

# Whole Genome Sequencing Analysis of SARS-CoV-2 Household Transmission in Chiang Mai, Thailand During 2022–2024



Premmarin Inmonthian, Pitaya Suebtam, Patumrat Sripan,  
Nantawan Wangsaeng, Nang Kham-Kjing, Aphirak Pinasu,  
Napatsakorn Kohklang, Jeannie Collins, Moira Spyer,  
Chaisiri Angkurawaranon, Nicole Ngo-Giang-Huong, Woottichai Khamduang



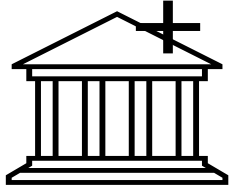
Funded by the  
European Union

The VERDI project (101045989) is funded by the European Union. Views and opinions expressed are however those of the VERDI Consortium only and do not necessarily reflect those of the European Union or the European Health and Digital Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

# Outline

- Context, rationale, and objective
- Methods
- Results
- Conclusion

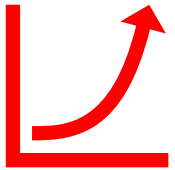
# Context of Thailand in 2022



- The Ministry of Public Health declared COVID-19 endemic



- Lockdown and isolation measures were lifted



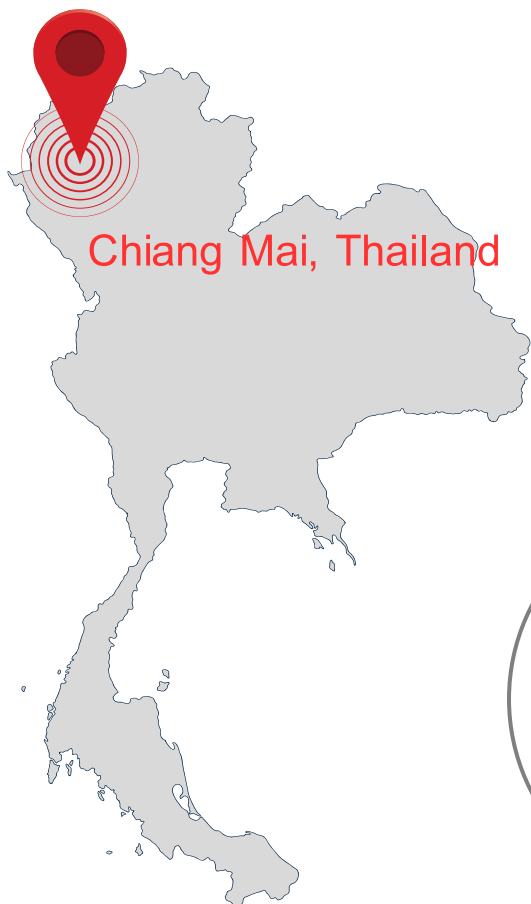
- Omicron spread in Thailand
  - Omicron variants are **highly transmissible** and immune escape potential (Alam, M. S., 2023)
  - Despite high national vaccination coverage, infections continued to occur.

# Rationale of household study

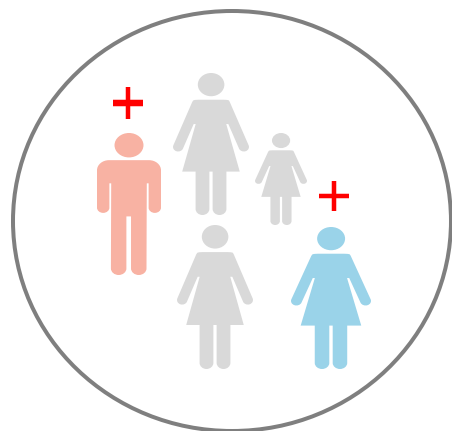
- Households are the major site for **SARS-CoV-2 transmission**.
  - Close contact, shared spaces, and inconsistent adherence to preventive measures
  - High risk of **secondary infections**
- Most studies on transmission are based on detection of SARS-CoV2 RNA
- The route of transmission was not clearly identified

# Objective

To analyze transmission of SARS-CoV-2 in households using whole genome sequencing and phylogenetic analysis



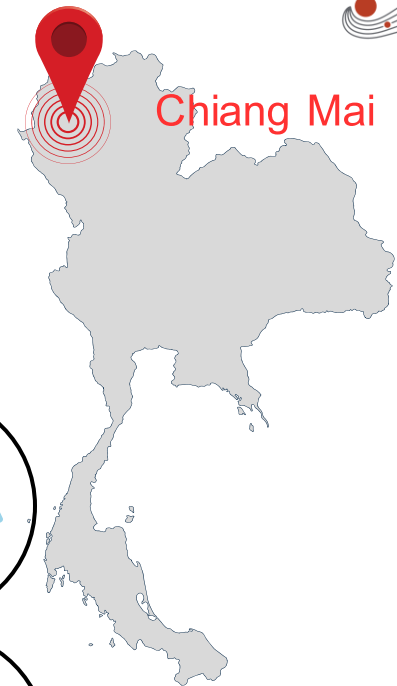
- Prospective observational study in Chiang Mai, Thailand
- Between July 2022 and May 2024
- To assess the occurrence of SARS-CoV-2 in household contacts and estimate the key transmission parameters.



## Household enrollment criteria

- ✓ Adult or child tested SARS-CoV-2 antigen positive within 48 hours (index case)
- ✓ Household has **≥2 members**, including the index case
- ✓ At least **1 child <18 years**
- ✓ At least **one member uninfected** at enrollment date
- ✓ Household has access to **smartphone, tablet, or phone** with internet
- ✓ **All members** provided informed consent or assent

# The Thailand VERDI-RECOVER study procedures

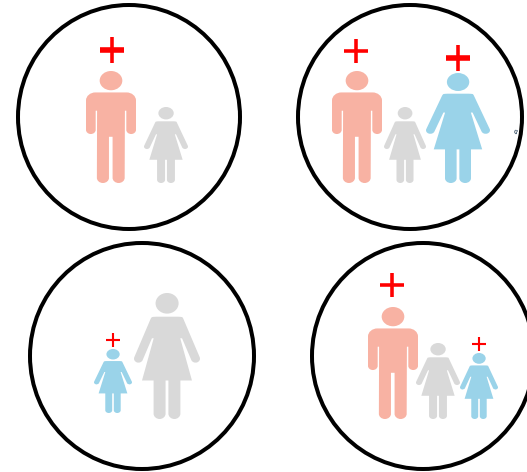


Chiang Mai



Mobile unit

- Informed consent form, questionnaire
- NP/T swabs-> Ag testing
- Blood draw-> serology



- Day 21: closing
- Blood draw
  - Questionnaire

Household members: fill daily diary/symptoms - antigen test every other day

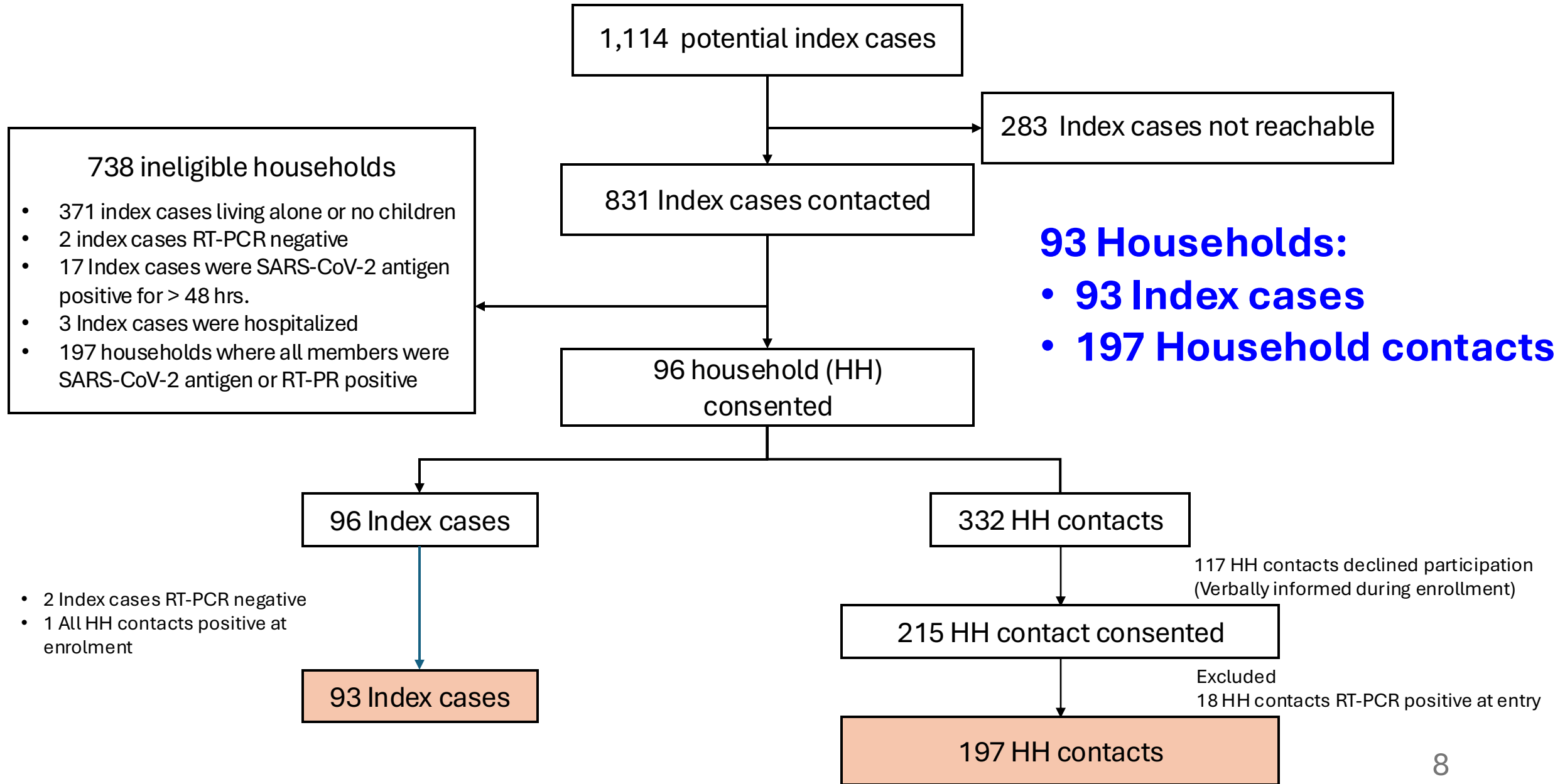
If positive SARS-CoV-2 antigen, follow-up 14 days



- Blood draw
- NP/T swabs

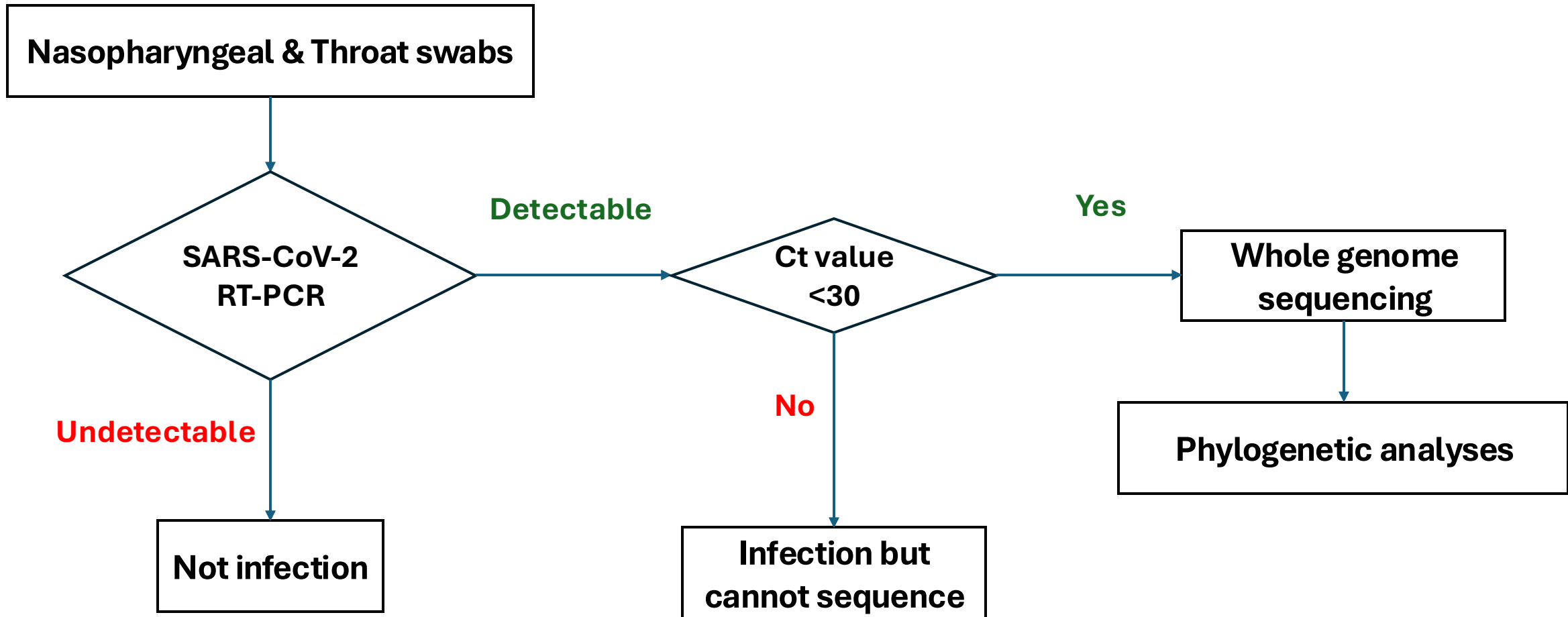
- Closing
- Blood draw
  - Questionnaire

# Participant Enrollment





# Molecular laboratory workflow



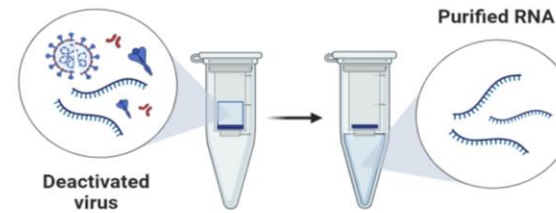
# Detection of SARS-CoV-2 RNA in nasopharyngeal swabs

## ■ RNA extraction

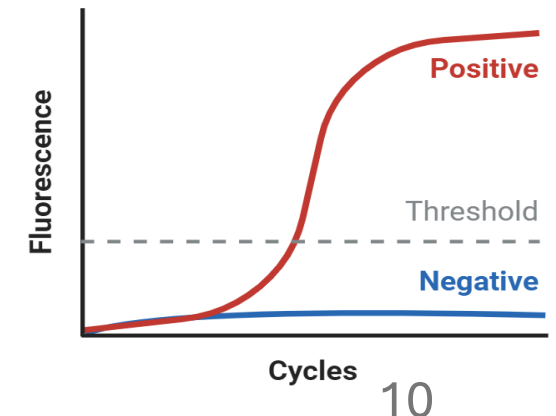
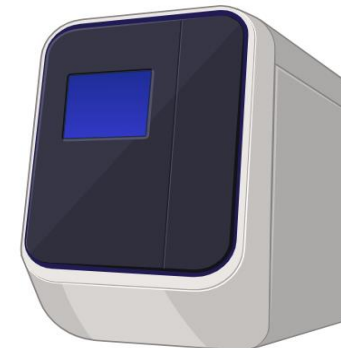
- Magnetic bead-based purification  
( Zybio Nucleic Acid Extraction Kit B200-32T,China )

## ■ RT-PCR assay for N and S genes

- SARS-CoV-2 Nucleic Acid Detection Kit  
(Tellgen Corp., Shanghai, China)
- Criterion for detectable RNA **Ct ≤ 40**



**Sample was considered positive if RT-PCR for N and S genes were positive**



# Whole genome sequencing using Nanopore technologies.

- ARTIC SARS-CoV-2 v4, 2-pool amplicon-based strategy
- Reverse transcription and multiplex PCR (ARTIC v5.3.2 primers)
- Library preparation: Native barcoding (SQK-NBD114-24) and adapter ligation using ONT kits
- Sequencing conducted on MinION Mk1C platforms



# Bioinformatic and phylogenetic analysis

Raw signal

Pod5

- Base calling
- Adapter trimming
- demultiplex

*Dorado base caller*

Raw reads

(sup)

**Cleaning**

- Reads: 400-600 bp
- Phred score: >10

*fastp*

Clean reads

(sup)

**Aligning**

- Wuhan.Hu\_1 (NC.045512.2)

*dorado aligner*  
*Samtools sort*

Mapped reads

(BAM)

**consensus**

Consensus reads

(fasta)

- *samtools mpileup*
  - Phred score 50 mean 99.999 accuracy
- *ivar consensus*
  - Coverage: >10

Aligned multiple sequences

multiple sequence alignment

- *mafft*

Phylogenetic tree

Nextstrain

Metadata

- Name/label
- Type of virus
- Date of collection

visualization

R

ggplot and ggtree

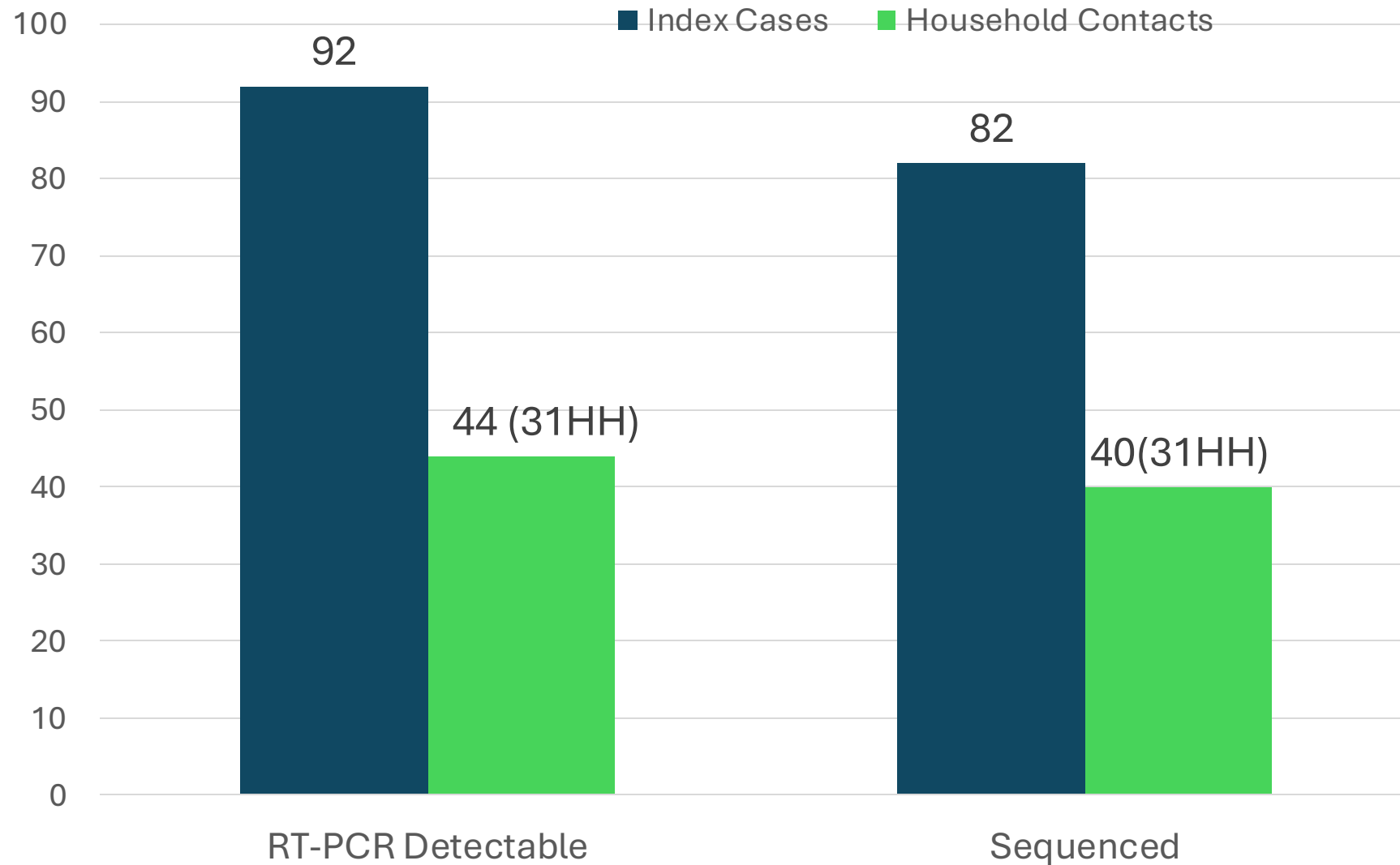
Auspice.us

# Results

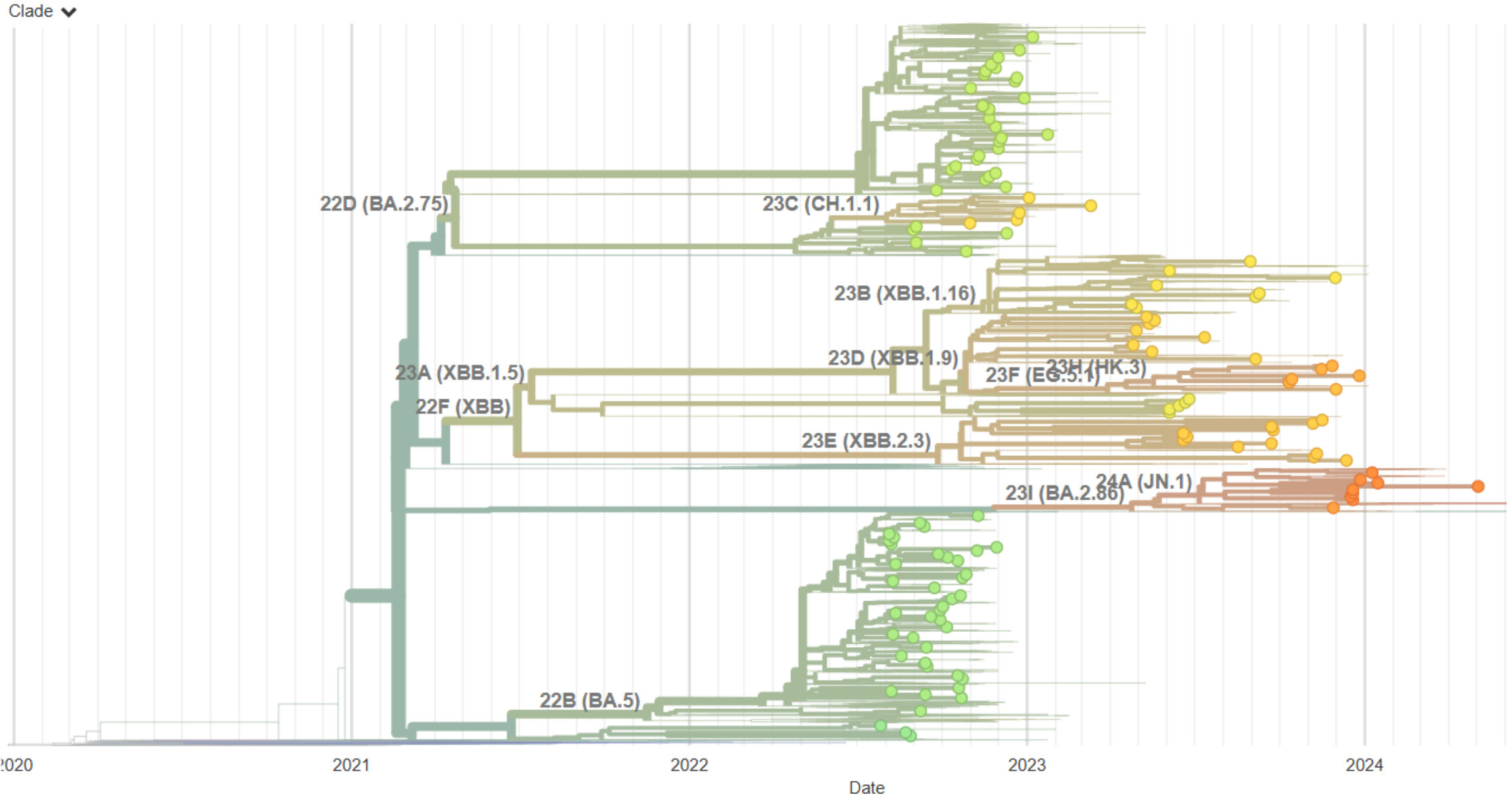
# Sample testing and sequencing overview

93 index cases 197 contacts

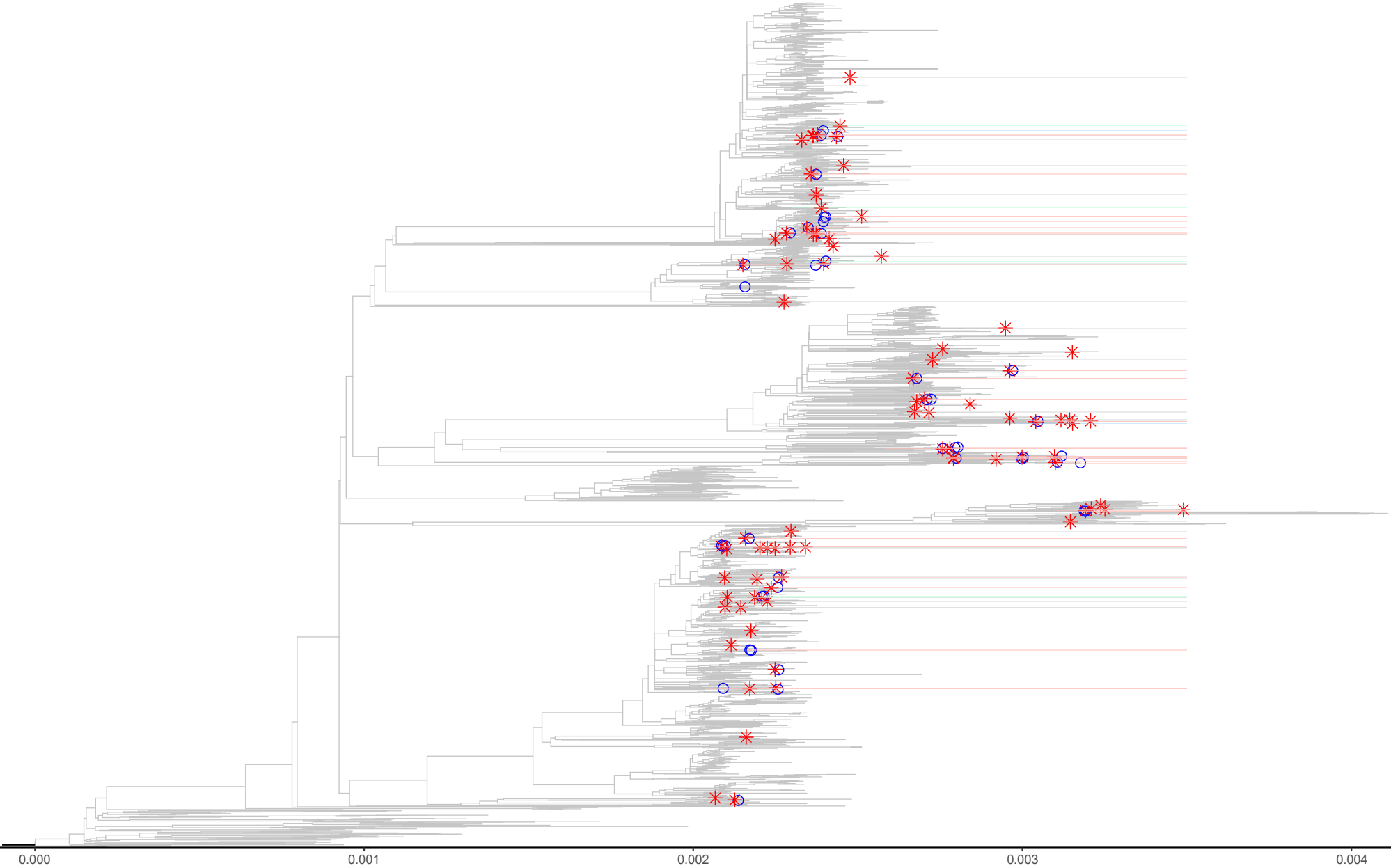
## Testing and Sequencing Results



# Phylogeny of SARS-CoV-2 Genomes from Chiang Mai Households (2022–2024)



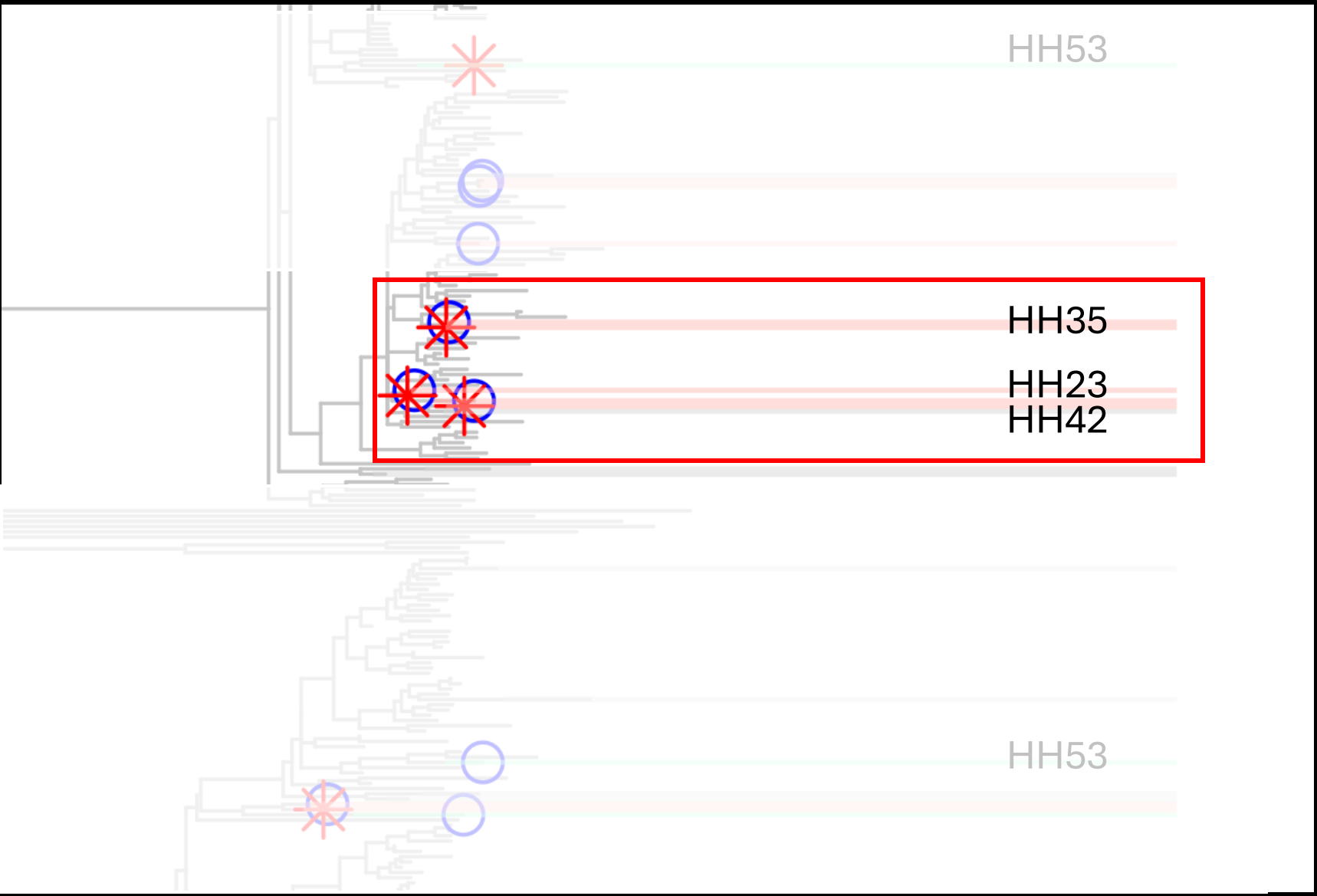
# Phylogenetic tree (R ggplot, ggtree)



- Case
- Contact
  - \* Index
- Route of transmission
- Household infection
  - Household transmission
  - Indetermine
  - No new infection

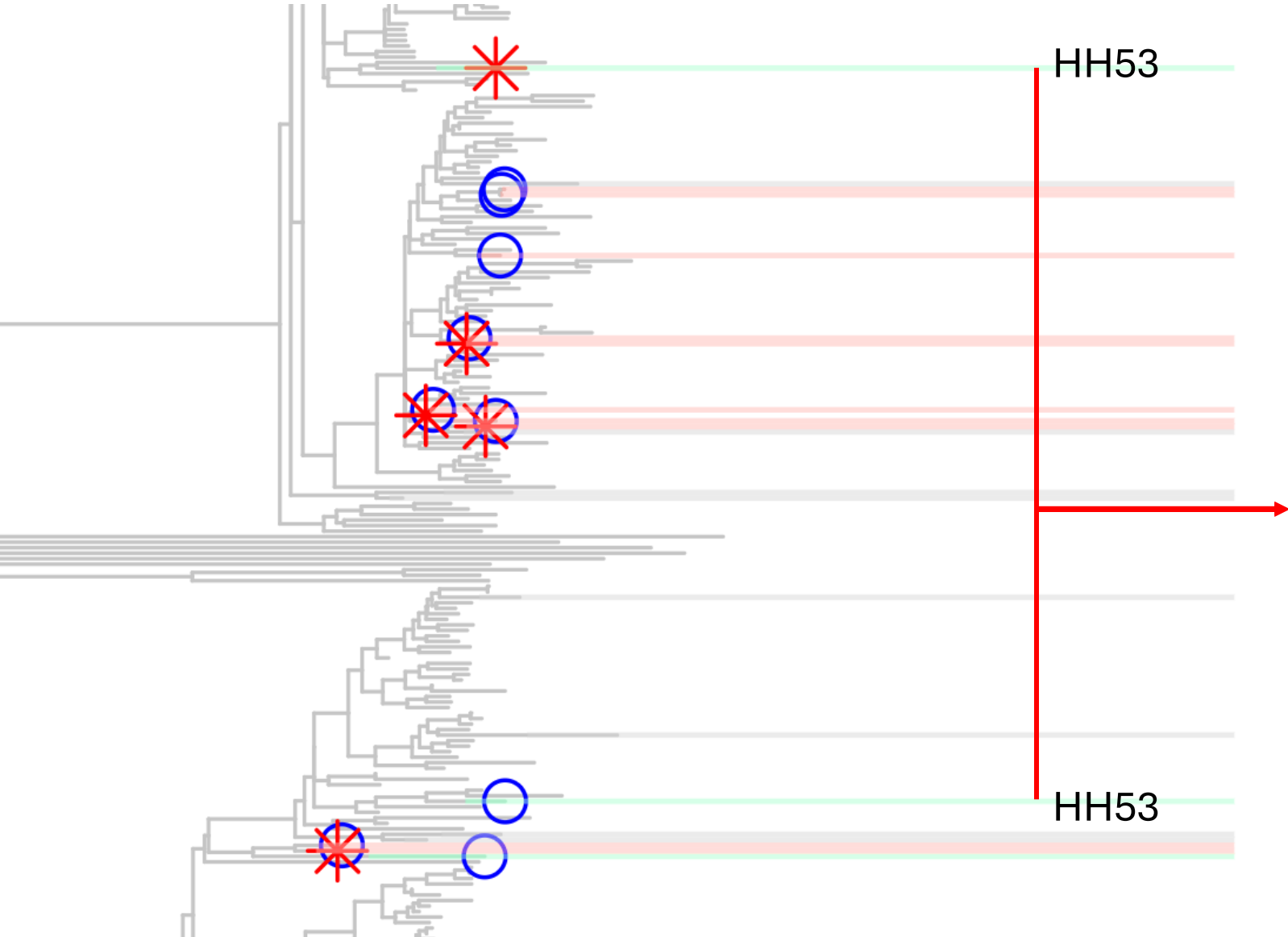


# Phylogenetic tree (R ggplot, ggtree)



**Intra-household  
transmission**

# Phylogenetic tree (R ggplot, ggtree)



**Community acquired transmission**

- 44 HH contacts across 31 households were infected.
  - **36 HH contacts (26 households)**, were infected through **intra-household transmission**.
  - **3 HH contacts (3 households)** were **community acquired infection**
  - 5 HH contacts (2 households) had no sequences available, either index or contacts

# Conclusion

- Intra- household transmission was the predominant source of SARS-CoV-2 infections
- SARS-CoV-2 genomic data is crucial in identifying infection routes and tracking the genetic evolution of the virus.
- Genomic data help to inform the development of targeted public health interventions.

# Acknowledgement

- Participating households
- VERDI study team
- Provincial health office



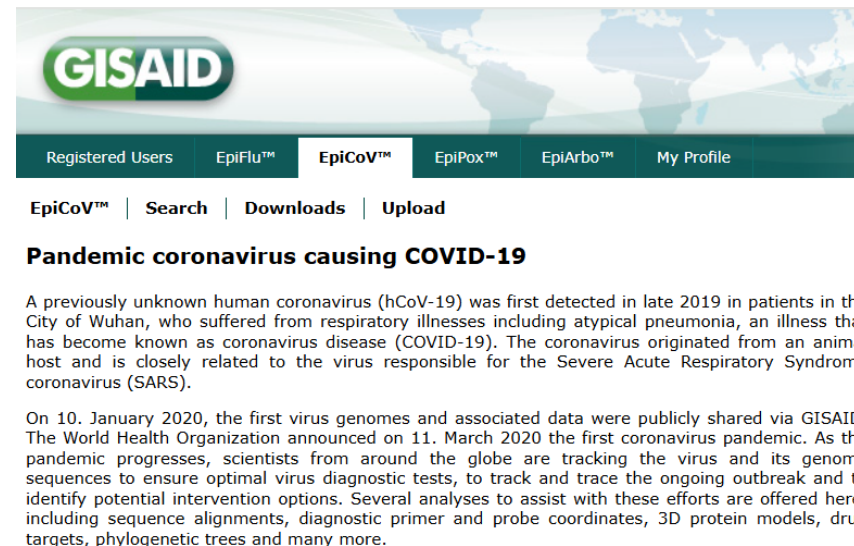
Funded by  
the European Union

The VERDI project grant number 101045989



# Northern sequences and metadata

- Available from GISAID EpiCoV database
- Downloaded: 5 March 2025
- Criteria:
  - Host: Human
  - Completed genome (nt > 29,000)
  - Collection date: 1 July 2022 – 31 March 2024



The screenshot shows the GISAID EpiCoV database interface. At the top, there is a navigation bar with the GISAID logo and a world map. Below the logo, there are tabs for 'Registered Users', 'EpiFlu™', 'EpiCoV™', 'EpiPox™', 'EpiArbo™', and 'My Profile'. The 'EpiCoV™' tab is selected. Below the navigation bar, there are links for 'EpiCoV™ | Search | Downloads | Upload'. The main content area is titled 'Pandemic coronavirus causing COVID-19'. Below the title, there is a paragraph of text describing the virus and its discovery. At the bottom of the screenshot, there is a grid of icons representing various tools and data visualizations, including Audacity, AudacityInstant, BLAST, CoVizu<sup>e</sup>, and a world map.

**Pandemic coronavirus causing COVID-19**

A previously unknown human coronavirus (hCoV-19) was first detected in late 2019 in patients in the City of Wuhan, who suffered from respiratory illnesses including atypical pneumonia, an illness that has become known as coronavirus disease (COVID-19). The coronavirus originated from an animal host and is closely related to the virus responsible for the Severe Acute Respiratory Syndrome coronavirus (SARS).

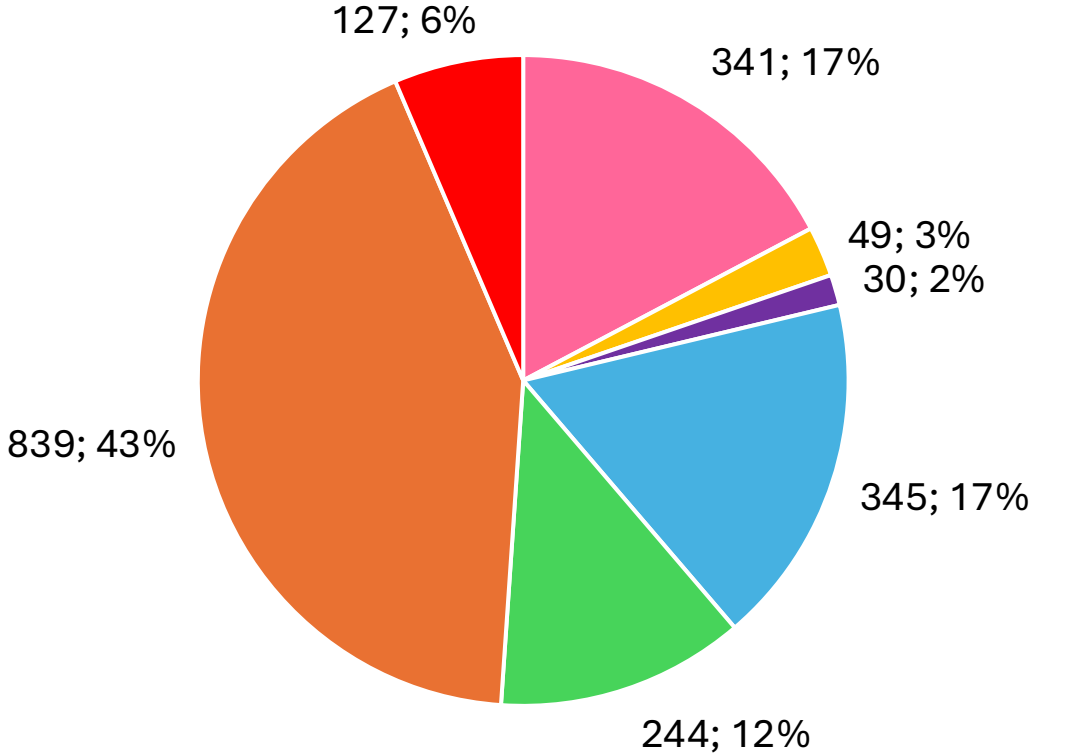
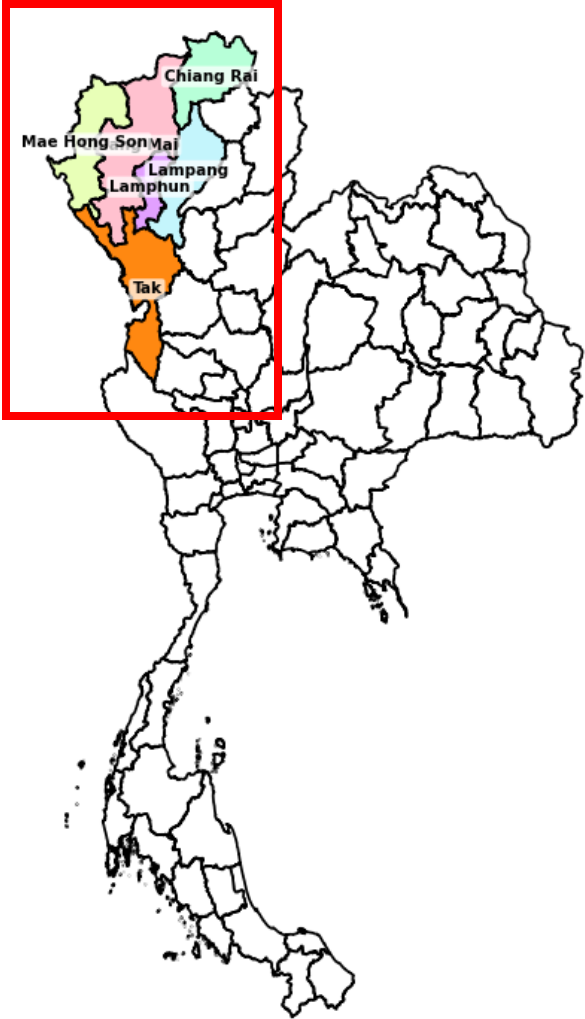
On 10. January 2020, the first virus genomes and associated data were publicly shared via GISAID. The World Health Organization announced on 11. March 2020 the first coronavirus pandemic. As the pandemic progresses, scientists from around the globe are tracking the virus and its genome sequences to ensure optimal virus diagnostic tests, to track and trace the ongoing outbreak and to identify potential intervention options. Several analyses to assist with these efforts are offered here, including sequence alignments, diagnostic primer and probe coordinates, 3D protein models, drug targets, phylogenetic trees and many more.



<https://gisaid.org/>

# Samples inclusion

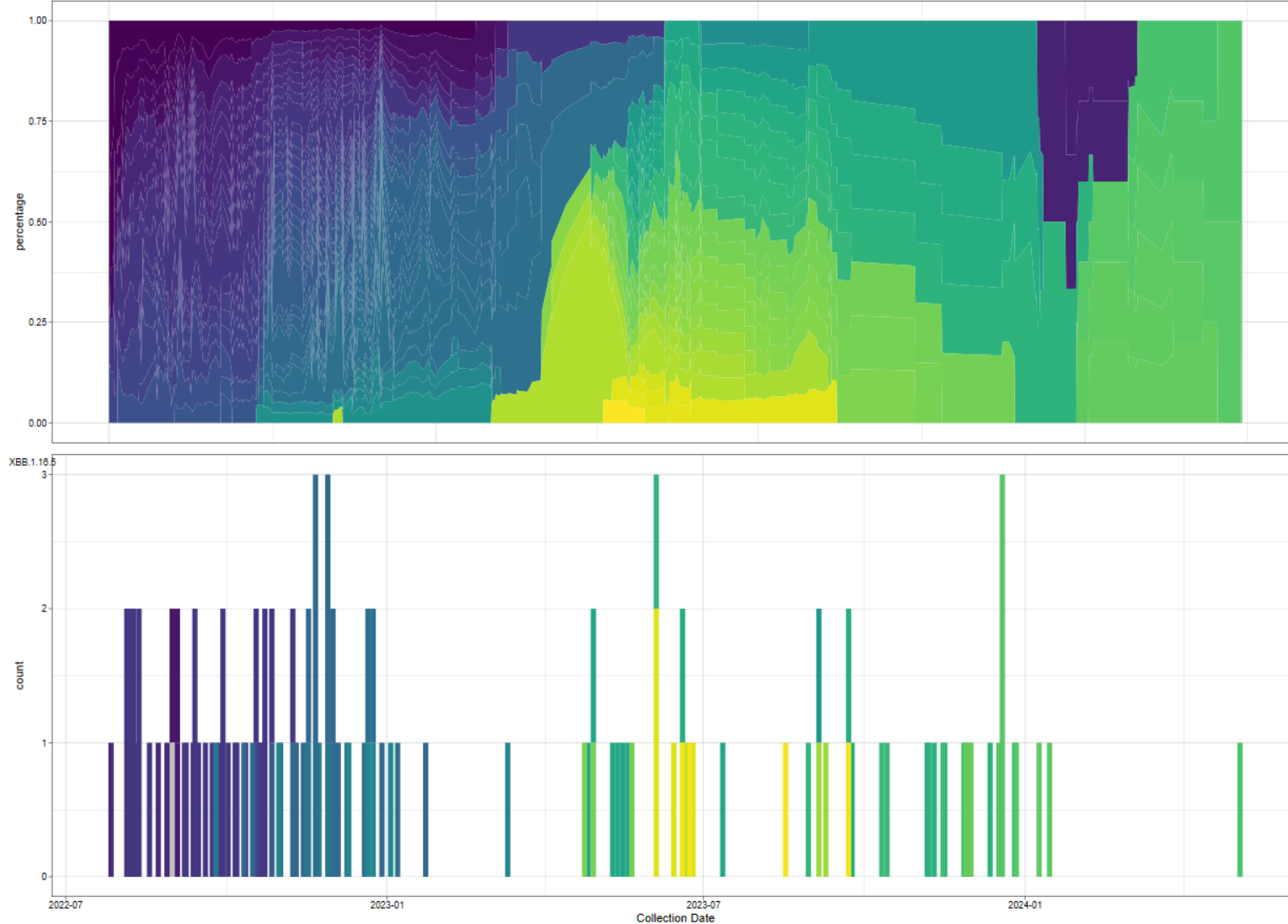
Thailand



- Chiang Mai
- Mae Hong Son
- Lamphun
- Lampang
- Chiang Rai
- Tak
- VERDI

- July 2022 – May 2024
- 1,975 sequences

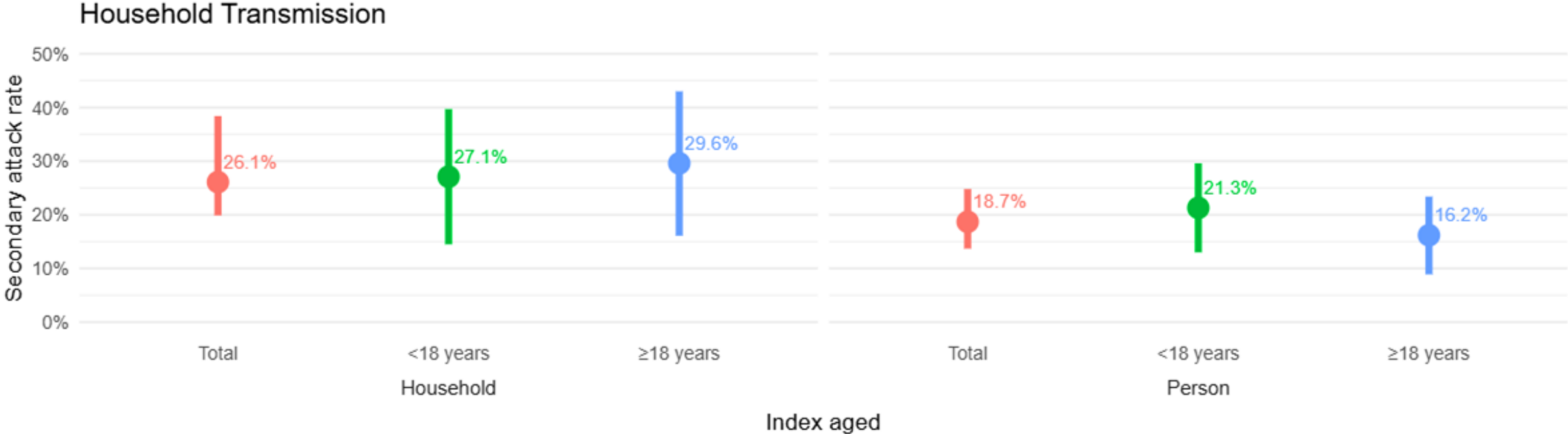
# The evolutionary of SARS-CoV-2 Omicron sub-lineages in Chiang Mai



- Dynamic viral evolution is seen older lineages (BA.2, ba.5) faded as newer sub-lineages (BN, XBB, JN) emerged.
- Household transmission patterns mirrored regional lineage shifts, confirming the role of intra-household spread in overall epidemiology.
- XBB.1.16 and JN.1 were dominant in later household clusters, consistent with their increased fitness and global rise.



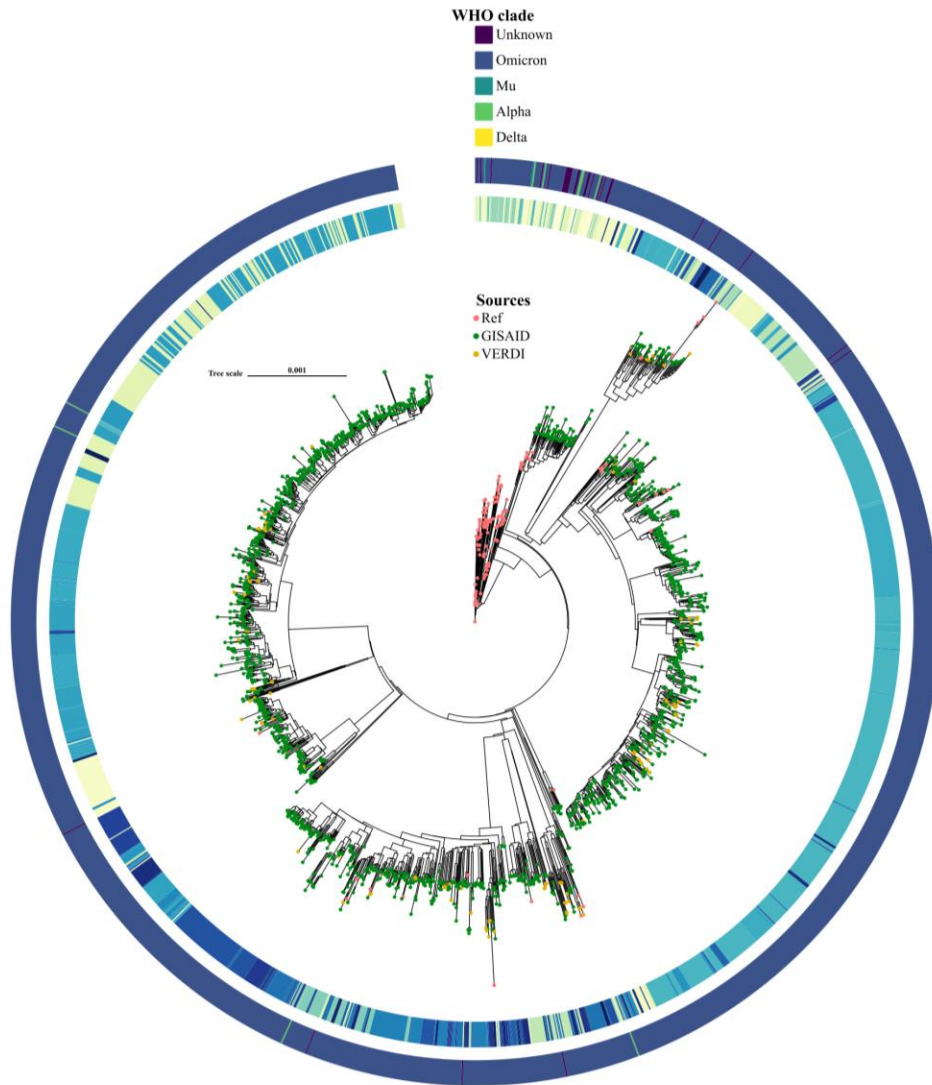
# SARS CoV-2 Sequence relationships indicate the household transmission



Not found significant association between household transmission and age of index case

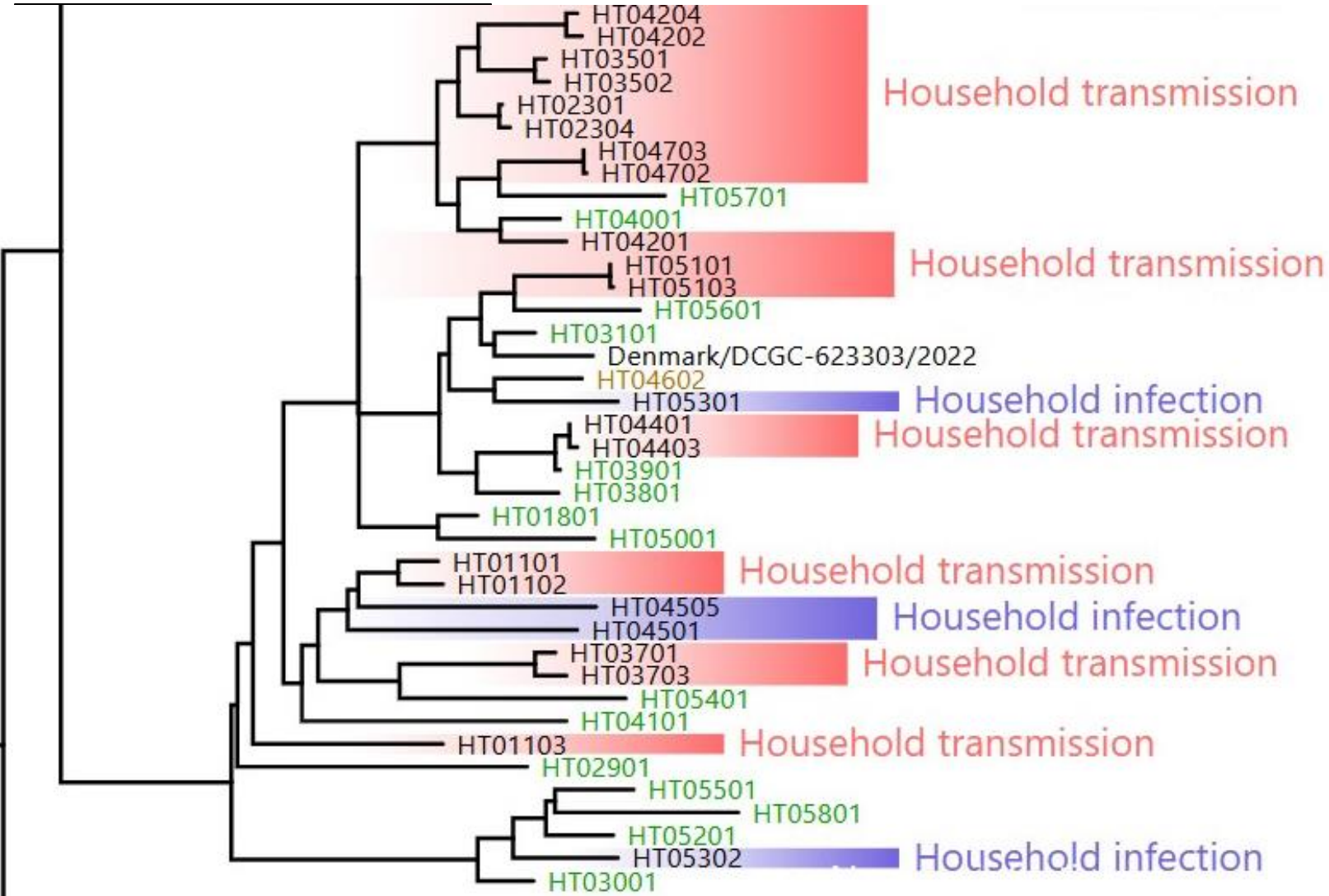
Person: 36/193  
Households: 26/92

# Phylogenetic Tree of SARS-CoV-2 Lineages



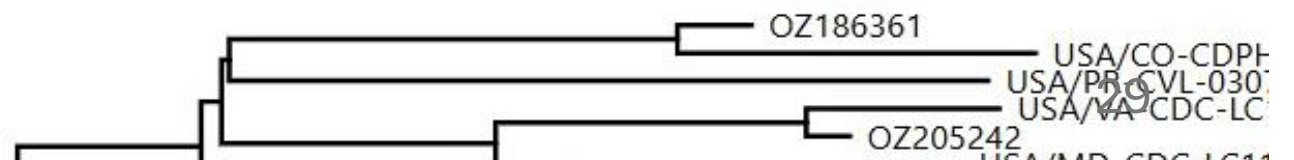
- **Multiple introductions** of the virus into the community during the study period
- **Household-level transmission events** from a single introduction

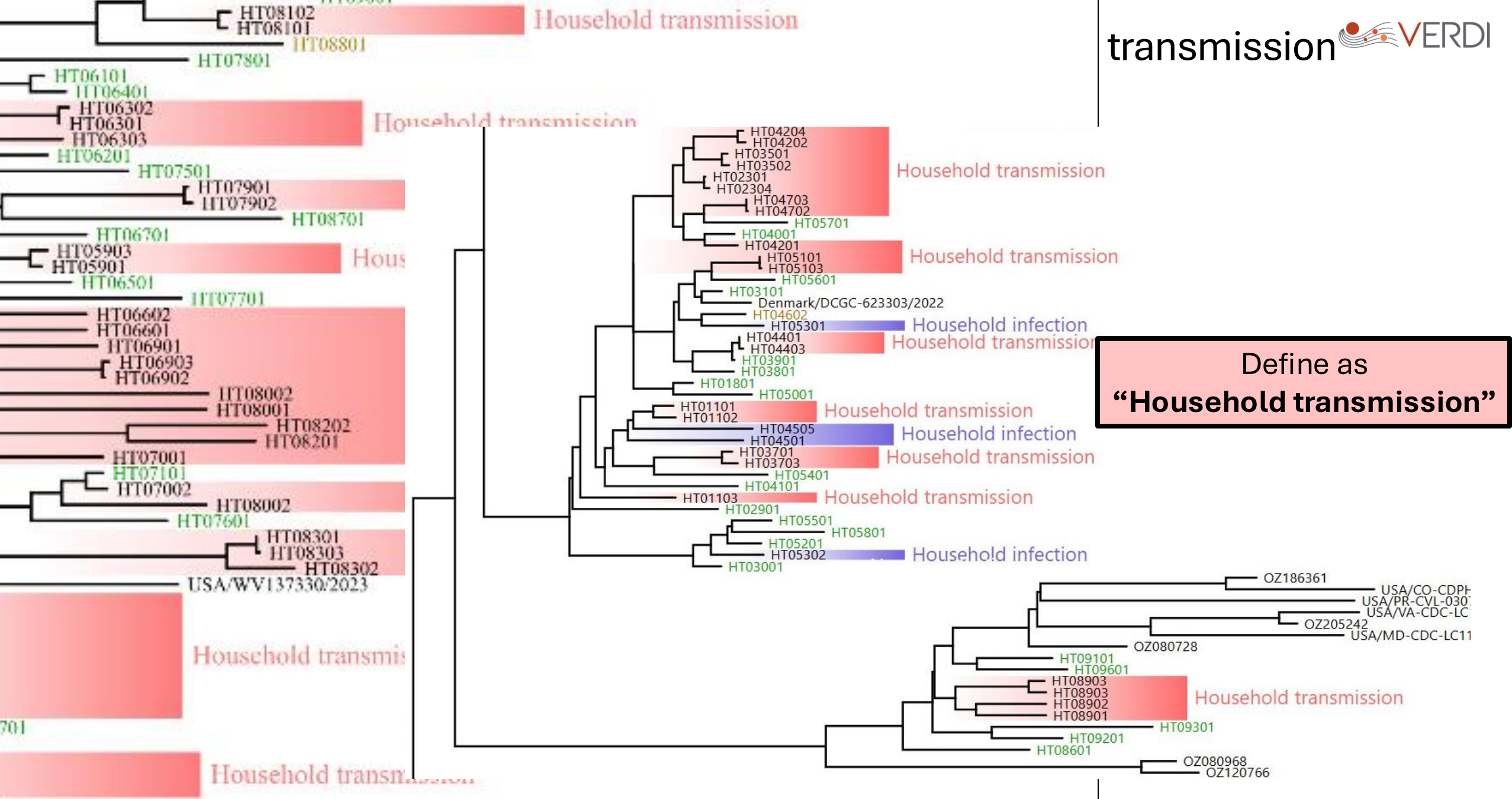
# SARS CoV-2 Sequence relationships indicate the household transmission



**Individuals: 36/193**  
**Households: 26/31**

Define as “Household infection” or “Community transmission”





Define as  
**"Household transmission"**