

Rapid Analysis of Microbes Workshop

Pathogenic organisms can rapidly spread, either from a contaminated source through a population, or from person to person contact. Understanding both the evolutionary relationship of novel sequences as well as their functional capabilities is essential in dealing with virulent organisms. Next generation sequencing provides the initial step in providing the raw material that can then be put into context through the annotation and analysis of the sequences. NCBI has created an annotation pipeline that is increasingly being used by external sequencing centers, government agencies, and academic researchers, for the analysis of prokaryotes including those associated with disease outbreaks. Rapid automated analysis of prokaryotic genomes will provide the ability to quickly determine the relatedness of a novel genome to existing genomes, determine conserved and accessory sequences, and place these into both a functional and evolutionary context and may enable a common origin to be assigned and NCBI is reconfiguring the analytical tools present to be able to handle the increasing number of submissions, which in some cases has reached over 100 genomes from one center. Many thousands more are expected. This workshop includes a community of researchers investigating pathogenic prokaryotes and will highlight the current capabilities and generate a discussion on the requirements for the rapid analysis of prokaryotic pathogens: what current resources including databases and analytic tools exist at NCBI or externally, what resources can be shared, what resources may need to be developed, what standards such as biological metadata are required or need to be developed in order to facilitate the rapid sequence, annotation, and analysis of pathogens during outbreaks that will enable scientists to rapidly react and potentially deal with biological threats in real time.

Rapid Analysis of Microbes Workshop

1. Resources available

Databases

- NCBI

- other databases

Tools

- NCBI

- other tools

2. Resources required

3. Standards and data submission for pathogens

Metadata submission for pathogens

- NCBI Bioproject/Biosample

- batch submissions

Sequence submission during outbreak

- different sequencing platforms/assembly states with different inherent errors

Annotation standards during outbreak

Archival submission after outbreak

4. Analysis

Phylogenetic/evolutionary

- SNP identification

- trees/networks

Functional annotation

5. Rapid dissemination of information