Introduction

The Virus Pathogen Database and Analysis Resource (ViPR, www.viprbrc.org), sponsored by the National Institute of Allergy and Infectious Diseases, serves as a single publicly-accessible repository of integrated datasets and analysis tools for 14 different virus families to support wet-bench virology researchers focusing on the development of diagnostics, prophylactics, vaccines, and treatments for these pathogens.

ViPR Supports 14 Virus Families

- Arenaviridae, Bunyaviridae, Caliciviridae, Coronaviridae, Filoviridae, Flaviviridae, Hepeviridae, Herpesviridae, Paramyxoviridae, Picornaviridae, Poxviridae, Reoviridae, Rhabdoviridae, and Togaviridae.

ViPR Intergrates Data from External Sources and Generates Novel Data from Internal Computational Pipelines

- GeneBank sequence records
- Enriched gene and protein annotations from GeneBank, UniProtKB and ViPR annotation pipelines
- Gene Ontology (GO) classifications
- Protein domains and motifs
- Protein Ortholog Groups
- Protein Databank (PDB) 3D protein structures
- Experimentally-determined and computationally-predicted immune epitopes
- Clinical data
- Host factor data generated from the NIAID Systems Virology projects and the ViPR-funded Driving Biological Projects
- Additional data derived from computational algorithms

ViPR Provides Analysis and Visualization Tools

- Genome Browser
- Genome Annotator
- BLAST Sequence Similarity Search
- Multiple Sequence Alignment
- Phylogenetic Tree Construction
- PCR Primer Design
- Sequence Feature Variant Type (SFVT) Analysis
- Metadata-driven Comparative Genomics Analysis
- SNP Analysis
- 3D Protein Structures with Sequence Feature or Epitope Highlights

ViPR provides personal Workbench for data storage and sharing

Sequence Feature Variant Type (SFVT)

- Sequence Features (SFs): characterized structural, functional, immune epitope, or sequence alteration regions of a protein manually curated from UniProt, GenBank, and the Immune Epitope Database and then validated by expert researchers.
- Variant Type (VT): Polymorphisms in each Sequence Feature are identified as “Variant Types” of the Sequence Feature.
- Available for hepatitis C virus subtype 1a, Dengue virus type 1-4, and Orthopoxvirus (Vaccinia) in ViPR.

ViPR includes in silico prediction tools for the following categories:

- Comparative Genomics
- Surface Plasmon Resonance
- Immune Epitopes
- Protein Structure
- Metabolome

ViPR provides a unique platform for researchers to explore virus genomic sequences and associated data in the context of the disease they cause, including the interactions between viruses and their host.

Summary

ViPR combines the strength of a relational database with a suite of bioinformatics integrated tools to support everything from basic sequence and structural analyses to genotype-phenotype studies and host-virus interaction studies. The uniqueness of ViPR lies in:

- integrating data from various sources
- generating novel data from custom computational pipelines
- capturing unique data on host responses to virus infections
- encouraging the analysis of the comprehensive data contained within the system
- combining the available tools to quickly perform complex analytical workflows
- allowing data sharing and storage with collaborators

Acknowledgements

We would like to thank the primary data providers for the data that was used throughout this study. We also recognize the scientific and technical personnel responsible for supporting and developing ViPR, which has been wholly supported with federal funds from the NIH/NAID (MD012008/38 to R.H.S.).

References