Title: *de novo* genome assembly and annotation with an emphasis on phylogenetic and population genetic studies

Presenters: Jacob Landis, Fay-Wei Li, Adrian Powell, Bikash Shrestha, and Suzy Strickler

Length: 9 am EST - 5 pm EST

The costs of generating a draft genome for a study organism of choice are declining each year and more tools are becoming available to help in the endeavor. However, generating the appropriate sequence data is often not a trivial matter and currently, there are many options when it comes to analysis pipelines. The focus of this workshop is to work through the necessary steps for genome assembly and annotation when a closely related reference genome is not available. The workshop will cover assembling genomes using Illumina, Oxford Nanopore, and hybrid assemblers; as well as methods for extracting high-quality DNA suitable for ONT sequencing and library preparation, especially in species that possess secondary compounds which can be especially problematic for nanopore sequencing. Downstream applications/analyses that can be performed after assembly will also be covered while highlighting that a perfect assembly is not necessary to answer many evolutionary questions. Using the CyVerse platform, participants will get hands-on practice with all the necessary steps using either the supplied test data or their own data if previously generated.

<div align="center">Topics/layout</div>

Prerequisites and baseline skills:
- Unix command-line skills (we will review these at the start)
- Familiarly with some basic programming
- CyVerse account
- CoGe account

Syllabus
- Introductions and introduction to genome sequencing and assembly 9 am - Suzy
- Setting hardware/software baseline skills 9:15 (45min) - Adrian
- Generate suitable DNA for different platforms: tips, tricks, and experimental design 10 am (30min) - Fay-Wei
- Break 10:30 (15min)
- Sequencing data and genome assembly options 10:45 am (1.75h) - Jacob
- Break for lunch 12:30 am (1h)
- Genome annotation 1:30 (1.5h) - Suzy and Bikash
- Break 3:00 pm (15min)
- Downstream evolution analyses: genome management, genome visualization, comparative genomics, reference-based SNP calling, transcriptomics, and phylogenomics 3:15 (2h) - Adrian