5th International Viromics Workshop

Organizing Committee:
Evelien Adriaenssens, Ben Bolduc, Joanne Emerson, Manja Marz, Simon Roux, Matt Sullivan, Ben Temperton, and Arvind Varsani. Coordination by Olivier Zablocki.

Co-organizer and sponsor: The Gordon and Betty Moore Foundation (link)

Additional support was provided by the NSF Science and Technology Center for Chemical Currencies of a Microbial Planet (C-CoMP; link), the NSF EMERGE Biology Integration Institute (link), and the OSU Center of Microbiome Science (link).

Practical information

Venue: Ohio State University, Ohio Union, 3rd floor, Reflection Room. The full address of the Ohio Union is 1739 N High St, Columbus, OH 43210 (Map). The workshop will run from Wednesday May 11 to Friday May 13.

Getting around: Uber and Lyft will get you everywhere you need, with average fare around the city of ~$10-$15. A ride share trip to the closest airport (John Glenn International – CMH) to the Ohio Union, goes between $20-$35. There is also a bus service running around Columbus as well as to/from the airport. The bus service is called COTA (website here), and you can conveniently pay for rides (~$2) using the 'universal' 'Transit' app directly from your smartphone. The workshop venue is right in front of one of the main bus lines (Line 1).

COVID-19: *Masks are optional* in most indoor spaces on The Ohio State University campuses. While Ohio State students, faculty and staff are required to be vaccinated against COVID-19, visitors attending Ohio State events, COVID-19 vaccinations are strongly recommended but are not required.

Sightseeing, eats and bars: Ohio's state capital, Columbus, is a vibrant city teeming with things to see and do. There is a strong experimental cuisine scene here, as well as countless microbreweries and coffee shops to experience. Given the short time frame, we encourage you to explore the Short North—the trendiest side of town that has a great diversity of stores and place to eat and drink. Most places are along High Street, and you can easily walk from the Union (where the workshop is held) directly into High Street. Here is a good starting point to get you acquainted with the area: https://www.experiencecolumbus.com/neighborhoods/short-north/.

AGENDA

WEDNESDAY, MAY 11

8:00-9:00 Breakfast

9:00-11:30 Welcome, Viromics mini-Symposium

9:00 – 9:15 *Welcome address*

Matthew Sullivan, OSU

9:15 – 9:45 *Viromics overview*

Simon Roux, JGI

9:45 – 10:15 *Oceans overview*

Ahmed Zayed, OSU

10:15-10:45 Coffee Break

10:45 – 11:15 *Soils overview*

Joanne Emerson, UC Davis

11:15 – 11:45 *Host-associated overview (plants, insect, animal)*

Arvind Varsani, ASU

--- PUBLIC SYMPOSIUM ENDS ---

11:45-1:30 Lunch for workshop attendees

Lunch plus FLASH TALKS from workshop attendees

(1 minute, 1 slide - in speed mode, travel awards from C-CoMP)

1:30 - 2 Overview of the viromics pipeline (Simon Roux)

- How to get the best data - tips from the lab
- Overview of standard viromics analysis pipeline - MIUViG paper
- Introduction of the different topics covered during this workshop

2 - 3:30 Virus identification (Karthik Anantharaman, Simon Roux)

- Concepts behind virus identification, pros and cons of most used available tools (VIBRANT, VirSorter2, DeepVirFinder), link to hands-on webinar for details
- Evaluating uncultivated virus genome quality (Check V)

3:30 - 4 Coffee Break

4-5:30 Social/networking time

Optional campus tour (30'- 45' long); "unmeeting" time.

6:30 Dinner reservation - El Vaquero (Olentangy River Road, see map). Dinner costs are covered in the workshop.

THURSDAY, MAY 12

8:30-9:00 Breakfast

9-10:30 Annotation of viral genomes (Varsani, Tisza, Gittrich)

- Overview (Arvind Varsani opens) - 5”
- Concepts behind viral genome annotation, pros and cons of current comprehensive tools - CenoteTaker2 (Mike Tisza) - 20” / general phage annotation via multiple tools w/ pitfall examples seeded prior from the organizing comm. (Marissa Gittrich) - 45”
- hands-on webinar for details on scalable microbial genome annotation

10:30-11:00 Coffee Break

11:00-12:30 Promises and pitfalls of identifying AMGs in fragmented genomes

- Current VS2 SOPs (Adjie Pratama) and VIBRANT SOPs (Karthik Anantharaman)
- Panel discussion around tricky use cases (Joanne moderates, with Adjie and Karthik)

12:30-1:30 Lunch

1:30-3:00 Databases (Ben Bolduc, Simon Roux, Manja Marz)

- CyVerse / iMicrobe / NMDC / IMG/VR / MGnifY - 30” total
- Panel discussion / What is coming from the EVBC

3-3:30 Coffee break

3:30-5 Linking viruses to hosts – *in silico* (Simon Roux)

- In silico tools overview and pros/cons with focus on aggregate, score-based methods (Roux)

5-5:30 -- Picture Time!

6:30 Dinner in the Short North, self-organized and self-financed, but we will provide a “menu” of options

FRIDAY, MAY 13

8:30-9:00 Breakfast

9-10:30 -- Virus classification vs taxonomy (Evelien Adriaenssens)

- Challenges in viral taxonomy and the consequences for viromics
- What are the pros and cons of most-used tools (VIPTree, GraVity, vContact2), link to hands-on webinar for details
- What is different for the different types of viruses (RNA vs DNA viruses)?

10:30-11:00 Coffee break

11-12:30 -- Improving on short-read assembled viral genomes (Ben Temperton)

- Overview of long-reads, proximity ligation, single virus genomics - promises/pitfalls
- Focus on long-reads application
- Link to hands-on webinar for details

12:30-1:30 Lunch

1:30-3:00 The power of population-based inference (Ann Gregory)
- The vOTU / population and the theory behind these
- Link to webinars on ecological statistics, from simple to sophisticated

3:00-3:30 Coffee break

3:30-5:00 Linking viruses to hosts – experimentally (multiple, see below)
- Polonies (Natalie Sawaya)
- EpicPCR (Sarah Preheim)
- Viral tagging / phageFISH / HiC (Sullivan)

5:00 PM Workshop conclusion!

--- ADJOURN, NO FORMAL DINNER ORGANIZED, "UNMEETING" TIME ---

--- OPTIONAL WEEKEND ACTIVITIES IN AND AROUND COLUMBUS, please let us know if staying and you would like local guides ---

Suggested Reading

Databases

- Roux, S., Páez-Espino, D., Chen, I.M.A., Palaniappan, K., Ratner, A., Chu, K., Reddy, T.B.K., Nayfach, S., Schulz, F., Call, L. and Neches, R.Y., 2021. IMG/VR v3: an integrated ecological and evolutionary framework for interrogating genomes of uncultivated viruses. Nucleic acids research, 49(D1), pp.D764-D775. LINK

Virus Taxonomy

- Aiewsakun, P. and Simmonds, P., 2018. The genomic underpinnings of eukaryotic virus taxonomy: creating a sequence-based framework for family-level virus classification. Microbiome, 6(1), pp.1-24. LINK

- Bin Jang, H., Bolduc, B., Zablocki, O., Kuhn, J.H., Roux, S., Adriaenssens, E.M., Brister, J.R., Kropinski, A.M., Krupovic, M., Lavigne, R. and Turner, D., 2019. Taxonomic assignment of uncultivated prokaryotic virus genomes is enabled by gene-sharing networks. Nature biotechnology, 37(6), pp.632-639. LINK

- Nishimura, Y., Yoshida, T., Kuronishi, M., Uehara, H., Ogata, H. and Goto, S., 2017. ViPTree: the viral proteomic tree server. Bioinformatics, 33(15), pp.2379-2380. LINK

- Zayed, A.A., *et al*., 2022. Cryptic and abundant marine viruses at the evolutionary origins of Earth's RNA virome. Science, 376(6589), pp.156-162. LINK

Virus Identification

- Guo, J., Bolduc, B., Zayed, A.A., Varsani, A., Dominguez-Huerta, G., Delmont, T.O., Pratama, A.A., Gazitúa, M.C., Vik, D., Sullivan, M.B. and Roux, S., 2021. VirSorter2: a multi-classifier, expert-guided approach to detect diverse DNA and RNA viruses. Microbiome, 9(1), pp.1-13. LINK

- Kieft, K., Zhou, Z. and Anantharaman, K., 2020. VIBRANT: automated recovery, annotation and curation of microbial viruses, and evaluation of viral community function from genomic sequences. Microbiome, 8(1), pp.1-23. LINK

- Ren, J., Song, K., Deng, C., Ahlgren, N.A., Fuhrman, J.A., Li, Y., Xie, X., Poplin, R. and Sun, F., 2020. Identifying viruses from metagenomic data using deep learning. Quantitative Biology, 8(1), pp.64-77. LINK

- Tisza, M.J., Belford, A.K., Dominguez-Huerta, G., Bolduc, B. and Buck, C.B., 2021. Cenote-Taker 2 democratizes virus discovery and sequence annotation. Virus evolution, 7(1), p.veaa100. LINK

Benchmarking/SOP papers

- Nayfach, S., Camargo, A.P., Schulz, F., Eloe-Fadrosh, E., Roux, S. and Kyrpides, N.C., 2021. CheckV assesses the quality and completeness of metagenome-assembled viral genomes. Nature biotechnology, 39(5), pp.578-585. LINK

- Pratama, A.A., Bolduc, B., Zayed, A.A., Zhong, Z.P., Guo, J., Vik, D.R., Gazitúa, M.C., Wainaina, J.M., Roux, S. and Sullivan, M.B., 2021. Expanding standards in viromics: in silico evaluation of dsDNA viral genome identification, classification, and auxiliary metabolic gene curation. PeerJ, 9, p.e11447. LINK

- Roux, S., *et al*., 2019. Minimum information about an uncultivated virus genome (MIUViG). Nature biotechnology, 37(1), pp.29-37. LINK

- Roux, S., Emerson, J.B., Eloe-Fadrosh, E.A. and Sullivan, M.B., 2017. Benchmarking viromics: an in silico evaluation of metagenome-enabled estimates of viral community composition and diversity. PeerJ, 5, p.e3817. LINK

- Duhaime, M.B. and Sullivan, M.B., 2012. Ocean viruses: rigorously evaluating the metagenomic sample-to-sequence pipeline. Virology, 434(2), pp.181-186. LINK

- Hurwitz, B.L., Hallam, S.J. and Sullivan, M.B., 2013. Metabolic reprogramming by viruses in the sunlit and dark ocean. Genome biology, 14(11), pp.1-14. LINK

- Solonenko, S.A., Ignacio-Espinoza, J.C., Alberti, A., Cruaud, C., Hallam, S., Konstantinidis, K., Tyson, G., Wincker, P. and Sullivan, M.B., 2013. Sequencing platform and library preparation choices impact viral metagenomes. BMC genomics, 14(1), pp.1-12. LINK

Commentaries & Reviews

- Bolduc, B., Youens-Clark, K., Roux, S., Hurwitz, B.L. and Sullivan, M.B., 2017. iVirus: facilitating new insights in viral ecology with software and community data sets imbedded in a cyberinfrastructure. The ISME journal, 11(1), pp.7-14. LINK
- Bolduc, B., Zablocki, O., Guo, J., Zayed, A.A., Vik, D., Dehal, P., Wood-Charlson, E.M., Arkin, A., Merchant, N., Pett-Ridge, J. and Roux, S., 2021. iVirus 2.0: Cyberinfrastructure-supported tools and data to power DNA virus ecology. ISME Communications, 1(1), pp.1-8. LINK
- Coclet, C. and Roux, S., 2021. Global overview and major challenges of host prediction methods for uncultivated phages. Current Opinion in Virology, 49, pp.117-126. LINK
- Hurwitz, B.L., Ponsero, A., Thornton Jr, J. and U'Ren, J.M., 2018. Phage hunters: Computational strategies for finding phages in large-scale 'omics datasets. Virus research, 244, pp.110-115. LINK
- Lefeuvre, P., Martin, D.P., Elena, S.F., Shepherd, D.N., Roumagnac, P. and Varsani, A., 2019. Evolution and ecology of plant viruses. Nature Reviews Microbiology, 17(10), pp.632-644. LINK
- Sommers, P., Chatterjee, A., Varsani, A. and Trubl, G., 2021. Integrating viral metagenomics into an ecological framework. Annual Review of Virology, 8, pp.133-158. LINK
- Varsani, A., Lefeuvre, P., Roumagnac, P. and Martin, D., 2018. Notes on recombination and reassortment in multipartite/segmented viruses. Current opinion in virology, 33, pp.156-166. LINK

Ecology

- Dayaram, A., Galatowitsch, M.L., Argüello-Astorga, G.R., van Bysterveldt, K., Kraberger, S., Stainton, D., Harding, J.S., Roumagnac, P., Martin, D.P., Lefeuvre, P. and Varsani, A., 2016. Diverse circular replication-associated protein encoding viruses circulating in invertebrates within a lake ecosystem. Infection, Genetics and Evolution, 39, pp.304-316. LINK
- Sakowski, E.G., Arora-Williams, K., Tian, F., Zayed, A.A., Zablocki, O., Sullivan, M.B. and Preheim, S.P., 2021. Interaction dynamics and virus–host range for estuarine actinophages captured by epicPCR. Nature Microbiology, 6(5), pp.630-642. LINK
- Deng, L., Ignacio-Espinoza, J.C., Gregory, A.C., Poulos, B.T., Weitz, J.S., Hugenholtz, P. and Sullivan, M.B., 2014. Viral tagging reveals discrete

populations in Synechococcus viral genome sequence space. Nature, 513(7517), pp.242-245. LINK

- Jang, H.B., Chittick, L., Li, Y.F., Zablocki, O., Sanderson, C.M., Carrillo, A., van den Engh, G. and Sullivan, M.B., 2022. Viral tag and grow: a scalable approach to capture and characterize infectious virus–host pairs. ISME Communications, 2(1), pp.1-11. LINK

Population genetics

- Rosselló-Mora, R. and Amann, R., 2001. The species concept for prokaryotes. FEMS microbiology reviews, 25(1), pp.39-67. LINK

- Konstantinidis, K.T. and Tiedje, J.M., 2005. Genomic insights that advance the species definition for prokaryotes. *Proceedings of the National Academy of Sciences*, *102*(7), pp.2567-2572. LINK

- Wiedenbeck, J. and Cohan, F.M., 2011. Origins of bacterial diversity through horizontal genetic transfer and adaptation to new ecological niches. *FEMS microbiology reviews*, *35*(5), pp.957-976. LINK

- Cadillo-Quiroz, H., Didelot, X., Held, N.L., Herrera, A., Darling, A., Reno, M.L., Krause, D.J. and Whitaker, R.J., 2012. Patterns of gene flow define species of thermophilic Archaea. *PLoS biology*, *10*(2), p.e1001265. LINK

- Gregory, A.C., Solonenko, S.A., Ignacio-Espinoza, J.C., LaButti, K., Copeland, A., Sudek, S., Maitland, A., Chittick, L., Dos Santos, F., Weitz, J.S. and Worden, A.Z., 2016. Genomic differentiation among wild cyanophages despite widespread horizontal gene transfer. *BMC genomics*, *17*(1), pp.1-13. LINK