MEETING OF THE STUDY GROUP CHAIRS AND EXECUTIVE COMMITTEE OF THE INTERNATIONAL COMMITTEE ON TAXONOMY OF VIRUSES

1–2 February 2016

Hinxton Hall, Wellcome Trust Genome Campus

Aims of the meeting

Questions

- What are the problems that virus classification is facing?
- What is the ICTV going to do about them?

Aims

- To inform on current plans to develop ICTV resources
- To listen to views on the ICTV
- To see ways forward in the context of wider taxonomy

Threads

- Explanations of the ICTV's present situation and future plans
- A session devoted to discussion of views on the ICTV
- Discussion of topics of general interest in viral taxonomy
- Plenary lectures that help us look over the battlements
- Opportunities for informal discussions

ICTV resources and organisation

A wide range of members

- Executive Committee (EC)
- Life members
- National representatives
- Subcommittees, including the 83 Study Group (SG) chairs

Involving hundreds of virologists

- 6-12 members in each SG

Representing the top expertise in each field

- ICTV Report publication
- Maintenance of the Master Species List (MSL)
- Rigorous, formalised, democratic procedures for assigning taxa

And yet…

The ICTV thread

1. The current situation

- History
- Operations
- Resources
- Challenges

2. Redesigning the ICTV Report

- Change
- Practicalities
- Prototype
- Implications

3. Development of bioinformatic tools

- Alignments
- Evolutionary analysis
- Wikis, etc.
- Tools

4. The future

- Metagenomics
- Taxonomical assignments
- Nomenclature
- Databases and tools

THE CURRENT SITUATION

- Development of the ICTV: a historical perspective
  Mike Adams
- Maintaining and updating the ICTV taxonomy
  Andrew King
- ICTV resources: the MSL, the ICTV Database, the 9th Report
  Elliot Lefkowitz
- The challenges ahead
  Andrew Davison

DEVELOPMENT OF THE ICTV: A HISTORICAL PERSPECTIVE

1966 ICTV 2016

Mike Adams

The beginnings

- The International Committee on Nomenclature of Viruses (ICNV) was established by the International Association of Microbiological Societies during its 9th Congress in Moscow, July 1966
- Wherever practicable, each country was represented by a member nominated by the national microbiological society (43/46 at the first meetings)
- An elected Executive Committee (EC) proposed rules and created subcommittees to propose genera and families for the major virus groups (insect, vertebrate, plant and bacteria). Specialist Study Groups (SGs) were created within each subcommittee.
- The aim was to produce an international and universal classification system

Planned virus taxonomy ‘life-cycle’

- SGs discuss and prepare proposals
- EC discusses proposals and negotiates changes with the SGs
- Proposals finalized at an EC meeting held in association with the triennial International Congress of Virology (ICV)
- Proposals presented to the entire ICTV at a special session of the ICV
- EC prepares and publishes a Report of the latest taxonomy, incorporating the changes

Organisational development

- ICTV is now a standing committee of the Virology Division of IUMS
- ICTV consists of the EC, national representatives, SG chairs and life members (~150 in total)
- Subcommittee remits have expanded and changed
- The EC meets annually; wherever possible, proposals move to acceptance in 6-9 months
- Ratification votes are conducted by email; plenary sessions at ICV are for reports and elections to the EC
- Annual changes to taxonomy are incorporated into the website and announced in Arch Virol (Virology Division News)

The first ICNV votes (26 July 1966)

- 1. The committee considers that an international nomenclature for viruses is desirable (carried unanimously)
- 2. An effort should be made towards a latinized binomial nomenclature (carried with 2 against)
- 3. If and when latinized binomials are introduced, the existing names should be retained whenever feasible (carried with 1 against)

Some of the first ICNV rules (July 1966)

- 6. The law of priority shall not be observed
- 7. New sigla shall not be introduced
- 8. No person's name shall be used
- 10. For pragmatic purposes, the species is considered to be collections of viruses with like characters
- Note 2 - Diacritic signs are not used in names or in specific epithets in virology. In names or epithets derived from words with such diacritic signs, the signs must be suppressed and the letters transcribed as follows:

  1 – ä, ö and ü become respectively ae, oe, ue (German)

  2 – é, è and ê become e (French)

  3 – ø (ö), aê (ä) and å become oe, ae, and aa, respectively (Scandinavian)

Virus taxonomy in practice

- Initial work classified viruses (tobacco mosaic virus, rabies virus, etc.) into genera and a few families, based on particle morphology, pathology, transmission and serological properties
- There was little agreement about the definition of a species and uncertainties about the appropriate taxonomic level for many virus groups; this probably explains why latinized binomials were not adopted
- Species were only universally adopted (and made the subject of voting) from 1999, often using the existing virus names

50 years of effort...

First Report (81 pages, published in 1971)

- 2 families
- 43 groups or genera
- 290 viruses listed as members

Latest taxonomy (ratified February 2015)

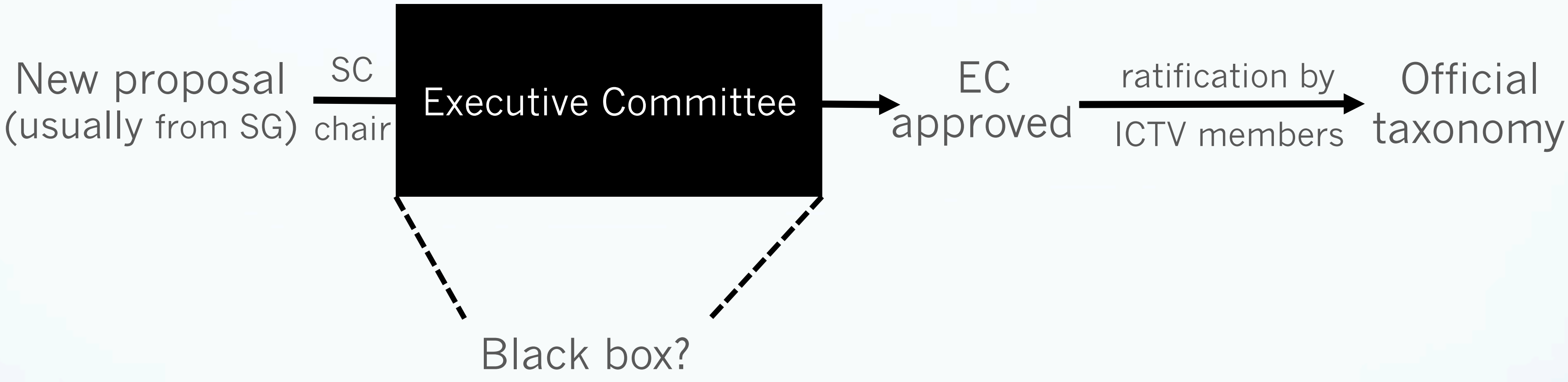
- 7 orders
- 104 families
- 23 subfamilies
- 505 genera
- 3186 species

MAINTAINING AND UPDATING THE ICTV TAXONOMY

1966 2016

Andrew King

Decision-making: approving taxonomic proposals (TPs)


New proposal
(usually from SG)

SC
chair

Executive Committee

EC
approved

ratification by
ICTV members

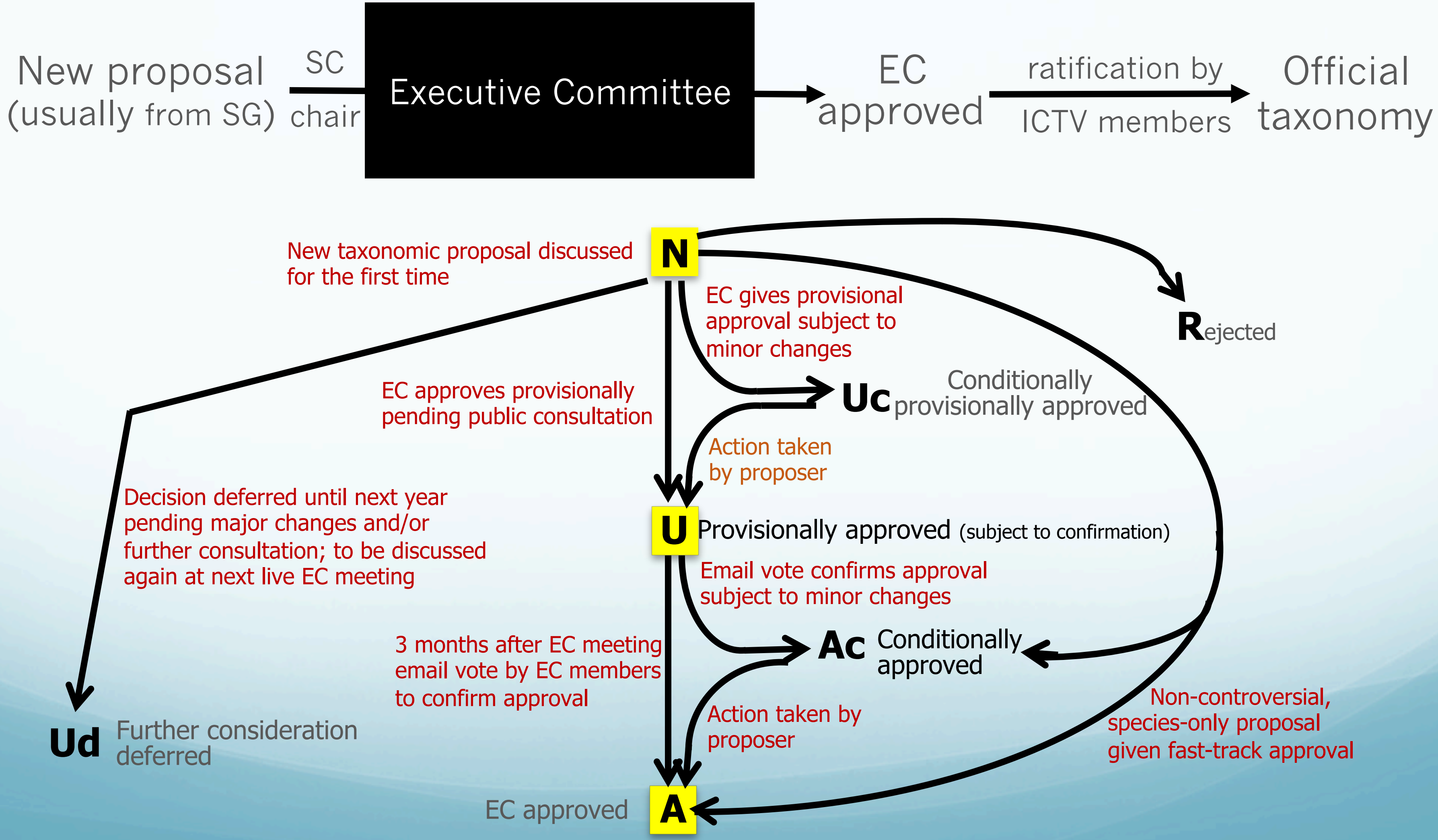
Official
taxonomy

Black box?

But taxonomic proposals are

- Posted on ICTV's discussion pages, where they are open to public comment
- Considered by the EC on a minimum of two occasions, at least 3 months apart, so that any online comments can be taken into account

Decision-making: the black box opened



New proposal (usually from SG)

SC chair

Executive Committee

EC approved

ratification by ICTV members

Official taxonomy

N New taxonomic proposal discussed for the first time

EC gives provisional approval subject to minor changes

Uc Conditionally provisionally approved

Action taken by proposer

EC approves provisionally pending public consultation

U Provisionally approved (subject to confirmation)

Email vote confirms approval subject to minor changes

Ac Conditionally approved

Action taken by proposer

3 months after EC meeting email vote by EC members to confirm approval

A EC approved

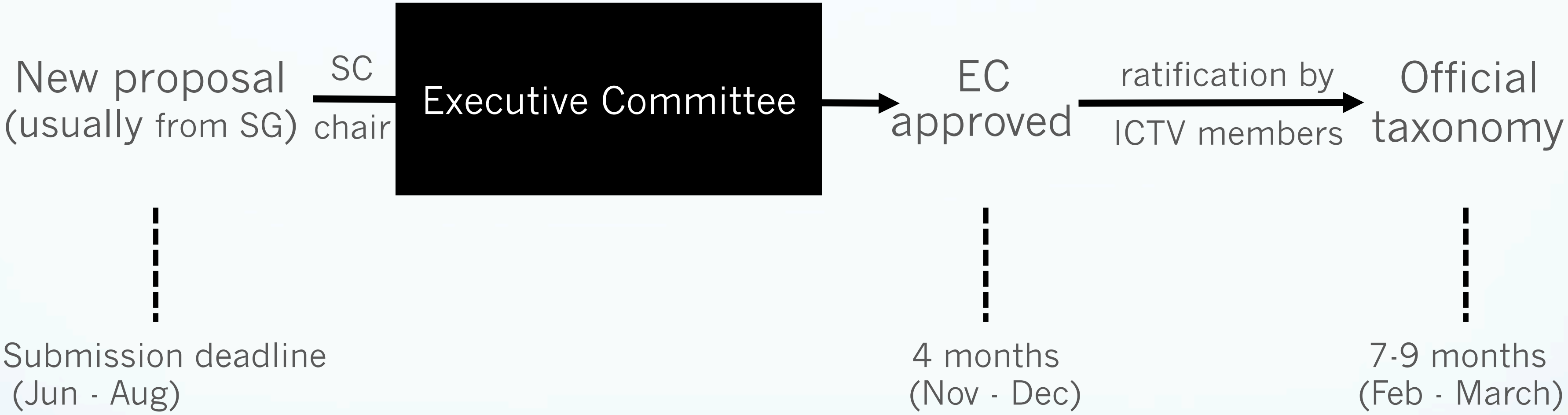
Rejected

Non-controversial, species-only proposal given fast-track approval

Decision deferred until next year pending major changes and/or further consultation; to be discussed again at next live EC meeting

Ud Further consideration deferred

Decision-making: time course* from submission to ratification



New proposal (usually from SG)

SC chair

Executive Committee

EC approved

ratification by ICTV members

Official taxonomy

Submission deadline (Jun - Aug)

4 months (Nov - Dec)

7-9 months (Feb - March)

*In 2015, the EC deferred a decision on only 4 of 131 new proposals

Mechanics: the TP template

taxonomic actions

Existing taxonomy

New taxonomy

These are the 18 taxonomic actions provided for in the template

- Create and name species
  - (Each with designated type virus)
- Create genus
  - Name genus
  - Designate type species of genus
- Create subfamily
  - Name subfamily
  - Assign genera to subfamily
- Create family
  - Name family
  - Assign subfamilies to family
  - Assign genera to family
- Create order
  - Name order
  - Assign families to order
  - Assign unassigned genera to order
- Remove (abolish) taxa
- Reassign taxa
- Rename taxa

Mechanics: the Secretary steps in

Existing taxonomy → taxonomic actions → New taxonomy

These are the 18 taxonomic actions provided for in the template

- Create and name species
  - (Each with designated type virus)
- Create genus
  - Name genus
  - Designate type species of genus
- Create subfamily
  - Name subfamily
  - Assign genera to subfamily
- Create family
  - Name family
  - Assign subfamilies to family
  - Assign genera to family
- Create order
  - Name order
  - Assign families to order
  - Assign unassigned genera to order
- Remove (abolish) taxa
- Reassign taxa
- Rename taxa

→

The Proposals Secretary

1. Codes the proposal document
2. Codes each action
3. Posts the document online
4. Extracts taxo-actions into an Excel spreadsheet
5. Curates the proposal docs
6. Keeps all the above updated

Mechanics: the coding convention

filename:2015.009a-gM.N.v2.Sunviridae

2015	Year initially submitted
.009	9th proposal submitted to this subcommittee
a-g	Taxonomic actions proposed

2015.009**a**M	create species
2015.009**b**M	create genus
2015.009**c**M	name genus
2015.009**d**M	assign type species
2015.009**e**M	create family
2015.009**f**M	name family
2015.009**g**M	assign genus to family

M	Subcommittee (animal ds and RNA– viruses)
.N	Status code (posted in ‘New’ folder)
.v2	Version 2 of this proposal
.Sunviridae	Description (proposed family *Sunviridae*)

Mechanics: the master TP spreadsheet

Proposal **(filename)**	**Short Title**	**Authors** (corresponding author's email in brackets)	**action code**	**proposed taxonomic action**	**post-EC47 status**	**pre-E-vote status**	**post-E-vote status**	**pre-ratification**	**SC**
2015.054a-dB.U.v3. Septima3virus	create genus *Septima3virus* including 5 new species within the family *Siphoviridae*, order *Caudovirales*	Kropinski AM (Phage.Canada@gmail.com)	2015.054aB	create 5 species, (*Pseudomonas virus 73, Pseudomonas virus, Kakheti25, Pseudomonas virus Ab26* and *Burkholderia virus KL1*), in genus *Septima3virus* (new) in the family *Siphoviridae*	U	U	A	A	B
			2015.054bB	create a new genus in the family *Siphoviridae*, order *Caudovirales*	U	U	A	A	B
			2015.054cB	name the genus *Septima3virus*	U	U	A	A	B
			2015.054dB	designate *Pseudomonas virus 73* as type species of the new genus	U	U	A	A	B
2015.001aD.A.v 2. Betabaculovirus_sp	create species *Agrotis segetum granulovirus* in the genus *Betabaculovirus*, family *Baculoviridae*	Alletti GG, Wennmann JT, Carstens EB, Jehle JA (Johannes.Jehle@jki.bund.de)	2015.001aD	create species *Agrotis segetum granulovirus* in the genus *Betabaculovirus*, family *Baculoviridae*	A	A	A	A	D
2015.002a,bD.A.v2. Alphabaculovirus_sp,ren	in genus *Alphabaculovirus* (family *Baculoviridae*) create species *Agrotis segetum nucleopolyhedrovirus B* and change the name of species *Agrotis segetum nucleopolyhedrovirus* to *Agrotis segetum nucleopolyhedrovirus A*	Wennmann JT, Alletti GG, Carstens EB, Jehle JA (Johannes.Jehle@jki.bund.de)	2015.002aD	create species *Agrotis segetum nucleopolyhedrovirus B* in genus *Alphabaculovirus,* family Baculoviridae	Ac	Ac	Ac	A	D
			2015.002bD	in genus *Alphabaculovirus* change the name of species *Agrotis segetum nucleopolyhedrovirus* to *Agrotis segetum nucleopolyhedrovirus A*	Ac	Ac	Ac	A	D
2015.003a,bD.A.v1. Kappatorquevirus_sp,ren	In genus *Kappatorquevirus* (family *Anelloviridae*), create species *Torque teno sus virus k2b* and change the name of species *Torque teno sus virus k2* to *Torque teno sus virus k2a*	Kekarainen T (tuija.kekarainen@irta.cat)	2015.003aD	in genus *Kappatorquevirus* (family *Anelloviridae*), create species *Torque teno sus virus k2b*	A	A	A	A	D
			2015.003bD	in genus *Kappatorquevirus* (family *Anelloviridae*), change the name of species *Torque teno sus virus k2* to *Torque teno sus virus k2a*	A	A	A	A	D
2015.004a-eD.U.v2. Toursvirus	create genus *Toursvirus* in the family *Ascoviridae* and move species *Diadromus pulchellus ascovirus 4a* from genus *Ascovirus* to genus *Toursvirus*	Asgari S (s.asgari@uq.edu.au), Bideshi D, Bigot Y, Cheng X-W, Federici BA	2015.004aD	create a new genus in family *Ascoviridae*	Uc	U	A	A	D
			2015.004bD	name the genus *Toursvirus*	Uc	U	A	A	D
			2015.004cD	designate *Diadromus pulchellus ascovirus 4a* as type species of the new genus	Uc	U	A	A	D
			2015.004dD	remove species *Diadromus pulchellus ascovirus 4a* from genus *Ascovirus*	Uc	U	A	A	D
			2015.004eD	re-assign the above species to genus *Toursvirus*	Uc	U	A	A	D
2015.005aD.A.v1. Aviadenovirus_4sp	create 4 species in the genus *Aviadenovirus*, family *Adenoviridae*	Győző Kaján G (kajan.gyozo@agrar.mta.hu), Podgorski I, Harrach B	2015.005aD	create 4 species (*Duck aviadenovirus B, Pigeon aviadenovirus A, Turkey aviadenovirus C* and *Turkey aviadenovirus D*) in the genus *Aviadenovirus*, family *Adenoviridae*	A	A	A	A	D
2015.006aD.N.v1. Simplexvirus-sp	create a species in genus *Simplexvirus*, subfamily *Alphaherpesvirinae*, family *Herpesviridae*	Davison A (andrew.davison@glasgow.ac.uk), Eberle R	2015.006aD	create a species (name: *Panine herpesvirus 3* OR *Panine alphaherpesvirus 3*, depending on approval of proposal 2015.010aD.N.v1.Herpesvirales_spren) in genus *Simplexvirus*, subfamily *Alphaherpesvirinae*, family *Herpesviridae*	Ac	Ac	Ac	Ac	D

Keeping taxonomy updated: a wish list

Taxonomic proposals: TP template is complex, prescriptive and demanding

How to
- make proposing less of a deterrent?
- give more credit to authors of proposals?
- encourage SGs in other ways to submit TPs?

Document handling: manual, error prone, labour intensive
document transmission almost entirely by email
proposal docs, TP spreadsheet, NOT shared

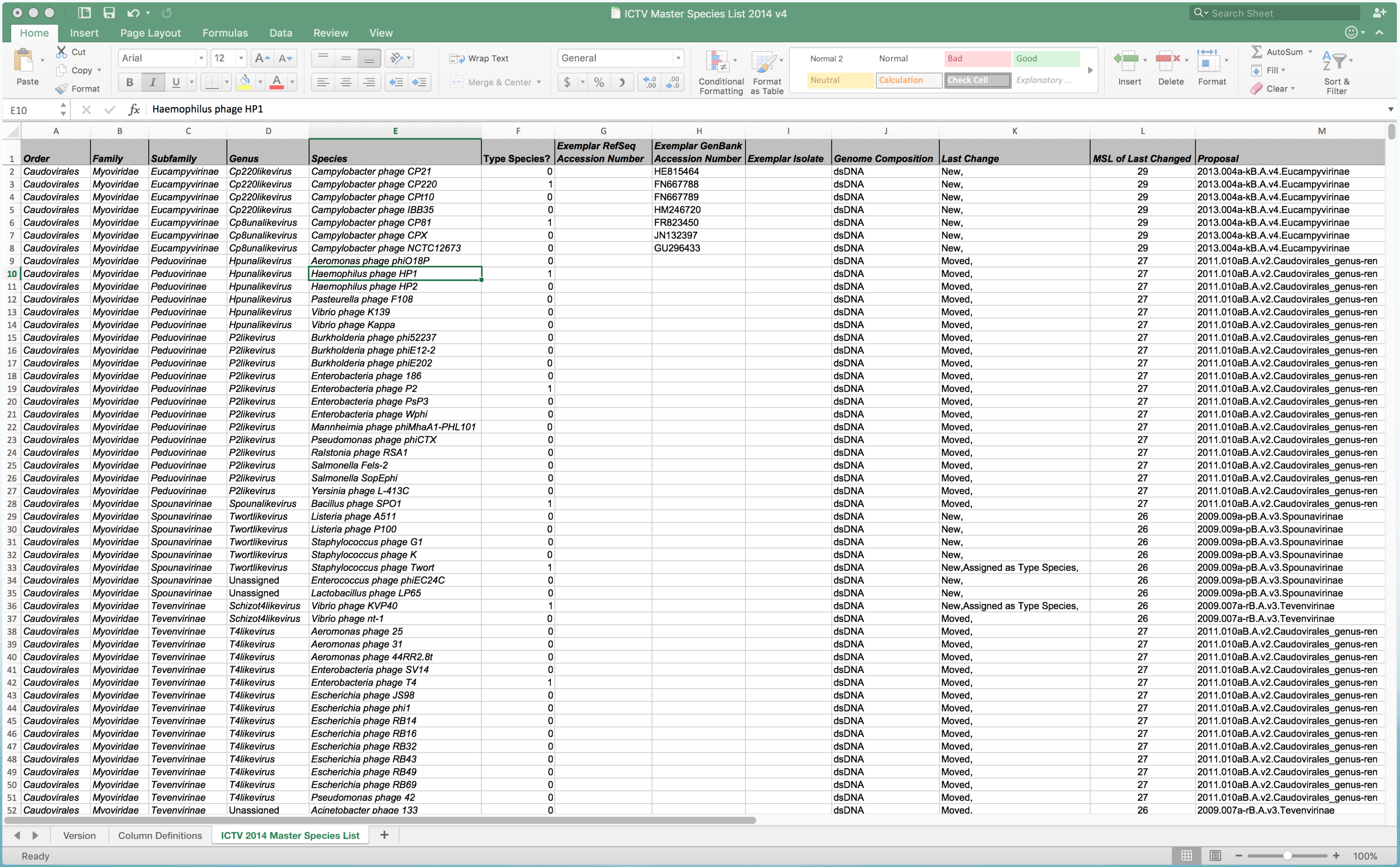
Need to be able to
- submit and update proposals online
- extract, update, share information automatically
- share proposal documents
- update MSL and ICTV Report automatically

ICTV RESOURCES: THE MSL, THE ICTV DATABASE AND THE 9TH REPORT

1966 ICTV 2016

Elliot Lefkowitz

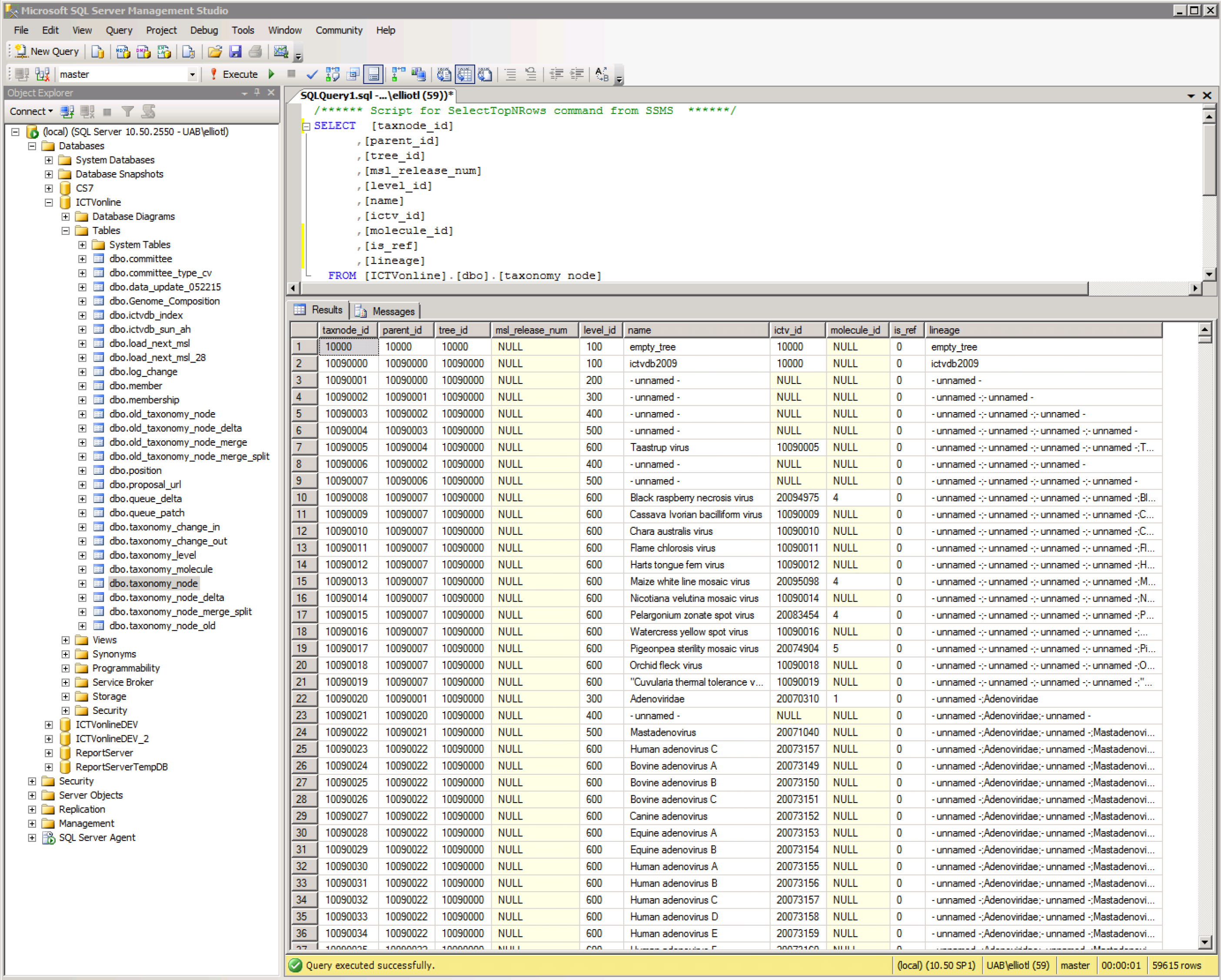
The Master Species List (MSL)



Haemophilus phage HP1

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	**Order**	**Family**	**Subfamily**	**Genus**	**Species**	**Type Species?**	**Exemplar RefSeq Accession Number**	**Exemplar GenBank Accession Number**	**Exemplar Isolate**	**Genome Composition**	**Last Change**	**MSL of Last Changed**	**Proposal**
2	Caudovirales	Myoviridae	Eucampyvirinae	Cp220likevirus	Campylobacter phage CP21	0		HE815464		dsDNA	New,	29	2013.004a-kB.A.v4.Eucampyvirinae
3	Caudovirales	Myoviridae	Eucampyvirinae	Cp220likevirus	Campylobacter phage CP220	1		FN667788		dsDNA	New,	29	2013.004a-kB.A.v4.Eucampyvirinae
4	Caudovirales	Myoviridae	Eucampyvirinae	Cp220likevirus	Campylobacter phage CPt10	0		FN667789		dsDNA	New,	29	2013.004a-kB.A.v4.Eucampyvirinae
5	Caudovirales	Myoviridae	Eucampyvirinae	Cp220likevirus	Campylobacter phage IBB35	0		HM246720		dsDNA	New,	29	2013.004a-kB.A.v4.Eucampyvirinae
6	Caudovirales	Myoviridae	Eucampyvirinae	Cp8unalikevirus	Campylobacter phage CP81	1		FR823450		dsDNA	New,	29	2013.004a-kB.A.v4.Eucampyvirinae
7	Caudovirales	Myoviridae	Eucampyvirinae	Cp8unalikevirus	Campylobacter phage CPX	0		JN132397		dsDNA	New,	29	2013.004a-kB.A.v4.Eucampyvirinae
8	Caudovirales	Myoviridae	Eucampyvirinae	Cp8unalikevirus	Campylobacter phage NCTC12673	0		GU296433		dsDNA	New,	29	2013.004a-kB.A.v4.Eucampyvirinae
9	Caudovirales	Myoviridae	Peduovirinae	Hpunalikevirus	Aeromonas phage phiO18P	0				dsDNA	Moved,	27	2011.010aB.A.v2.Caudovirales_genus-ren
10	Caudovirales	Myoviridae	Peduovirinae	Hpunalikevirus	Haemophilus phage HP1	1				dsDNA	Moved,	27	2011.010aB.A.v2.Caudovirales_genus-ren
11	Caudovirales	Myoviridae	Peduovirinae	Hpunalikevirus	Haemophilus phage HP2	0				dsDNA	Moved,	27	2011.010aB.A.v2.Caudovirales_genus-ren
12	Caudovirales	Myoviridae	Peduovirinae	Hpunalikevirus	Pasteurella phage F108	0				dsDNA	Moved,	27	2011.010aB.A.v2.Caudovirales_genus-ren
13	Caudovirales	Myoviridae	Peduovirinae	Hpunalikevirus	Vibrio phage K139	0				dsDNA	Moved,	27	2011.010aB.A.v2.Caudovirales_genus-ren
14	Caudovirales	Myoviridae	Peduovirinae	Hpunalikevirus	Vibrio phage Kappa	0				dsDNA	Moved,	27	2011.010aB.A.v2.Caudovirales_genus-ren
15	Caudovirales	Myoviridae	Peduovirinae	P2likevirus	Burkholderia phage phi52237	0				dsDNA	Moved,	27	2011.010aB.A.v2.Caudovirales_genus-ren
16	Caudovirales	Myoviridae	Peduovirinae	P2likevirus	Burkholderia phage phiE12-2	0				dsDNA	Moved,	27	2011.010aB.A.v2.Caudovirales_genus-ren
17	Caudovirales	Myoviridae	Peduovirinae	P2likevirus	Burkholderia phage phiE202	0				dsDNA	Moved,	27	2011.010aB.A.v2.Caudovirales_genus-ren
18	Caudovirales	Myoviridae	Peduovirinae	P2likevirus	Enterobacteria phage 186	0				dsDNA	Moved,	27	2011.010aB.A.v2.Caudovirales_genus-ren
19	Caudovirales	Myoviridae	Peduovirinae	P2likevirus	Enterobacteria phage P2	1				dsDNA	Moved,	27	2011.010aB.A.v2.Caudovirales_genus-ren
20	Caudovirales	Myoviridae	Peduovirinae	P2likevirus	Enterobacteria phage PsP3	0				dsDNA	Moved,	27	2011.010aB.A.v2.Caudovirales_genus-ren
21	Caudovirales	Myoviridae	Peduovirinae	P2likevirus	Enterobacteria phage Wphi	0				dsDNA	Moved,	27	2011.010aB.A.v2.Caudovirales_genus-ren
22	Caudovirales	Myoviridae	Peduovirinae	P2likevirus	Mannheimia phage phiMhaA1-PHL101	0				dsDNA	Moved,	27	2011.010aB.A.v2.Caudovirales_genus-ren
23	Caudovirales	Myoviridae	Peduovirinae	P2likevirus	Pseudomonas phage phiCTX	0				dsDNA	Moved,	27	2011.010aB.A.v2.Caudovirales_genus-ren
24	Caudovirales	Myoviridae	Peduovirinae	P2likevirus	Ralstonia phage RSA1	0				dsDNA	Moved,	27	2011.010aB.A.v2.Caudovirales_genus-ren
25	Caudovirales	Myoviridae	Peduovirinae	P2likevirus	Salmonella Fels-2	0				dsDNA	Moved,	27	2011.010aB.A.v2.Caudovirales_genus-ren
26	Caudovirales	Myoviridae	Peduovirinae	P2likevirus	Salmonella SopEphi	0				dsDNA	Moved,	27	2011.010aB.A.v2.Caudovirales_genus-ren
27	Caudovirales	Myoviridae	Peduovirinae	P2likevirus	Yersinia phage L-413C	0				dsDNA	Moved,	27	2011.010aB.A.v2.Caudovirales_genus-ren
28	Caudovirales	Myoviridae	Spounavirinae	Spounalikevirus	Bacillus phage SPO1	1				dsDNA	Moved,	27	2011.010aB.A.v2.Caudovirales_genus-ren
29	Caudovirales	Myoviridae	Spounavirinae	Twortlikevirus	Listeria phage A511	0				dsDNA	New,	26	2009.009a-pB.A.v3.Spounavirinae
30	Caudovirales	Myoviridae	Spounavirinae	Twortlikevirus	Listeria phage P100	0				dsDNA	New,	26	2009.009a-pB.A.v3.Spounavirinae
31	Caudovirales	Myoviridae	Spounavirinae	Twortlikevirus	Staphylococcus phage G1	0				dsDNA	New,	26	2009.009a-pB.A.v3.Spounavirinae
32	Caudovirales	Myoviridae	Spounavirinae	Twortlikevirus	Staphylococcus phage K	0				dsDNA	New,	26	2009.009a-pB.A.v3.Spounavirinae
33	Caudovirales	Myoviridae	Spounavirinae	Twortlikevirus	Staphylococcus phage Twort	1				dsDNA	New,Assigned as Type Species,	26	2009.009a-pB.A.v3.Spounavirinae
34	Caudovirales	Myoviridae	Spounavirinae	Unassigned	Enterococcus phage phiEC24C	0				dsDNA	New,	26	2009.009a-pB.A.v3.Spounavirinae
35	Caudovirales	Myoviridae	Spounavirinae	Unassigned	Lactobacillus phage LP65	0				dsDNA	New,	26	2009.009a-pB.A.v3.Spounavirinae
36	Caudovirales	Myoviridae	Tevenvirinae	Schizot4likevirus	Vibrio phage KVP40	1				dsDNA	New,Assigned as Type Species,	26	2009.007a-rB.A.v3.Tevenvirinae
37	Caudovirales	Myoviridae	Tevenvirinae	Schizot4likevirus	Vibrio phage nt-1	0				dsDNA	Moved,	26	2009.007a-rB.A.v3.Tevenvirinae
38	Caudovirales	Myoviridae	Tevenvirinae	T4likevirus	Aeromonas phage 25	0				dsDNA	Moved,	27	2011.010aB.A.v2.Caudovirales_genus-ren
39	Caudovirales	Myoviridae	Tevenvirinae	T4likevirus	Aeromonas phage 31	0				dsDNA	Moved,	27	2011.010aB.A.v2.Caudovirales_genus-ren
40	Caudovirales	Myoviridae	Tevenvirinae	T4likevirus	Aeromonas phage 44RR2.8t	0				dsDNA	Moved,	27	2011.010aB.A.v2.Caudovirales_genus-ren
41	Caudovirales	Myoviridae	Tevenvirinae	T4likevirus	Enterobacteria phage SV14	0				dsDNA	Moved,	27	2011.010aB.A.v2.Caudovirales_genus-ren
42	Caudovirales	Myoviridae	Tevenvirinae	T4likevirus	Enterobacteria phage T4	1				dsDNA	Moved,	27	2011.010aB.A.v2.Caudovirales_genus-ren
43	Caudovirales	Myoviridae	Tevenvirinae	T4likevirus	Escherichia phage JS98	0				dsDNA	Moved,	27	2011.010aB.A.v2.Caudovirales_genus-ren
44	Caudovirales	Myoviridae	Tevenvirinae	T4likevirus	Escherichia phage phi1	0				dsDNA	Moved,	27	2011.010aB.A.v2.Caudovirales_genus-ren
45	Caudovirales	Myoviridae	Tevenvirinae	T4likevirus	Escherichia phage RB14	0				dsDNA	Moved,	27	2011.010aB.A.v2.Caudovirales_genus-ren
46	Caudovirales	Myoviridae	Tevenvirinae	T4likevirus	Escherichia phage RB16	0				dsDNA	Moved,	27	2011.010aB.A.v2.Caudovirales_genus-ren
47	Caudovirales	Myoviridae	Tevenvirinae	T4likevirus	Escherichia phage RB32	0				dsDNA	Moved,	27	2011.010aB.A.v2.Caudovirales_genus-ren
48	Caudovirales	Myoviridae	Tevenvirinae	T4likevirus	Escherichia phage RB43	0				dsDNA	Moved,	27	2011.010aB.A.v2.Caudovirales_genus-ren
49	Caudovirales	Myoviridae	Tevenvirinae	T4likevirus	Escherichia phage RB49	0				dsDNA	Moved,	27	2011.010aB.A.v2.Caudovirales_genus-ren
50	Caudovirales	Myoviridae	Tevenvirinae	T4likevirus	Escherichia phage RB69	0				dsDNA	Moved,	27	2011.010aB.A.v2.Caudovirales_genus-ren
51	Caudovirales	Myoviridae	Tevenvirinae	T4likevirus	Pseudomonas phage 42	0				dsDNA	Moved,	27	2011.010aB.A.v2.Caudovirales_genus-ren
52	Caudovirales	Myoviridae	Tevenvirinae	Unassigned	Acinetobacter phage 133	0				dsDNA	Moved,	26	2009.007a-rB.A.v3.Tevenvirinae

The ICTV Database

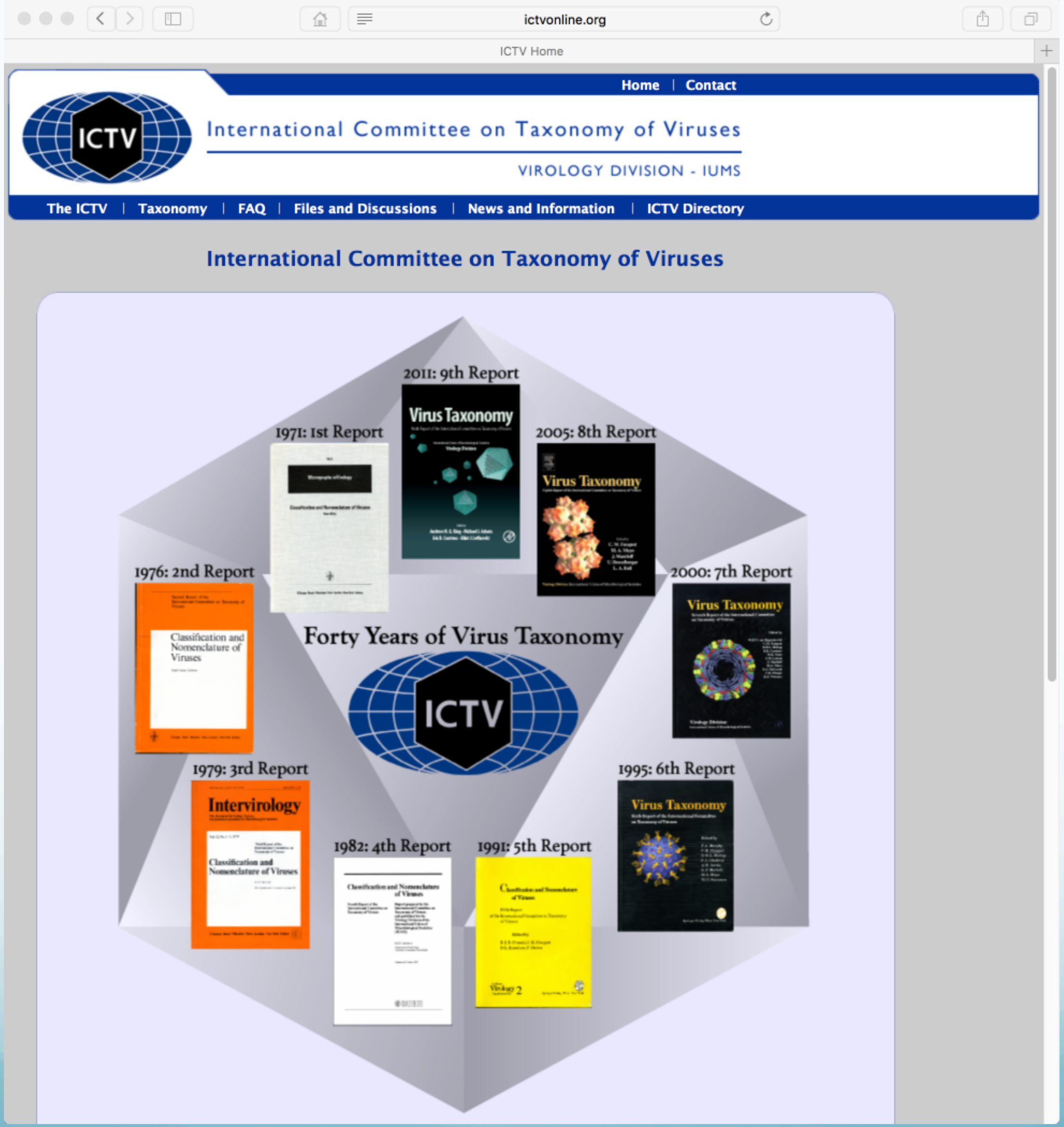


```
/****** Script for SelectTopNRows command from SSMS  ******/
SELECT  [taxnode_id]
      ,[parent_id]
      ,[tree_id]
      ,[msl_release_num]
      ,[level_id]
      ,[name]
      ,[ictv_id]
      ,[molecule_id]
      ,[is_ref]
      ,[lineage]
  FROM [ICTVonline].[dbo].[taxonomy_node]
```

	taxnode_id	parent_id	tree_id	msl_release_num	level_id	name	ictv_id	molecule_id	is_ref	lineage
1	10000	10000	10000	NULL	100	empty_tree	10000	NULL	0	empty_tree
2	10090000	10090000	10090000	NULL	100	ictvdb2009	10000	NULL	0	ictvdb2009
3	10090001	10090000	10090000	NULL	200	- unnamed -	NULL	NULL	0	- unnamed -
4	10090002	10090001	10090000	NULL	300	- unnamed -	NULL	NULL	0	- unnamed -;- unnamed -
5	10090003	10090002	10090000	NULL	400	- unnamed -	NULL	NULL	0	- unnamed -;- unnamed -;- unnamed -
6	10090004	10090003	10090000	NULL	500	- unnamed -	NULL	NULL	0	- unnamed -;- unnamed -;- unnamed -;- unnamed -
7	10090005	10090004	10090000	NULL	600	Taastrup virus	10090005	NULL	0	- unnamed -;- unnamed -;- unnamed -;- unnamed -;T...
8	10090006	10090002	10090000	NULL	400	- unnamed -	NULL	NULL	0	- unnamed -;- unnamed -;- unnamed -
9	10090007	10090006	10090000	NULL	500	- unnamed -	NULL	NULL	0	- unnamed -;- unnamed -;- unnamed -;- unnamed -
10	10090008	10090007	10090000	NULL	600	Black raspberry necrosis virus	20094975	4	0	- unnamed -;- unnamed -;- unnamed -;- unnamed -;Bl...
11	10090009	10090007	10090000	NULL	600	Cassava Ivorian bacilliform virus	10090009	NULL	0	- unnamed -;- unnamed -;- unnamed -;- unnamed -;C...
12	10090010	10090007	10090000	NULL	600	Chara australis virus	10090010	NULL	0	- unnamed -;- unnamed -;- unnamed -;- unnamed -;C...
13	10090011	10090007	10090000	NULL	600	Flame chlorosis virus	10090011	NULL	0	- unnamed -;- unnamed -;- unnamed -;- unnamed -;Fl...
14	10090012	10090007	10090000	NULL	600	Harts tongue fern virus	10090012	NULL	0	- unnamed -;- unnamed -;- unnamed -;- unnamed -;H...
15	10090013	10090007	10090000	NULL	600	Maize white line mosaic virus	20095098	4	0	- unnamed -;- unnamed -;- unnamed -;- unnamed -;M...
16	10090014	10090007	10090000	NULL	600	Nicotiana velutina mosaic virus	10090014	NULL	0	- unnamed -;- unnamed -;- unnamed -;- unnamed -;N...
17	10090015	10090007	10090000	NULL	600	Pelargonium zonate spot virus	20083454	4	0	- unnamed -;- unnamed -;- unnamed -;- unnamed -;P...
18	10090016	10090007	10090000	NULL	600	Watercress yellow spot virus	10090016	NULL	0	- unnamed -;- unnamed -;- unnamed -;- unnamed -;...
19	10090017	10090007	10090000	NULL	600	Pigeonpea sterility mosaic virus	20074904	5	0	- unnamed -;- unnamed -;- unnamed -;- unnamed -;Pi...
20	10090018	10090007	10090000	NULL	600	Orchid fleck virus	10090018	NULL	0	- unnamed -;- unnamed -;- unnamed -;- unnamed -;O...
21	10090019	10090007	10090000	NULL	600	"Cuvularia thermal tolerance v...	10090019	NULL	0	- unnamed -;- unnamed -;- unnamed -;- unnamed -;"...
22	10090020	10090001	10090000	NULL	300	Adenoviridae	20070310	1	0	- unnamed -;Adenoviridae
23	10090021	10090020	10090000	NULL	400	- unnamed -	NULL	NULL	0	- unnamed -;Adenoviridae;- unnamed -
24	10090022	10090021	10090000	NULL	500	Mastadenovirus	20071040	NULL	0	- unnamed -;Adenoviridae;- unnamed -;Mastadenovi...
25	10090023	10090022	10090000	NULL	600	Human adenovirus C	20073157	NULL	0	- unnamed -;Adenoviridae;- unnamed -;Mastadenovi...
26	10090024	10090022	10090000	NULL	600	Bovine adenovirus A	20073149	NULL	0	- unnamed -;Adenoviridae;- unnamed -;Mastadenovi...
27	10090025	10090022	10090000	NULL	600	Bovine adenovirus B	20073150	NULL	0	- unnamed -;Adenoviridae;- unnamed -;Mastadenovi...
28	10090026	10090022	10090000	NULL	600	Bovine adenovirus C	20073151	NULL	0	- unnamed -;Adenoviridae;- unnamed -;Mastadenovi...
29	10090027	10090022	10090000	NULL	600	Canine adenovirus	20073152	NULL	0	- unnamed -;Adenoviridae;- unnamed -;Mastadenovi...
30	10090028	10090022	10090000	NULL	600	Equine adenovirus A	20073153	NULL	0	- unnamed -;Adenoviridae;- unnamed -;Mastadenovi...
31	10090029	10090022	10090000	NULL	600	Equine adenovirus B	20073154	NULL	0	- unnamed -;Adenoviridae;- unnamed -;Mastadenovi...
32	10090030	10090022	10090000	NULL	600	Human adenovirus A	20073155	NULL	0	- unnamed -;Adenoviridae;- unnamed -;Mastadenovi...
33	10090031	10090022	10090000	NULL	600	Human adenovirus B	20073156	NULL	0	- unnamed -;Adenoviridae;- unnamed -;Mastadenovi...
34	10090032	10090022	10090000	NULL	600	Human adenovirus C	20073157	NULL	0	- unnamed -;Adenoviridae;- unnamed -;Mastadenovi...
35	10090033	10090022	10090000	NULL	600	Human adenovirus D	20073158	NULL	0	- unnamed -;Adenoviridae;- unnamed -;Mastadenovi...
36	10090034	10090022	10090000	NULL	600	Human adenovirus E	20073159	NULL	0	- unnamed -;Adenoviridae;- unnamed -;Mastadenovi...

```
Query executed successfully.    (local) (10.50 SP1) | UAB\elliotl (59) | master | 00:00:01 | 59615 rows
```

The ICTV website



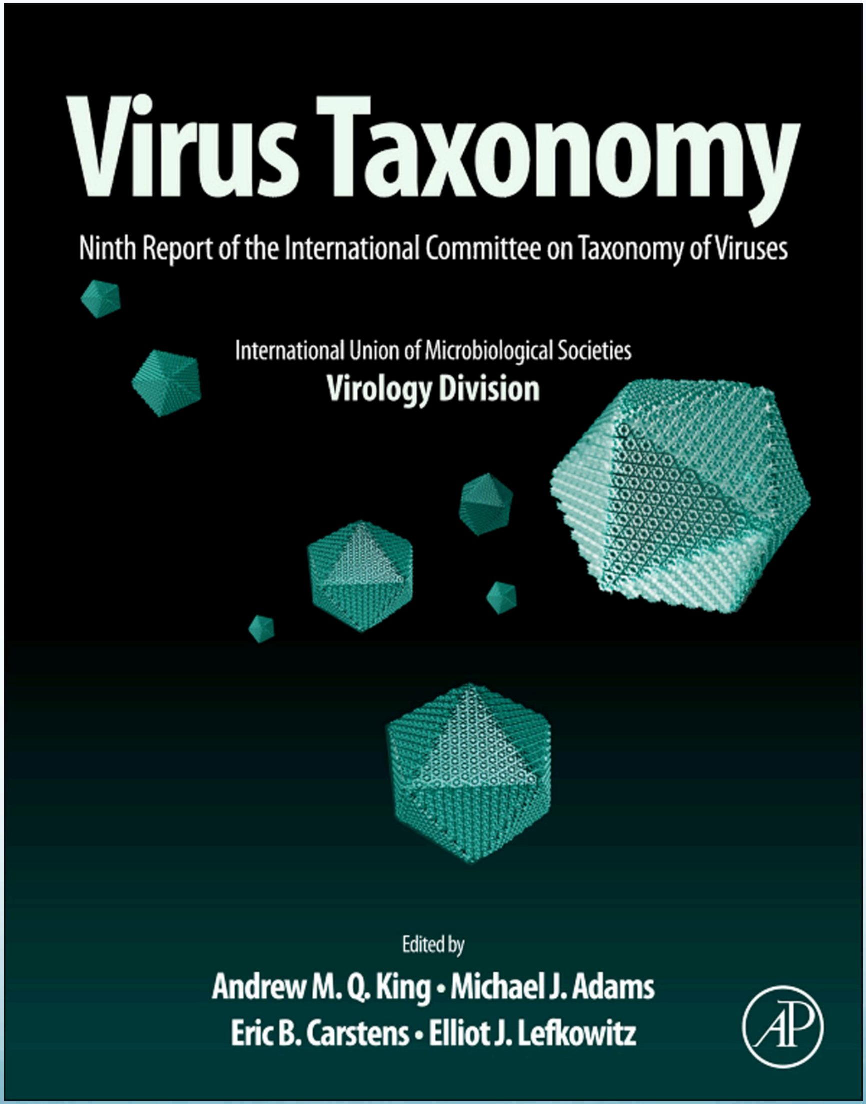
International Committee on Taxonomy of Viruses

VIROLOGY DIVISION - IUMS

International Committee on Taxonomy of Viruses

Forty Years of Virus Taxonomy

The 9th Report



Virus Taxonomy

Ninth Report of the International Committee on Taxonomy of Viruses

International Union of Microbiological Societies
Virology Division

Edited by
Andrew M. Q. King • Michael J. Adams
Eric B. Carstens • Elliot J. Lefkowitz

THE CHALLENGES AHEAD

1966 ICTV 2016

Andrew Davison

The Wellcome Trust grant

Challenges

- Publishing the next Report
- Linking with other databases
- Involving the community in tackling key questions

Objectives

- Updating resources and making them freely available to all
- Establishing clearer methods for classifying viruses
- Engaging with the community to resolve fundamental issues

Plan

- Making the next Report freely available online in a citable form
- Providing resources: links, methods, alignments, TP submission
- Holding focused meetings

The Wellcome Trust grant

Resources

- Oxford post (1.0): Report, alignments, metagenomic data
- Glasgow post (0.5; Richard Orton): links, alignments, methods, TP submission
- Meetings

Oversight

- Grantholders: Peter Simmonds, Andrew Davison, Stuart Siddell
- Management committee: grantholders plus ICTV officers [Sasha Gorbalenya (Vice-President), Mike Adams (Business Secretary), Andrew King (Proposals Secretary), Elliot Lefkowitz (Data Secretary, Editor-in-Chief)]

Particular recognition of the role of SG chairs

REDESIGNING THE ICTV REPORT

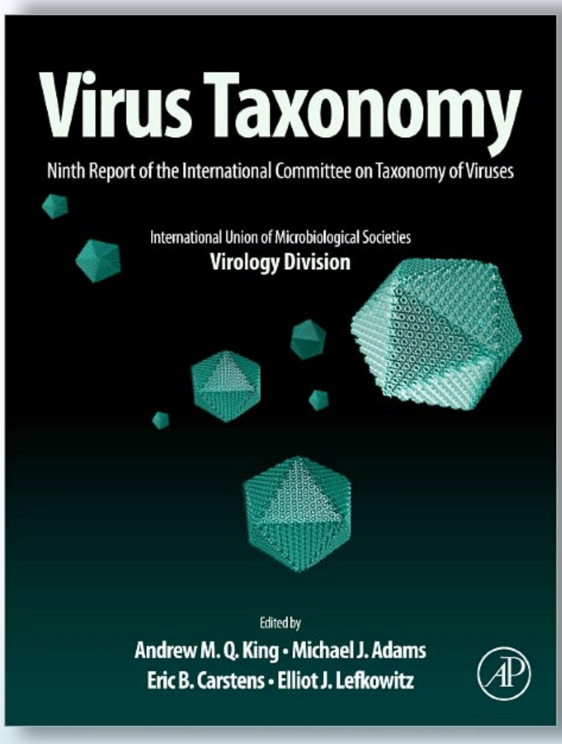
- Limitations of the 9th Report: the need for change
  Stuart Siddell
- The practicalities of publishing online and being cited
  Peter Simmonds
- A prototype of the new Report
  Elliot Lefkowitz
- Implications for Study Groups
  Andrew Davison

LIMITATIONS OF THE 9TH REPORT: THE NEED FOR CHANGE

1966 2016

Stuart Siddell

Limitations of the 9th Report



Virus Taxonomy

Ninth Report of the International Committee on Taxonomy of Viruses

International Union of Microbiological Societies

Virology Division

Edited by

Andrew M. Q. King · Michael J. Adams

Eric B. Carstens · Elliot J. Lefkowitz

Infrequent

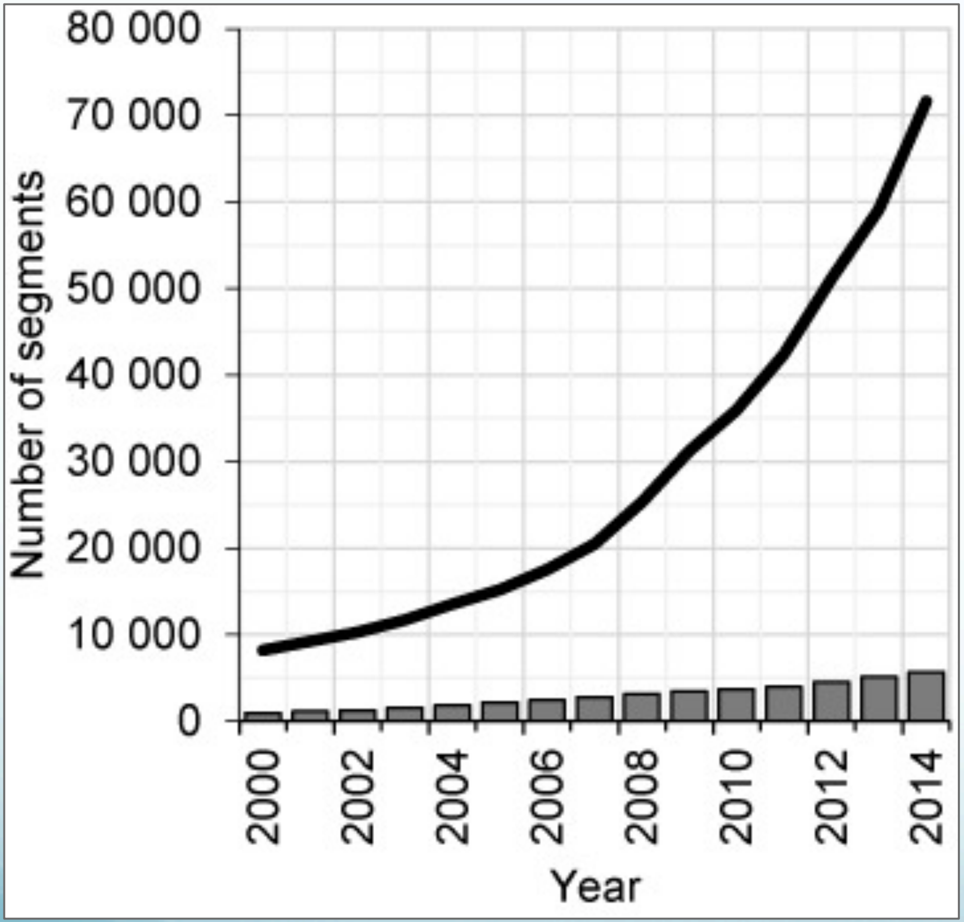
No open access

No links to databases

Citation of authors

£255, 1344 pages, 3.5 kg

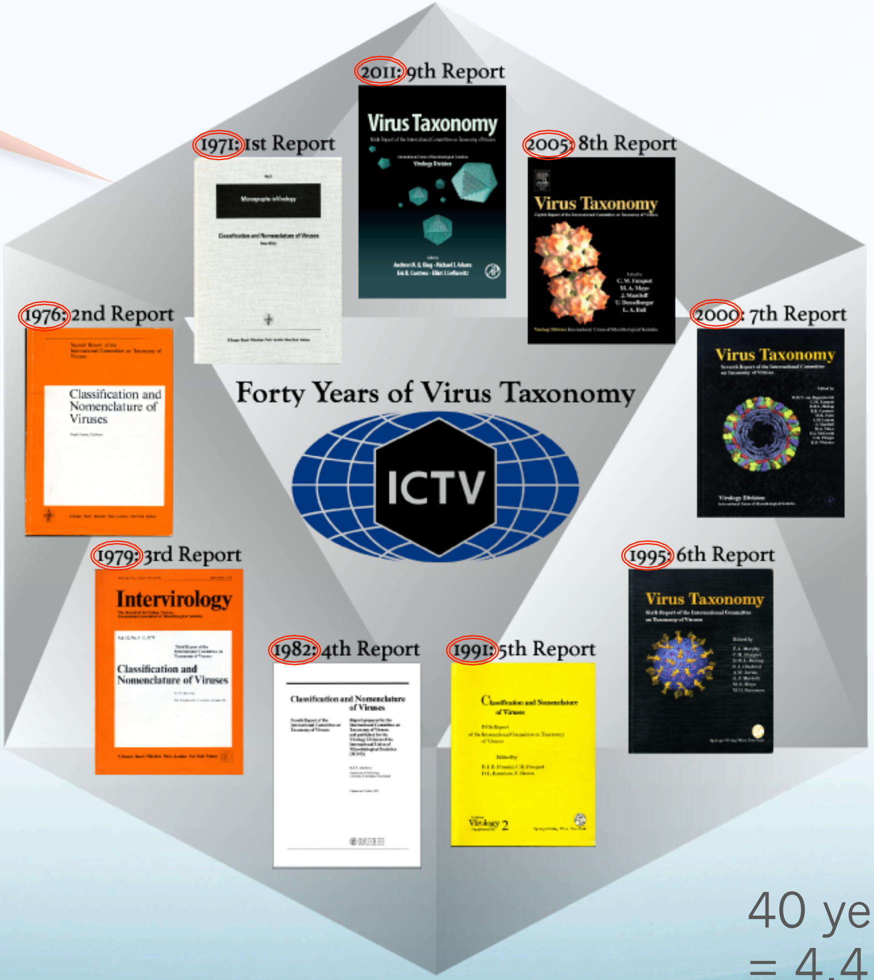
Validated virus and viroid segments



The numbers of validated virus and viroid segments available in INSDC databases are depicted by the black line, and the numbers of RefSeq virus and viroid segments by the gray columns. **Brister et al. (2015) NCBI viral genomes resource. Nucleic Acids Research 43: D571-D577.** The graphic is used to illustrate the pace of virus discovery, and does not mean that the EC considers sequence alone as the basis of virus classification.

How often is the Report published?

Infrequent



40 years, 9 reports
= 4.4 years on average

Why does it take so long?

- Simultaneous updates on 109 chapters for 6 orders, 87 families, 349 genera and 2284 species
- Laborious editing to produce some degree of uniformity with regard to figures, tables and diagrams
- Production and publication

No open access

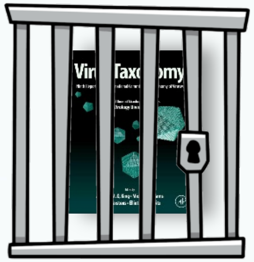
Note that the complete online Report is not for sale. Only individual chapters are available to purchase (at $31.50 each). The complete Report is available as an e-book, but this lacks supplementary material.

How is the Report linked?

ViralZone

INSDC

- GenBank
- EMBL-EBI
- DDBJ



ICTV
Master Species List
(MSL)

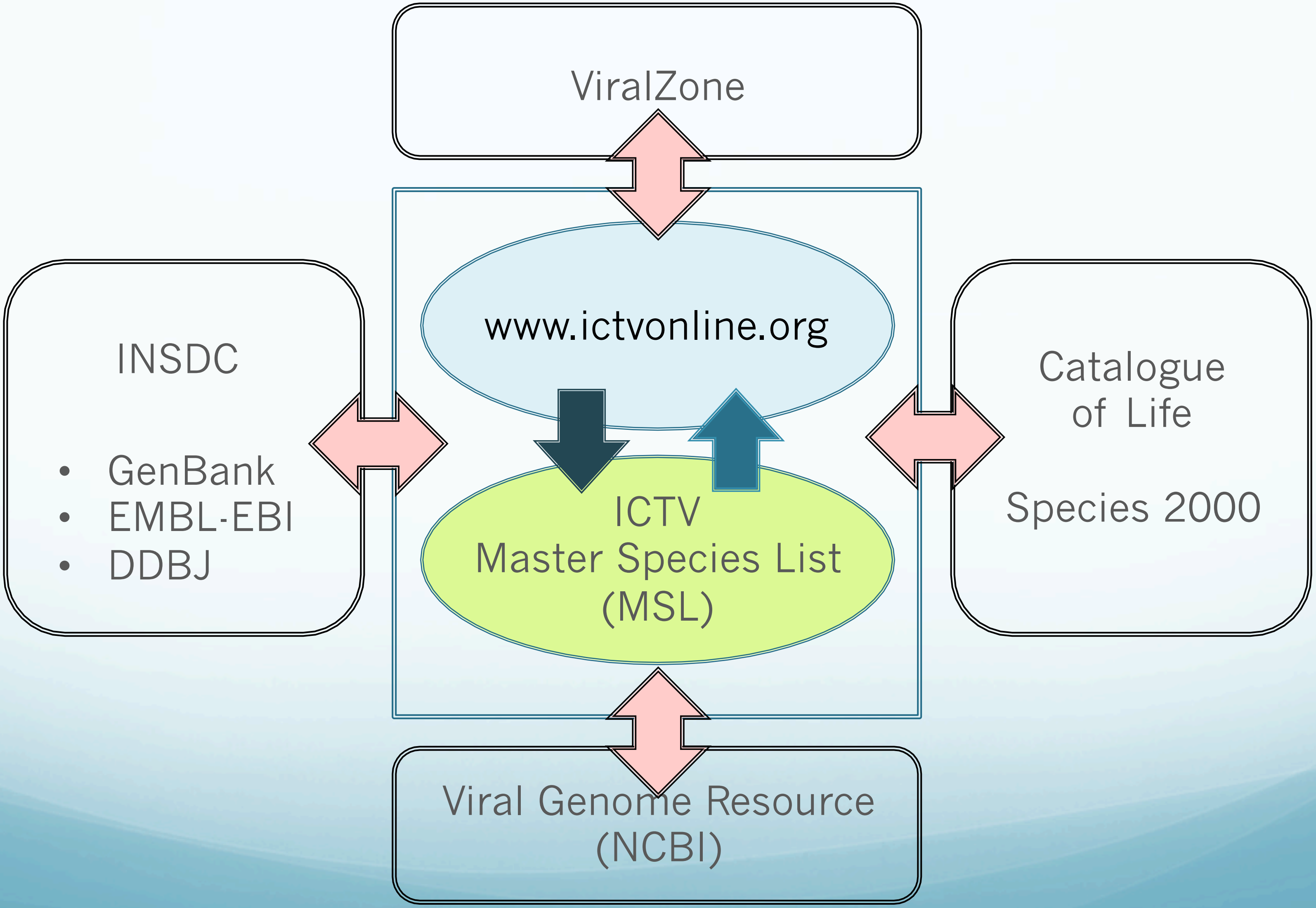
Catalogue
of Life

Species 2000

Viral Genome Resource
(NCBI)

No links to
databases

How will the Report be linked?



ViralZone

www.ictvonline.org

ICTV
Master Species List
(MSL)

INSDC

- GenBank
- EMBL-EBI
- DDBJ

Catalogue
of Life

Species 2000

Viral Genome Resource
(NCBI)

What about impact?

Google Scholar, 9th Report	
Adenoviridae	88
Baculoviridae	49
Nimaviridae	7
Geminiviridae	122
Parvoviridae	62
Caulimoviridae	13
Paramyxoviridae	19
Rhabdoviridae	106
Ophioviridae	12
Varicosavirus	5
Nidovirales	31
Arteriviridae	26
Coronaviridae	94
Roniviridae	5
Dicistroviridae	6
Iflaviridae	5
Picornaviridae	201
Secoviridae	40
Alphaflexiviridae	13
Betaflexiviridae	40
Bromoviridae	16
Caliciviridae	30
Flaviviridae	68
Hepeviridae	84
Potyviridae	100
Umbravirus	5
Mean	**48**
Citations of whole book	**1097**

- The ICTV Reports are not included in the National Library of Medicine catalogue
- Therefore, it is not possible to search authors or chapters in PubMed
- Other searches yield very patchy results
- This makes citation much more difficult, decreasing the impact of the chapters in the Report and failing to acknowledge the contribution of authors

Citation of authors

THE PRACTICALITIES OF PUBLISHING ONLINE AND BEING CITED

1966 ICTV 2016

Peter Simmonds

Current

The ICTV Report

- Last published in 2012
- Increasingly divergent from the ICTV classification
- Restricted availability, especially in developing countries

Usage by the virology community

- Erratic collection of citations
- Not a primary source of taxonomic information for most

Useability

- Provides helpful overviews of virus families and orders
- Poor or absent linkage to MSL and other databases

Planned

On-line publication

- Electronic format (HTML) and PDF
- Fully linked to MSL, resources and external databases

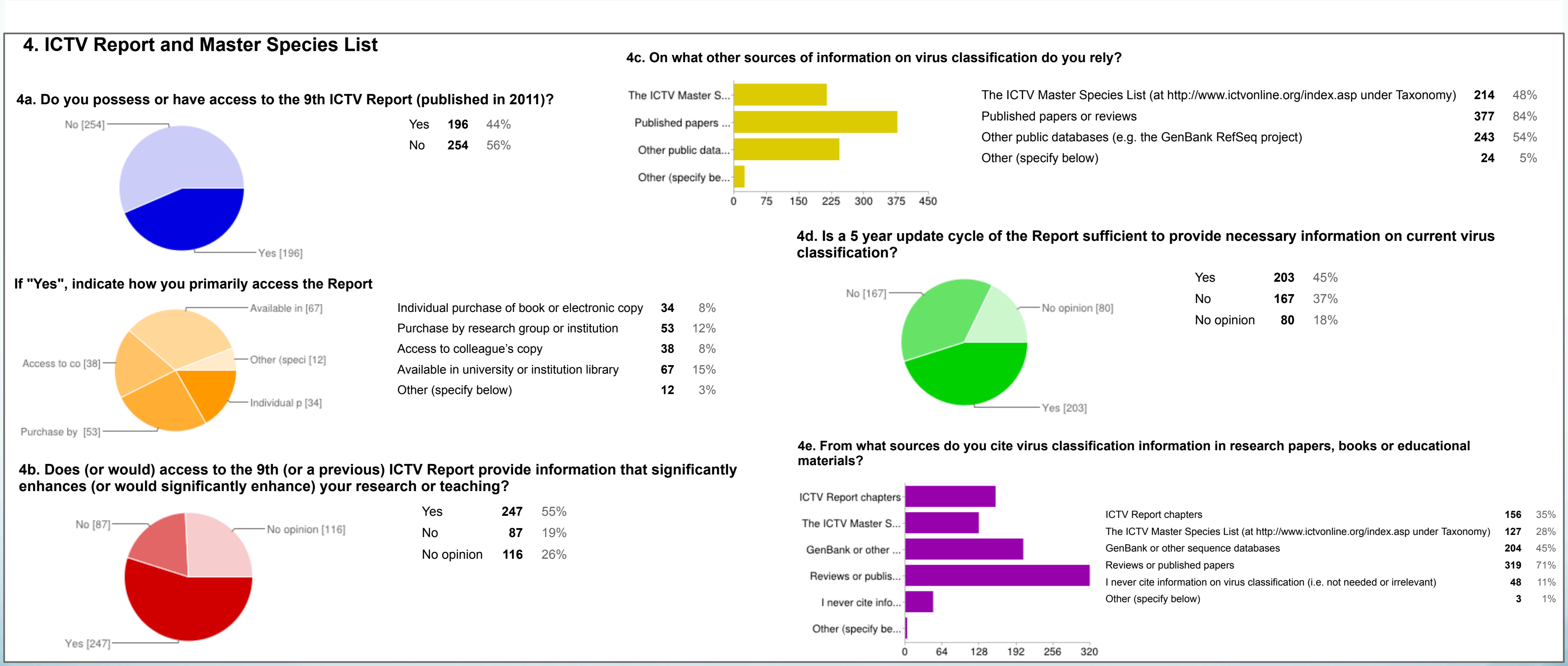
Updating

- Updated whenever taxonomy of a chapter (family, order) is changed
- Cannot diverge from the MSL

Citation

- Chapter summaries published in an open access journal
- Allows accurate citation metrics to be collected
- Contains links to online chapter, resources, SG wikis

Survey of >500 virologists (2014)



4. ICTV Report and Master Species List

4a. Do you possess or have access to the 9th ICTV Report (published in 2011)?

Yes	**196**	44%
No	**254**	56%

If "Yes", indicate how you primarily access the Report

Individual purchase of book or electronic copy	**34**	8%
Purchase by research group or institution	**53**	12%
Access to colleague's copy	**38**	8%
Available in university or institution library	**67**	15%
Other (specify below)	**12**	3%

4b. Does (or would) access to the 9th (or a previous) ICTV Report provide information that significantly enhances (or would significantly enhance) your research or teaching?

Yes	**247**	55%
No	**87**	19%
No opinion	**116**	26%

4c. On what other sources of information on virus classification do you rely?

The ICTV Master Species List (at http://www.ictvonline.org/index.asp under Taxonomy)	**214**	48%
Published papers or reviews	**377**	84%
Other public databases (e.g. the GenBank RefSeq project)	**243**	54%
Other (specify below)	**24**	5%

4d. Is a 5 year update cycle of the Report sufficient to provide necessary information on current virus classification?

Yes	**203**	45%
No	**167**	37%
No opinion	**80**	18%

4e. From what sources do you cite virus classification information in research papers, books or educational materials?

ICTV Report chapters	**156**	35%
The ICTV Master Species List (at http://www.ictvonline.org/index.asp under Taxonomy)	**127**	28%
GenBank or other sequence databases	**204**	45%
Reviews or published papers	**319**	71%
I never cite information on virus classification (i.e. not needed or irrelevant)	**48**	11%
Other (specify below)	**3**	1%

Current

The ICTV Report

- Last published in 2012
- Increasingly divergent from the ICTV classification
- Restricted availability, especially in developing countries

Usage by the virology community

- Erratic collection of citations
- Not a primary source of taxonomic information for most

Useability

- Provides helpful overviews of virus families and orders
- Poor or absent linkage to MSL and other databases

Planned

On-line publication

- Electronic format (HTML) and PDF
- Fully linked to MSL, resources and external databases

Updating

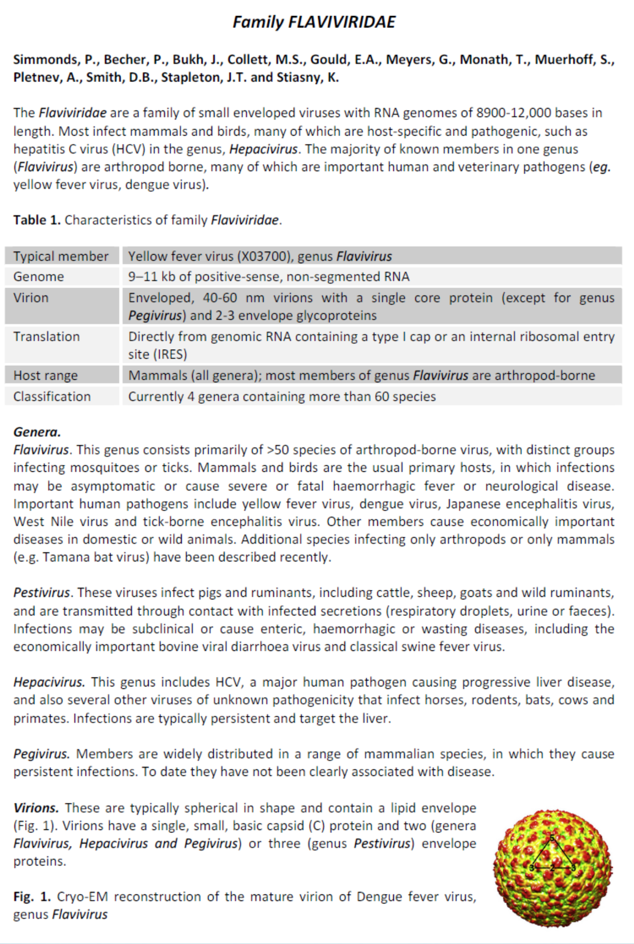
- Updated whenever taxonomy of a chapter (family, order) is changed
- Cannot diverge from the MSL

Citation

- Chapter summaries published in an open access journal
- Allows accurate citation metrics to be collected
- Contains links to online chapter, resources, SG wikis

Report chapter summaries

- On-line publication
- Open access
- PMC numbers assigned
- Co-authored by SG members
- Summary format
- Informative overview of group
- Longer publication cycle than chapters
- Citation
- The summary is the citation object
- Brings together links to other resources
- Visibility



Family FLAVIVIRIDAE

Simmonds, P., Becher, P., Bukh, J., Collett, M.S., Gould, E.A., Meyers, G., Monath, T., Muerhoff, S., Pletnev, A., Smith, D.B., Stapleton, J.T. and Stiasny, K.

The *Flaviviridae* are a family of small enveloped viruses with RNA genomes of 8900-12,000 bases in length. Most infect mammals and birds, many of which are host-specific and pathogenic, such as hepatitis C virus (HCV) in the genus, ***Hepacivirus***. The majority of known members in one genus (*Flavivirus*) are arthropod borne, many of which are important human and veterinary pathogens (*eg.* yellow fever virus, dengue virus).

Table 1. Characteristics of family *Flaviviridae*.

Typical member	Yellow fever virus (X03700), genus *Flavivirus*
Genome	9–11 kb of positive-sense, non-segmented RNA
Virion	Enveloped, 40-60 nm virions with a single core protein (except for genus *Pegivirus*) and 2-3 envelope glycoproteins
Translation	Directly from genomic RNA containing a type I cap or an internal ribosomal entry site (IRES)
Host range	Mammals (all genera); most members of genus *Flavivirus* are arthropod-borne
Classification	Currently 4 genera containing more than 60 species

Genera.

Flavivirus. This genus consists primarily of >50 species of arthropod-borne virus, with distinct groups infecting mosquitoes or ticks. Mammals and birds are the usual primary hosts, in which infections may be asymptomatic or cause severe or fatal haemorrhagic fever or neurological disease. Important human pathogens include yellow fever virus, dengue virus, Japanese encephalitis virus, West Nile virus and tick-borne encephalitis virus. Other members cause economically important diseases in domestic or wild animals. Additional species infecting only arthropods or only mammals (e.g. Tamana bat virus) have been described recently.

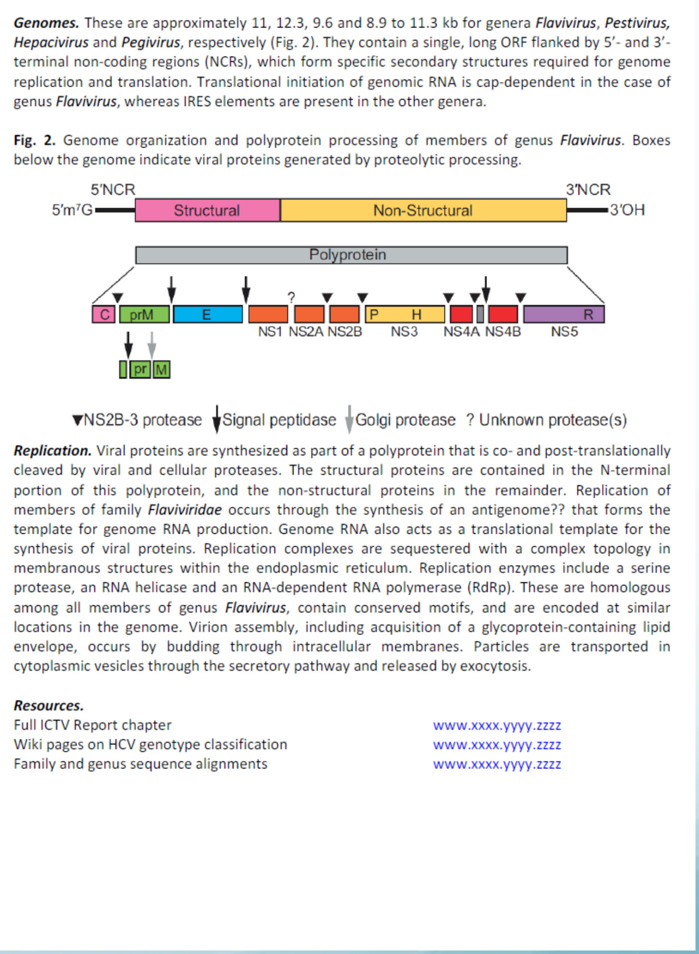
Pestivirus. These viruses infect pigs and ruminants, including cattle, sheep, goats and wild ruminants, and are transmitted through contact with infected secretions (respiratory droplets, urine or faeces). Infections may be subclinical or cause enteric, haemorrhagic or wasting diseases, including the economically important bovine viral diarrhoea virus and classical swine fever virus.

Hepacivirus. This genus includes HCV, a major human pathogen causing progressive liver disease, and also several other viruses of unknown pathogenicity that infect horses, rodents, bats, cows and primates. Infections are typically persistent and target the liver.

Pegivirus. Members are widely distributed in a range of mammalian species, in which they cause persistent infections. To date they have not been clearly associated with disease.

Virions. These are typically spherical in shape and contain a lipid envelope (Fig. 1). Virions have a single, small, basic capsid (C) protein and two (genera *Flavivirus, Hepacivirus and Pegivirus*) or three (genus *Pestivirus*) envelope proteins.

Fig. 1. Cryo-EM reconstruction of the mature virion of Dengue fever virus, genus *Flavivirus*



Genomes. These are approximately 11, 12.3, 9.6 and 8.9 to 11.3 kb for genera *Flavivirus*, *Pestivirus*, *Hepacivirus* and *Pegivirus*, respectively (Fig. 2). They contain a single, long ORF flanked by 5'- and 3'-terminal non-coding regions (NCRs), which form specific secondary structures required for genome replication and translation. Translational initiation of genomic RNA is cap-dependent in the case of genus *Flavivirus*, whereas IRES elements are present in the other genera.

Fig. 2. Genome organization and polyprotein processing of members of genus *Flavivirus*. Boxes below the genome indicate viral proteins generated by proteolytic processing.

Replication. Viral proteins are synthesized as part of a polyprotein that is co- and post-translationally cleaved by viral and cellular proteases. The structural proteins are contained in the N-terminal portion of this polyprotein, and the non-structural proteins in the remainder. Replication of members of family ***Flaviviridae*** occurs through the synthesis of an antigenome?? that forms the template for genome RNA production. Genome RNA also acts as a translational template for the synthesis of viral proteins. Replication complexes are sequestered with a complex topology in membranous structures within the endoplasmic reticulum. Replication enzymes include a serine protease, an RNA helicase and an RNA-dependent RNA polymerase (RdRp). These are homologous among all members of genus *Flavivirus*, contain conserved motifs, and are encoded at similar locations in the genome. Virion assembly, including acquisition of a glycoprotein-containing lipid envelope, occurs by budding through intracellular membranes. Particles are transported in cytoplasmic vesicles through the secretory pathway and released by exocytosis.

Resources.

Full ICTV Report chapter	www.xxxx.yyyy.zzzz
Wiki pages on HCV genotype classification	www.xxxx.yyyy.zzzz
Family and genus sequence alignments	www.xxxx.yyyy.zzzz

Impact

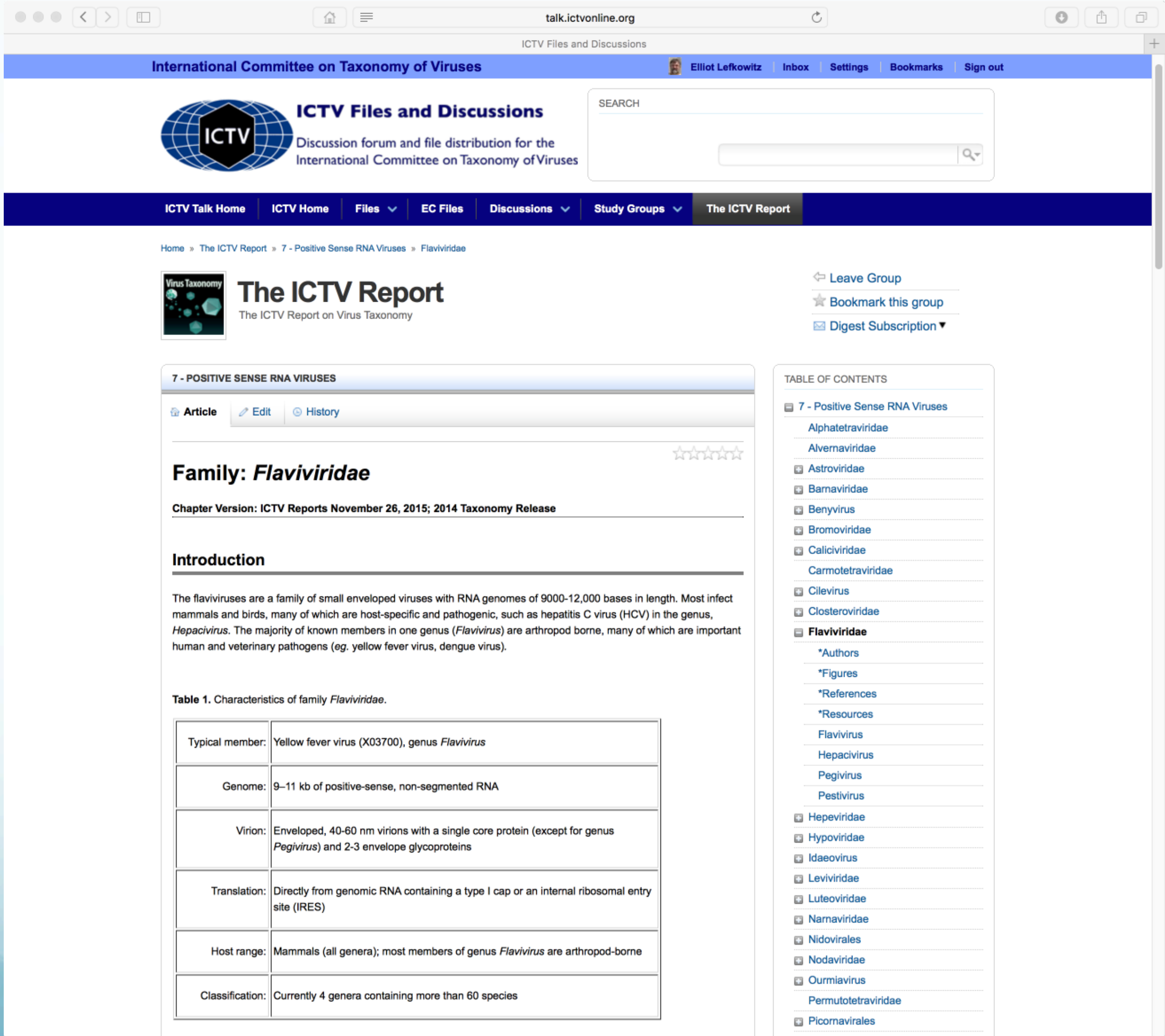
- Authoritative descriptions of virus families and orders
- An accessible and highly visible source of virus taxonomic information and overviews of virus properties
- Through bioinformatic links, the starting point for data compilation, sequence analysis and taxonomical decisions
- Enhanced visibility in the scientific community

A PROTOTYPE OF THE NEW REPORT

1966 ICTV 2016

Elliot Lefkowitz

A prototype of the new Report



The ICTV Report

The ICTV Report on Virus Taxonomy

Family: *Flaviviridae*

Chapter Version: ICTV Reports November 26, 2015; 2014 Taxonomy Release

Introduction

The flaviviruses are a family of small enveloped viruses with RNA genomes of 9000-12,000 bases in length. Most infect mammals and birds, many of which are host-specific and pathogenic, such as hepatitis C virus (HCV) in the genus, *Hepacivirus*. The majority of known members in one genus (*Flavivirus*) are arthropod borne, many of which are important human and veterinary pathogens (*eg.* yellow fever virus, dengue virus).

Table 1. Characteristics of family *Flaviviridae*.

Typical member:	Yellow fever virus (X03700), genus *Flavivirus*
Genome:	9–11 kb of positive-sense, non-segmented RNA
Virion:	Enveloped, 40-60 nm virions with a single core protein (except for genus *Pegivirus*) and 2-3 envelope glycoproteins
Translation:	Directly from genomic RNA containing a type I cap or an internal ribosomal entry site (IRES)
Host range:	Mammals (all genera); most members of genus *Flavivirus* are arthropod-borne
Classification:	Currently 4 genera containing more than 60 species

IMPLICATIONS FOR STUDY GROUPS

1966 ICTV 2016

Andrew Davison

Study Group involvement

The Report

- Batches of ~30 chapters and summaries to appear in early 2017, early 2018 and early 2019
- Publication of summaries in an open access journal
- Summaries to be published at about the same time as chapters
- Initial preference for the most highly citable chapters
- Volunteers for the first ~30 chapters for this autumn!

Other things

- Willingness to be consulted on development of other resources

Recognition that all this will be two-way process

DEVELOPMENT OF TOOLS

- Alignment construction and standardisation
  Donald Smith
- Methods for evolutionary analysis
  Richard Orton
- Study Group wikis and other specialised information
  Peter Simmonds
- Bioinformatic tools
  Sasha Gorbalenya

ALIGNMENT CONSTRUCTION AND STANDARDISATION

1966 ICTV 2016

Donald Smith

Why provide sequence alignments?

Phylogeny / sequence relatedness required for virus classification

- Proposed taxon assignments have to be consistent with phylogeny
- Assignments can be based on sequence relationships (e.g. *Papillomaviridae)*

Sequence relationships are dependent on

- Sequence type (amino acid or nucleotide)
- Divergence
- Alignment method, exclusion of non-alignable regions

Provision of standardised alignments

- Avoid generating *de novo* for taxonomy proposals
- Ensure appropriate genome region is analysed
- Alignment data ensures objectivity and reproducibility

Alignment issues

HOW?

- Manual/Clustal/Muscle/BLAST

WHAT?

- Nucleotide/amino acid
- Representative/type sequences

WHERE?

- Whole/subgenomic regions
- Concatenated or separate

ACCESSIBLE?

- Provision through ICTV website
- Alignment formats/annotation

Example

Hepeviridae

Positive-sense RNA virus

Orthohepevirus genus, 4 species

- A – hepatitis E virus (HEV; human, pig)
- B – avian HEV (chicken)
- C – rat HEV (rat, ferret)
- D – bat HEV (bat)

Piscihepevirus genus, 1 species

- A – trout HEV (fish)

Choice of alignment method

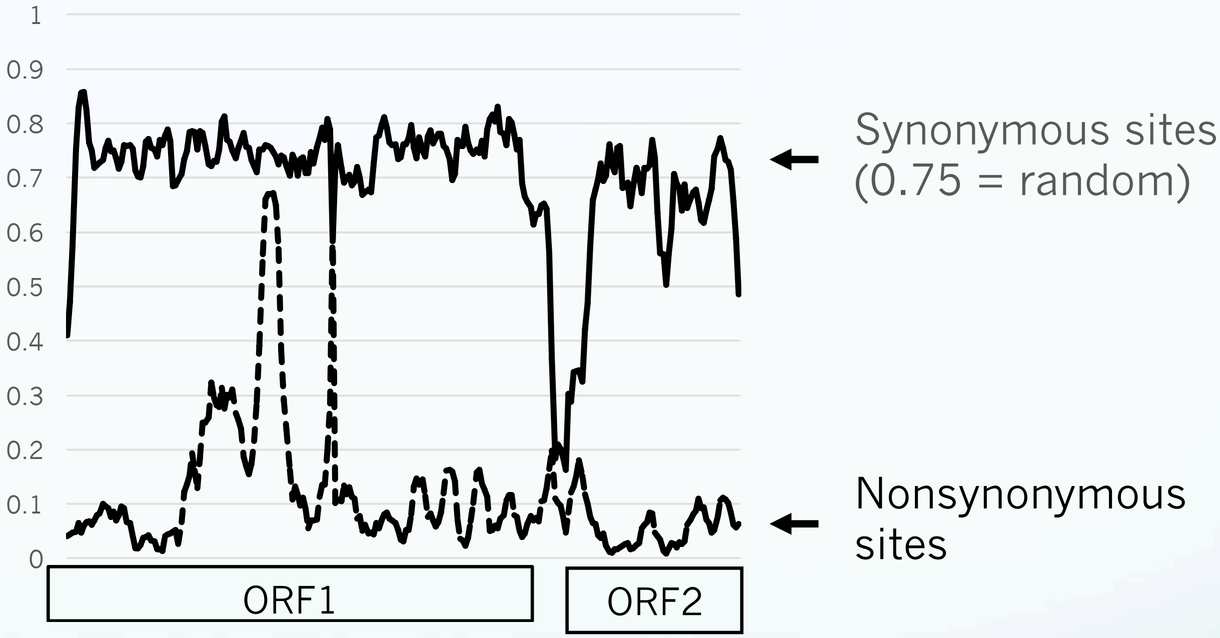
HOW?	Nucleotide or amino acid
	Manual / Clustal / Muscle / BLAST
WHAT?	⊙ Penalty for insertion
	⊙ Penalty for size of insertion
WHERE?	Different numbers of sequences in groups?
	Effects of adding new sequences
ACCESSIBLE?	

Genome scan – *Orthohepevirus A*

HOW?

WHAT?

WHERE?



ACCESSIBLE?

- If synonymous saturated – use aa/nonsynonymous
- If nonsynonymous saturated – use protein structure?

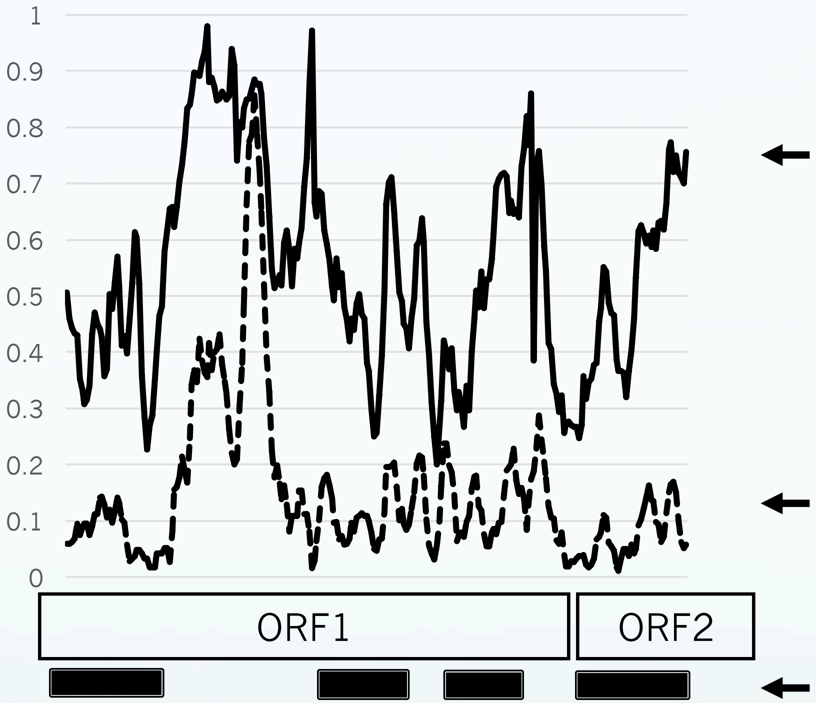
Genome aa scan – *Orthohepevirus A*

HOW?

WHAT?

WHERE?

ACCESSIBLE?



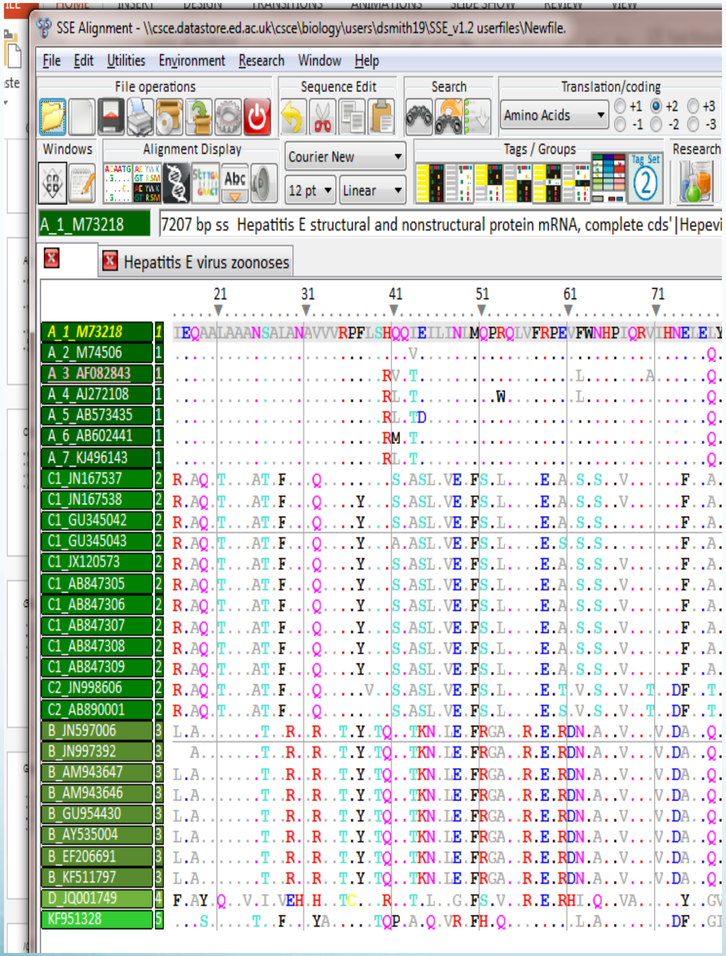
Between *Orthohepevirus A, B, C and D*

Within *Orthohepevirus A*

Conserved regions

- Identify conserved regions
- Do they give identical relationships?

Taxonomy of genus *Orthohepevirus*



SSE Alignment - \\csce.datastore.ed.ac.uk\csce\biology\users\dsmith19\SSE_v1.2 userfiles\Newfile.

Amino Acids

Courier New

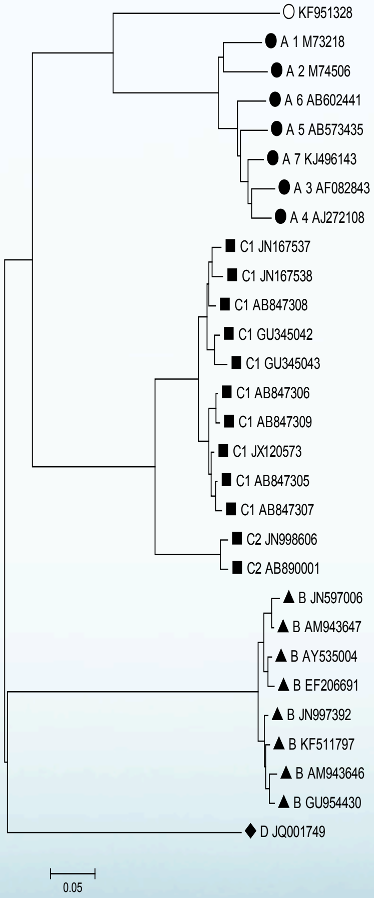
12 pt

Linear

A_1_M73218 | 7207 bp ss Hepatitis E structural and nonstructural protein mRNA, complete cds'|Hepevi

Hepatitis E virus zoonoses

```
                  21        31        41        51        61        71
A_1_M73218   1 IEQAALAAANSALANAVVVRPFLSHQQIEILINLMQPRQLVFRPEVFWNHPIQRVIHNELELY
A_2_M74506   1 ..........................V.......................................Q.
A_3_AF082843 1 ........................RV.T.................L.......A......Q.
A_4_AJ272108 1 ........................RL.T.........W........L................Q.
A_5_AB573435 1 ........................RL.TD.....................................Q.
A_6_AB602441 1 ........................RM.T......................................Q.
A_7_KJ496143 1 ........................RL.T......................................Q.
C1_JN167537  2 R.AQ.T...AT.F...Q.........S.ASL.VE.FS.L....E.A.S.S..V......F..A.
C1_JN167538  2 R.AQ.T...AT.F...Q....Y...S.ASL.VE.FS.L....E.A.S.S..V......F..A.
C1_GU345042  2 R.AQ.T...AT.F...Q....Y...S.ASL.VE.FS.L....E.A.S.S.........F..A.
C1_GU345043  2 R.AQ.T...AT.F...Q....Y...A.ASL.VE.FS.L....E.S.S.S.........F..A.
C1_JX120573  2 R.AQ.T...AT.F...Q....Y...S.ASL.VE.FS.L....E.A.S.S..V......F..A.
C1_AB847305  2 R.AQ.T...AT.F...Q....Y...S.ASL.VE.FS.L....E.A.S.S..V......F..A.
C1_AB847306  2 R.AQ.T...AT.F...Q....Y...S.ASL.VE.FS.L....E.A.S.S..V......F..A.
C1_AB847307  2 R.AQ.T...AT.F...Q....Y...S.ASL.VE.FS.L....E.A.S.S..V......F..A.
C1_AB847308  2 R.AQ.T...AT.F...Q....Y...S.ASL.VE.FS.L....E.A.S.S..V......F..A.
C1_AB847309  2 R.AQ.T...AT.F...Q....Y...S.ASL.VE.FS.L....E.A.S.S..V......F..A.
C2_JN998606  2 R.AQ.T...AT.F...Q.....V..S.ASL.VE.FS.L....E.T.V.S..V..T..DF..T.
C2_AB890001  2 R.AQ.T...AT.F...Q........S.ASL.VE.FS.L....E.S.V.S..V..T..DF..T.
B_JN597006   3 L.A.......T..R..R..T.Y.TQ..TKN.LE.FRGA..R.E.RDN.A..V...V.DA..Q.
B_JN997392   3   A.......T..R..R..T.Y.TQ..TKN.LE.FRGA..R.E.RDN.A..V...V.DA..Q.
B_AM943647   3 L.A.......T..R..R..T.Y.TQ..TKN.LE.FRGA..R.E.RDN.A..V...V.DA..Q.
B_AM943646   3 L.A.......T..R..R..T.Y.TQ..TKN.LE.FRGA..R.E.RDN.A..V...V.DA..Q.
B_GU954430   3 L.A.......T..R..R..T.Y.TQ..TKN.LE.FRGA..R.E.RDN.A..V...V.DA..Q.
B_AY535004   3 L.A.......T..R..R..T.Y.TQ..TKN.LE.FRGA..R.E.RDN.A..V...V.DA..Q.
B_EF206691   3 L.A.......T..R..R..T.Y.TQ..TKN.LE.FRGA..R.E.RDN.A..V...V.DA..Q.
B_KF511797   3 L.A.......T..R..R..T.Y.TQ..TKN.LE.FRGA..R.E.RDN.A..V...V.DA..Q.
D_JQ001749   4 F.AY.Q..V.I.VEH.H..TC...R..T.L..G.FS.V..R.E.RHI.Q..VA.....Y..GV
KF951328     5 ...S.....T..F...YA.....TQP.A.Q.VR.FH.Q..........L.A........DF..GI
```



Host	Species
Moose	?
Human Pig, wild boar Camel	A
Rat Ferret	C
Chicken	B
Bat	D

Amino acid sequence alignment

ORF1 aa 28-389

Providing alignments

Sequence representation

- Limited to exemplar sequences or examplar sequences clearly identified
- May include provisionally or unclassified sequences from SG

Alignment format

- Importable by standard bioinformatic packages
- Primary labelling by accession numbers, with additional information of taxonomic assignment, strain, host, etc.

Maintained by SGs

- Regularly updated to include newly assigned taxa
- Downloadable for a DOI specified in Report/published summary

The perfect alignment...

HOW? → Method of alignment stated

- Reproducible

WHAT? → Characters used justified

- Saturation at synonymous sites?

WHERE? → Choice of regions to compare

- Variability across genome assessed

ACCESSIBLE? → Alignment available on ICTV website

- Taxonomic conclusions verifiable
- Incorporate new variants

METHODS FOR EVOLUTIONARY ANALYSIS

1966 ICTV 2016

Richard Orton

Introduction

Richard Orton

- Bioinformatician
- Viral Genomics & Bioinformatics group (Andrew Davison)
- Centre for Virus Research, University of Glasgow

ICTV Wellcome Trust Grant

- A database for the universal classification of viruses
- Peter Simmonds, Andrew Davison, Stuart Siddell

Working on

- Online ICTV Report
- Online standardized alignments and trees
- Online tools for preparing and submitting TPs
- Cross-links with sequence databases such as GenBank, ENA, ATCC, DDBJ, Species 2000, Catalogue of Life

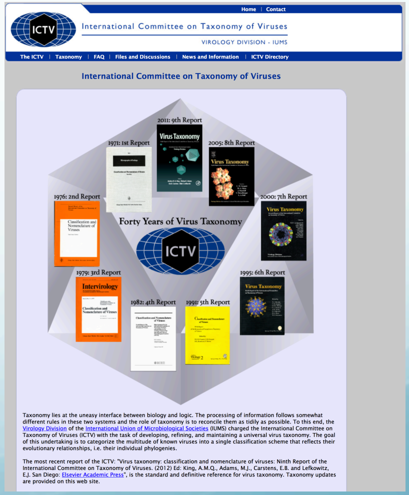
Help SGs

- Report tools
- Proposal tools
- Bioinformatics tools

CVR
Medical Research Council
University of Glasgow
Centre for Virus Research

University of Glasgow

wellcome trust



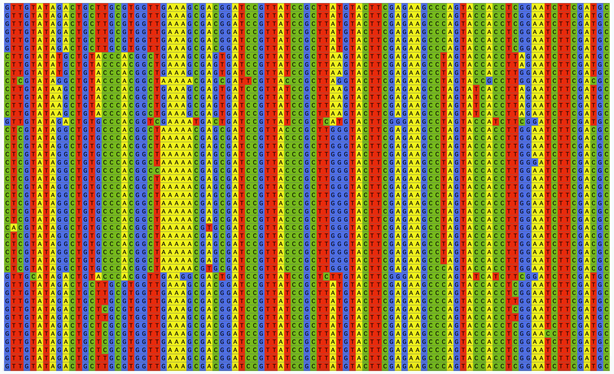
International Committee on Taxonomy of Viruses

Taxonomy lies at the uneasy interface between biology and logic. The processing of information follows somewhat different rules in these two systems and the role of taxonomy is to reconcile them as tidily as possible. To this end, the Virology Division of the International Union of Microbiological Societies (IUMS) charged the International Committee on Taxonomy of Viruses (ICTV) with the task of developing, refining, and maintaining a universal virus taxonomy. The goal of this undertaking is to categorize the multitude of known viruses into a single classification scheme that reflects their evolutionary relationships, i.e. their individual phylogenies.

The most recent report of the ICTV: "Virus taxonomy: classification and nomenclature of viruses: Ninth Report of the International Committee on Taxonomy of Viruses. (2012) Ed: King, A.M.Q., Adams, M.J., Carstens, E.B. and Lefkowitz, E.J. San Diego: Elsevier Academic Press", is the standard and definitive reference for virus taxonomy. Taxonomy updates are provided on this web site.

http://talk.ictvonline.org

Evolutionary analysis (1)



Distance-based

- Create pairwise distance matrix: neighbour-joining: fast and large

Maximum parsimony

- Tree with the smallest number of evolutionary events to explain the observed sequence data

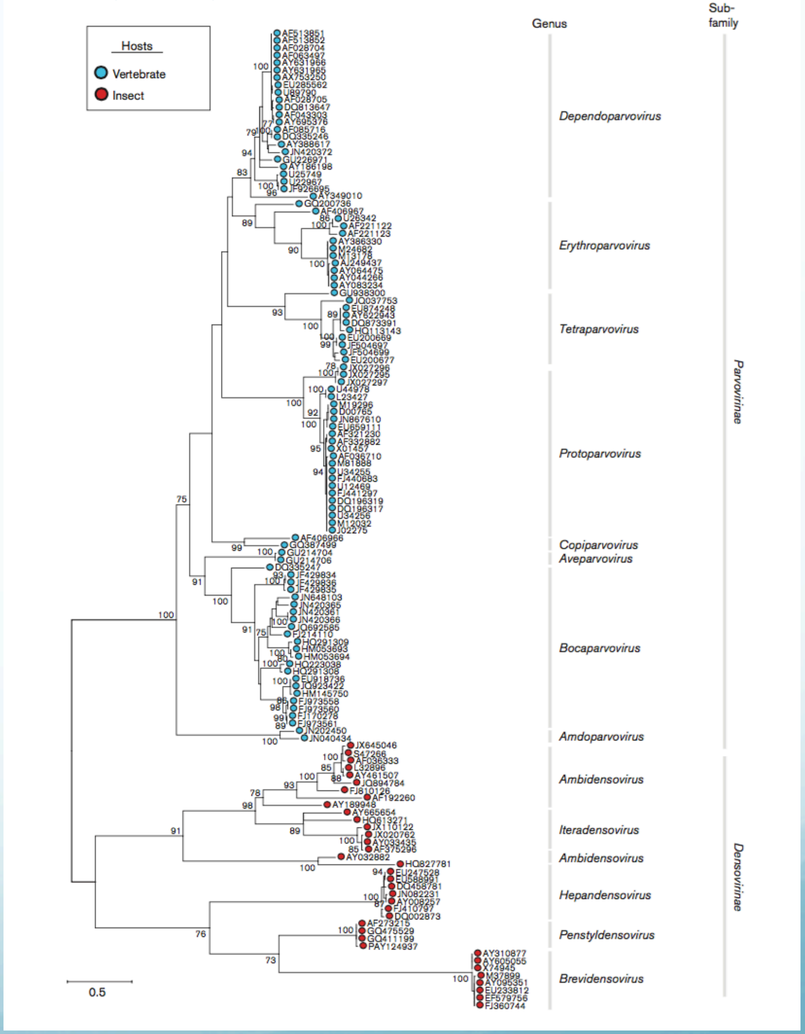
Maximum likelihood

- Model-based (e.g. JC, HKY, GTP … LG, WAG, JTT)
- Statistically assigns probabilities to possible phylogenetic trees

Bayesian

- Similar to maximum likelihood: model-based
- Bayesian statistics, prior and posterior sampling, MCMC, to produce the most likely tree given the data

Phylogeny and classification of *Parvoviridae*



Simmonds (2015) J Gen Virol 96:1193-1206

Evolutionary analysis (2)

Online

- ICTV website
- Image and tree (e.g. Newick format)

Clear methods

- Type (e.g. NJ, ML …)
- Substitution model
- Number of bootstraps
- The alignment – protein, nucleotide, gene, genome

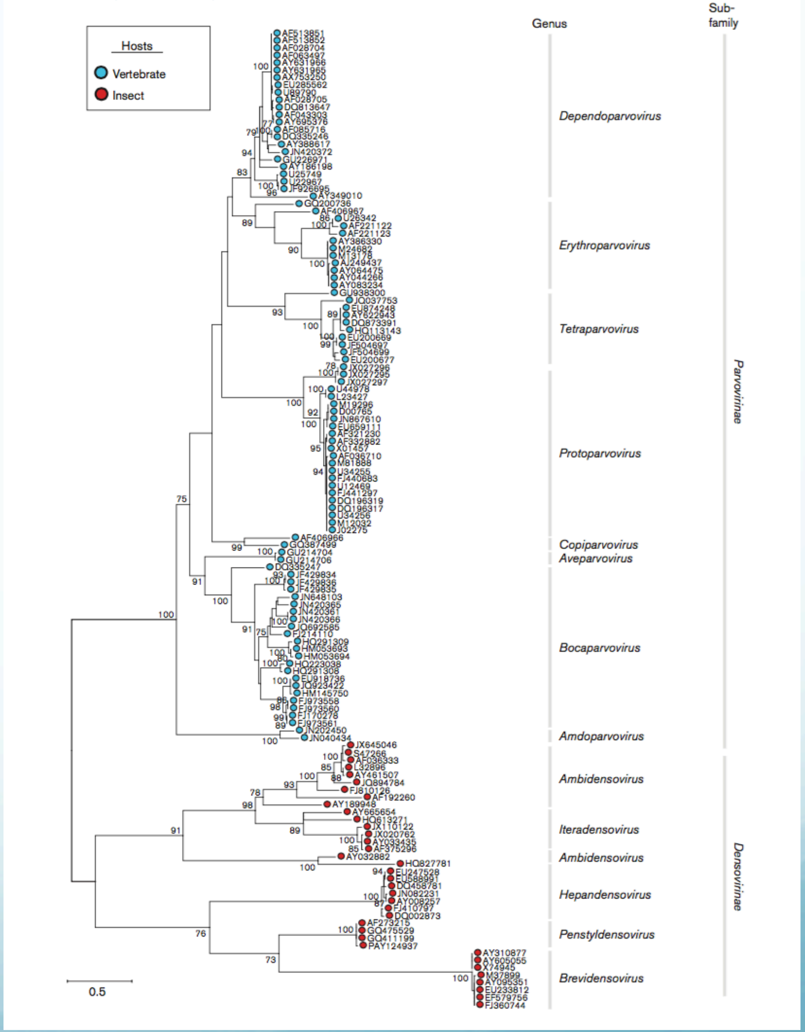
Format

- Scale bar
- Bootstrap values
- Standard format/visual

Distance cutoffs

- SGs current criteria for new species definition
- E.g. distance thresholds

Phylogeny and classification of *Parvoviridae*



Simmonds (2015) J Gen Virol 96:1193-1206

Applications

Sequences are playing an increasingly important role in virus classification

Comparison of new sequences against existing ones can aid virus classification

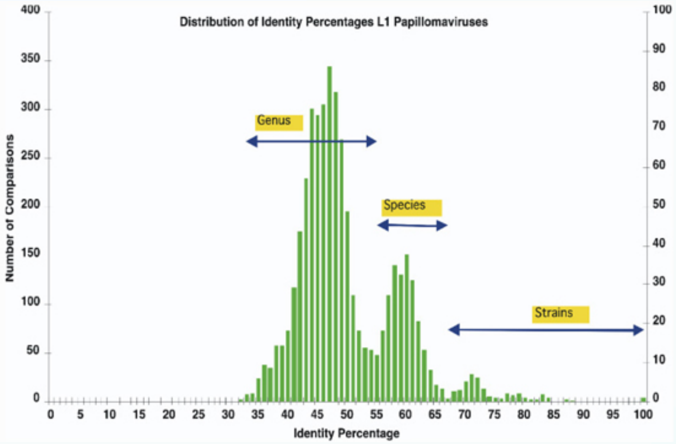
PASC
- PAirwise Sequence Comparison
- Bao et al (2014) Arch Virol

DeMARC
- DivErsity pArtitioning by hieRarchical Clustering
- Lauber et al. (2012) J Virol

PUmPER
- Phylogenies Updated PERpetually
- Izquierdo-Carrasco et al. (2014) Bioinf

SDT
- Sequence Demarcation Tool
- Muhire et al. (2014) PLoS One

ViCTree
- Virus Classification using phylogenetic Trees
- Modha et al. CVR University of Glasgow



Fauquet et al. (2006) Kansas School Naturalist Vol 53 No 1, April 2006

Click on the chart to get a list of GIs for each bar or submit sequences (as accessions, GIs or in FASTA format) to find how they fit within this family.

Sequence:

No file selected.

Genera selection:
- ☑ Check/Uncheck All
- ☑ Bdellomicrovirus
- ☑ Chlamydiamicrovirus
- ☑ Microvirus
- ☑ Spiromicrovirus
- ☑ Unassigned

Microviridae taxonomy at ICTV and NCBI

Taxonomy changes:

Merge species above
Separate species below
Merge genera above
Separate genera below

Split genomes:

Split genomes below

Bao et al. (2014) Arch Viol

Parvoviridae Study Group

Susan Cotmore
SG Chair

Andrew
Davison

Sejal
Modha

Joseph
Hughes

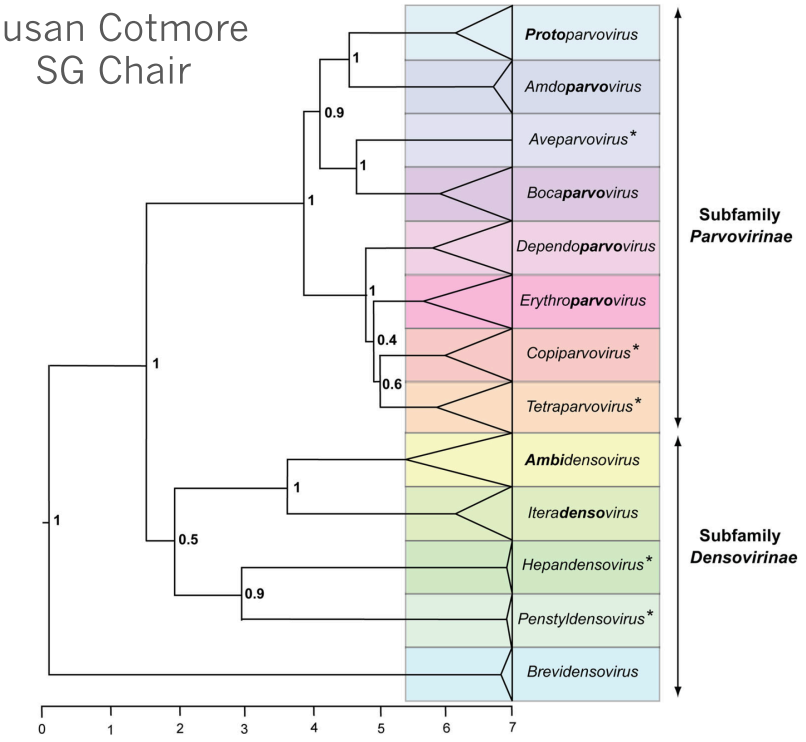
Parvoviridae family

- *Parvovirinae* subfamily – viruses that infect vertebrate hosts
- *Densovirinae* subfamily – viruses that infect arthropod hosts

Tool for use in viral classification

Automatically produce distance matrices and phylogenetic trees for a set of related sequences derived from GenBank

Update existing alignments and trees in an iterative fashion



Cotmore et al. (2014) Arch Virol 159:1239-1247

ViCTree pipeline

Example*: Densovirinae* subfamily
- NS1 protein sequences

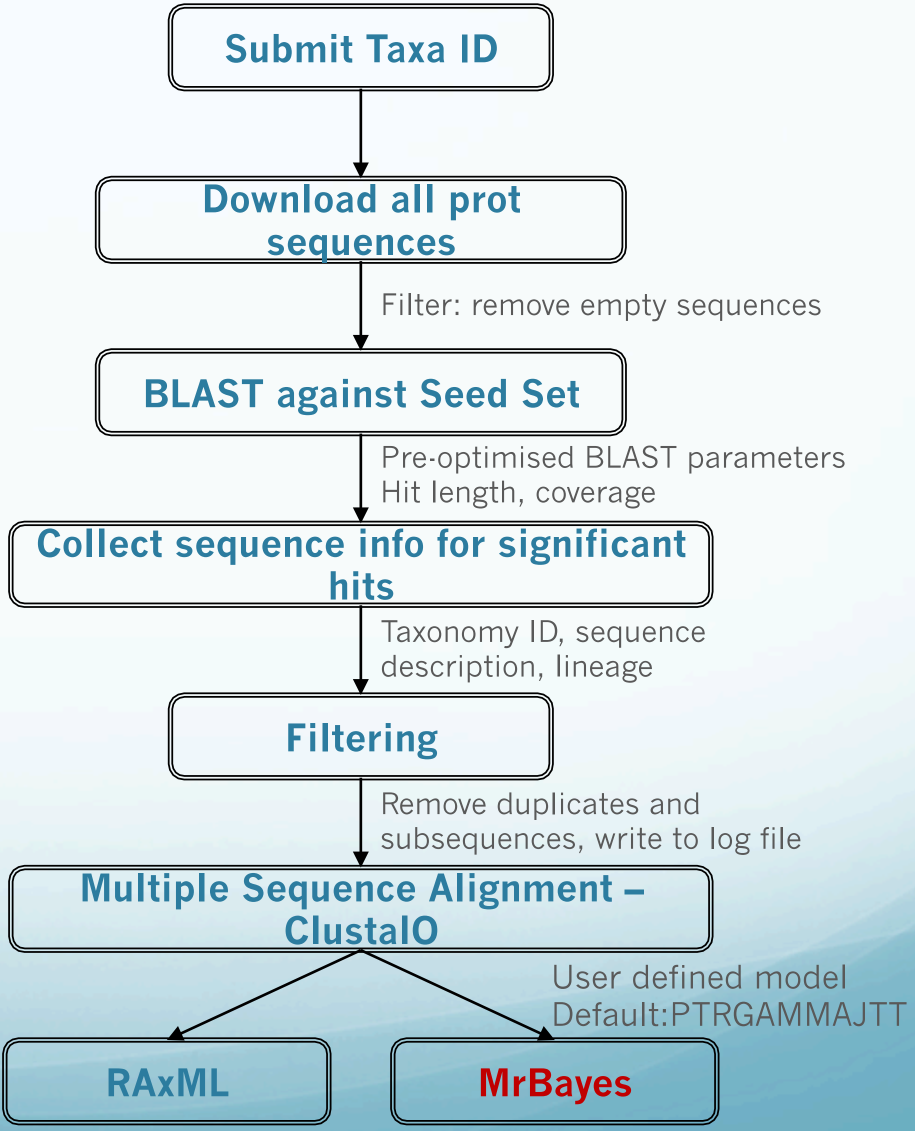
Seed Set: a set of sequences that encapsulates the diversity of the family
- E.g. at least 1 for each known species
- Can add novel (non-GenBank) seqs

BLAST parameters
- Optimised for each family

Input
- Taxa ID: txid40120
- Seed Set: 22 NS1 prot seqs: GI in header

Output
- 710 raw *Densovirinae* protein seqs
- Metadata file (GI, accession, species)
- 82 unique NS1 seqs
- Metadata file
- Multiple sequence alignment
- Distance matrix
- Phylogenetic trees

Applicable to other families



Submit Taxa ID

Download all prot sequences

Filter: remove empty sequences

BLAST against Seed Set

Pre-optimised BLAST parameters
Hit length, coverage

Collect sequence info for significant hits

Taxonomy ID, sequence description, lineage

Filtering

Remove duplicates and subsequences, write to log file

Multiple Sequence Alignment – ClustalO

User defined model
Default:PTRGAMMAJTT

RAxML

MrBayes

Densovirinae ViCTree

Phylogenetic tree

- Cluster picker [Ragonnet-Cronin et al (2013) BMC Bioinformatics 14:317]
- Cut-offs for within cluster genetic distance and bootstrap support are selected by the user

Pairwise distance matrix

- ViCTree integrates the pairwise distance matrix with the maximum likelihood tree to aid species identification
- User defines distance cutoff for family

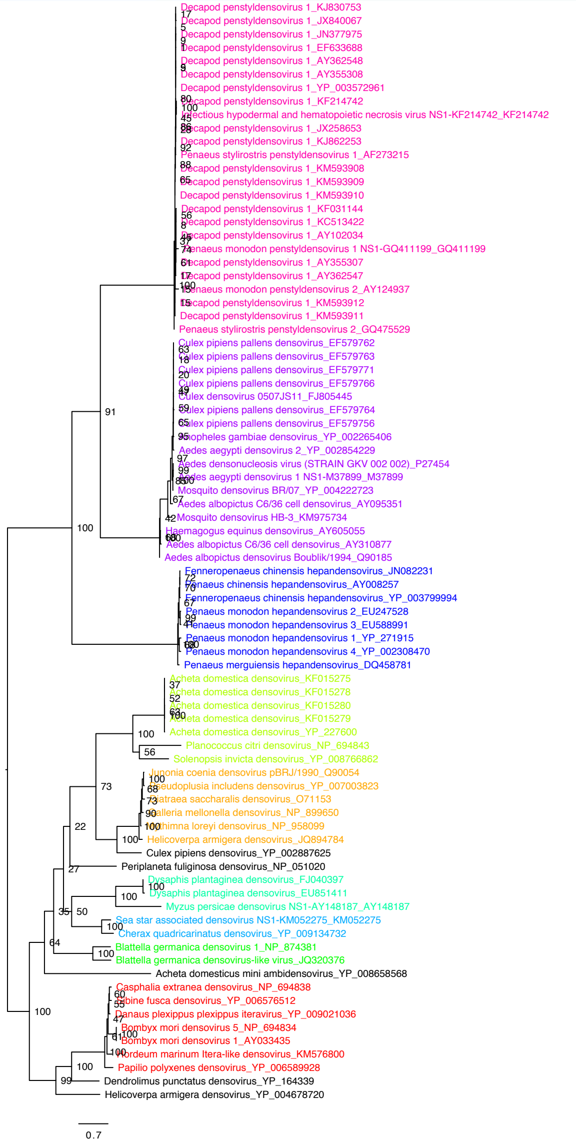
Web-based visualisation

- Sliding window for distance: clusters highlighted dynamically

Version Control

- Github
- Pipeline and code
- Alignments and trees

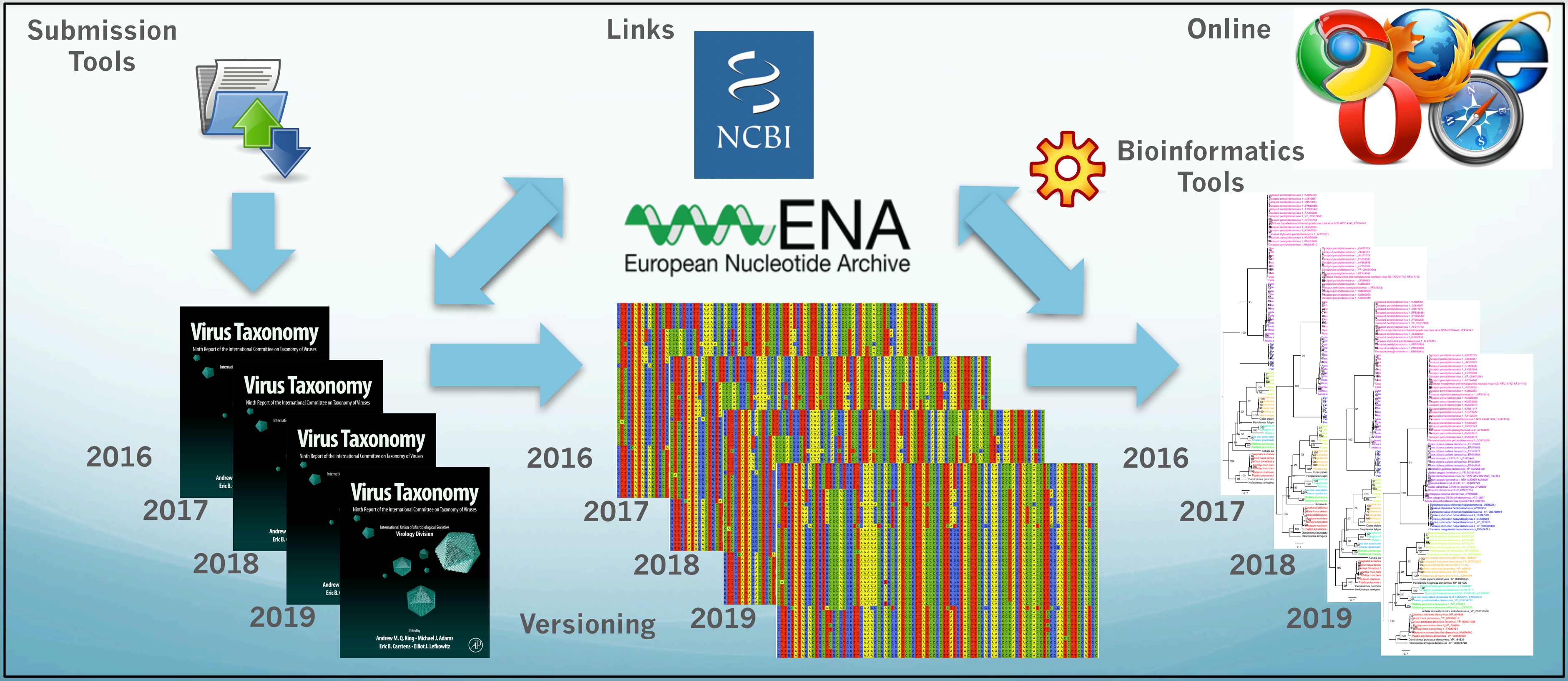
Tool for helping in new species identification



Future

Please feel free to contact us with suggestions or ideas for improvements and tools

- Submission tools
- Online format
- Linked databases
- Bioinformatics tools



Submission Tools

Links

Online

Bioinformatics Tools

2016
2017
2018
2019

2016
2017
2018
2019

Versioning

2016
2017
2018
2019

STUDY GROUP WIKIS AND OTHER SPECIALISED INFORMATION

1966 ICTV 2016

Peter Simmonds

Study Group wikis

A framework for information not contained in the Report chapters

- Other taxonomic levels
- Discussion of classification decisions
- Sequences of currently unclassified (tentative) family members

Accessible and editable by the virology community

- Online discussions of resource or classification issues

Permanence and authority

- A permanent component of the ICTV website
- Actively maintained by the SGs
- Repository for external data requiring visibility in the community

Study Group wikis

A framework for information not contained in the Report chapters

- Other taxonomic levels *HCV Genotypes, HEV types within species A*
- Discussion of classification decisions
- Sequences of currently unclassified (tentative) family members

Accessible and editable by the virology community

- Online discussions of resource or classification issues

Permanence and authority

- A permanent component of the ICTV website
- Actively maintained by the SGs
- Repository for external data requiring visibility in the community

Adoption by Study Groups

- Repository for sequence data, tentative family members, alignments, programs and literature
- Content variable
- Optional
- Cooperation and cross-linking with other websites and collaborative resources

BIOINFORMATIC TOOLS

1966 ICTV 2016

Sasha Gorbalenya

LEIDEN UNIVERSITY MEDICAL CENTER

DEmARC: Software for Virus Taxonomy

Alexander (Sasha) E. Gorbalenya

Department of Medical Microbiology, Leiden University Medical Center;
Faculty of Bioengineering & Bioinformatics, Lomonosov Moscow State University

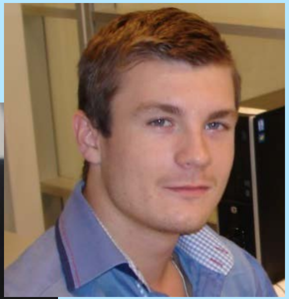
Team

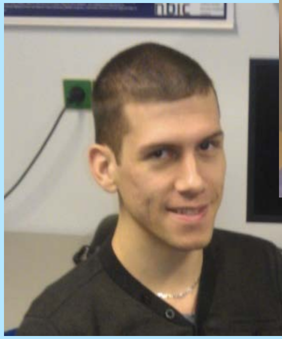

Alexander Kravchenko


Chris Lauber


Dmitry Samborskiy


Anastasia Gulyaeva


Matvey Zakharov


Erik Hoogendoorn


Igor Sidorov


Andrey Leontovich

Tree of Life depicts evolution of all life forms <u>but viruses</u>

“ scientists consider viruses as NOT living, it is WHY they are not in the tree of life”

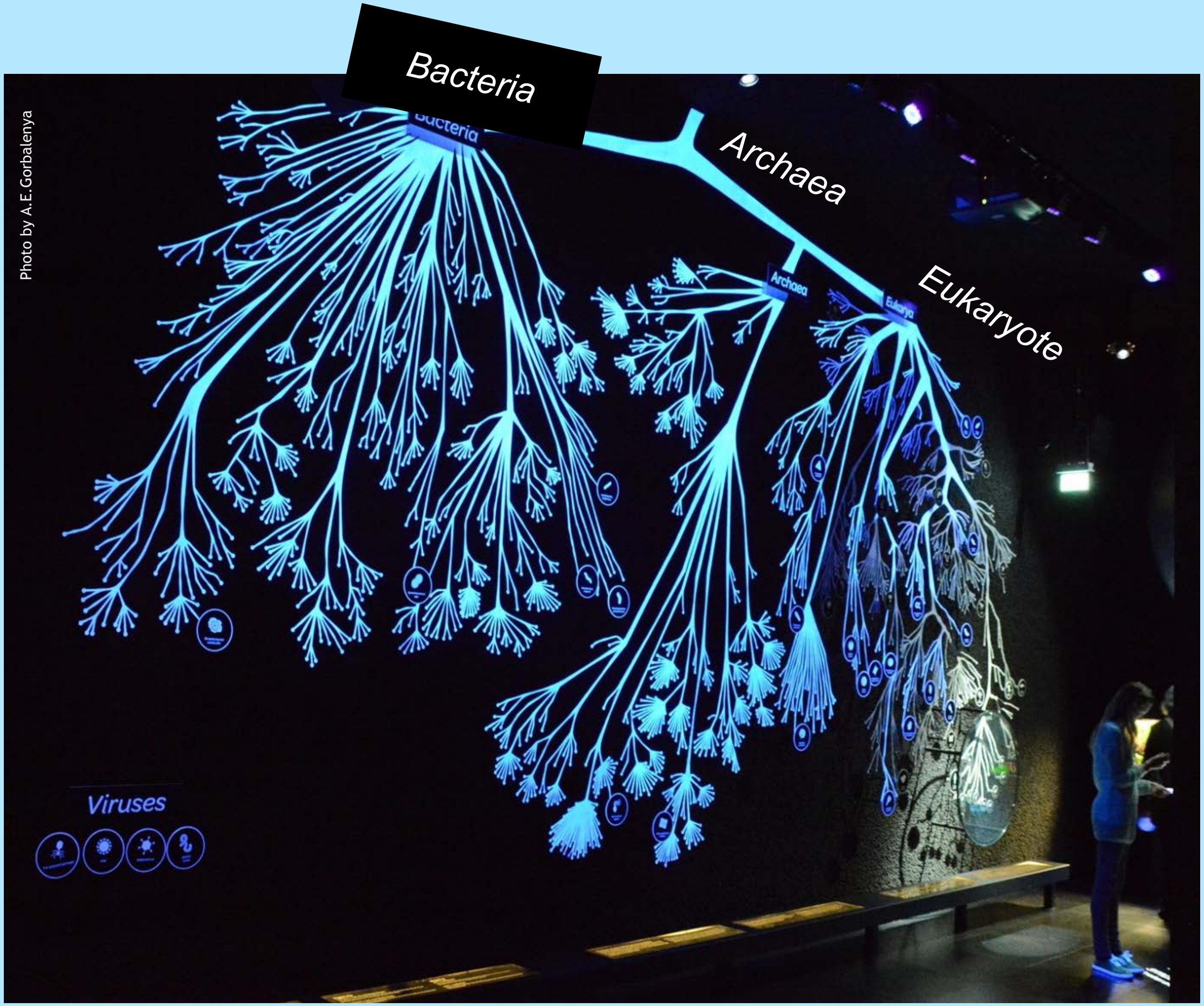
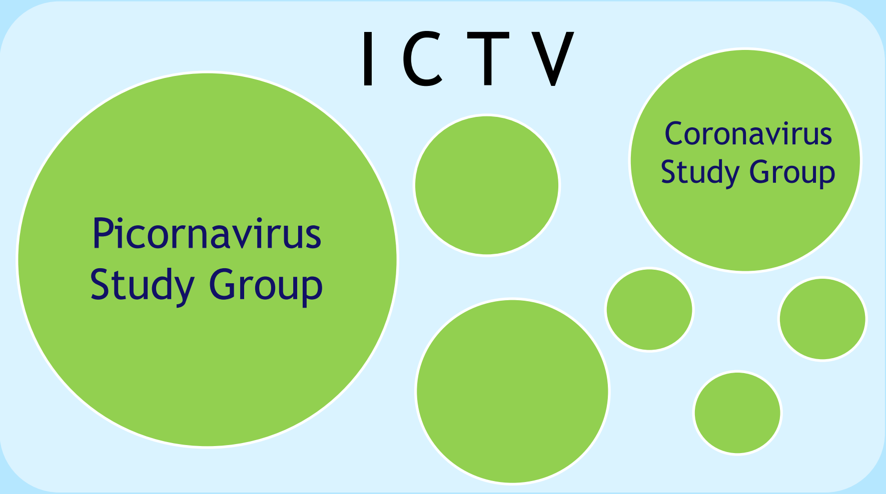


Photo by A.E.Gorbalenya

Microbe Museum, Amsterdam

Virus Taxonomy is produced by expert virologists

virus taxa	order family sub-family genus species
institution	**I**nternational **C**ommitee on **T**axonomy of **V**iruses
flexible framework	• expert-mediated • polythetic species demarcation criteria • time consuming

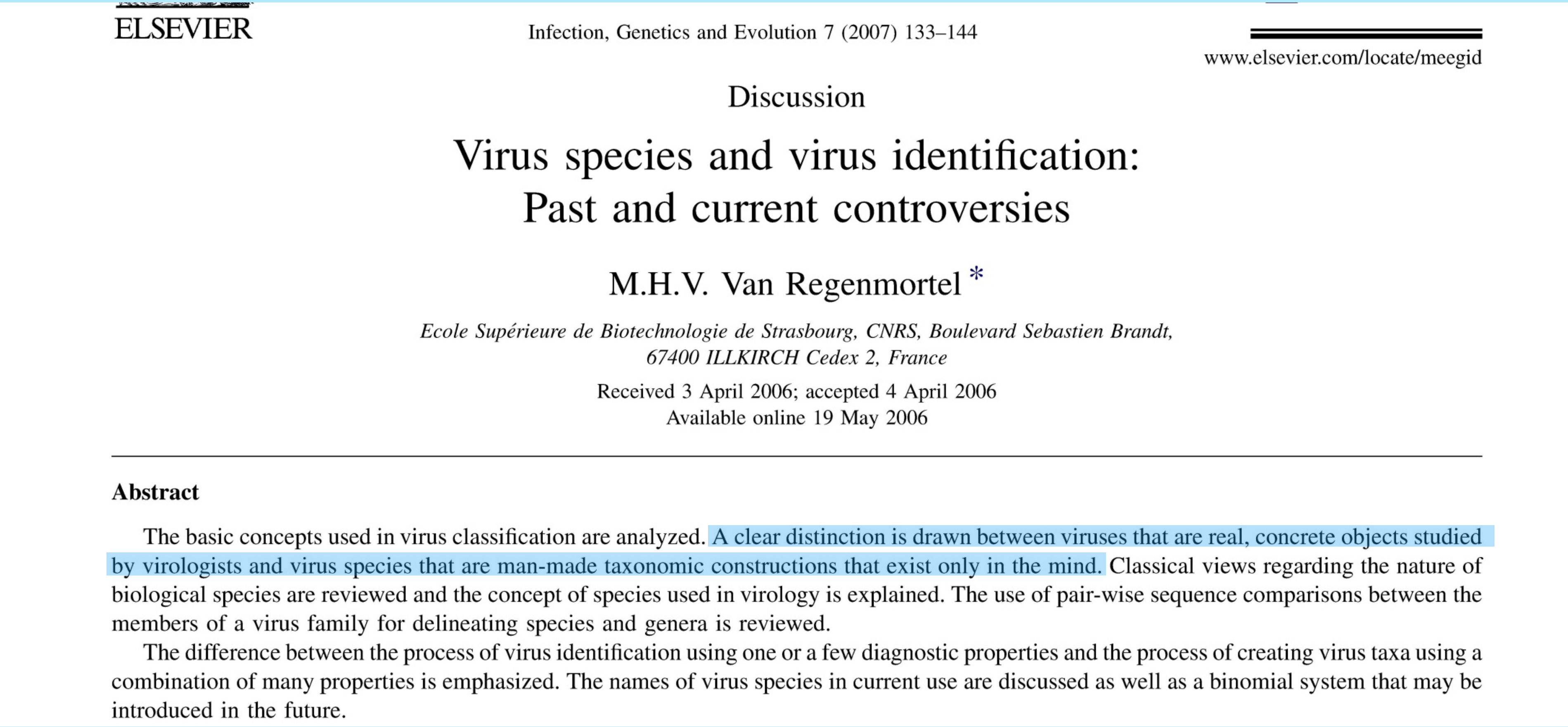


ICTV

Picornavirus Study Group

Coronavirus Study Group

Virus species are man-made taxonomic constructions



ELSEVIER

Infection, Genetics and Evolution 7 (2007) 133–144

www.elsevier.com/locate/meegid

Discussion

Virus species and virus identification: Past and current controversies

M.H.V. Van Regenmortel *

Ecole Supérieure de Biotechnologie de Strasbourg, CNRS, Boulevard Sebastien Brandt,
67400 ILLKIRCH Cedex 2, France

Received 3 April 2006; accepted 4 April 2006
Available online 19 May 2006

Abstract

The basic concepts used in virus classification are analyzed. A clear distinction is drawn between viruses that are real, concrete objects studied by virologists and virus species that are man-made taxonomic constructions that exist only in the mind. Classical views regarding the nature of biological species are reviewed and the concept of species used in virology is explained. The use of pair-wise sequence comparisons between the members of a virus family for delineating species and genera is reviewed.

The difference between the process of virus identification using one or a few diagnostic properties and the process of creating virus taxa using a combination of many properties is emphasized. The names of virus species in current use are discussed as well as a binomial system that may be introduced in the future.

> *“A clear distinction is drawn between viruses that are real, concrete objects studied by virologists and virus species that are man-made constructions that exist only in the mind.”*

Virus Taxonomy Framework: Benefits & Cost

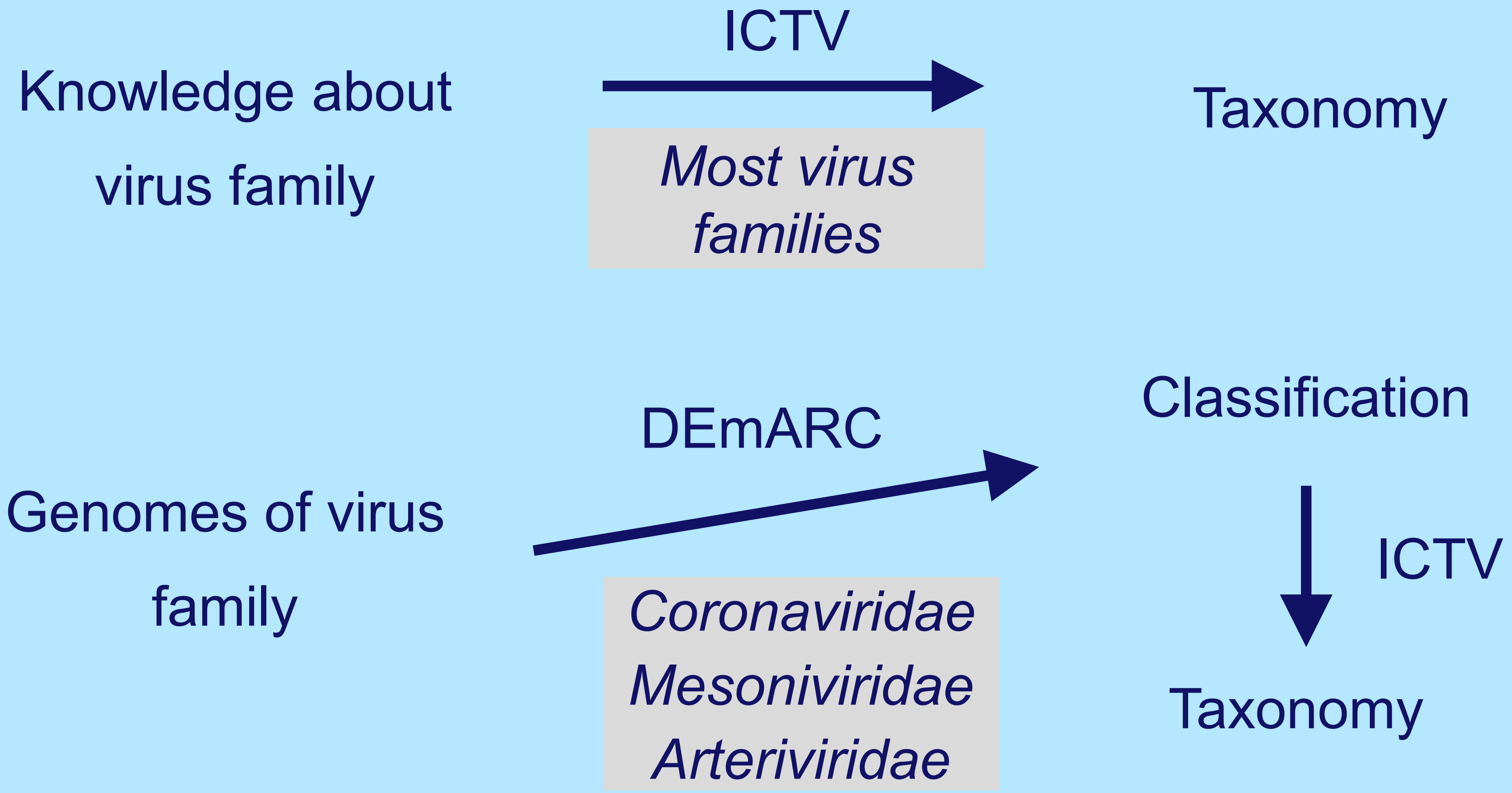
Benefits

- Any virus may be classified

Cost

- Across-taxa consistency is lacking;
- Limited biological insight

Traditional & Genome-based Approaches To Virus Taxonomy



Knowledge about virus family

ICTV

Taxonomy

Most virus families

Genomes of virus family

DEmARC

Classification

ICTV

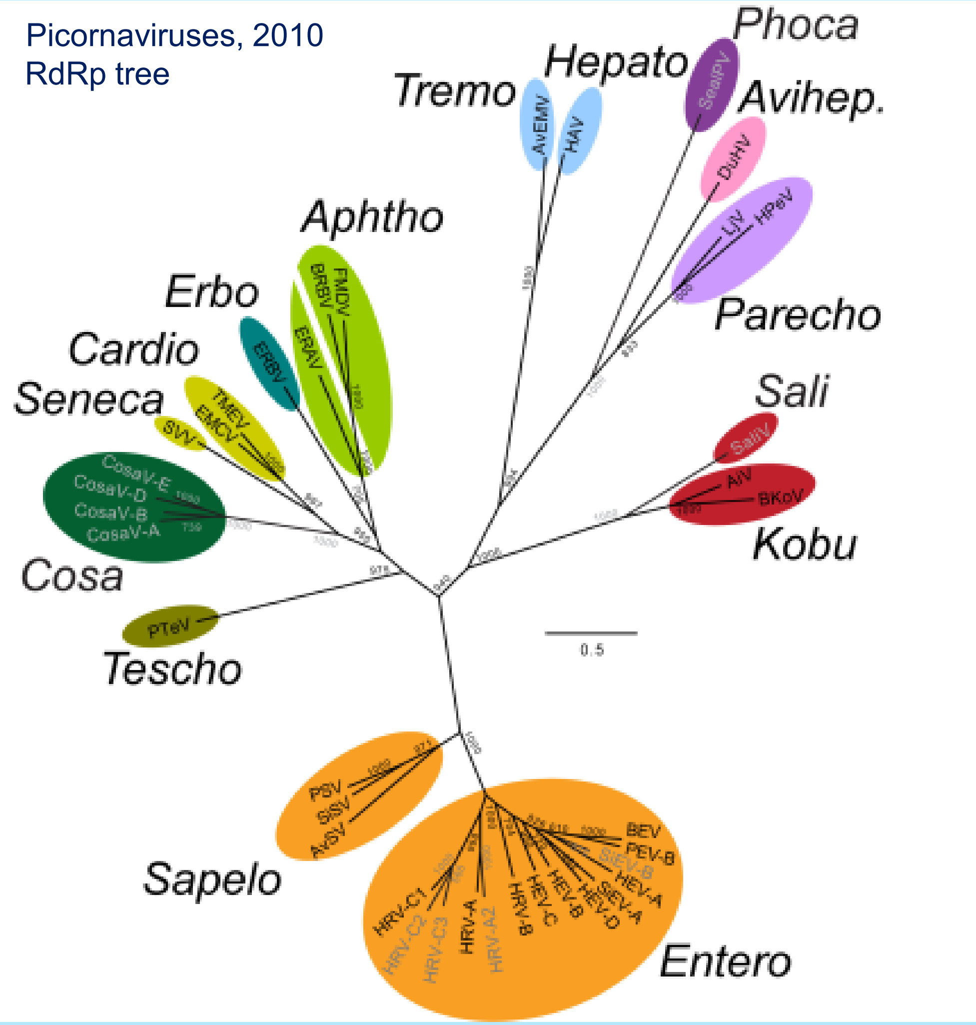
Taxonomy

Coronaviridae
Mesoniviridae
Arteriviridae

Critical factors in decision making:

Quantity & quality of our knowledge (sequences)
Approach to process information

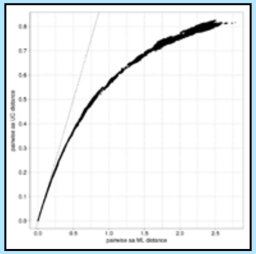
Virus Taxonomy: experts define clusters (taxa) in phylogeny

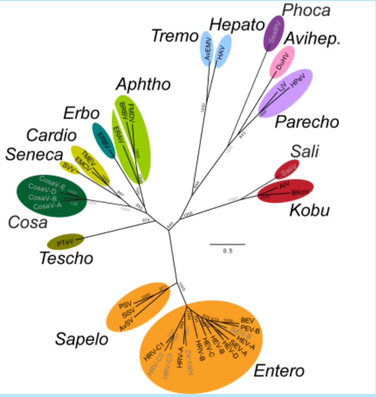


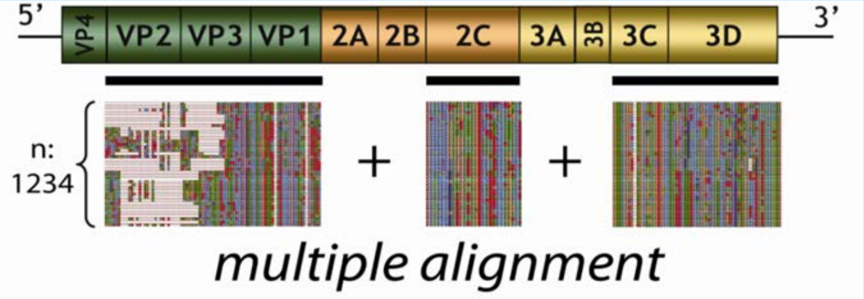
Lauber & Gorbalenya (2012) JVI, 86: 3905

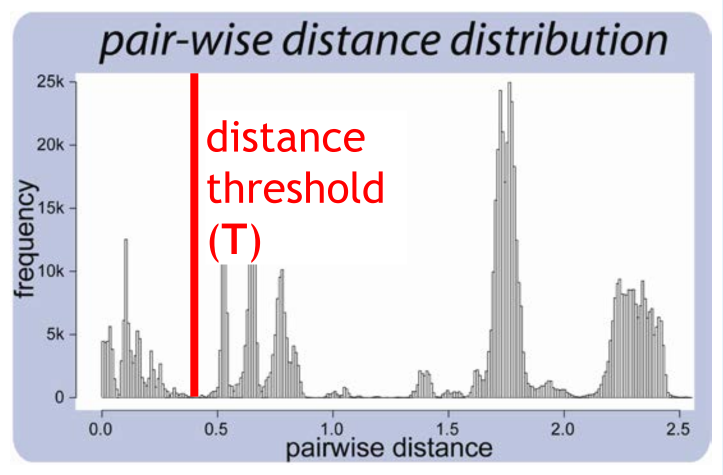
Could clusters be defined in a consistent quantitative manner across the entire family?

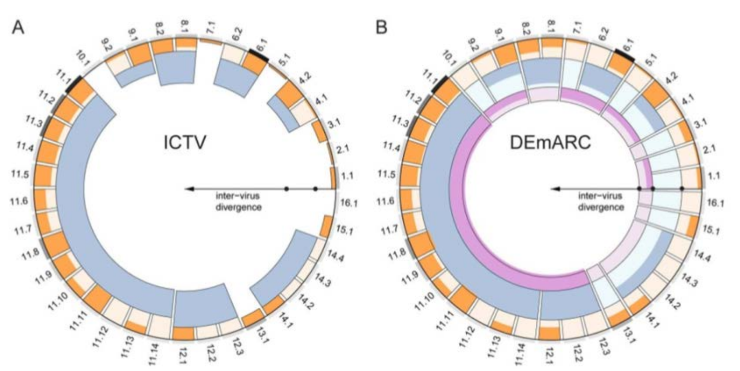
*DEmARC – **D**iv**E**rsity p**A**rtitioning by hie**R**archical **Cl**ustering*









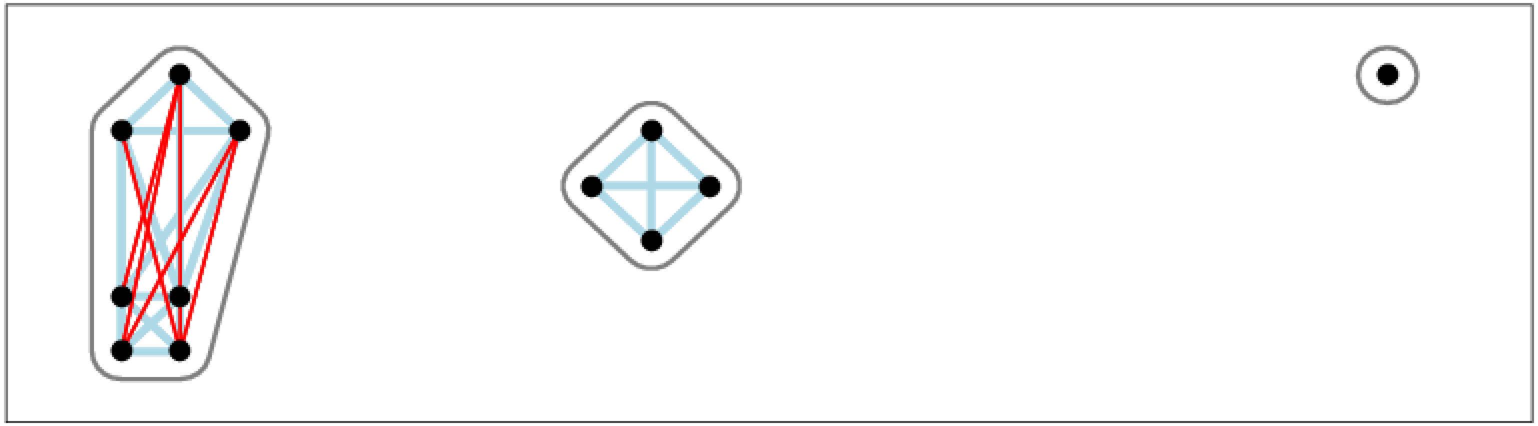


D>T, 1

D<T, 0

Lauber & Gorbalenya (2012)
JVI, 86: 3890 & 3905

DEmARC: cost function for defining rank demarcation



$$\mathrm{CC}(t) = \sum_{e_v} \frac{l(e_v) - t}{t}$$

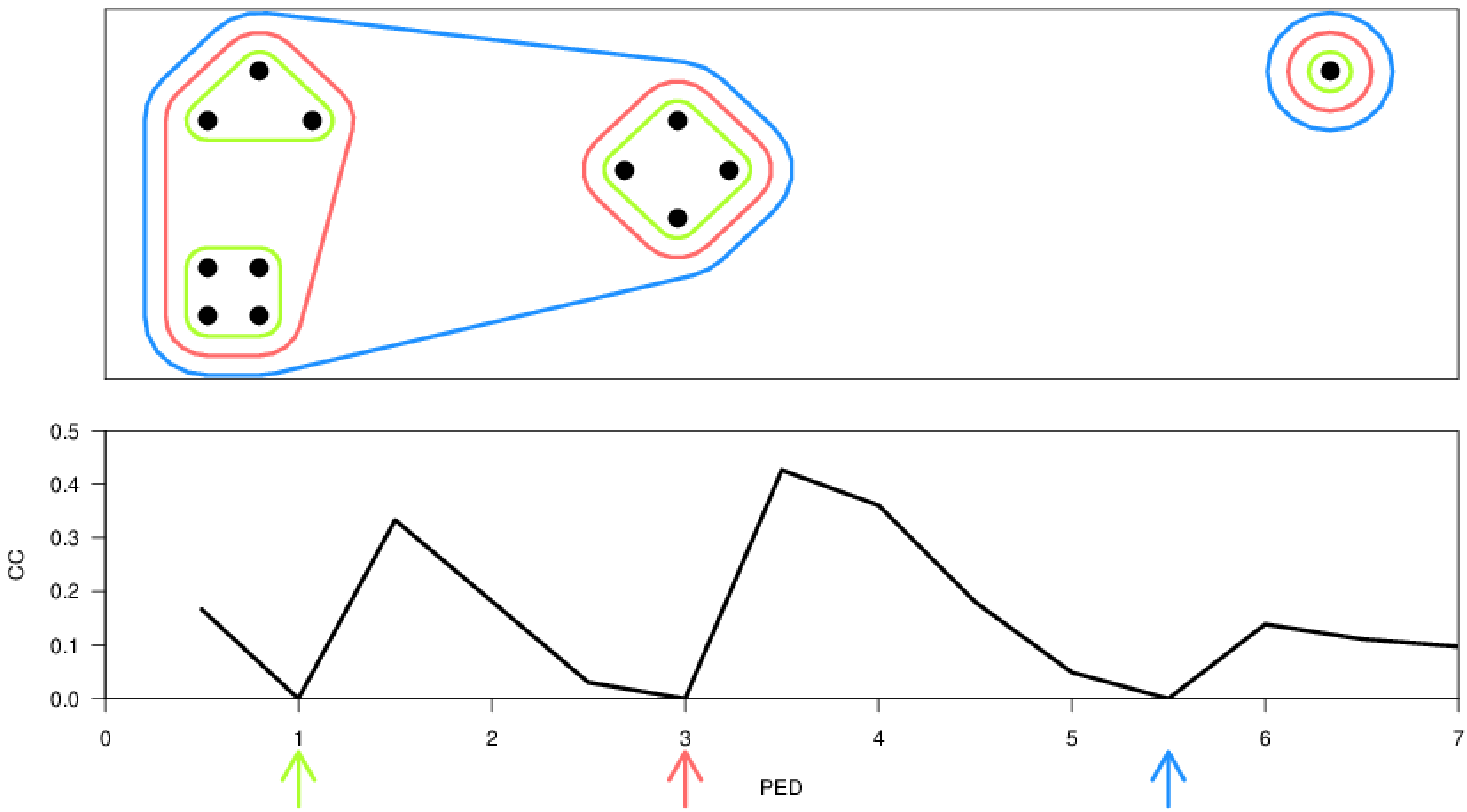
●	sequence
—	PED ≤ threshold(*en*)
○	cluster
—	PED > threshold(*ev*)

t − threshold
l(e) − edge length

Lauber & Gorbalenya, *J. Virology* **86(7)**, 3890–3904 (2012).

“movie” by A. Gulyaeva

DEmARC: cost function for defining rank demarcation



Lauber & Gorbalenya*, J. Virology* **86(7)**, 3890–3904 (2012).

“movie” by A. Gulyaeva

DEmARC: major variables of pipeline

Dataset (April 2014):

Viruses: all full genome sequenced (**720**)

Proteins: conserved ns and str proteins (**20**)

Alignment: curated Muscle/Clustal /HMMER

Processing (DEmARC 1.3):

Measure: pairwise ML evolutionary distance

Clustering: Single-linkage

Criterion: minimization of weighted violation of clustering

Thresholds: ranking by cost and “persistence”

Post-processing:

Quality controls: clustering & tree topology violations

Shown is example of devising coronavirus taxonomy; Lauber et al Gorbalenya, *in preparation*

DEmARC-produced classification closely recovers the ICTV taxonomy of picornaviruses

match

new

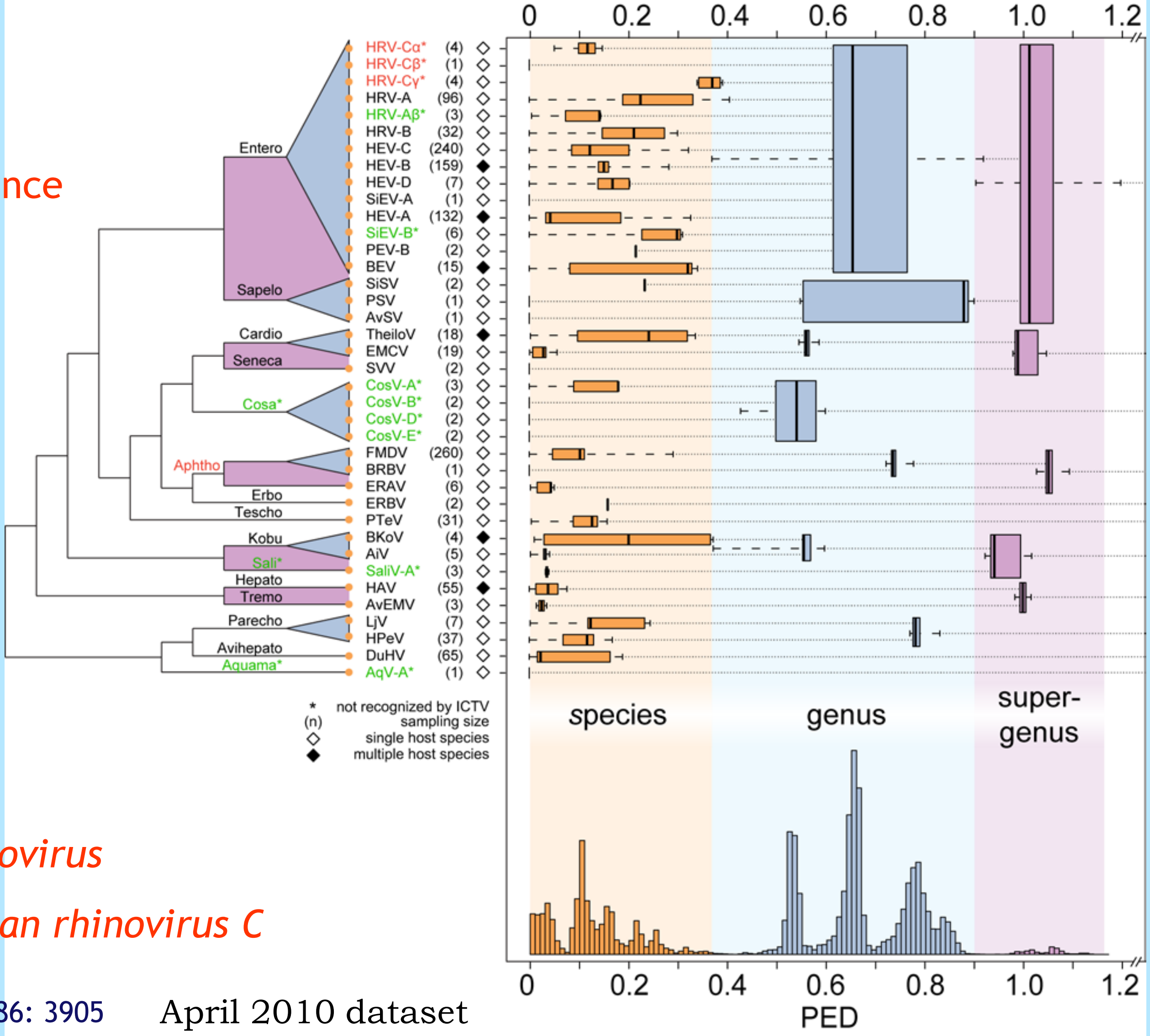
difference

27 out of 28 **Species**

15 out of 16 **Genera**

are recovered

1) “super-genus” level
2) split of genus *Aphthovirus*
3) split of species *Human rhinovirus C*



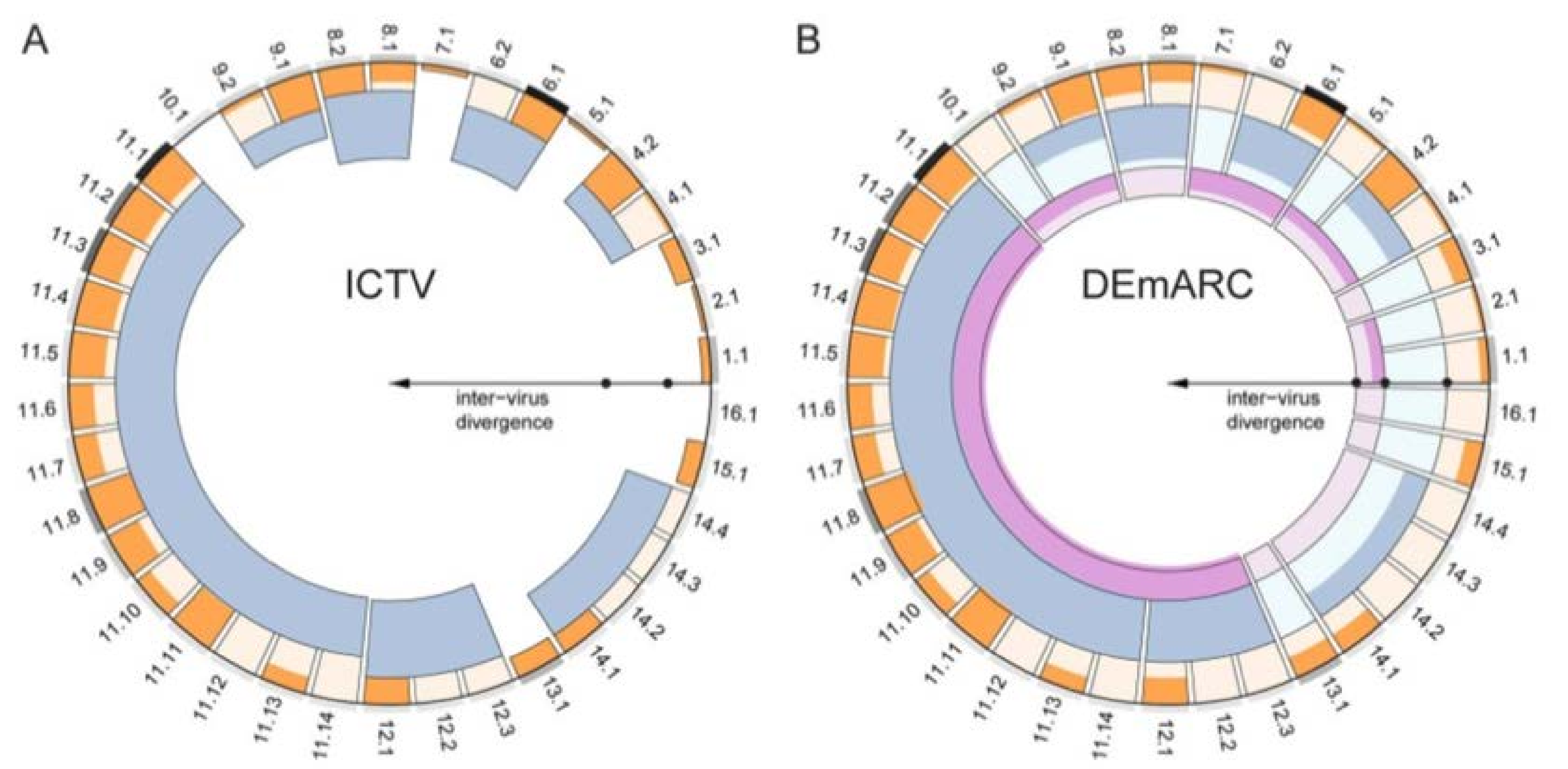
Lauber & Gorbalenya (2012) JVI, 86: 3905

April 2010 dataset

DEmARC-based classification facilitates decision making and offers taxonomy with biological insight

Lauber & Gorbalenya (2012)
JVI, 86: 3905

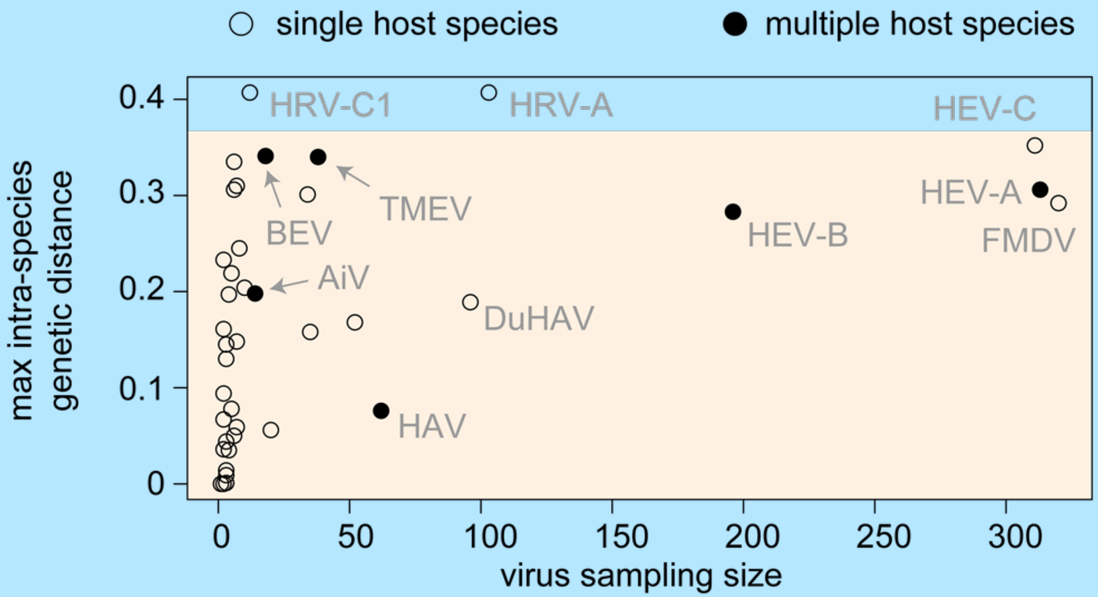
Picornaviruses, 2010



Genus- vs Family-based criteria
Man-made construction vs Evolutionary-based structure

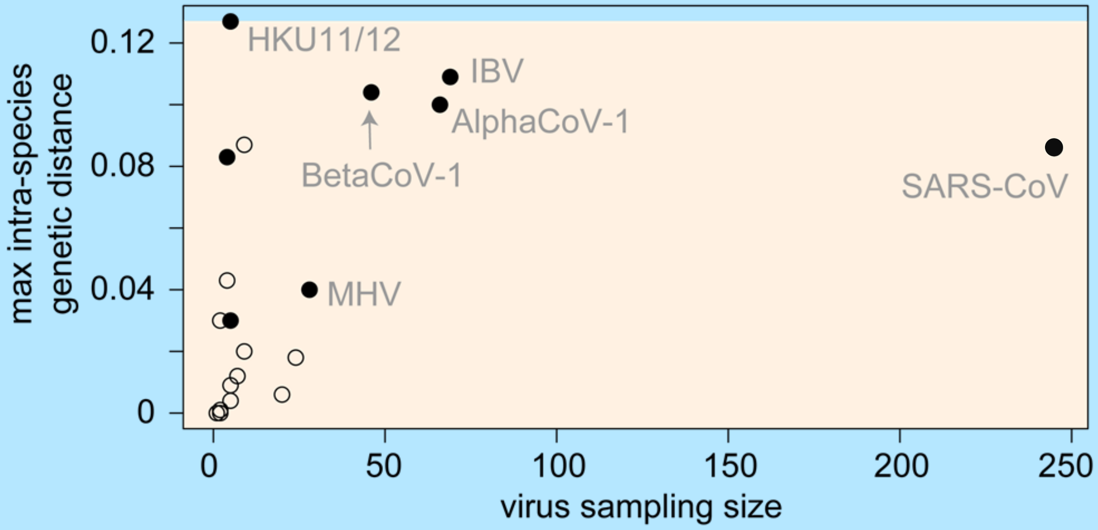
Intra-species divergence: limit, host, pathogenicity

Picornaviruses
51 species
23 genera



April 2013 datasets

Coronaviruses
27 species
4 genera



Lauber & Gorbalenya (2012) JVI, 86: 3905
Lauber et al., & Gorbalenya, in preparation

DEmARC package

- Devises hierarchical classification of genomes
- Classifies *all* viruses in group
- Realized in R

- Includes scripts for
  - processing distance matrices
  - plotting results with biological insight
- Available for download at
  - *http://veb.lumc.nl/VEB/tools.cgi*

- Being advanced
- WEB-site is under construction

Thank You!

THE FUTURE

- Can viruses be classified by sequence data alone?
  Andrew Davison
- Taxonomical assignments, procedures and scales
  Andrew King
- Is it time to consider the adoption of binomial nomenclature?
  Mike Adams
- Databases and bioinformatic tools
  Elliot Lefkowitz

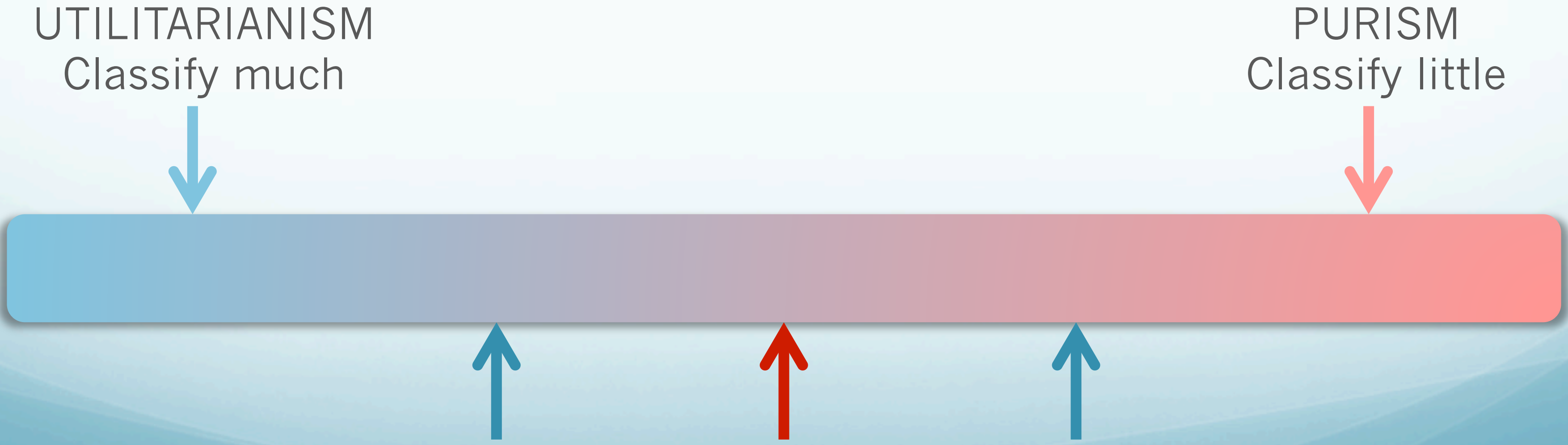
CAN VIRUSES BE CLASSIFIED BY SEQUENCE DATA ALONE?

1966 ICTV 2016

Andrew Davison

Why?

- Not “can” but “should”
- We need to cope with metagenomic data
- We need to remain relevant to the community
- If we don’t do it, someone else will



UTILITARIANISM
Classify much

PURISM
Classify little

Why not?

The species definition

- Previous: “A virus species is defined as a polythetic class of viruses that constitutes a replicating lineage and occupies a particular ecological niche”
- Polythetic: relating to or sharing a number of characteristics which occur commonly in members of a group or class, but none of which is essential for membership of that group or class
- Present: “A species is a monophyletic group of viruses whose properties can be distinguished from those of other species by multiple criteria”
- “Comment: The criteria by which different species within a genus are distinguished shall be established by the appropriate Study Group. These criteria may include, but are not limited to, natural and experimental host range, cell and tissue tropism, pathogenicity, vector specificity, antigenicity, and the degree of relatedness of their genomes or genes.”

Biology emerges fundamentally from the genome

The current stance of the ICTV

- "The EC has agreed to accept species proposals based only on sequence data with certain safeguards. These include evidence that the sequences are effectively complete, that correct assembly has been verified, and that the sequence is indeed viral in origin. Sequences would have to be analysed to provide evidence of taxonomic placement that did not exclusively rely on a simple measure of genetic difference. Also, the EC encourages Study Groups to submit proposals that would make a convincing case for recognizing taxa of any rank on a genome-only basis." (2013)
- "It was agreed to continue the current policy of considering species proposals based on deep sequencing where there was information on the host and sufficient similarity with well-characterized viruses to be confident that the genome was correctly assembled and coding-complete. No distinctive taxon names would be applied by ICTV. The issues involved would be further explored at a special topic meeting funded as part of the Wellcome Trust Grant." (2015)

Reaching a clear position

The role of SGs

- ICTV is a grass-roots organisation
- The EC is reluctant to work in a top-down way, recognising that the needs of virus families vary
- This can make clear positions difficult to reach and agree

Meeting on The Impact Of Metagenomic Sequencing On Viral Classification

- Involving ~12 experts and ~12 EC members
- 9-11 June 2016, Boston
- Publication of proceedings

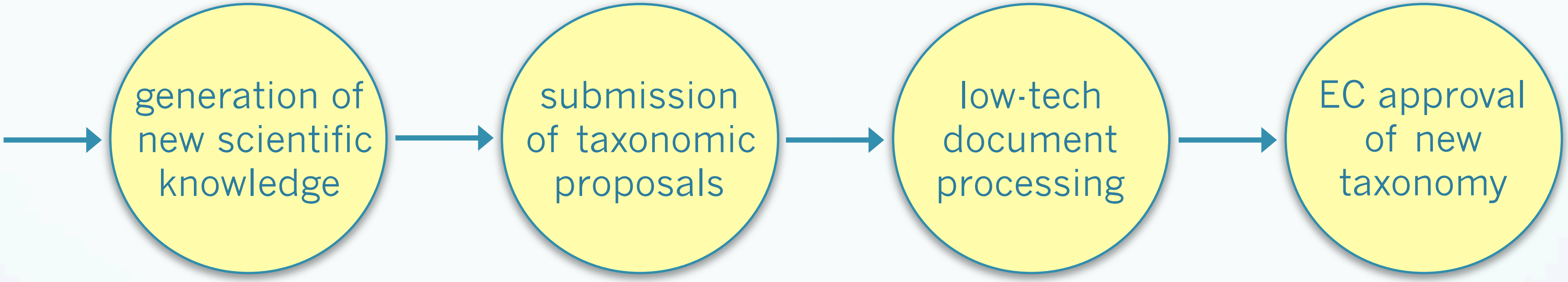
Feeding into the next EC meeting, 22-24 August, Budapest

TAXONOMICAL ASSIGNMENTS, PROCEDURES AND SCALES

1966 ICTV 2016

Andrew King

What limits the rate of development of virus taxonomy?



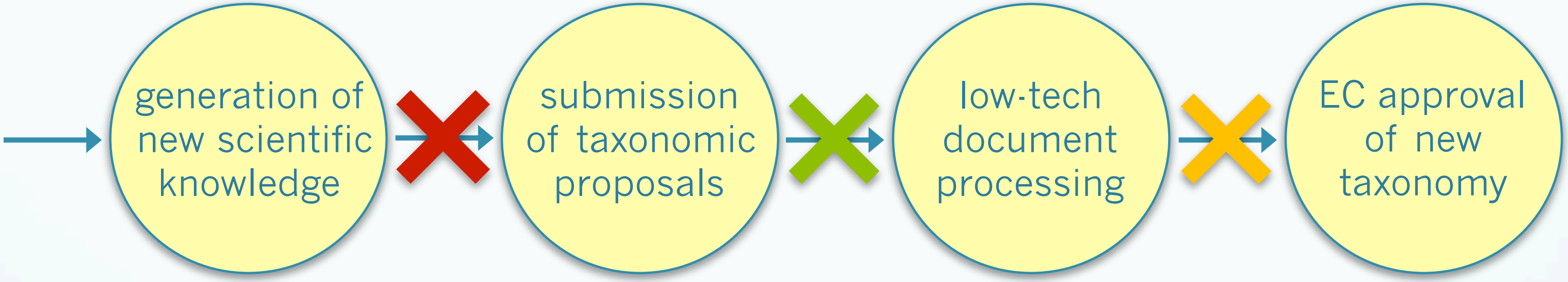
generation of new scientific knowledge

submission of taxonomic proposals

low-tech document processing

EC approval of new taxonomy

Rate-limiting steps in taxonomic development

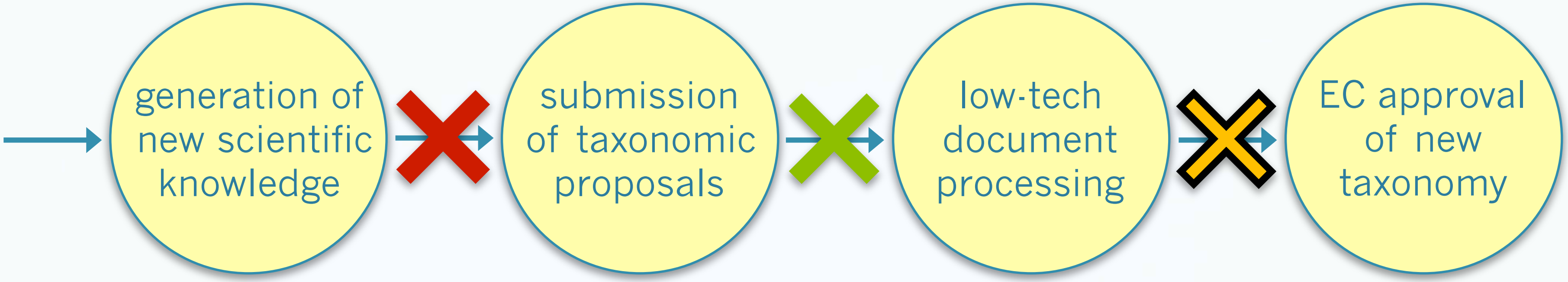


generation of new scientific knowledge

submission of taxonomic proposals

low-tech document processing

EC approval of new taxonomy



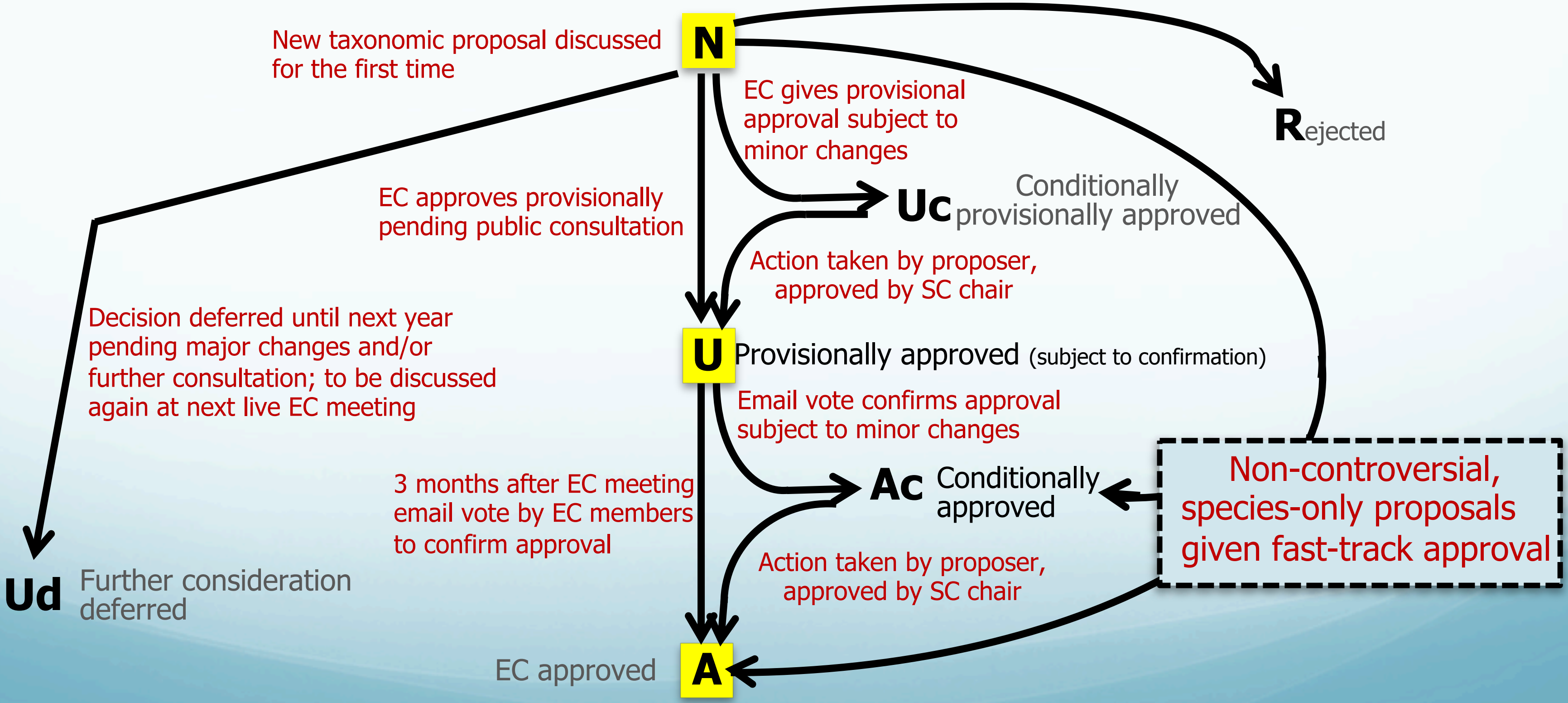
generation of new scientific knowledge

submission of taxonomic proposals

low-tech document processing

EC approval of new taxonomy

EC decision-making: now



N New taxonomic proposal discussed for the first time

Rejected

EC gives provisional approval subject to minor changes

Uc Conditionally provisionally approved

EC approves provisionally pending public consultation

Action taken by proposer, approved by SC chair

Decision deferred until next year pending major changes and/or further consultation; to be discussed again at next live EC meeting

U Provisionally approved (subject to confirmation)

Email vote confirms approval subject to minor changes

3 months after EC meeting email vote by EC members to confirm approval

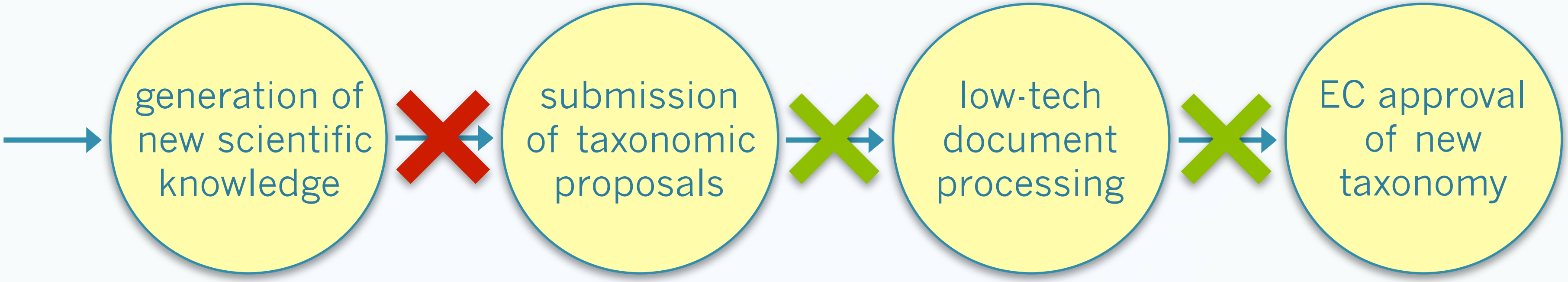
Ac Conditionally approved

Non-controversial, species-only proposals given fast-track approval

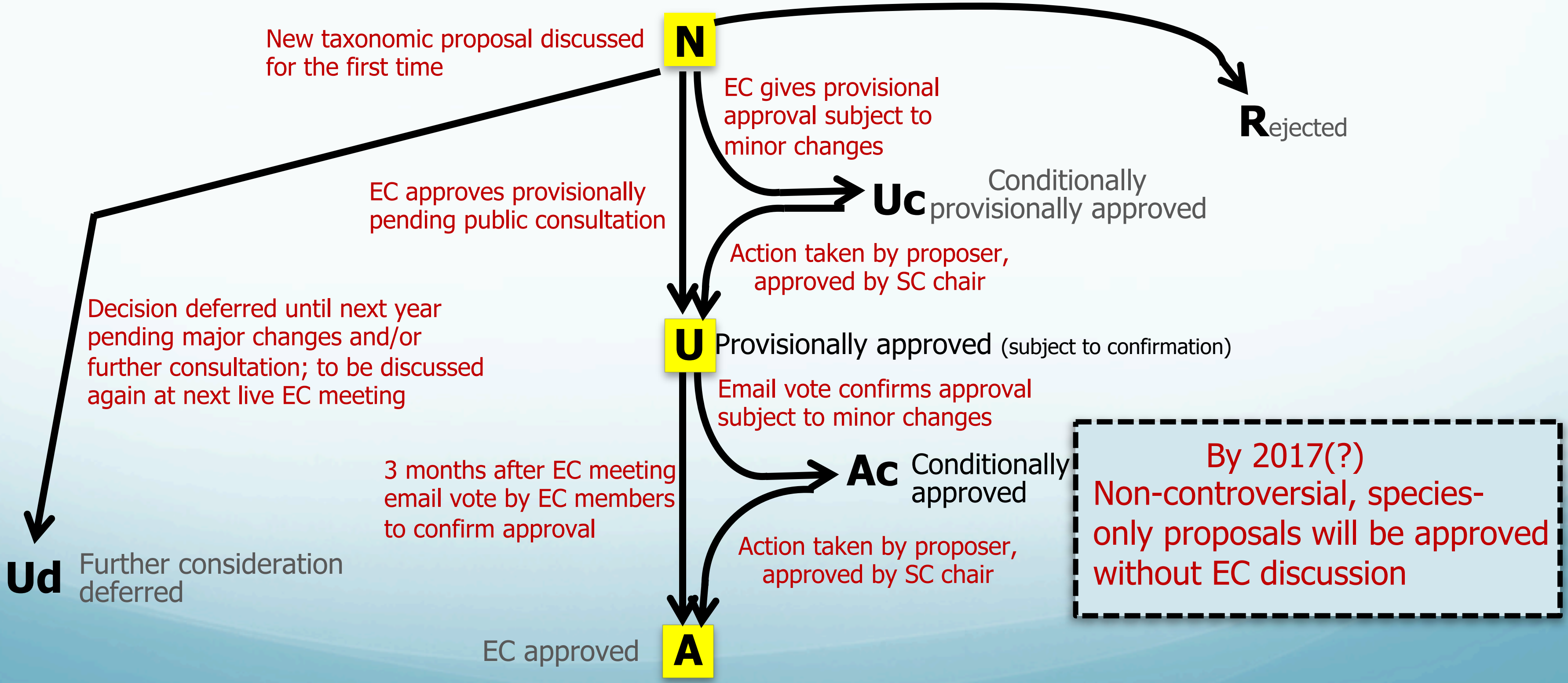
Action taken by proposer, approved by SC chair

Ud Further consideration deferred

EC approved **A**

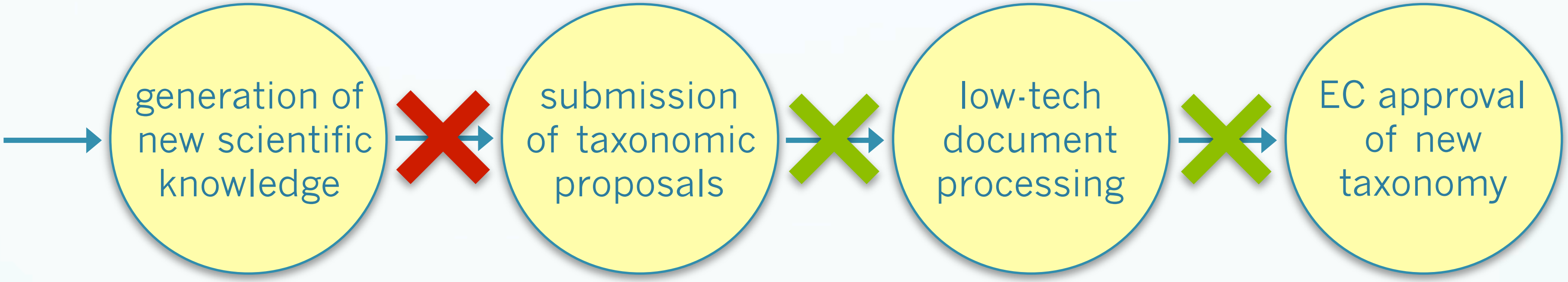


EC decision-making: now



By 2017(?)
Non-controversial, species-only proposals will be approved without EC discussion

Rate-limiting steps in taxonomic development



generation of new scientific knowledge

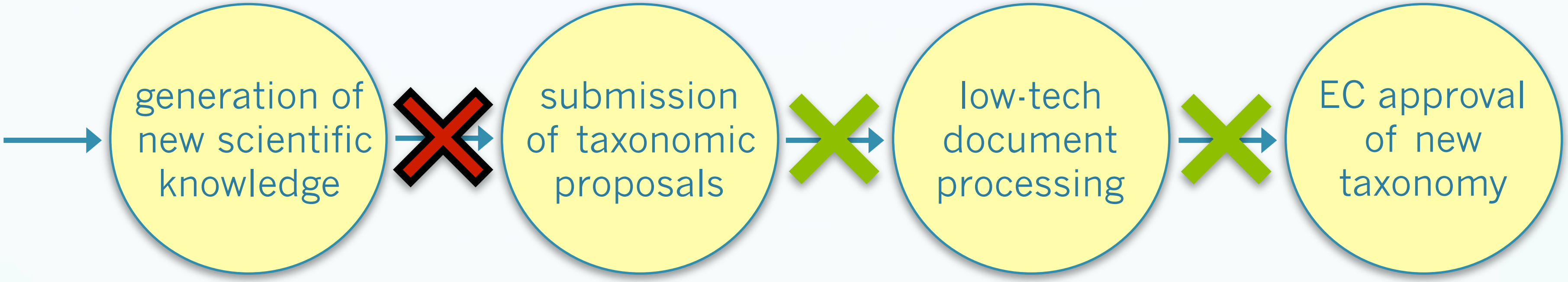
submission of taxonomic proposals

low-tech document processing

EC approval of new taxonomy

By 2017(?), submission of taxonomic proposals will be web-based

Rate-limiting steps in taxonomic development



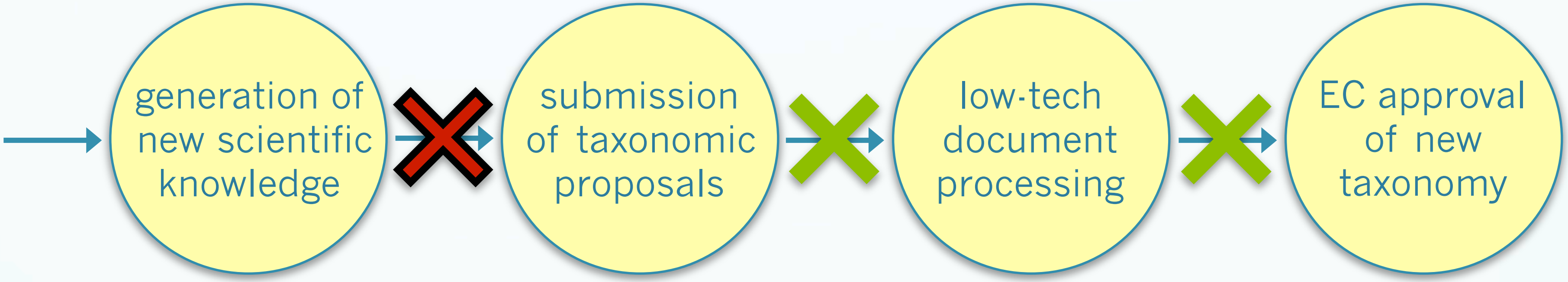
generation of new scientific knowledge

submission of taxonomic proposals

low-tech document processing

EC approval of new taxonomy

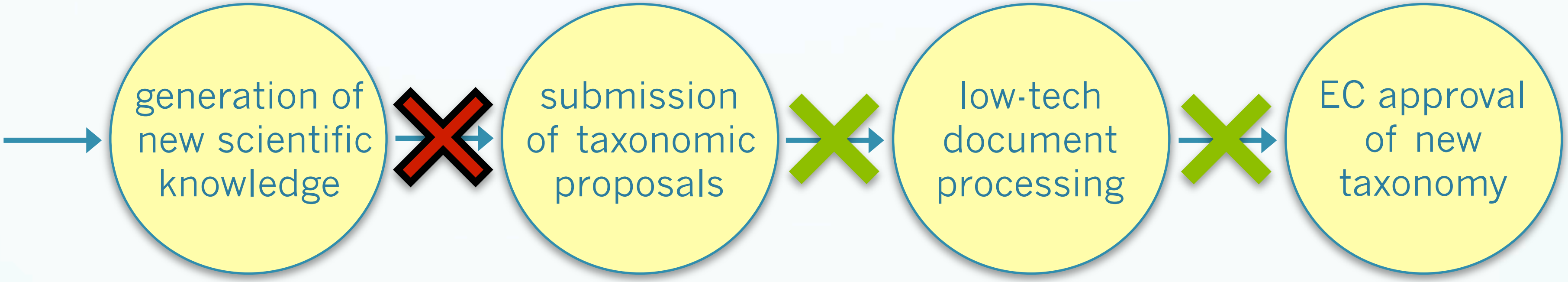
Rate-limiting steps in taxonomic development



Taxonomic proposals: TP template is complex, prescriptive and demanding

How to
- make proposing less of a deterrent?
- give more credit to authors of proposals?
- encourage/help SGs in other ways to submit TPs?

Rate-limiting steps in taxonomic development

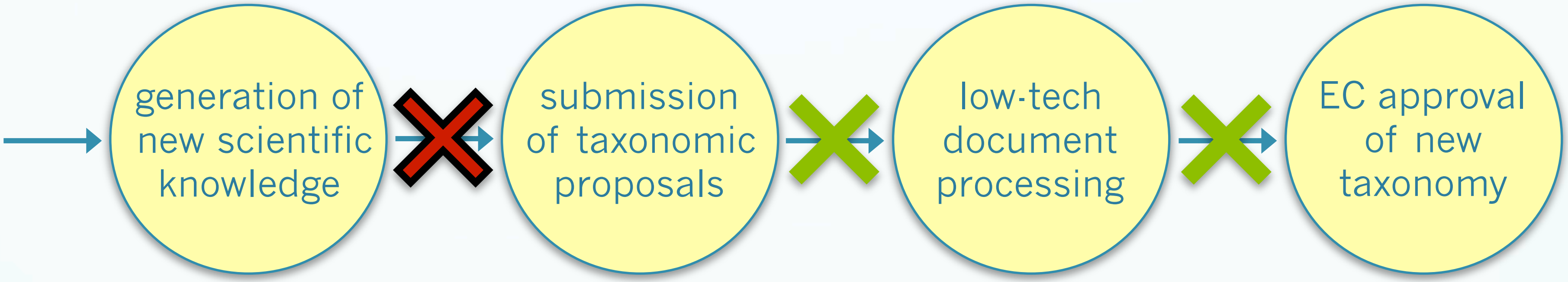


Taxonomic proposals: TP template is complex, prescriptive and demanding

How to
- make proposing less of a deterrent?
- give more credit to authors of proposals?
- encourage/help SGs in other ways to submit TPs?

Need for
- more SGs to represent the ICTV in new areas

Rate-limiting steps in taxonomic development



generation of new scientific knowledge

submission of taxonomic proposals

low-tech document processing

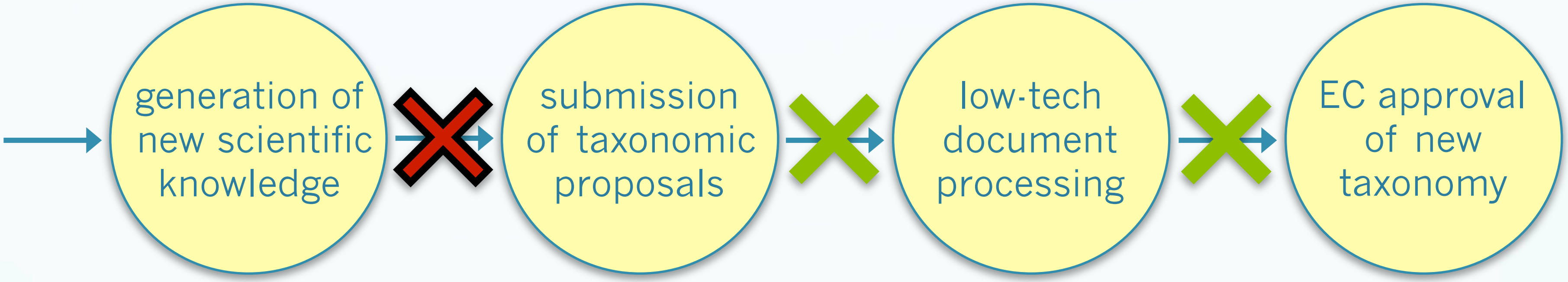
EC approval of new taxonomy

Help for SGs

Online tools:

- Reference sequence alignments
- VicTree for identifying taxonomic cut-offs
- Wikis for sharing information within and between SCs

Rate-limiting steps in taxonomic development



generation of new scientific knowledge

submission of taxonomic proposals

low-tech document processing

EC approval of new taxonomy

Help for SGs

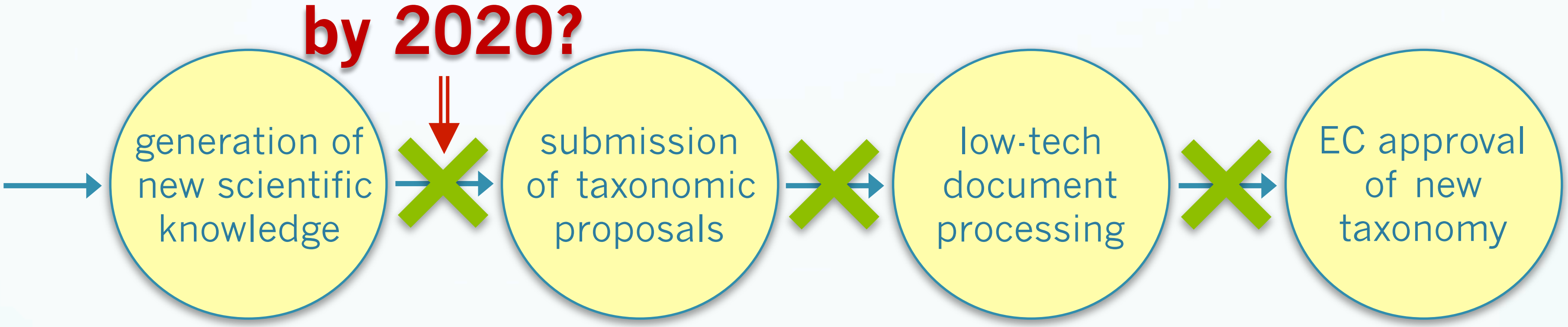
Online tools:

- Reference sequence alignments
- VicTree for identifying taxonomic cut-offs
- Wikis for sharing information within and between SCs

The human touch:

- Better succession planning and mentoring

Rate-limiting steps in taxonomic development



by 2020?

generation of new scientific knowledge

submission of taxonomic proposals

low-tech document processing

EC approval of new taxonomy

Help for SGs

Online tools:

- Reference sequence alignments
- VicTree for identifying taxonomic cut-offs
- Wikis for sharing information within and between SCs

The human touch:

- Better succession planning and mentoring

IS IT TIME TO CONSIDER THE ADOPTION OF BINOMIAL NOMENCLATURE?

1966 ICTV 2016

Mike Adams

The confusion in nomenclature (1)

Differences in the styles of virus species names:

Mycobacterium phage TM4	
Human herpesvirus 1	Order
Bovine ephemeral fever virus	
Mammalian 1 bornavirus	Genus
Potato yellow dwarf virus	
Alphacoronavirus 1	Genus
Cardiovirus A	Genus
Potato virus X	
Rhizosolenia setigera RNA virus 01	
Human mastadenovirus C	Genus
Autographa californica multiple nucleopolyhedrovirus	?
Drosophila X virus	
Sapporo virus	
Melolontha melolontha entomopoxvirus	Subfamily

The confusion in nomenclature (2)

The virus versus species issue:

Andes virus

Andes virus	(ANDV)
Bermejo virus	(BMJV)
Lechiguanas virus	(LECV)
Maciel virus	(MCLV)
Oran virus	(ORNV)
Pergamino virus	(PRGV)

If these are biologically distinct,
they are strains or subspecies of
Andes virus

Bayou virus

Bayou virus	(BAYV)

Black Creek Canal virus

- etc.

The latinized binomial (Linnean) system

- Used almost universally in biology and understood by scientists, editors, etc., many of whom have not really understood the current ICTV system
- Would make integration of virus data with other biological databases much easier
- Would clearly distinguish the scientific name from the current names, which would become the common (vernacular) names; this distinction is easily understood and easily applied
- Would clearly be universal, naming viruses of prokaryotes, plants, vertebrates, etc., in exactly the same way

A possible way forward

Hantavirus andiensis

Andes virus	(ANDV)	[ssp. *andiensis*?]
Bermejo virus	(BMJV)	[ssp. *bermejenis*?]

Bymovirus hordei-luteus

Scientific name (international)

- barley yellow mosaic virus (English)
- Gerstengelbmosaik-virus (German)
- mosaïque jaune de l'orge (French)
- etc.

Vernacular names (local)

Of course, there are problems...

These would include:

- Names of genera and above would not be affected, but the names of all species would change
- It could not sensibly be introduced gradually
- As with all changes, it would likely encounter lively opposition
- Current species that are assigned to a family or subfamily but not to a genus would pose a difficulty (but there are not many...)

If there is a reasonable consensus:

- It should be done sooner rather than later
- ICTV could encourage the use of the prefix '*Candidatus* species' in publications referring to probable new species that had not yet been officially approved

DATABASES AND BIOINFORMATIC TOOLS

1966 ICTV 2016

Elliot Lefkowitz

What can the ICTV do to help?

- Discovery
- Annotation
- **Classification**
- **Taxonomic assignment**

Classification and taxonomic assignment

Publish unambiguous species demarcation criteria (SDC)

- Demarcation criteria determine approach and tools necessary for classification and taxonomic assignment
- Responsibility of each SG
- Publish on the ICTV website
- Link from the ICTV Report chapter

Provide classification tools and guidance

Requirements: taxon-specific criteria

- Morphological
- Physical
- Structural
- Biological
- Genomic organisation
- Sequence

Taxon-specific criteria

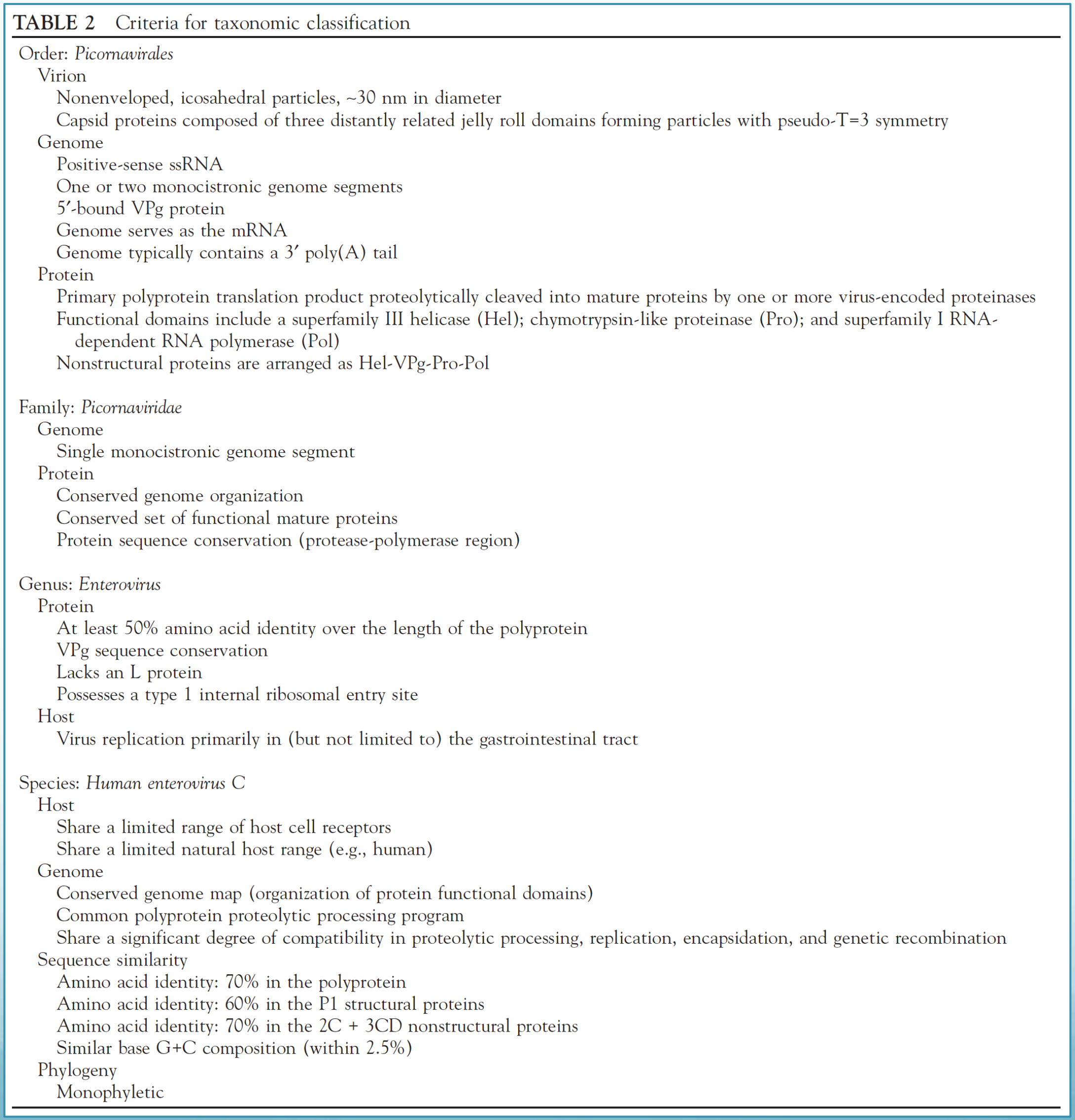


TABLE 2 Criteria for taxonomic classification

- Order: *Picornavirales*
  - Virion
    - Nonenveloped, icosahedral particles, ~30 nm in diameter
    - Capsid proteins composed of three distantly related jelly roll domains forming particles with pseudo-T=3 symmetry
  - Genome
    - Positive-sense ssRNA
    - One or two monocistronic genome segments
    - 5′-bound VPg protein
    - Genome serves as the mRNA
    - Genome typically contains a 3′ poly(A) tail
  - Protein
    - Primary polyprotein translation product proteolytically cleaved into mature proteins by one or more virus-encoded proteinases
    - Functional domains include a superfamily III helicase (Hel); chymotrypsin-like proteinase (Pro); and superfamily I RNA-dependent RNA polymerase (Pol)
    - Nonstructural proteins are arranged as Hel-VPg-Pro-Pol

- Family: *Picornaviridae*
  - Genome
    - Single monocistronic genome segment
  - Protein
    - Conserved genome organization
    - Conserved set of functional mature proteins
    - Protein sequence conservation (protease-polymerase region)

- Genus: *Enterovirus*
  - Protein
    - At least 50% amino acid identity over the length of the polyprotein
    - VPg sequence conservation
    - Lacks an L protein
    - Possesses a type 1 internal ribosomal entry site
  - Host
    - Virus replication primarily in (but not limited to) the gastrointestinal tract

- Species: *Human enterovirus C*
  - Host
    - Share a limited range of host cell receptors
    - Share a limited natural host range (e.g., human)
  - Genome
    - Conserved genome map (organization of protein functional domains)
    - Common polyprotein proteolytic processing program
    - Share a significant degree of compatibility in proteolytic processing, replication, encapsidation, and genetic recombination
  - Sequence similarity
    - Amino acid identity: 70% in the polyprotein
    - Amino acid identity: 60% in the P1 structural proteins
    - Amino acid identity: 70% in the 2C + 3CD nonstructural proteins
    - Similar base G+C composition (within 2.5%)
  - Phylogeny
    - Monophyletic

Requirements: sequence similarity

Sequence database

- Comprised of all ICTV species
- Defined genomic region
- Nucleic acid/protein
- Dependent on SDC

Requirements: pairwise comparison

Define a tool and parameters

- PASC
  - PAirwise Sequence Comparison
  - NCBI
- DEmARC
  - DivErsity pArtitioning by hieRarchical Clustering
  - Leiden University Medical Center, Leiden, The Netherlands
- Species Demarcation Tool
  - University of Western Cape, Cape Town, SA
- Other

Requirements: alignments

- Provide existing multiple sequence alignment (MSA) of approved isolate sequences and genomic region

- Provide guidelines for aligning new sequences to the MSA

Requirements: phylogenetic analysis

Algorithm

- e.g. Bayesian

Tool

- e.g. MrBayes

Parameters

- Substitution model

Visualisation

- Rooted phylogram

MEETING SUMMARY

1966 ICTV 2016

Andrew Davison & Peter Simmonds

To take forward

- Diversity within the ranks of the ICTV
- A more significant role for the National Representatives
- Educating the scientific community, especially journal editors, on the importance of taxonomy and the role of the ICTV
- Accept that the ICTV will always be behind the curve
- SGs have the option of using linked wikipages in the new online Report to at least provide listings (and perhaps other data) on nascent information
- Proposing new species needs to be made easier
- 50th anniversary publication incorporating information on the present meeting
- Mentoring of SG chairs
- Classifying viruses from metagenomic data
- Availability of sequence alignments and tools

Meeting summary

Discussion sessions

- Major concern about impact of metagenomic sequences
- Virus classification methods, many differences of opinions
- Are we doing virus classification or are we doing virus evolution?
- Generalism versus pragmatism
- Usability of current ICTV resources, taxonomy proposals

Short-term plans

- Distribution of meeting summary to all SG chairs and EC members
- Request for further opinions about ICTV developments
- Online Report, summaries and publication strategy
- Bioinformatic links and resources

Taking this forward in the medium term

- Templates and requirements for chapter updates and summaries
- Development of wiki pages for SGs, further examples
- Specification for standardised sequence alignments

THANK YOU AND BON VOYAGE!