at species, genus, or any other rank.	Naser Poursalavati	*live answered*	Curtis Hendrickson
It would be helpful for the NCBI to get a comprehensive straightforward one-to-one list of the old species name and the new binary species name (not as part of the MSL).	Detlef Leipe	We should be able to generate such a list fairly easily. Should we follow up with you on this, or can you recommend someone else?	Don Dempsey

Question	Asker Name	Answer	Answer Name
What program are usually employed to apply the viral demarcation criteria based on identity %?	Arthur Gruber	Sequence (Demarcation Tool) is one of the best choices – as its name implies, it was developed specifically for this purpose. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4178126/	Murilo Zerbini
At some point it has been considered adding to the master list also the host species where each virus, viroid or satellite RNA has been reported.	José Manuel Cambrón Crisantos	You can get general information about host from the VMR - but only at the level of vertebrate, plant etc. - not quite what you are after. One issue is that this would not be an exhaustive list of hosts, just of those used as exemplar or additional isolates on the VMR.	Donald Smith
In the process of proposal preparation, it is important to download from GenBank a non-redundant set of viral sequences related to the sequence to be classified. Most of the sequences are in duplicate, some others are incomplete. Is there any convenient tool to filter the data for downloading in such a way as to avoid duplicated and incomplete sequences and not to lose anything important?	Małgorzata Łobocka	The list of virus accessions in the VMR provides a good set of distinct sequences.	Curtis Hendrickson
Are there recommended best practices for sequence alignment and phylogenetic tree construction to yield the 'best' tree for taxonomic purposes?	John Hammond	*live answered*	Curtis Hendrickson

Question	Asker Name	Answer	Answer Name
Don, still a work in progress but the datasets tools installation info are here: https://www.ncbi.nlm.nih.gov/datasets/docs/v2/download-and-install/ and we have some taxonomy documentation here: https://www.ncbi.nlm.nih.gov/datasets/docs/v2/how-tos/taxonomy/taxonomy/	Conrad Schoch	Thanks, Conrad. I'm looking forward to finding out more about this.	Don Dempsey
Hi Malgorzata, The NCBI Virus resource allows searching for virus sequences using taxonomy, and there is also a filter for partial or complete sequences. We do not yet support reducing duplications of similar sequences, but we do have a "random download" to reduce the size of very large datasets. Please feel free to reach out to me at eneida.hatche@nih.gov, or visit us at https://www.ncbi.nlm.nih.gov/labs/virus/vssi/#/virus?SeqType_s=Nucleotide	Eneida Hatcher		
What is the usage license of the family reports, esp. of the images? The last families I looked up had a copyright. But I think I saw CreativeCommons 4.0 in the past years. I don't know if it would be possible to make this available in principle or as far as the authors allow.	Ernst Sauerwein	All the Report content can be reused with acknowledgement - images used from other sources can only be reused under the license of that source.	Donald Smith