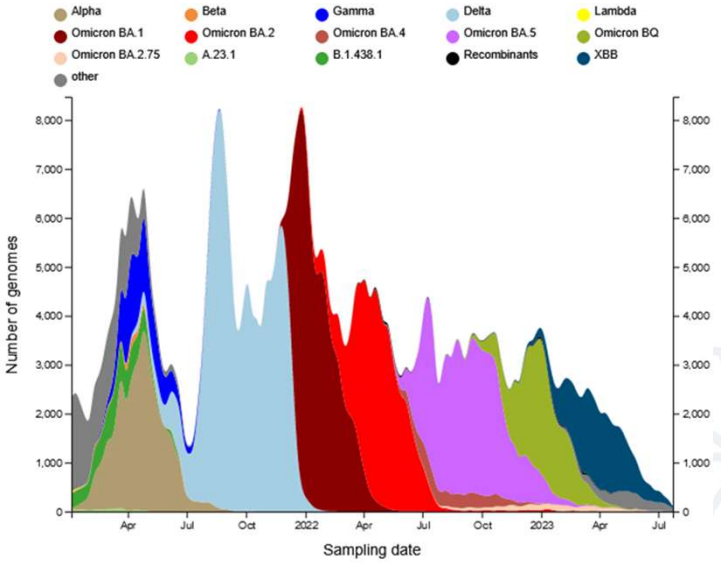
# DUOTANG

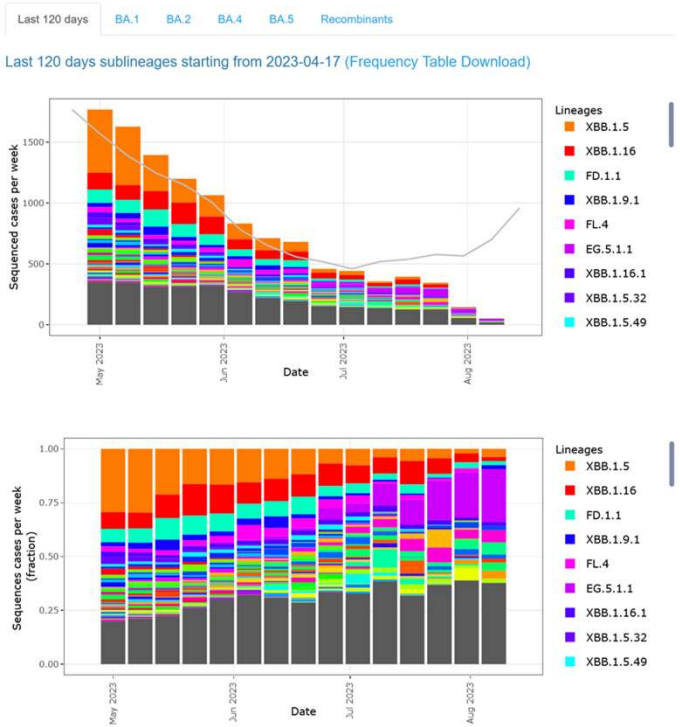
## A Genomic Epidemiology Analyses and Mathematical Modelling Notebook

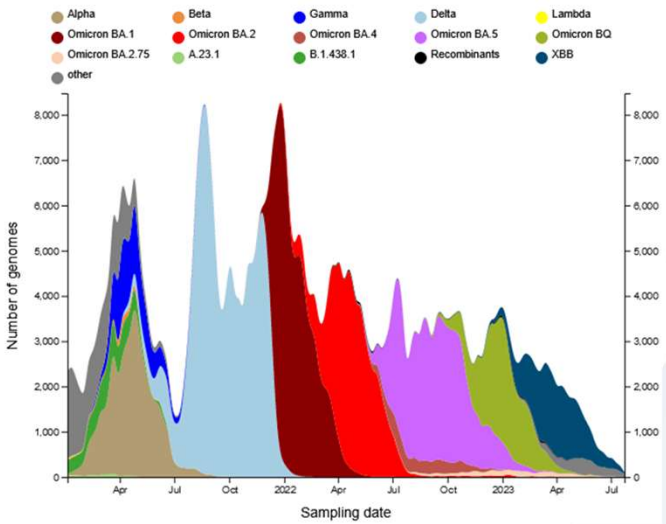
# Duotang

- Web notebook for genomic epidemiology analysis and mathematical modelling designed originally for SARS-CoV-2.
- Semi-automated report generation with expert review.
- Includes multiple, interactive, plots for analysis of frequency changes, estimating selection on variants, phylogenetics, and more.
- **Goal:** investigate SARS-CoV-2 evolution and spread within the country and to spur scientific discussion and communication with public health authorities.

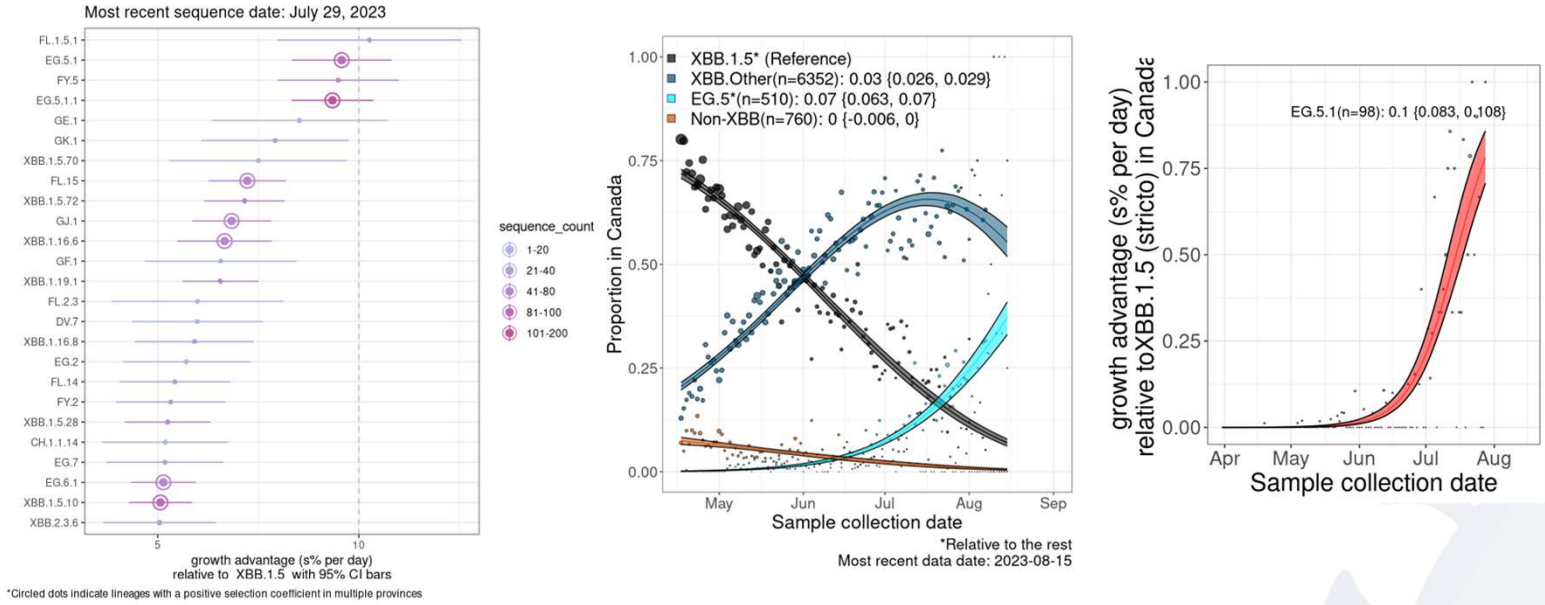


# Plots: Frequency Changes



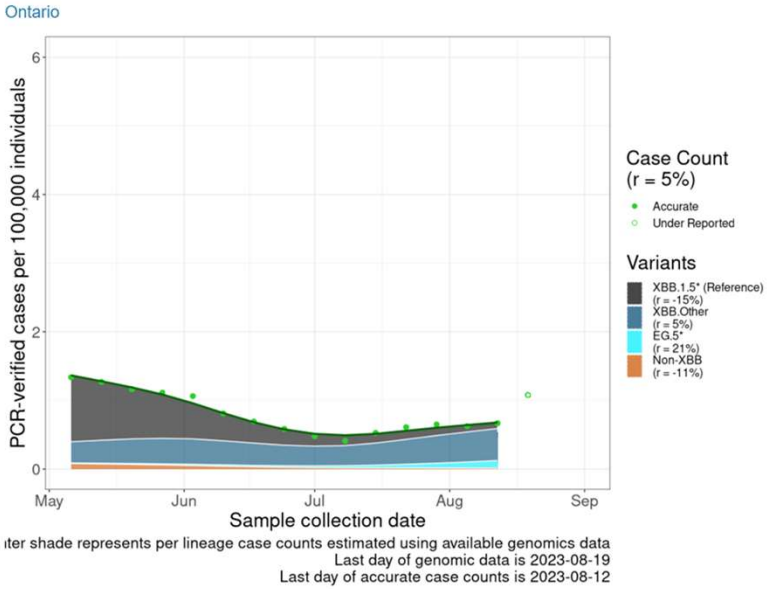



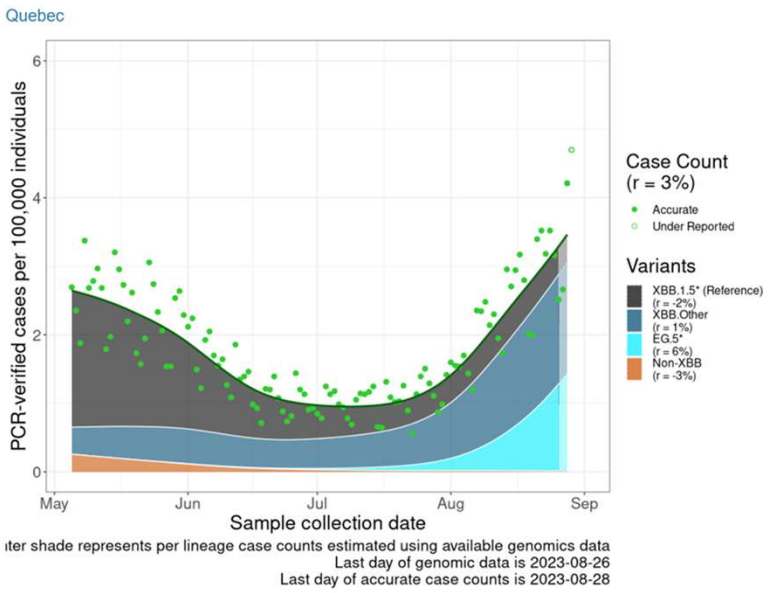
# Plots: Estimating Selection



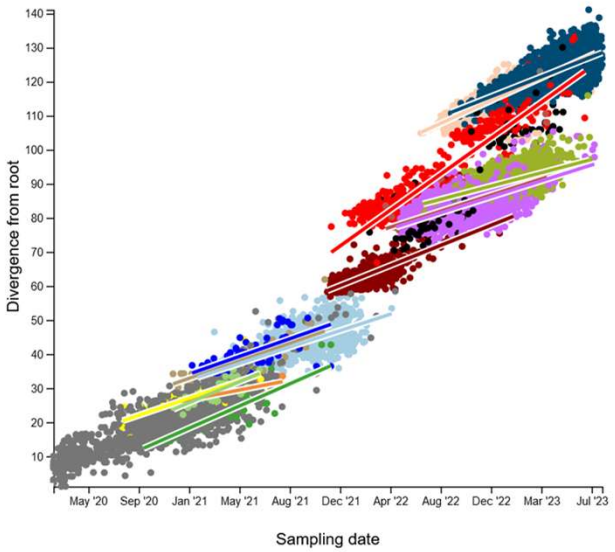

**148 named lineages have been circulating in Canada over the last month.** Measuring the selective advantage of each relative to XBB.1.5, the fastest growing lineages in multiple provinces and with sufficient data is EG.5.1 (s ~ 10%).

# Plots: Estimating Case Counts

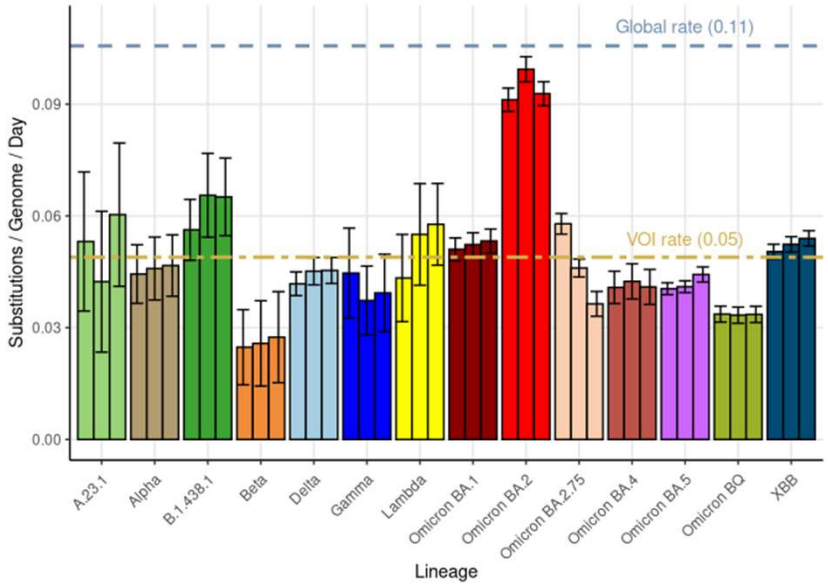


# Plots: Phylogeny and Estimating Substitution Rates



# Plots: Phylogeny and Estimating Substitution Rates



# Other Information

- Weekly updates on the current SARS-CoV-2 landscape in Canada (Current Situation)
- Searchable lineage descriptions and ancestry information
- Detailed methodologies and source information
- Archive of past notebooks
- Open source, freely available.

# Duotang as a Framework & Service

- Expansion of tools and methodologies learned from SARS-CoV-2 into other emerging viruses and data types.

- Enable rapid and automated deployment for visualization of custom datasets without requiring an in-house bioinformatician.

- Can serve as a framework for other jurisdictions interested in open, collaborative sequence data sharing and analysis.

# Usage Example

- •E.g. seeing BA.2 lineage contains multiple more divergent lineages based on the high substitution rate of the group.

- •Find robust growth rates not so alarming versus international tools
- •Alert both when there is a problem but also when there isn't a problem.

- •Identify lineages of note specifically in the Canadian context.

- Notes from chat:
- I like those use cases — when there is, and when there isn't, a problem
- Maybe a concrete example: we knew that [whatever] was a problem because .. but we could see that [whatever else] was not because [.. ]

- BA.2.86 is not growing really notably anywhere, so avoid referring to alarming growth

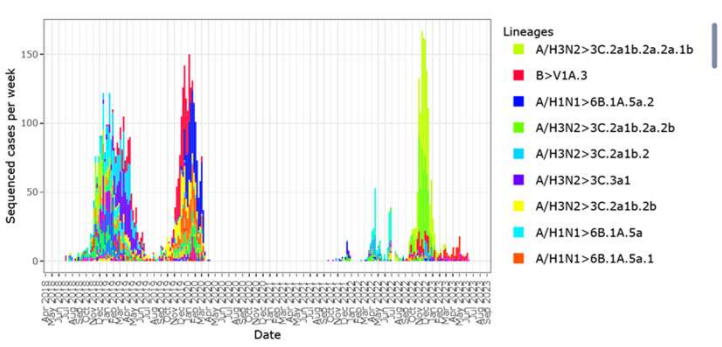
# Duotang Can Be Virus Agnostic

CoVaRR-Net

## CAMEO - Fluotang

Fluotang, a genomic epidemiology analyses and mathematical modelling notebook. For Flu!

### Influenza In Canada

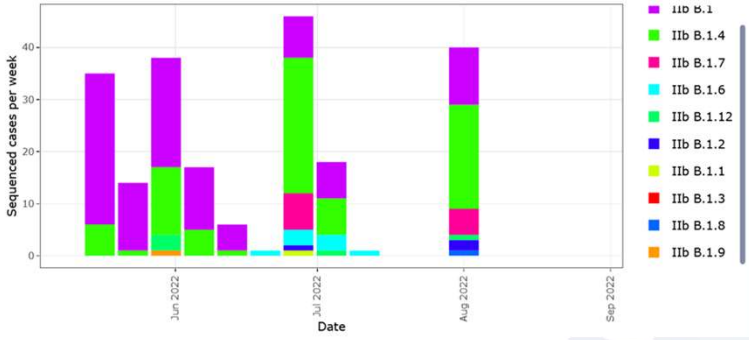



CoVaRR-Net

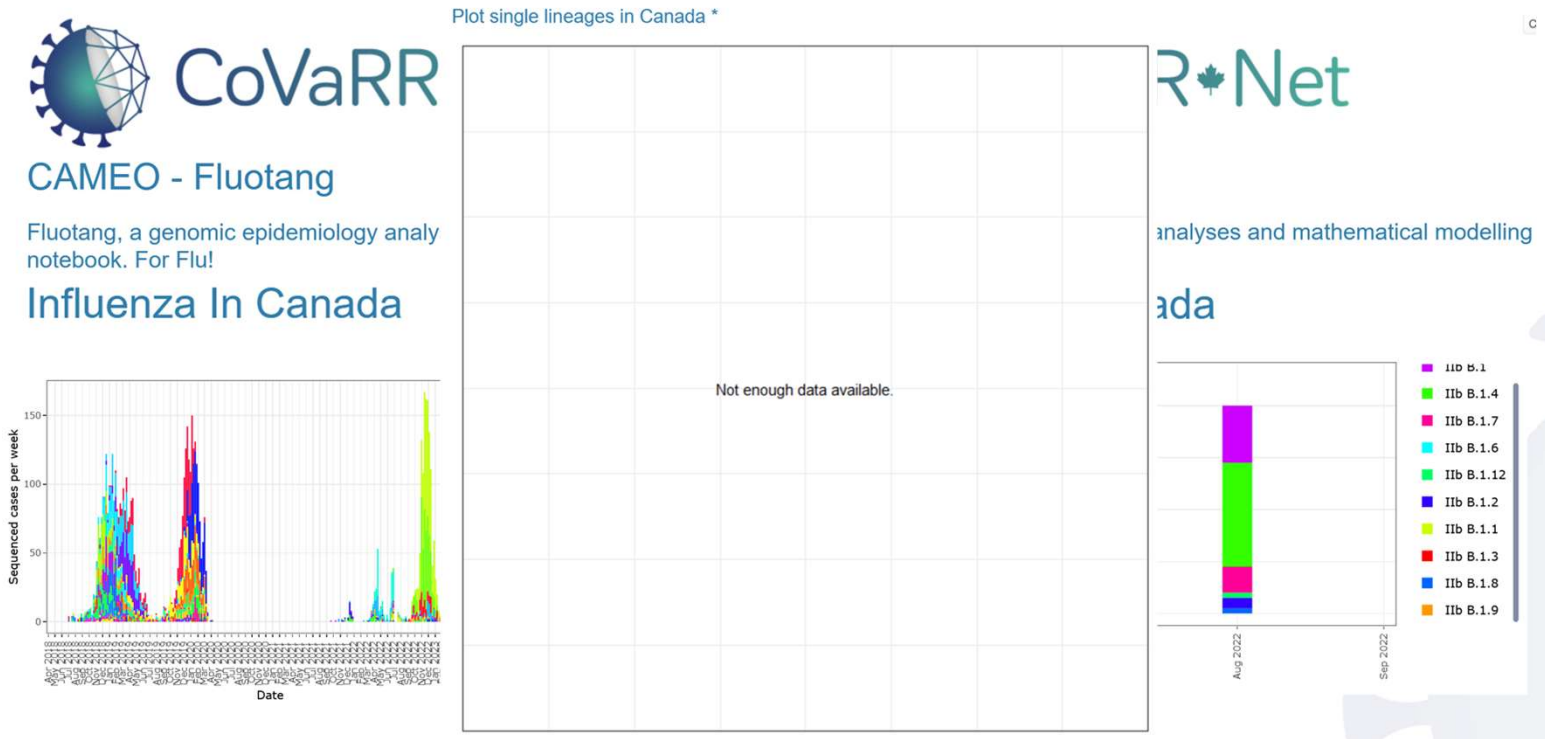
## CAMEO - MPuotang

MPuotang, a genomic epidemiology analyses and mathematical modelling notebook. For Monkey Pox!

### Monkey Pox In Canada



# Expansion Into Emerging Pathogens



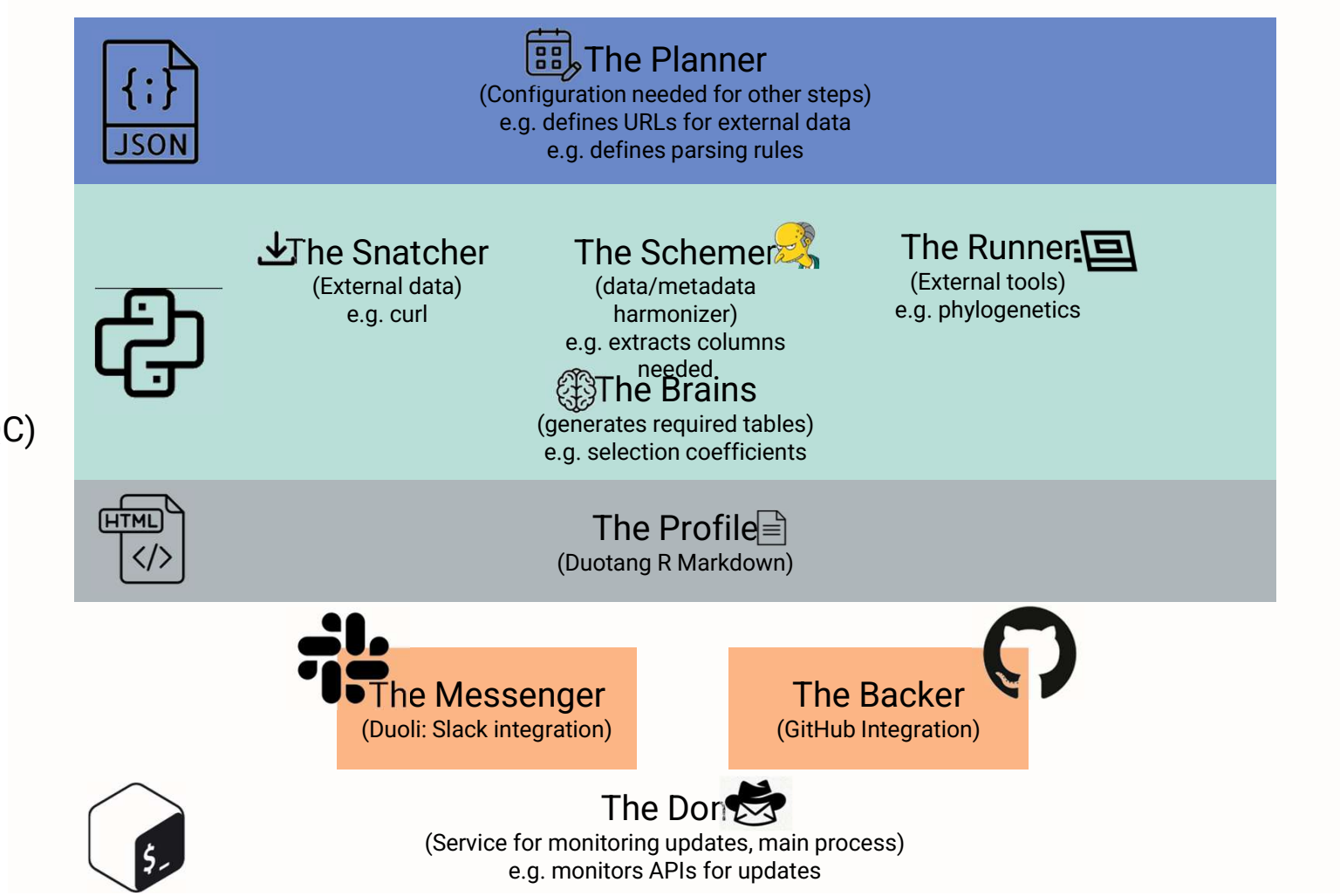
# Duotang Can Be Data Agnostic

# DuoGang (Not actual name)

## Modular framework for data-independent Duotang notebooks

Current supported inputs:

- COVID (VirusSeq, GISAID)
  - Duotang
- Flu (GISAID)
  - Fluotang
- Wastewater (Steven/Mock QC)
  - WWuotang



# Duotang Notebook Creation Time

- SARS-CoV-2: Years
- Wastewater: Months
- Influenza: Weeks
- Monkey Pox: Days
- Emerging pathogens & Custom datasets: Hours (?)