U.S. FOOD & DRUG ADMINISTRATION
Human Foods Program

# ViroTrakr: A Genomic Platform for Detection, Identification and Outbreak Investigation of Foodborne Viral Pathogens

**Zhihui Yang**
**Research Biologist**
**Virology & Parasitology Branch (VPB)**
**Division of Food and Environmental Safety (DFES)**
**Office of Applied Microbiology and Technology (OAMT)**
**Office of Laboratory Operations and Applied Science (OLOAS)**
**Human Foods Program (HFP)**

**2026 APHL GenomeTrakr Meeting**
**November 4, 2026**

# Introduction

U.S. FOOD & DRUG ADMINISTRATION
Human Foods Program

**What is ViroTrakr:**

- A genomic platform for foodborne viral pathogen surveillance
- An integrated framework for sequence data collection, sharing, and standardized bioinformatics analysis
- To support foodborne virus surveillance, source attribution and outbreak investigations

**ViroTrakr framework: (**three major components)

(1) Sequence data generation and submission

(2) A public genomic reference sequence database (housed at NCBI)

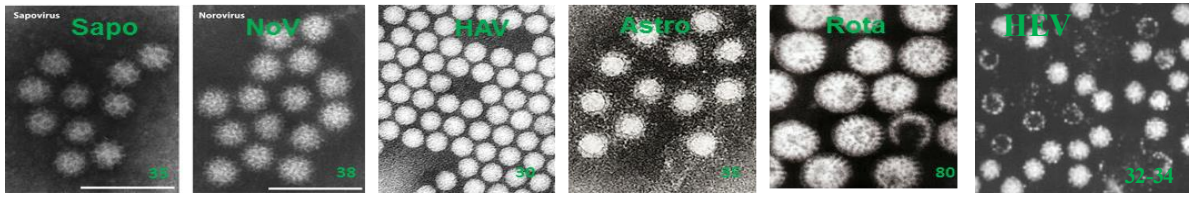
(3) Standardized bioinformatics workflows for genotyping and phylogenetic analysis

**Objectives and impact:**

ViroTrakr provides genomic surveillance infrastructure for the rapid detection, identification and traceback of foodborne viral pathogens. It enhances data sharing, supports outbreak investigation and public health response, and potentially help harmonize genomic surveillance across laboratories.

# Background

- ***The challenge*: foodborne viruses:**
  - Responsible for nearly 60% of foodborne illnesses in the US (CDC)
  - Broad transmission routes and carriers
  - Low infectious doses in humans
  - High genetic diversity and high mutation rates

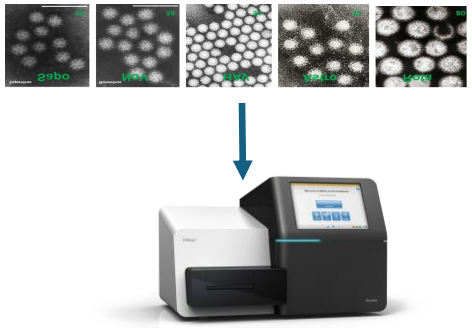



(Source: CDC)

- ***Need*: why accurate and sensitive identification matters:**
  - Strengthen foodborne virus surveillance
  - Accelerate traceback of viral foodborne outbreaks
  - Improve food safety

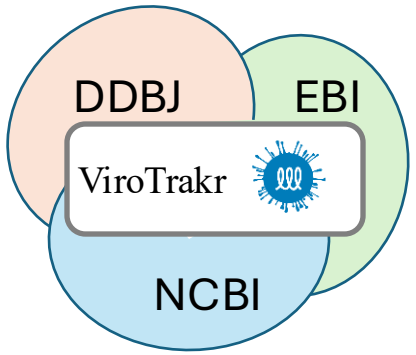
- ***Our approach*:**
  - Modeled on GenomeTrakr for bacterial pathogen identification and surveillance, develop a genomic platform for identifying and surveilling foodborne viruses.

# Core Components of ViroTrakr



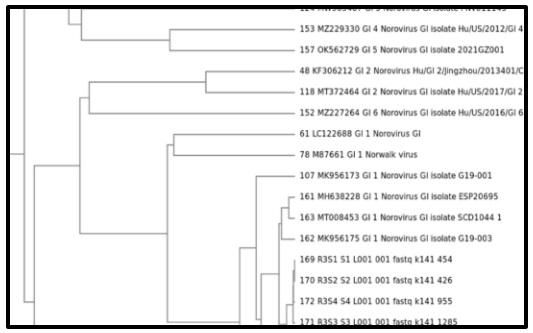
**Sequence data generation:**

- Food, water, environmental and clinical samples
- Whole-genome sequencing using next-generation sequencing platforms
- Submit sequence data to ViroTrakr




**Data collection and sharing:**

- ViroTrakr genomic database housed in NCBI
- Standardized protocol for submitting sequence data and metadata to ViroTrakr
- Publicly accessible data through international databases




**Standardized data analysis:**

- Genotype assignment
- Phylogenetic characterization
- Viral strain comparison and differentiation
- Support foodborne outbreak investigation and traceback

# ViroTrakr: foodborne viruses (ID 396739) - BioProject - NCBI (nih.gov)

U.S. FOOD & DRUG ADMINISTRATION
Human Foods Program

National Library of Medicine
National Center for Biotechnology Information

Log in

BioProject | BioProject | Search | Advanced | Help

Display Settings: | Send to:

## foodborne viruses

Accession: PRJNA396739 ID: 396739

**ViroTrakr**

Foodborne viruses are the leading cause of foodborne illness, surpassing all other pathogens combined, including bacteria. ViroTrakr, a genomic database initiated by HFP (formerly CFSAN) and hosted by NCBI, aims to (1) include sequences from a wide range of foodborne viruses (e.g., norovirus, hepatitis A virus, sapovirus) derived from clinical, food, and environmental samples, and (2) provide real-time reference sequences to support phylogenetic analysis and epidemiological investigations of foodborne illnesses. Similar to GenomeTrakr for bacterial pathogen surveillance, ViroTrakr has the potential to be used for foodborne viral pathogen surveillance. The SOPs for ViroTrakr Sequence Data Submission and the ViroTrakr Sequence Data Analysis workflow have been developed and published, and are accessible through the links under "Related Resources" below. For technical assistance regarding ViroTrakr, please contact: ViroTrakr@fda.hhs.gov

| | |
|---|---|
| Accession | PRJNA396739 |
| Type | Umbrella project |
| Publications | 1. Yang Z *et al.*, "Whole-Genome Sequencing-Based Confirmatory Methods on RT-qPCR Results for the Detection of Foodborne Viruses in Frozen Berries.", *Food Environ Virol*, 2024 Jun;16(2):225-240<br>2. Raymond P *et al.*, "Detection and Sequencing of Multiple Human Norovirus Genotypes from Imported Frozen Raspberries Linked to Outbreaks in the Province of Quebec, Canada, in 2017.", *Food Environ Virol*, 2022 Mar;14(1):40-58 |
| Submission | Registration date: 1-Aug-2017<br>**OARSA/CFSAN/FDA** |
| Related Resources | • NCBI ViroTrakr Data Submission Protocol for foodborne virus surveillance<br>• NCBI ViroTrakr Data Analysis Protocol: Genotyping and phylogenetic analysis_ViroTrakr workflows |
| Relevance | Foodborne viruses |

**Related information**
Assembly
BioSample
Data projects
Full text in PMC
Genome
Genomic RNA
Nucleotide
Protein
PubMed
SRA

**Related Resources**
NCBI ViroTrakr Data Submission Protocol for foodborne virus surveillance
NCBI ViroTrakr Data Analysis Protocol: Genotyping and phylogenetic analysis_ViroTrakr workflows

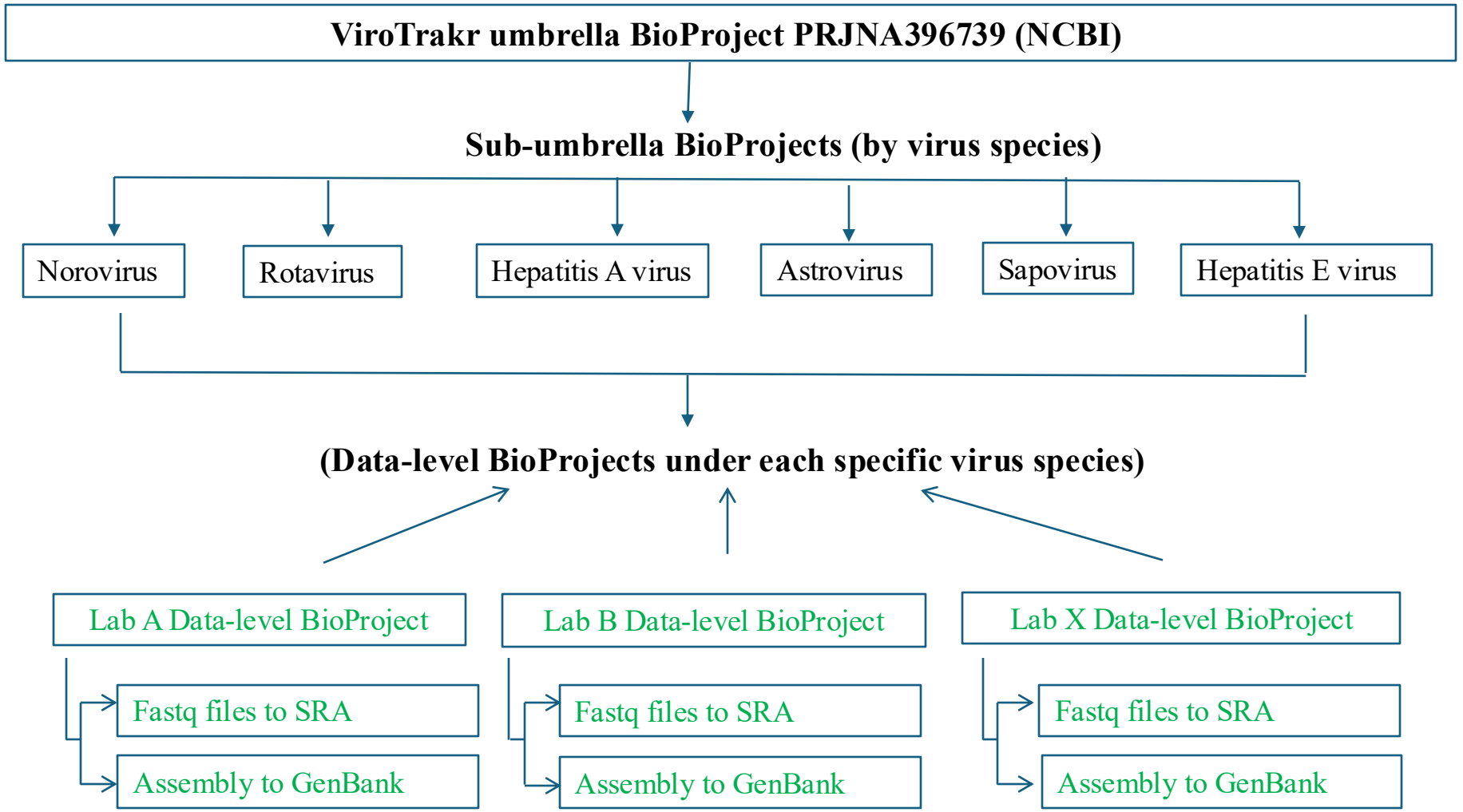
**Recent activity**

| Umbrella project | 6 |
|---|---|

| BioProject accession | Name | Title |
|---|---|---|
| PRJNA817227 | Metagenomic sequence data of astroviruses | Metagenomic sequence data of astroviruses (OARSA/CFSAN/FDA) |
| PRJNA433976 | Metagenomic sequence data of hepatitis A viruses | Metagenomic sequence data of hepatitis A viruses (OARSA/CFSAN/FDA) |
| PRJNA817228 | Metagenomic sequence data of hepatitis E virus | Metagenomic sequence data of hepatitis E virus (OARSA/CFSAN/FDA) |
| PRJNA433975 | Metagenomic sequence data of human noroviruses | Metagenomic sequence data of human noroviruses (OARSA/CFSAN/FDA) |
| PRJNA817226 | Metagenomic sequence data of rotaviruses | Metagenomic sequence data of rotaviruses (OARSA/CFSAN/FDA) |
| PRJNA433977 | Metagenomic sequence data of sapovirus | Metagenomic sequence data of sapovirus (OARSA/CFSAN/FDA) |

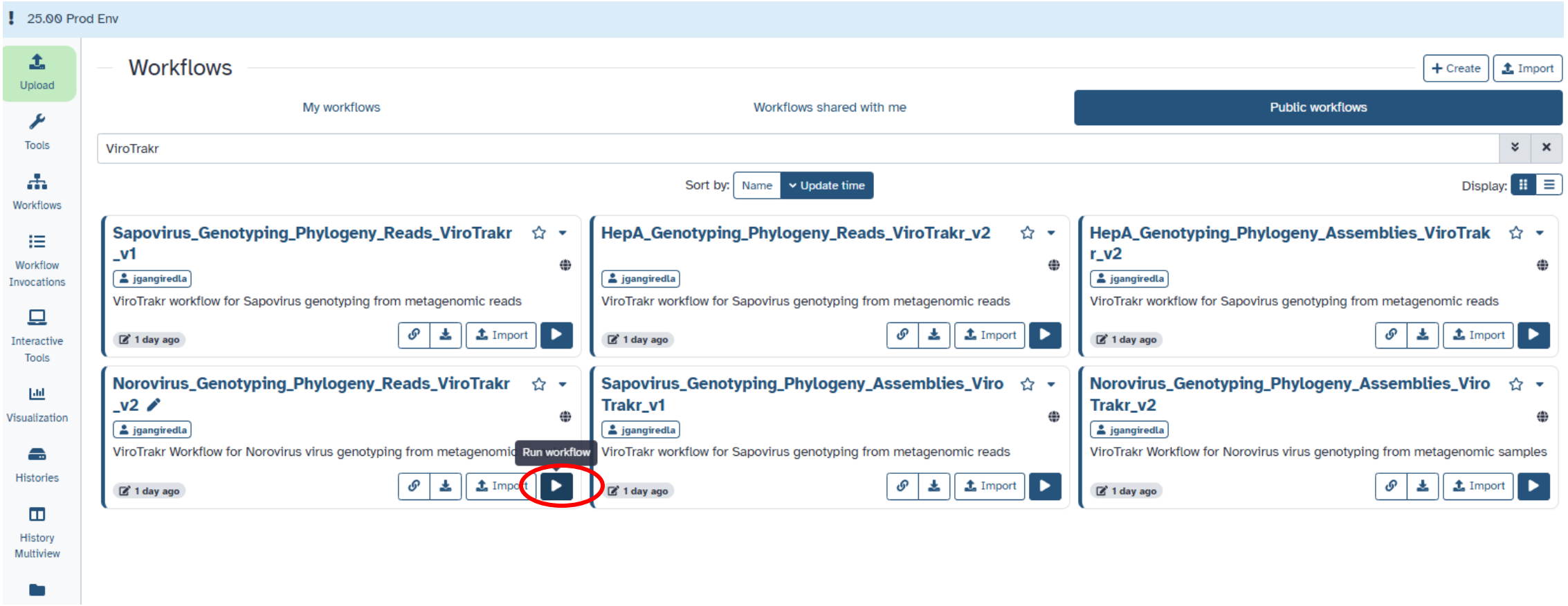
# ViroTrakr Database Structure

U.S. FOOD & DRUG ADMINISTRATION
Human Foods Program

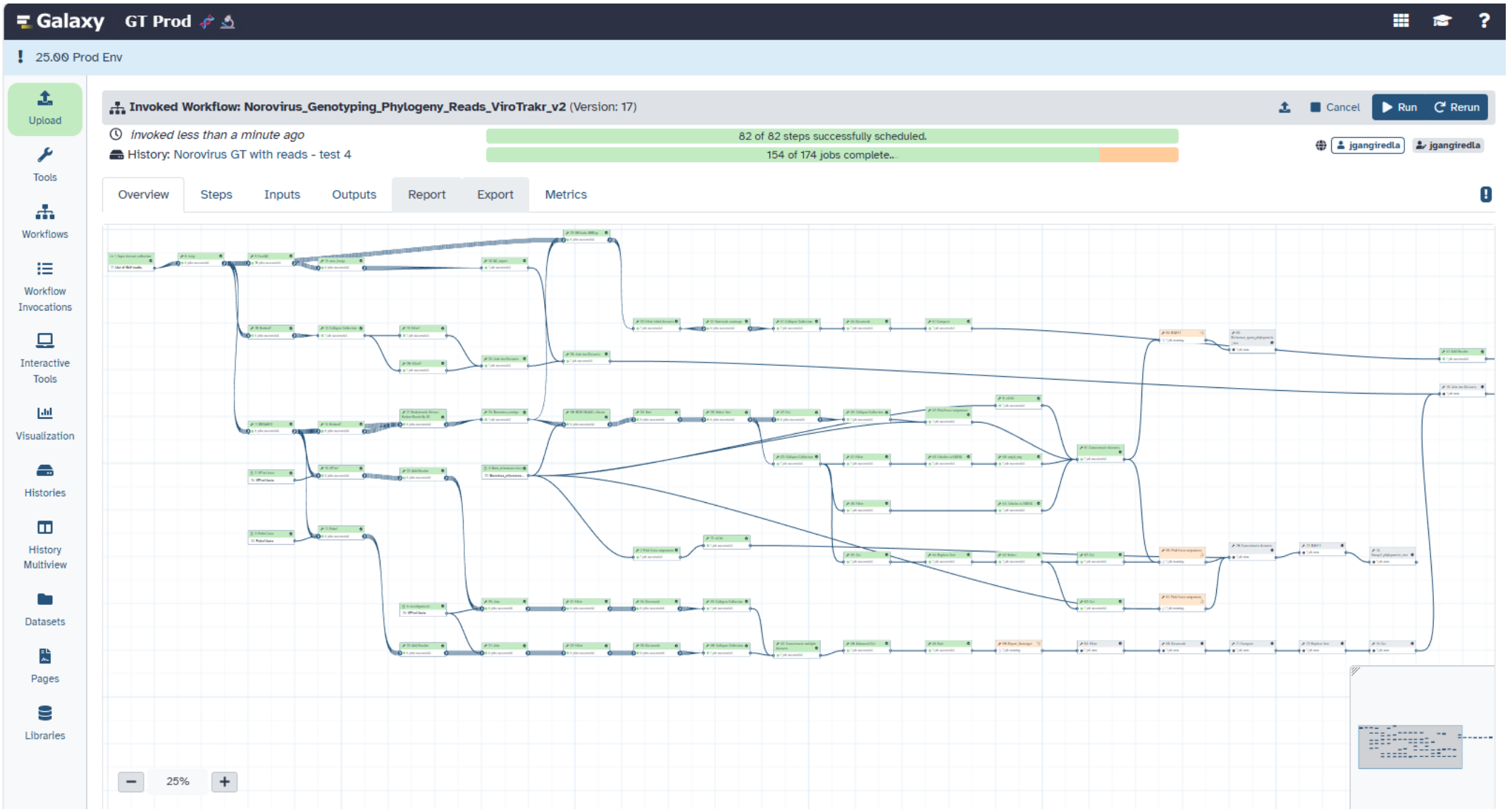



***Note:*** *Steps in the ViroTrakr submission process are highlighted in green.*

# ViroTrakr Data Analysis Workflows in GalaxyTrakr Platform



# Example of a Real-time Workflow Status



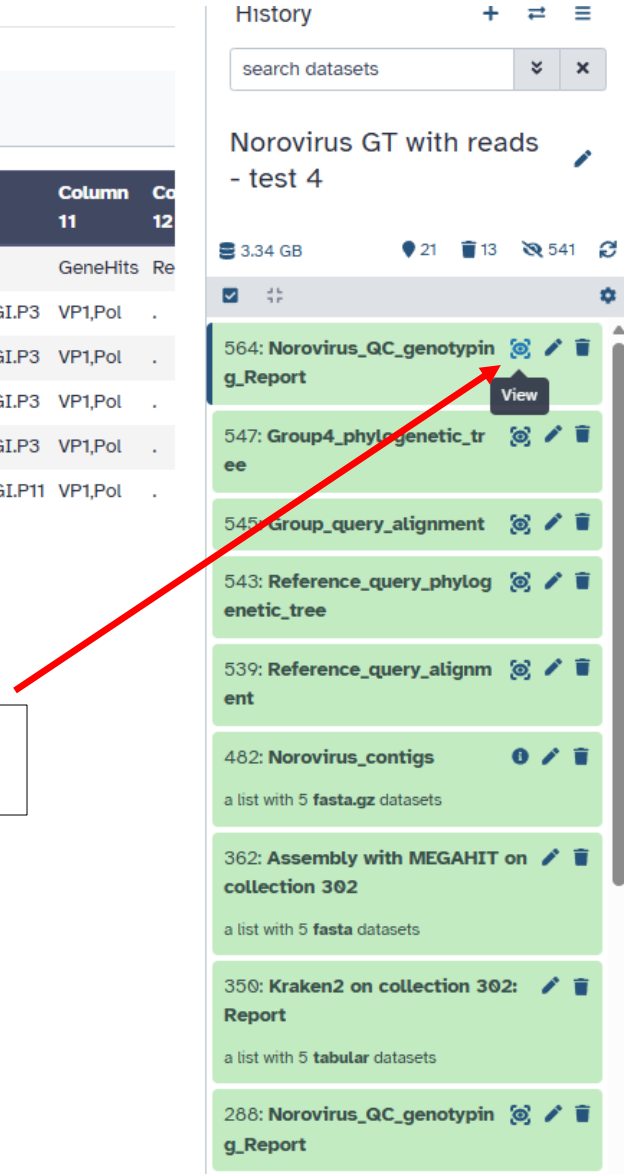
# Example ViroTrakr Genotyping Report

U.S. FOOD & DRUG ADMINISTRATION
Human Foods Program

564: **Norovirus_QC_genotyping_Report** ok

Preview | Visualize | Details | Edit

| Column 1 | Column 2 | Column 3 | Column 4 | Column 5 | Column 6 | Column 7 | Column 8 | Column 9 | Column 10 | Column 11 | Co 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Sample | contigs | length_assembly | Average_covergeX | Total_Reads | Avg_length | Avg_Q | viral_reads% | Norovirus_reads% | Genotype | GeneHits | Re |
| R051225S1_S1_L001_001 | 1 | 7767 | 2986 | 1072924 | 73.1 | 34.41 | 18.48 | 17.32 | NorovirusGI_GI.3,NorovirusGI_GI.P3 | VP1,Pol | . |
| R051225S2_S2_L001_001 | 1 | 7810 | 4889 | 1153610 | 73.3 | 34.33 | 28.07 | 26.40 | NorovirusGI_GI.3,NorovirusGI_GI.P3 | VP1,Pol | . |
| R051225S3_S3_L001_001 | 1 | 7817 | 8414 | 1075406 | 82.2 | 34.68 | 49.78 | 49.07 | NorovirusGI_GI.3,NorovirusGI_GI.P3 | VP1,Pol | . |
| R051225S4_S4_L001_001 | 1 | 7819 | 4657 | 933362 | 85.0 | 34.69 | 31.02 | 30.40 | NorovirusGI_GI.3,NorovirusGI_GI.P3 | VP1,Pol | . |
| R051225S5_S5_L001_001 | 1 | 7687 | 752 | 1364484 | 73.3 | 32.43 | 3.97 | 1.50 | NorovirusGI_GI.6,NorovirusGI_GI.P11 | VP1,Pol | . |

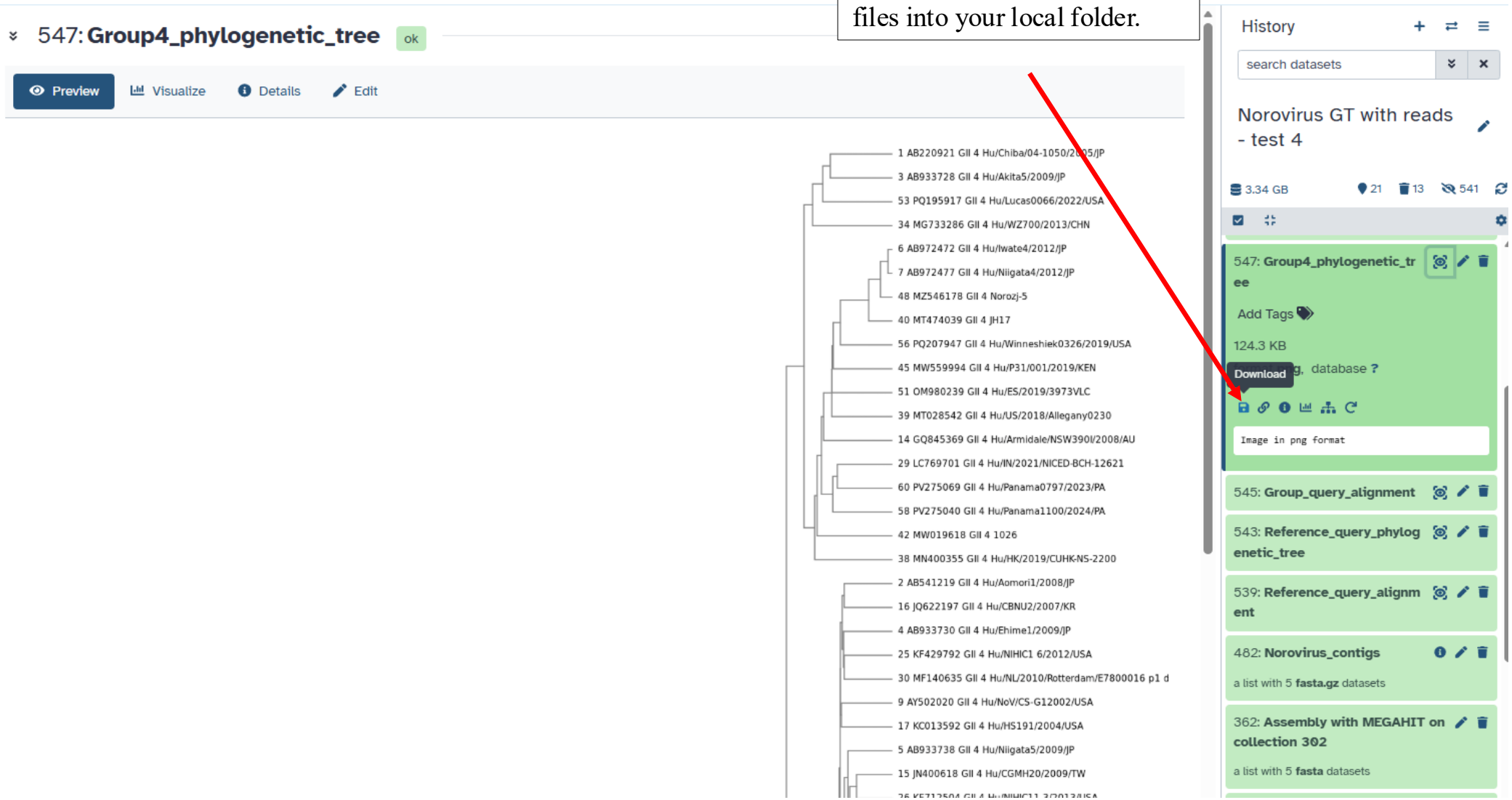



Click the **eye** icon to view results for each sample.

# Example ViroTrakr Phylogenetic Analysis Report

U.S. FOOD & DRUG ADMINISTRATION
Human Foods Program

Click on the **floppy** icon to download and save the results files into your local folder.



# Data Generation

- **No specific protocol requirements for generating sequencing data for ViroTrakr submissions**

- **Challenges and Efforts**
  - Low virus amounts in contaminated food
  - Foodborne viruses are generally non-culturable
  - Potential co-contamination and high viral genetic diversity
  - Need for sensitive and accurate detection assays

- **Ongoing work:**

Preliminary virus-panel sequencing results for norovirus-spiked frozen raspberries

| NoV-Spiked Berry Samples | RNA Input | Virus Panel WGS | Method Comparison |
|---|---|---|---|
| GI (4 distinct strains) | ~20 cps/15 uL (~1.3 cps/uL) | 100% | 100% |
| GII (4 distinct strains) | ~20 cps/15 uL (~1.3 cps/uL) | 100% | 80% |

# Summary and Continuing Efforts

## Summary

- Established ViroTrakr, a genomic platform for foodborne viral pathogens detection and identification, to strengthen genomic surveillance and support outbreak investigation and trackback.
- ViroTrakr workflows, particularly the norovirus workflow, has been adopted by stakeholders.
- Preliminary data indicate improved sensitivity of our new WGS method for detecting norovirus in food samples.

## Future directions

- Expand standardized bioinformatics workflows to cover all six foodborne viruses included in ViroTrakr
- Continue evaluating and optimizing, as needed, our new WGS for additional virus species in food samples
- Increase participation from national and international contributors and collaborators
- Expand the ViroTrakr database with additional genomic sequences and richer metadata

# Supporting Links and Resources

ViroTrakr NCBI BioProject (PRJNA396739): foodborne viruses (ID 396739) - BioProject - NCBI

ViroTrakr Submission Protocol: NCBI ViroTrakr submission protocol for foodborne virus surveillance_version 2 V.2

ViroTrakr genotyping and phylogenetic analysis workflows: Genotyping and phylogenetic analysis_ViroTrakr workflows

GalaxyTrakr publication (Jayanthi Gangiredla): GalaxyTrakr: a distributed analysis tool for public health whole genome sequence data accessible to non-bioinformaticians - PubMed

# Acknowledgements


ViroTrakr: (Technical Assistance: ViroTrakr@fda.hhs.gov)

Mark Mammel
Marc Allard
Jayanthi Gangiredla
Hugh Rand
Ruth Timme
James Pettengill
Maria Balkey
GenomeTrakr group; GenFS metadata group

Collaborative sequencing projects:

Collaborators (national and international) and ViroTrakr contributors
The Virology Team

Management: VPB, DFES, OAMT, OLOAS, HFP


**Thank you !**