

Abstract

The PathoSystems Resource Integration Center (PATRIC) at the Virginia Bioinformatics Institute (VBI) is one of eight Bioinformatics Resource Centers (BRC) funded by NIAID as part of its biodefense initiative. The BRCs will provide the scientific community with a web resource for research on biodefense and emerging or re-emerging diseases to facilitate the development of novel diagnostics, vaccine and therapeutics. As part of the BRC mission PATRIC will provide a comprehensive and centralized annotation resource for eight classes of microbial pathogens. It will present users with integrated access to genomic, proteomic, biochemical and microbiological information; experimental data and associated literature on each of the pathogen classes.

Introduction

The PathoSystems Resource Integration Center (PATRIC) was established at the Virginia Bioinformatics Institute (VBI), in July 2004. It is one of the eight Bioinformatics Resource Centers (BRC) that have been set up by the National Institute of Allergy and Infectious Disease (NIAID) with a mandate to integrate biological information on microbial pathogens, to aid research on disease detection and prevention, and on therapeutic solutions. The centers will focus on the NIAID/CDC category A-C pathogens and agents causing emerging or re-emerging disease that are relevant for biodefense. PATRIC annotates and curates genomic and related biological data for eight pathogenic microorganism classes, which include three bacterial and five viral classes. Hepatitis A and E are two of the five viruses that are included here. At present, for Hepatitis A, we have 16 genomes available with more on the way. Hepatitis E has 48 genomes available, 5 of which are reference genomes. The website address is: <http://patric.vbi.vt.edu>

Hepatitis A Reference Genome	HM-175 Wild Type
Hepatitis E Reference Genomes	Avian Hepatitis E virus Hepatitis E virus T1 Hepatitis E virus (clone pSHEV-3) Hepatitis E virus Hetian Xinjiang Uighur China Hepatitis E virus Mexican

PathoSystems

At PATRIC, we define the pathogen-host-environment triangle as a PathoSystem. The PathoSystem combines factors in the host, pathogen and environment (the "Disease Triangle") that, together are necessary and sufficient to specify the outcome of a host-pathogen encounter, which may range from microbe-clearance to full-blown disease. We believe that this holistic approach to studying a disease will yield valuable insights into the pathogenic mechanisms of these microbes. PATRIC will follow this "Pathosystems" approach in annotating and integrating biological data drawn from different sources.

PATRIC's Goals

In keeping with the BRC objectives, PATRIC has created a web-resource that provides access to genomic and associated data, and analysis tools on the selected pathogens and facilitate research in the areas of vaccines, diagnostics and therapeutics (Fig 3). We also welcome and expect the active participation of the research community in the development of this resource.

PATRIC Personnel

The core PATRIC team consists of Computational Biologists with a strong background in infectious disease, Proteomics, Bioinformatics, Software Development and a Database Analysis and Administration.

Data available at patric.vbi.vt.edu (some still under development)

- All Proteins the genome codes for:
- Function of each protein
- Experimental characterization
- Cellular Localization
- Structure (Secondary and/or 3D when available)
- Active sites reported
- Pathways Involved/Affected (maybe host)
- Phylogenetic trees of different strains (from Literature)
- Host Pathogenicity
- Protein interactions (Protein-Protein, Protein-drug, Protein-DNA, Protein-RNA, Protein-Small molecule)
- Regulatory features on genome sequence
- Epidemiology
- Literature: Author, Journal, Publication, etc.
- Current Status of Drug/Vaccine/Diagnostics
- Reasons for absence of a vaccine/drawbacks

Acknowledgements

This project is funded by NIAID / NIH contract **HHSN2662040035C** awarded to Bruno Sobral.

Our thanks to Yuri Khudiyakov, Organism Expert for HepA, XJ Meng, Organism Expert for HepE and June Mullins for help with poster preparation.

Integration of Organism Information

PATRIC is providing integrated access to diverse sets of biological data and information, along with a set of analysis tools through its user-friendly website.

Tools

Single and multi-genome sequence analysis tools; experimental data analysis tools

The Annotation Process

An annotation software infrastructure has been built to automate the annotation at various levels (nucleic acid, protein, pathways); it includes modules for comparative genomics; literature curation and annotating information of host-pathogen interactions (Fig 4).

A team of curators manually review and edit the output of the automated annotation. Curators are involved in working with members of the scientific community to ensure that their needs, in terms of data and tools, are incorporated at PATRIC. An organism expert (OE) has been assigned to each PathoSystem. The OE is a recognized authority on his/her assigned pathogen and is closely involved at each step of the annotation/curation process.

Data

Hepatitis A: We have used Hepatitis A HM-175 wild type strain as the reference genome and used GATU (Genome Annotation Transfer Utility-Tcherepanov et.al) to transfer coding features to the other Hepatitis A genomes. Non-coding features in the reference genome were annotated from literature, standardized and multiple sequence alignment program was used to transfer to other genomes (Fig 1, top). Phylogenetic trees have been created for all strains at VBI (Fig 2A).

Hepatitis E: Coding and non-coding features have been annotated from literature and standardized across all strains, including alternate translational start sites and potential overlapping features (Fig 1 bottom). Phylogenetic trees have been created for all strains at VBI (Fig 2B).

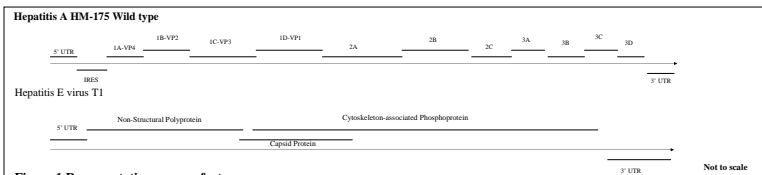


Figure 1 Representative genome features

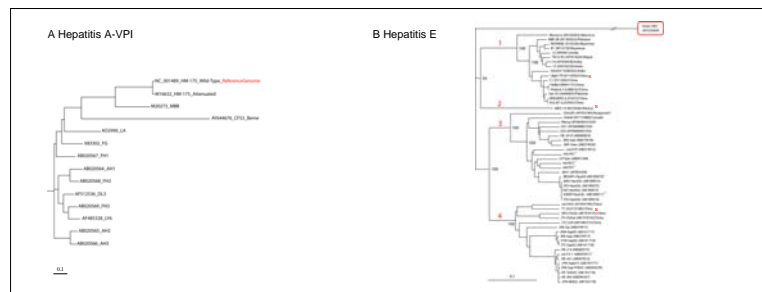


Figure 2 Phylogenetic trees for A. Hepatitis A (VP1 protein) and B. Hepatitis E (whole genome) for genomes at VBI

PATRIC Goals.

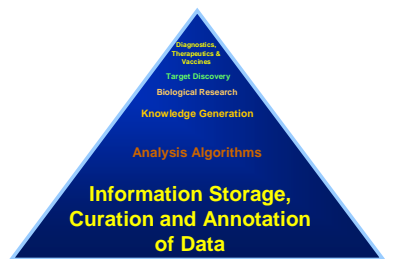


Figure 3 PATRIC goals

PATRIC Annotation Workflow.

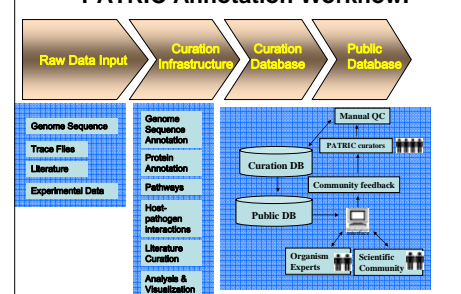


Figure 4 The PATRIC Annotation workflow

Summary

Through a user-friendly web interface, PATRIC will provide high quality information and tools to the scientific communities that study Hepatitis A, and Hepatitis E. Integrating manual genome / gene product curation, scientific literature, and a suite of comparative genomic tools will align with NIAID's mission to facilitate development of vaccine, therapeutic, and diagnostic agents. The project intends to remain flexible enough so that it can other include data and services for which there is a significant demand in the drugs/vaccines/diagnostic research community and welcomes collaborations to make this possible.

Reference: Tcherepanov VT, Ehlers A, Upton C. (2006) Genome Annotation Transfer Utility (GATU): Rapid annotation of viral genomes using a closely related reference genome. BMC Genomics Jun 13;7(1):150