Contribution ID: **10** Type: **not specified**

Workshop 6: Using the virus database VirJenDB to access metadata and sequences for bioinformatics analyses

Thursday 2 October 2025 10:45 (1h 30m)

Virus research depends on accurate sequence and metadata, but is often hampered by errors in user-submitted data and a lack of standardized metadata practices. Viral genomes are small and highly variable and current challenges in their organization are compounded by inconsistencies in related metadata (e.g., taxonomy, segment nomenclature, lifecycle definitions). VirJenDB (VJDB; virjendb.org) aims to provide a platform to address these challenges with a metadata-driven, integrative virus sequence database. Virus metadata and sequences from primary and secondary sources such as NCBI Virus, ENA, BVBRC, and community standards like MIGS-VI and ENA Sample Checklists have been merged and harmonized into an integrated dataset. The integrated dataset is used to enable development of metadata schema extensions and through different applications supports virus researchers throughout the data lifecycle. The VJDB dataset is accessible through the web portal, and provides input to bioinformatics webtools (future feature). In this demo, we will start with an overview of virus resources and demonstrate how to use currently available features of the VJDB portal: searching and refining results, downloading data, and exploring metadata schemas. Note that the demo content is subject to change depending on user demands.

Presenter: Dr CASSMAN, Noriko

Session Classification: From Lab Bench to Data Bench: Practical Research Data Management for Scientists