Animal Health Service AGAH

Swiss Institute of Bioinformatics

From OPENFLU to OPENFMD

a resource for automatic and curated nomenclature and tools for the FMD community

Filip Claes, Philippe Le Mercier, Dmitry Kuznetsov, Robin Liechti, Anne Gleizes, Ioannis Xenarios, Gwenaelle Dauphin

Food and Agriculture Organization of the United Nations
Animal Production and Health Division

Animal Health Service AGAH

Conclusions and recommendations

- Complete collection of FMDV (genetic) data is necessary for:
  - Surveillance of disease evolutionary trends
  - Scientific analysis
  - Infer gaps in epidemiological data
- OPENFLU-EMPRESi experience may be useful for FMDV community
- OPENFMD might become a central portal for the community to store and share FMDV data - and move towards harmonized nomenclature

Food and Agriculture Organization of the United Nations
Animal Production and Health Division

Animal Health Service AGAH

OpenFLU - an open database of influenza data

- http://openflu.vital-it.ch/
- Contains only public data
- All data are accessible for browsing
- All users are welcome
  - standard free web registration is needed to upload and manage own or laboratory data

Food and Agriculture Organization of the United Nations
Animal Production and Health Division

Animal Health Service AGAH

OpenFLU: an adequate data model for Influenza virus

- Sequences grouped per isolates
- Isolates accompanied by a comprehensive record of data for epidemiology:
  - Minimal data required: year, host species, country
  - Support for much more data: vaccination, patient status, resistance to antibiotics, animal domestication status, etc…
- Influenza-specific genetic knowledge and algorithms
  - Translation to viral proteins
  - Key mutations

Food and Agriculture Organization of the United Nations
Animal Production and Health Division

Animal Health Service AGAH

OpenFLU: data import/export

- Import:
  - GENBANK
  - USER SUBMISSIONS
- CURATION !
- Export: GenBank, data as a structured comment

Food and Agriculture Organization of the United Nations
Animal Production and Health Division

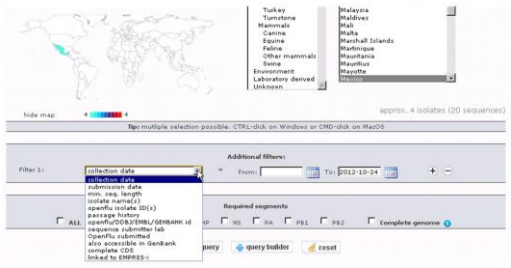
Animal Health Service AGAH

OpenFLU: easy to find your isolate

Food and Agriculture Organization of the United Nations
Animal Production and Health Division

Animal Health Service AGAH

OpenFLU: more search criteria



Food and Agriculture Organization of the United Nations
Animal Production and Health Division

Animal Health Service AGAH

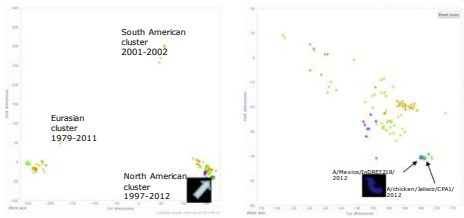
OpenFlu: analysis of the data found

Multiple sequence alignment, BLAST, phylogenetic tree, SSM

Food and Agriculture Organization of the United Nations
Animal Production and Health Division

Animal Health Service AGAH

OpenFLU example: SSM for H7N3

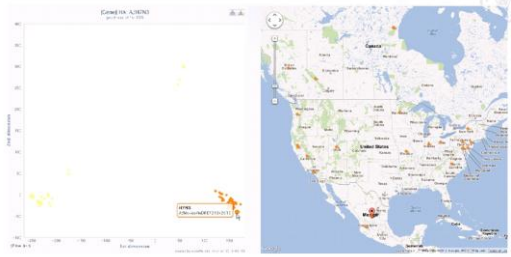


- Each sequence is a point on the map
- Similar sequences are close to each other

Food and Agriculture Organization of the United Nations
Animal Production and Health Division

Animal Health Service AGAH

OpenFLU example: SSM for H7N3



- Zoom into a geo region also selects and highlights corresponding sequences on SSM

Food and Agriculture Organization of the United Nations
Animal Production and Health Division

Animal Health Service AGAH

EMPRES-i Genetic Module

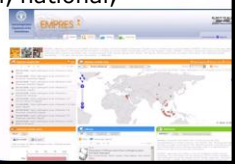
Linking epi and genetics

Food and Agriculture Organization of the United Nations
Animal Production and Health Division

Animal Health Service AGAH

Epidemiological information

- On outbreaks: date, location, species,…
- On surveillance
- Generated by field teams the day of the event, compiled and reported by vet services (epi unit) within a short time
- Stored in databases: local, national, regional, global



Food and Agriculture Organization of the United Nations
Animal Production and Health Division

Animal Health Service AGAH

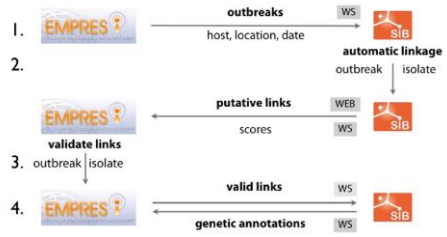
Virus information

- Generated by labs:
  - Detection of influenza virus, subtyping, RNA sequencing
  - Can take a few days (more)
- Sequence information deposited in a database
  - Genbank: universal repository
  - Other genetic databases, eg influenza specialized

Food and Agriculture Organization of the United Nations
Animal Production and Health Division

Animal Health Service AGAH

How does it work



Food and Agriculture Organization of the United Nations
Animal Production and Health Division

Animal Health Service AGAH

Food and Agriculture Organization of the United Nations
Animal Production and Health Division

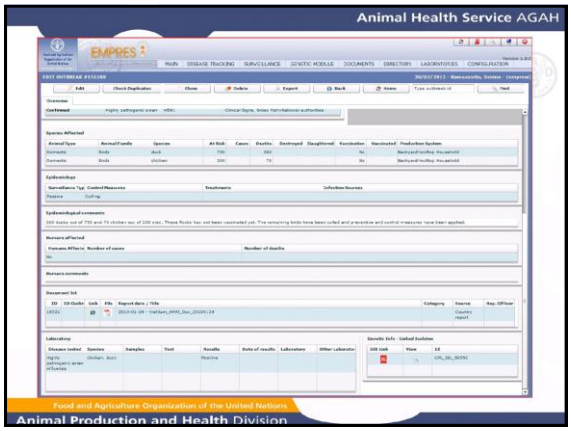
Animal Health Service AGAH

Food and Agriculture Organization of the United Nations
Animal Production and Health Division

Animal Health Service AGAH

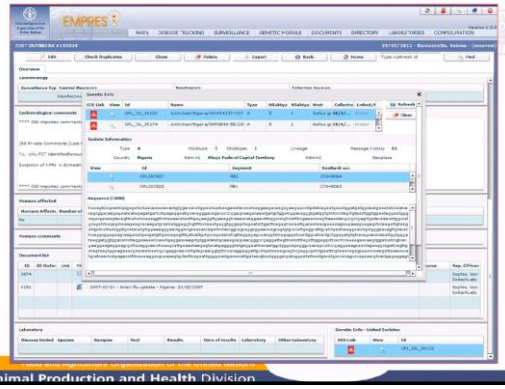
Food and Agriculture Organization of the United Nations
Animal Production and Health Division

Animal Health Service AGAH



Food and Agriculture Organization of the United Nations
Animal Production and Health Division

Animal Health Service AGAH



Animal Production and Health Division

Animal Health Service AGAH

OpenFMD: the project

- A public database that will
  - Store FMDV sequences and epidemiological data from various sources
    - Import from public repositories (like GenBank)
    - Direct user submission
  - Verify data consistency (e.g. control of user-submitted serotype by sequence similarity vs. collections of reference strain sequences
  - Eventually manually curate records
  - Analyse data using tools like phylo-trees or SSM
  - Serve all the curated and organized data together with analysis results back to the FMDV community
  - Towards a harmonized nomenclature (with parallel current names)

Food and Agriculture Organization of the United Nations
Animal Production and Health Division

Animal Health Service AGAH

OpenFMD: a live DB

- To be daily synchronized with all data sources
- GenBank (EFetch) allows data export in computer-friendly form (XML) suitable for daily incremental updates
- Precision of epidemiological data varies, but GenBank team collects and merges author's updates
- Authors might also correct/update data directly
- Already >4000 (full and partial) sequences available in GenBank

Food and Agriculture Organization of the United Nations
Animal Production and Health Division

Animal Health Service AGAH

OpenFMD: a sandbox trial

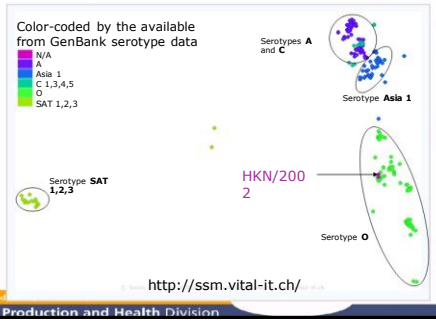
- FMDV specific data model
- Simplified partial GenBank import
- 257 full genome sequences
- All 7 serotypes
- 2 un-typed sequences
  - one is "HKN/2002"

SEROTYPE	SEROSUBTYPE	CNT
A		72
Asia 1		41
C	1,3,4,5	17
O		113
SAT	1	10
SAT	2	5
SAT	3	5
-N/A-		2

Food and Agriculture Organization of the United Nations
Animal Production and Health Division

Animal Health Service AGAH

OpenFMD: SSM try-out

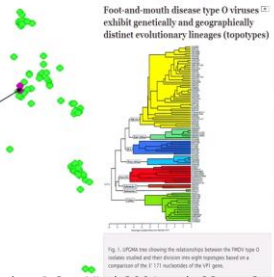


Animal Production and Health Division

Animal Health Service AGAH

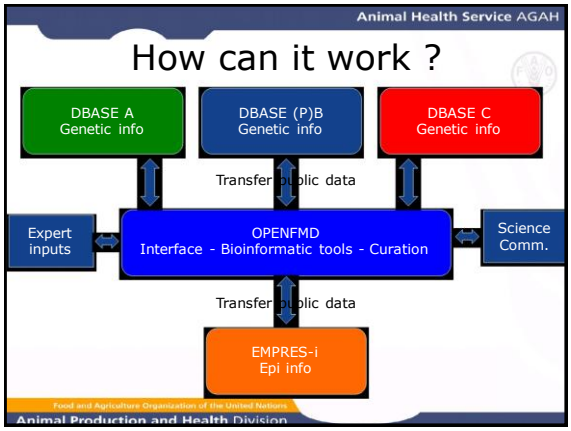
OpenFMD: a closer look on the map for the type O

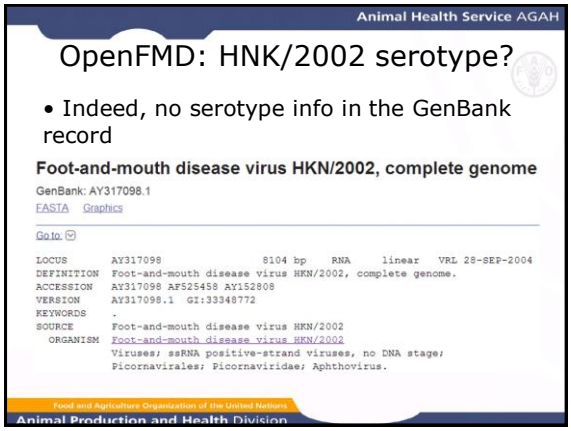
- Clusters (topotypes)
- No more precise annotation in the GenBank records
- No serotype for "HKN/2002" (pink)
- The serotype can be inferred from its neighbors, i.e. "O"

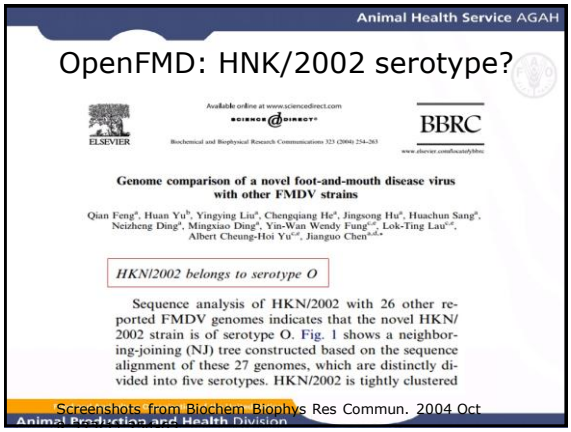


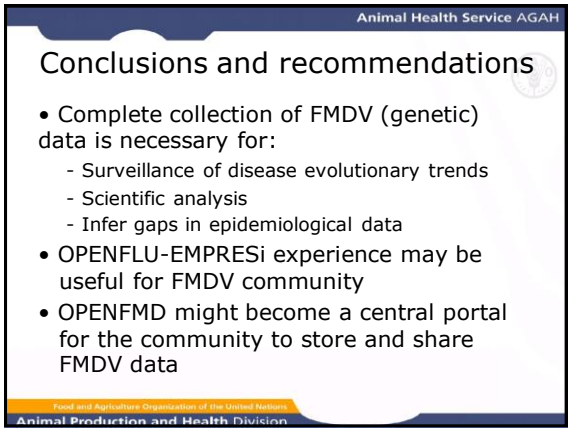
The phylogenetic tree from Samuel and Knowles, J Gen Virol 2001, vol. 82 no. 3 (609-621)

Food and Agriculture Organization of the United Nations
Animal Production and Health Division


Animal Health Service AGAH
How can it work ?
DBASE A
Genetic info
DBASE (P)B
Genetic info
DBASE C
Genetic info
Transfer public data
Expert
inputs
OPENFMD
Interface - Bioinformatic tools - Curation
Science
Comm.
Transfer public data
EMPRES-i
Epi info
Food and Agriculture Organization of the United Nations
Animal Production and Health Division


Animal Health Service AGAH
OpenFMD: HNK/2002 serotype?
• Indeed, no serotype info in the GenBank record
Foot-and-mouth disease virus HKN/2002, complete genome
GenBank: AY317098.1
FASTA Graphics
Go to:
LOCUS AY317098 8104 bp RNA linear VRL 28-SEP-2004
DEFINITION Foot-and-mouth disease virus HKN/2002, complete genome.
ACCESSION AY317098 AF525458 AY152808
VERSION AY317098.1 GI:33348772
KEYWORDS .
SOURCE Foot-and-mouth disease virus HKN/2002
ORGANISM Foot-and-mouth disease virus HKN/2002
Viruses; ssRNA positive-strand viruses, no DNA stage;
Picornavirales; Picornaviridae; Aphthovirus.
Food and Agriculture Organization of the United Nations
Animal Production and Health Division


Animal Health Service AGAH
OpenFMD: HNK/2002 serotype?
ELSEVIER
Available online at www.sciencedirect.com
SCIENCE DIRECT
Biochemical and Biophysical Research Communications 323 (2004) 254–263
BBRC
Genome comparison of a novel foot-and-mouth disease virus with other FMDV strains
Qian Feng, Huan Yu, Yingying Liu, Chengqiang He, Jingsong Hu, Huachun Sang, Neizheng Ding, Mingxiao Ding, Yin-Wan Wendy Fung, Lok-Ting Lau, Albert Cheung-Hoi Yu, Jianguo Chen
HKN/2002 belongs to serotype O
Sequence analysis of HKN/2002 with 26 other reported FMDV genomes indicates that the novel HKN/2002 strain is of serotype O. Fig. 1 shows a neighboring-joining (NJ) tree constructed based on the sequence alignment of these 27 genomes, which are distinctly divided into five serotypes. HKN/2002 is tightly clustered
Screenshots from Biochem Biophys Res Commun. 2004 Oct
Food and Agriculture Organization of the United Nations
Animal Production and Health Division


Animal Health Service AGAH
Conclusions and recommendations
• Complete collection of FMDV (genetic) data is necessary for:
- Surveillance of disease evolutionary trends
- Scientific analysis
- Infer gaps in epidemiological data
• OPENFLU-EMPRESi experience may be useful for FMDV community
• OPENFMD might become a central portal for the community to store and share FMDV data
Food and Agriculture Organization of the United Nations
Animal Production and Health Division