

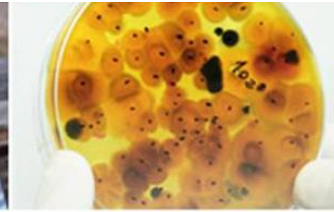
National Institute of Health (NIH) Dr. Ricardo Jorge, Lisbon, Portugal



INSaFLU-TELEVIR: an open web-based bioinformatics suite for metagenomic virus detection and routine genomic surveillance

<https://insaflu.insa.pt>

João Santos, PhD



National Institute of Health (NIH) Dr. Ricardo Jorge, Lisbon, Portugal

Background

Who we are

Genomic surveillance

INSaFLU

INSaFLU - TELEVIR

Demonstration

-- Home Page

-- Setup

-- Metagenomics

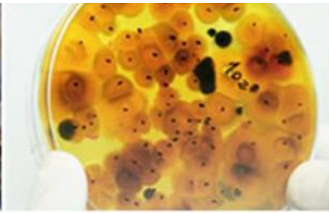
-- Documentation.

Input, Settings, Projects

Deployment & Results

Impact,
Added value

Demonstration
Results



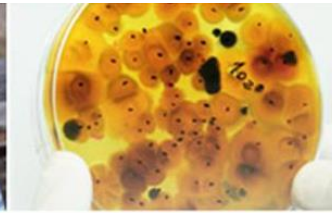
National Institute of Health (NIH) Dr. Ricardo Jorge, Lisbon, Portugal

Genomics and Bioinformatics Unit Department of Infectious Disease (DDI)

- detection and investigation of infectious outbreaks,
- investigation of emerging infectious agents,
- prediction of the resistance profile of microbial agents,
- prediction of the pathogenic potential of microbial agents.



Genomic Surveillance



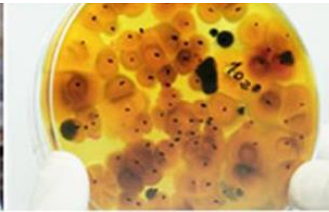
National Institute of Health (NIH) Dr. Ricardo Jorge, Lisbon, Portugal

Genomic surveillance is the process of constantly monitoring pathogens and analysing their genetic similarities and differences.

- expected seasonal epidemics – Influenza.
- unexpected pandemic threats – Zika, Ebola, Coronavirus.

this information is used to better understand how variants might impact public health.

Inform public policy



National Institute of Health (NIH) Dr. Ricardo Jorge, Lisbon, Portugal

Accessible

User Friendly

Fast

Open-Source

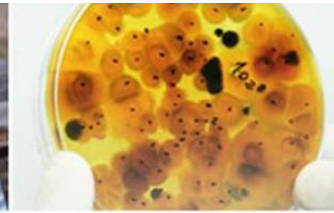
2018



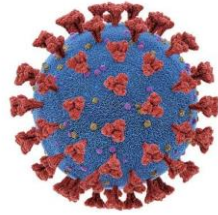
<https://insaflu.insa.pt>

INSaFLU: from “reads” to WGS-based first-line influenza laboratory surveillance

<https://doi.org/10.1186/s13073-018-0555-0>



National Institute of Health (NIH) Dr. Ricardo Jorge, Lisbon, Portugal



Adaptation to other pathogens (SARS-CoV-2)

Nextstrain



phylogeographical and temporal analysis



2018

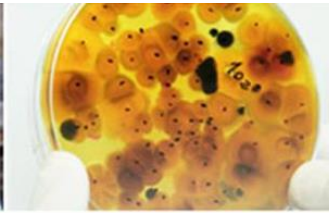
2021

2022

ONT, Ion Torrent

Genotype to Phenotype Analysis





Reference Based

*Targeted Genomic
Surveillance*

SARS-CoV-1:	2010
Ebola:	2013-2015,
West Nile :	2020 - 2023
SARS-CoV-2:	2020 - 2023
Ebola:	2022
Mpox :	2022-2023

Hypothesis free
diagnosis

Is there a Virus?

Which Virus?



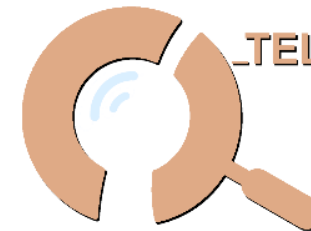
TELEVIR

Point-of-incidence toolbox for emerging virus threats

state of the art virus detection methods

Field
component

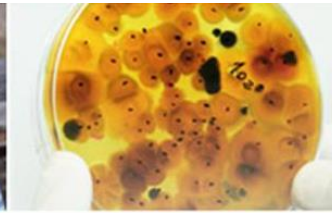
Dry-lab
component



TELEVIR pathogen detection module

December 21, 2022





National Institute of Health (NIH) Dr. Ricardo Jorge, Lisbon, Portugal



INSaFLU-TELEVIR: an open web-based bioinformatics suite for metagenomic virus detection and routine genomic surveillance


<https://insaflu.insa.pt>




INSaFLU-TELEVIR: an open web-based bioinformatics suite for metagenomic virus detection and routine genomic surveillance

<https://insaflu.insa.pt>

INPUTS

✓ **NGS data *** 

+

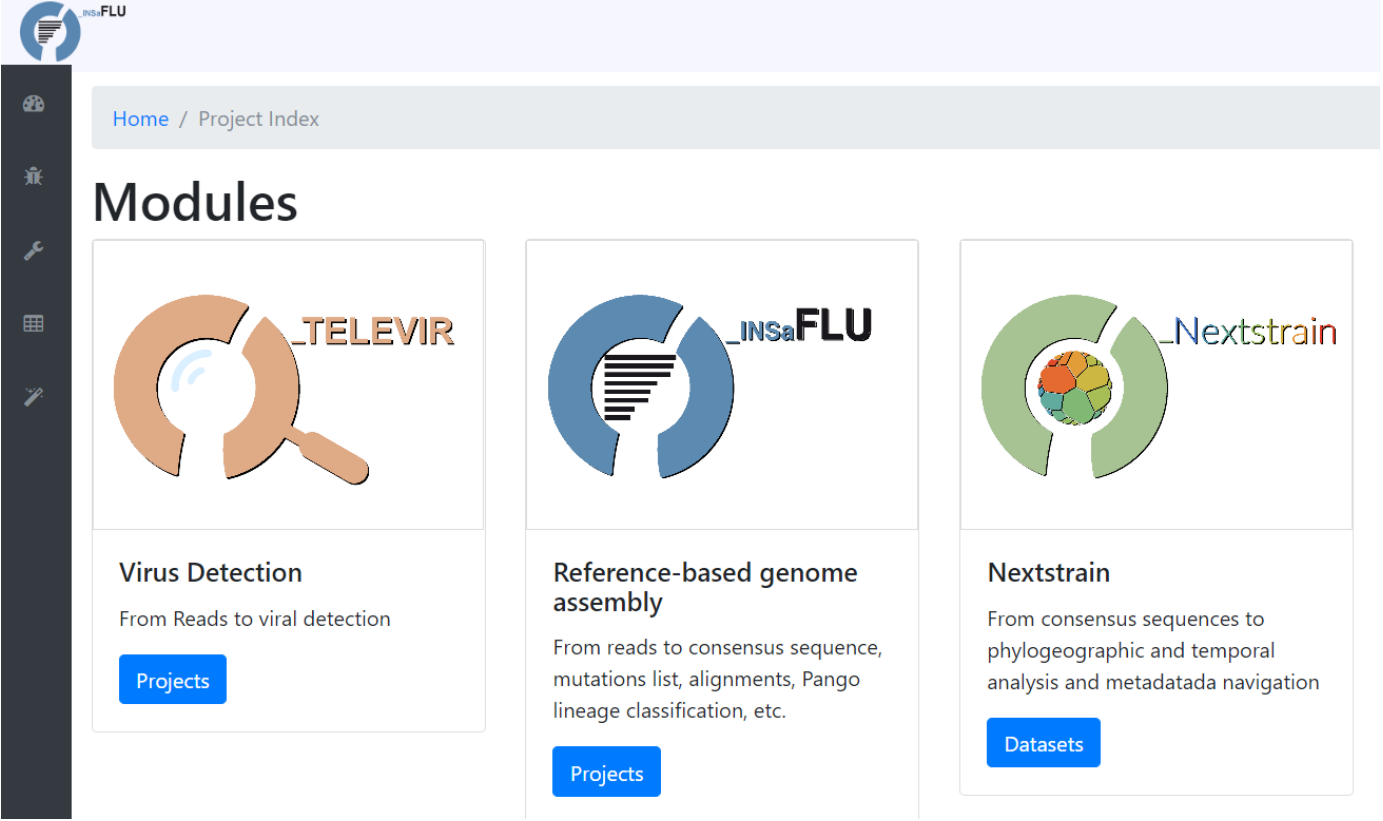
✓ **Sample metadata table** 
(tsv/csv format)

just Drag&Drop [fastq.gz](#) reads

**Compatible with distinct sequencing technologies:
Illumina, Ion Torrent and Oxford Nanopore*

Software parameters can be user-defined.
Settings and versions are recorded for backward traceability

Quality Control



Home / Project Index

Modules

- Virus Detection**
From Reads to viral detection
[Projects](#)
- Reference-based genome assembly**
From reads to consensus sequence, mutations list, alignments, Pango lineage classification, etc.
[Projects](#)
- Nextstrain**
From consensus sequences to phylogeographic and temporal analysis and metadata navigation
[Datasets](#)

New virus detection module (under Virus Detection)

Genome-based surveillance-oriented modules (under Reference-based genome assembly and Nextstrain)



INSaFLU-TELEVIR: an open web-based bioinformatics suite for metagenomic virus detection and routine genomic surveillance (influenza, SARS-CoV-2 and other pathogens)

<https://insaflu.insa.pt>

OUTPUTS

ROUTINE GENOMIC SURVEILLANCE

- ✓ Rapid classification of Human betacoronaviruses, Mpox, RSV and influenza (directly from reads)
- ✓ SARS-CoV-2 “Pango lineage” classification
- ✓ Intra-host minor variant detection (including flagging of putative mixed infections)
- ✓ Interactive coverage analysis (IGV)
- ✓ Mutation annotation and consensus sequence generation/curation
- ✓ Gene- and genome-based alignments and phylogenies
- ✓ Direct link to “Nextclade” navigation
- ✓ Genotype-phenotype screening





INSaFLU-TELEVIR: an open web-based bioinformatics suite for metagenomic virus detection and routine genomic surveillance (influenza, SARS-CoV-2 and other pathogens)

<https://insaflu.insa.pt>

OUTPUTS

ADVANCED PHYLOGEOGRAPHY

- ✓ Integrative Nextstrain phylogenetic, temporal and geographic analyses
- ✓ Flexible and scalable analysis
- ✓ Genetic modules for SARS-CoV-2, MPXV, seasonal influenza, RSV
- generic module for other viruses
- ✓ Dynamic metadata integration and visualization



Powered by  Nextstrain



<https://nextstrain.org/>

INSaFLU-TELEVIR development was co-funded by the European Commission on behalf of **OneHealth EJP**





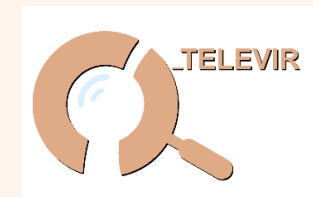
INSaFLU-TELEVIR: an open web-based bioinformatics suite for metagenomic virus detection and routine genomic surveillance (influenza, SARS-CoV-2 and other pathogens)

<https://insaflu.insa.pt>

OUTPUTS

VIRUS DETECTION (*TELEVIR module*)

- ✓ State of the art software
- ✓ Modular Pipeline
 - ✓ Multiple Classification methods
 - ✓ Multiple Viral databases
- ✓ Interactive Summary statistics and intuitive end-user reports



<https://insaflu.insa.pt>

INSaFLU-TELEVIR development is being co-funded by the European Commission on behalf of OneHealth EJP

TELE-Vir project: <https://onehealthejp.eu/jrp-tele-vir/>

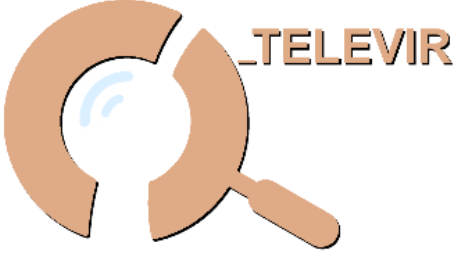




INSaFLU-TELEVIR: an open web-based bioinformatics suite for metagenomic virus detection and routine genomic surveillance (influenza, SARS-CoV-2 and other pathogens)

<https://insaflu.insa.pt>

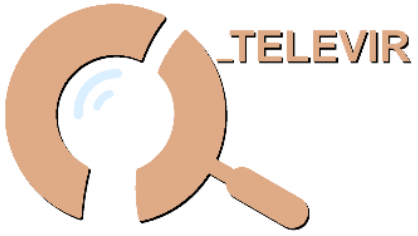
TELEVIR Module - Project TELEVIR (OneHealth EJP)



Virus Detection
From Reads to viral detection

Projects

The image shows a card for the TELEVIR module. At the top is the TELEVIR logo, which consists of a magnifying glass with a blue circular graphic inside the lens and the text "_TELEVIR" to its right. Below the logo, the text "Virus Detection" is displayed in a bold font, followed by "From Reads to viral detection" in a smaller font. At the bottom of the card is a blue button with the text "Projects" in white.



TELEVIR Module - Project TELEVIR

Virus detection from reads



Instituto Nacional de Saúde
Doutor Ricardo Jorge



Remove reads from host

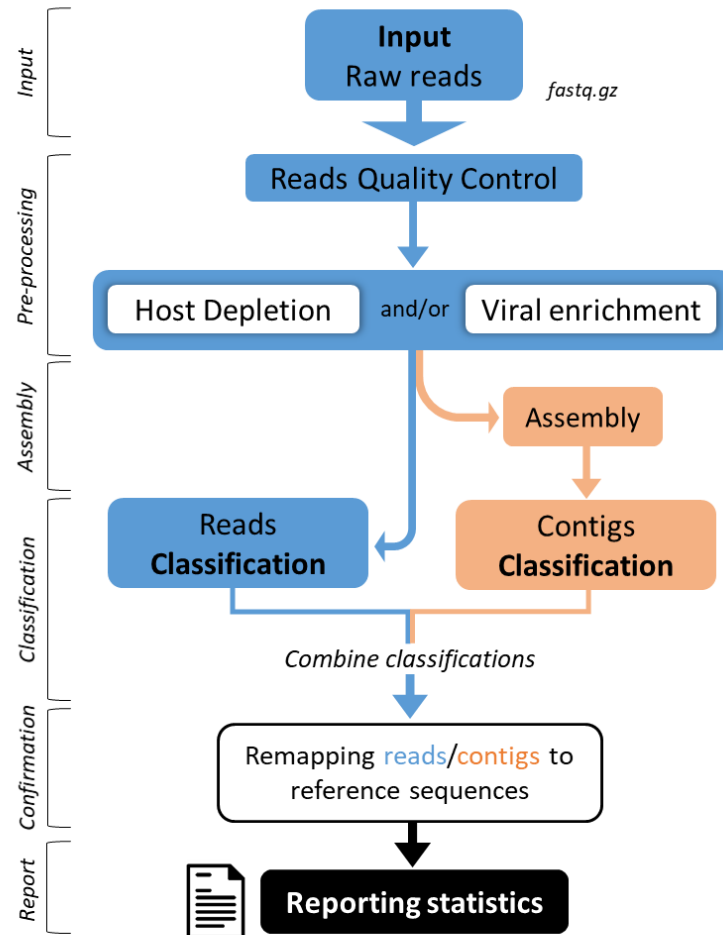
Keep the viral reads in the sample

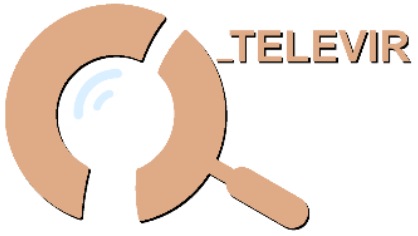
Assembly

Taxonomic classification of reads
and contigs

Reads/contigs are remapped against
reference genomes of the identified
taxa

Summary statistics of the identified hits
and intuitive end-user reports





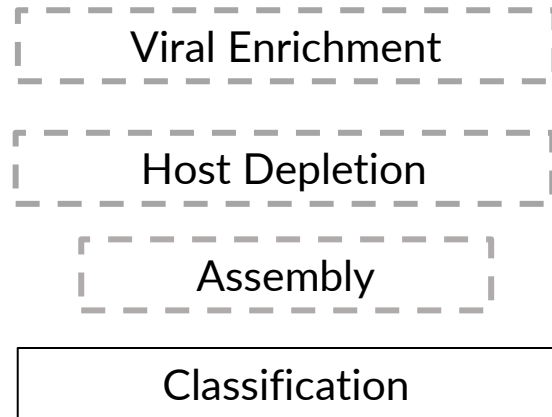
TELEVIR Module - Project TELEVIR

Virus detection from reads



A Modular Pipeline

Module




Software

Centrifuge, Kaiju, KrakenUniq, Kraken2

BWA, Minimap2, Bowtie

Spades (meta, metaviral), Raven

Centrifuge, Kaiju, KrakenUniq, Kraken2, Minimap2, FastViromeExplorer

 Optional modules

Cross validation is important

Integration of new tools

No one best approach
algorithm / database



Summary of the current INSaFLU main modules and outputs:

https://insaflu.readthedocs.io/en/latest/_downloads/a5fe7435b10e27d97fa58ef996fd8118/INSaFLU_current_outputs_26_10_2022.xlsx

Online tool: <https://insaflu.insa.pt>

Documentation / Tutorial: <https://insaflu.readthedocs.io/en/latest/>

Code: <https://github.com/INSaFLU/INSaFLU>

Easy local installation: <https://github.com/INSaFLU/docker>

Contact: vitor.borges@insa.min-saude.pt

TELEVIR Module - Project TELEVIR (OneHealth EJP)

Virus detection from reads

Run													
Pre-Processing													
Assembly													
Reads and Contigs Classification													
Remapping													
Pathogen Identification													
Multiple useful statistics in report											Download as CSV		
Description	Taxid	accID	Cov (%)	Depth	DepthC	Mapped reads	start prop (%)	mapped_prop (%)	Gaps	Windows Covered	class. success	mapping success	Warning
Influenza A virus (A/Korea/426/1968(H2N2)) segment 8, complete sequence	488241	NC_007380.1	98.57	10.17	11.42	16	0.15035	0.48222	4	3/3	reads	reads	
Influenza A virus (A/New York/392/2004(H3N2)) segment 1, complete sequence	335341	NC_007373.1	98.42	46.89	47.73	87	0.81752	2.62206	2	3/3	reads and contigs	reads	
Influenza A virus (A/Korea/426/1968(H2N2)) segment 1, complete sequence	488241	NC_007378.1	98.33	96.01	97.81	188	1.76659	5.66606	3	3/3	reads	reads	
Influenza A virus (A/New York/392/2004(H3N2)) segment 2, complete sequence	335341	NC_007372.1	97.57	418.13	439.14	963	9.04905	29.02351	2	3/3	reads and contigs	reads	

Horizontal coverage

Depth of coverage

N and % of reads mapped

TELEVIR Module - Project TELEVIR (OneHealth EJP)

Virus detection from reads

Description	Taxid	accID	Cov (%)	Depth	DepthC	Mapped reads	start prop (%)	mapped_prop (%)	Gaps	Windows Covered	class. success	mapping success	Warning
Influenza A virus (A/Korea/426/1968(H2N2)) segment 8, complete sequence	488241	NC_007380.1	98.57	10.17	11.42	16	0.15035	0.48222	4	3/3	reads	reads	
Influenza A virus (A/New York/392/2004(H3N2)) segment 1, complete sequence	335341	NC_007373.1	98.42	46.89	47.73	87	0.81752	2.62206	2	3/3	reads and contigs	reads	
Influenza A virus (A/Korea/426/1968(H2N2)) segment 1, complete sequence	488241	NC_007378.1	98.33	96.01	97.81	188	1.76659	5.66606	3	3/3	reads	reads	
Influenza A virus (A/New York/392/2004(H3N2)) segment 2, complete sequence	335341	NC_007372.1	97.57	418.13	439.14	963	9.04905	29.02351	2	3/3	reads and contigs	reads	

Database: /televir/mngs_benchmark/ref_fasta/refseq_viral.genome.fna.gz NCBI length: 2341 contig string: NC_007372.1

Mapping Coverage

IGV

.bam mapped reads (.fa)

.bai

Each hit is an “expand-and-collapse” panel that can be further visualized and explored

- Visualization of coverage throughout reference genome
- Mapped reads and contigs can be downloaded for further confirmations

etc



Summary of the current INSaFLU main modules and outputs:

https://insaflu.readthedocs.io/en/latest/_downloads/a5fe7435b10e27d97fa58ef996fd8118/INSaFLU_current_outputs_26_10_2022.xlsx

Online tool: <https://insaflu.insa.pt>

Documentation / Tutorial: <https://insaflu.readthedocs.io/en/latest/>

Code: <https://github.com/INSaFLU/INSaFLU>

Easy local installation: <https://github.com/INSaFLU/docker>

Contact: vitor.borges@insa.min-saude.pt