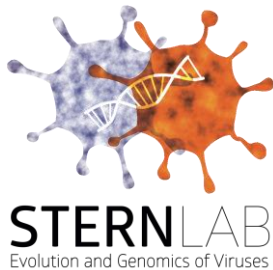


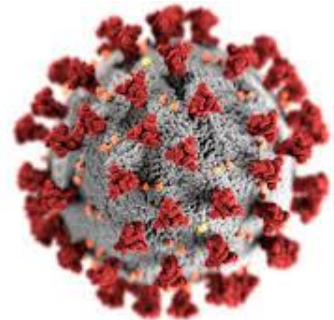
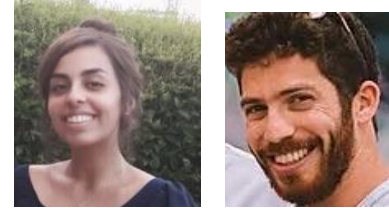
SARS-CoV2 sequencing reveals features of spread in Israel



TAU Library of Life Sciences
July 2 2020

COVID-19 collaboration

- Danielle Miller, Noam Harel, Omer Tirosh, Talia Kustin, Moran Meir
- Michael Martin, Katia Koelle (Emory)
- Amit Huppert (Gertner institute for epidemiology & health policy)
- Naama Kopelman (HiT)
- Nadav Sorek (Assuta Ashdod)
- Sharon Amit (Sheba hospital)
- Dana Wolf (Hadassa Ein Karem)
- Avi Peretz (Poriya medical center)
- Diana Roif-Kaminsky (Brazilay hospital)
- Yonat Shemer-Avni (Soroka hospital)
- Tal Katz (Technion Genome Center)



Motivation

- Gain a better understanding of the epidemiology of virus spread: incidence (asymptomatics), basic reproduction number (R_0)
- Traditional epidemiology is questionnaire-based: slow and may be biased
- Rapid outbreak requires rapid unbiased response: **sequencing**

A little on evolution of viruses

Mutations, mutations

- Mutations occur all the time in all viruses
- The vast majority of mutations we observe are **neutral**: no change to the function of the virus

nature
microbiology

Comment | Published: 18 February 2020

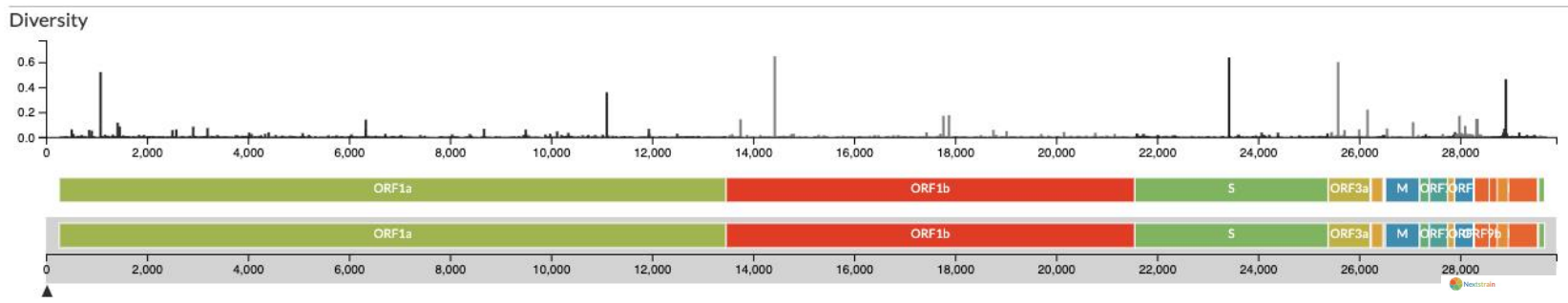
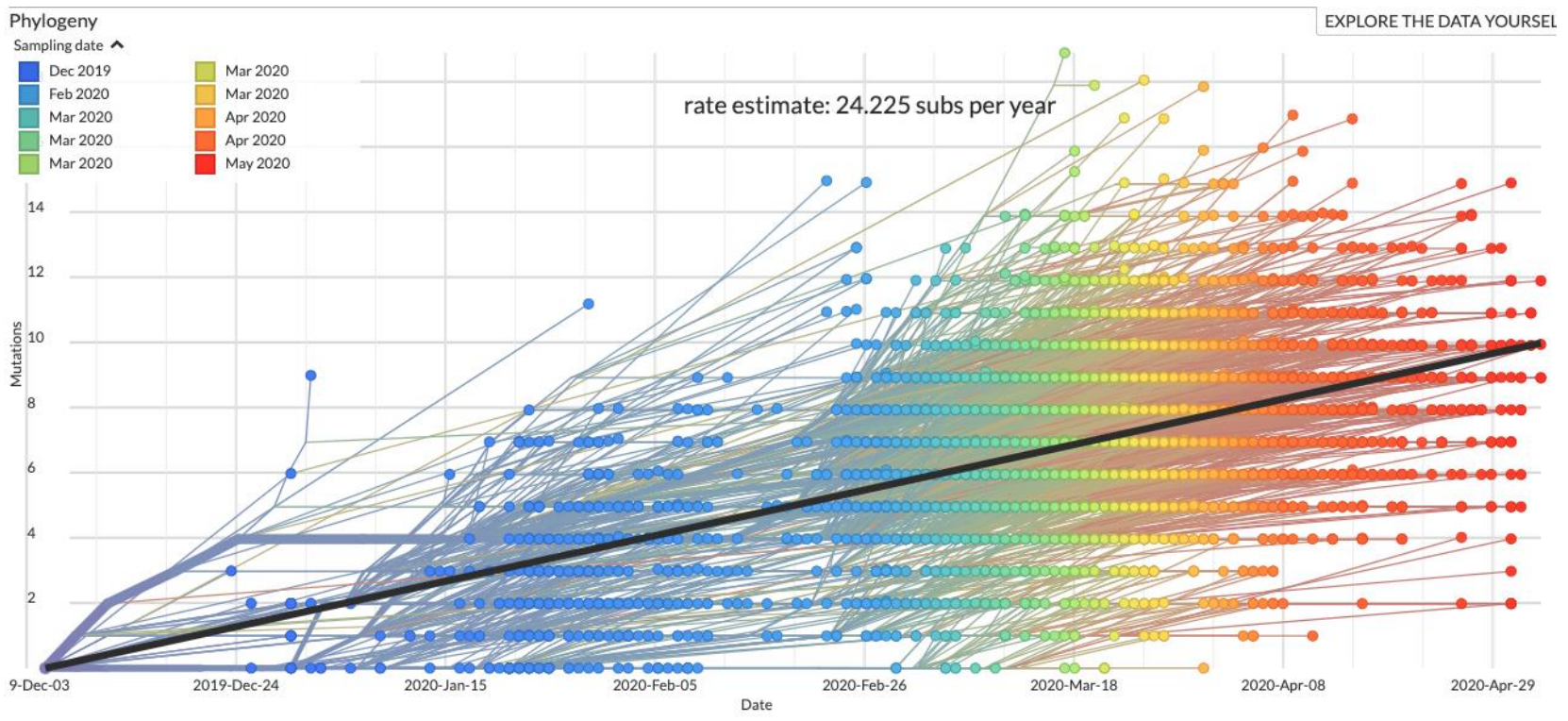
We shouldn't worry when a virus mutates during disease outbreaks

Nathan D. Grubaugh , Mary E. Petrone & Edward C. Holmes 

Nature Microbiology (2020) | [Cite this article](#)

Mutation. The word naturally conjures fears of unexpected and freakish changes. Ill-informed discussions of mutations thrive during virus outbreaks, including the ongoing spread of SARS-CoV-2. In reality, mutations are a natural part of the virus life cycle and rarely impact outbreaks dramatically.

Clock-like evolution of viruses



Evolutionary (fixation) rates of RNA viruses

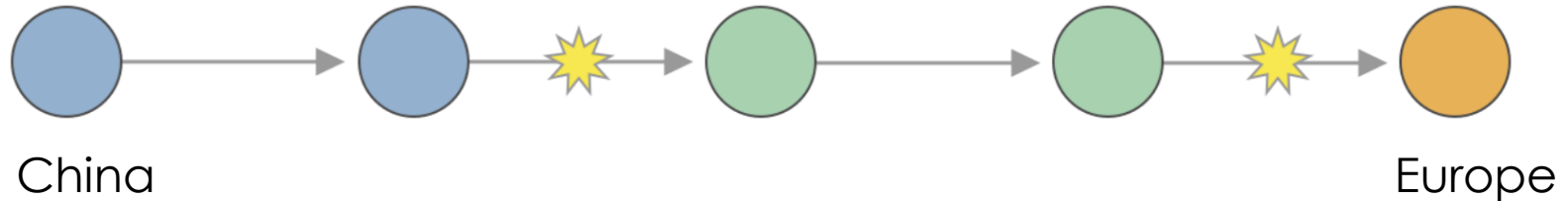
- SARS-CoV-2: ~1 mutation per 1,000 bases per year
- Influenza: ~2 mutations per 1,000 bases per year
- HIV: ~4 mutations per 1,000 bases per year

- Coronaviruses are the **only virus family** with proofreading – which means they mutate at a slower rate

SARS-CoV2 evolutionary rate

- We can use these mutations as “barcodes” to track viral spread

where, on average, we have 7 days from one infection to the next. As the virus transmits, it will mutate at this rate of two mutations per month. This means, that on average every other step in the transmission chain will have a mutation and so would look something like:



Heroic effort #1

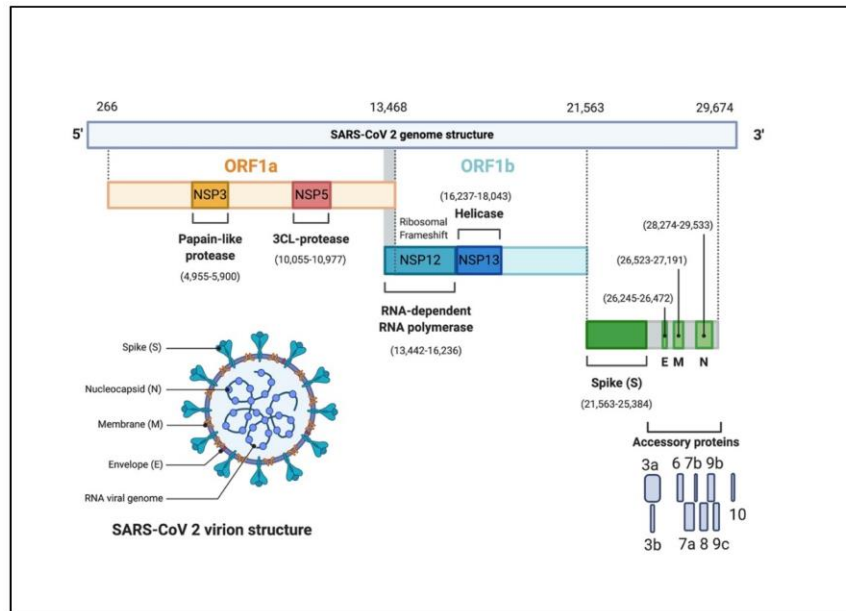
- Collected ~212 samples from SARS-CoV2 patients
- Leftover RNA from nasopharyngeal swabs
- Criteria:
 1. Different dates March-April
 2. No relatives
 3. From six hospitals spanning geography of Israel



→ **Random representative sample**

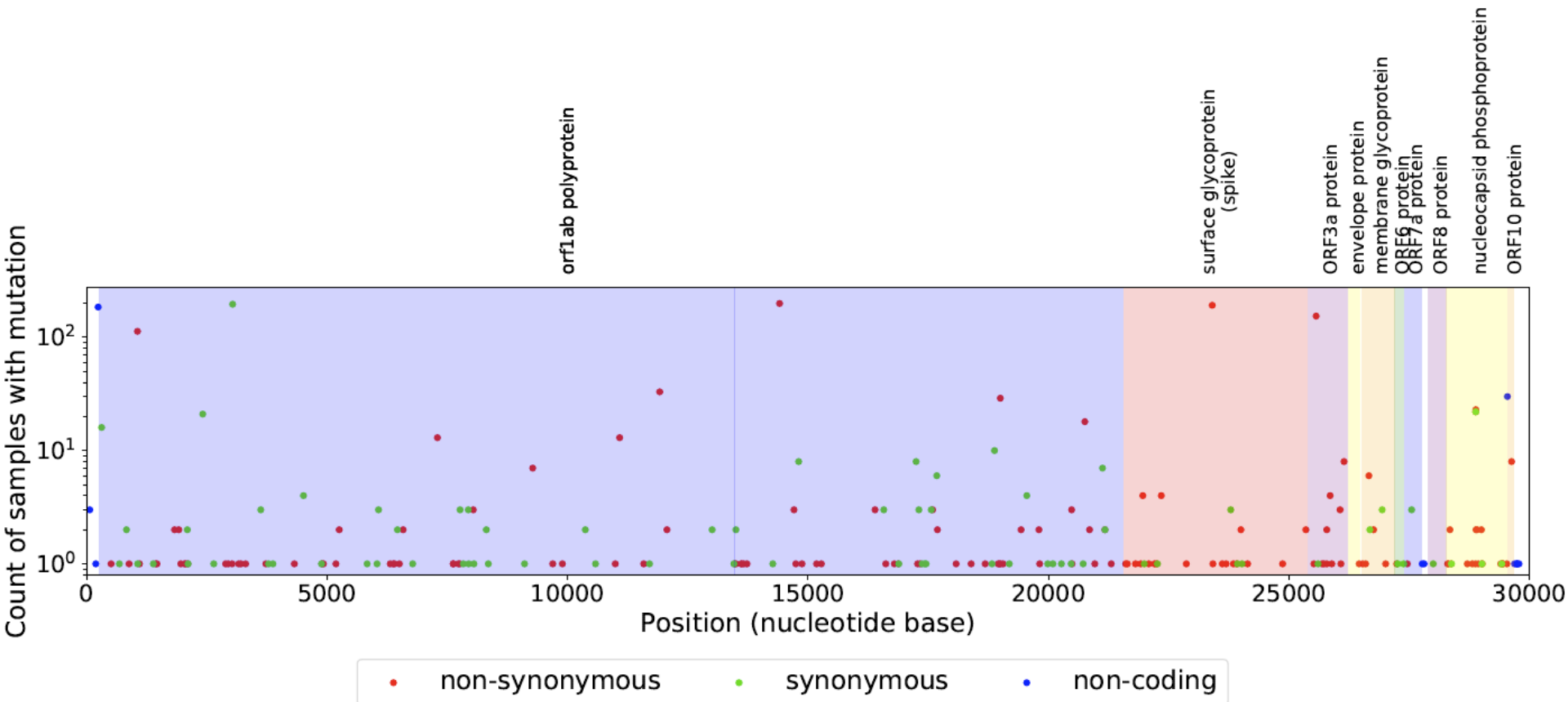
Heroic effort #2

- Full genome sequence of the virus successfully sequenced from 212 samples via next-gen sequencing



What did we learn from the sequences?

Map of mutations shows neutral pattern



- Mutations spread across genome
- 72 synonymous (silent)
141 non-synonymous (amino-acid altering)

Eight sequences harbour deletions

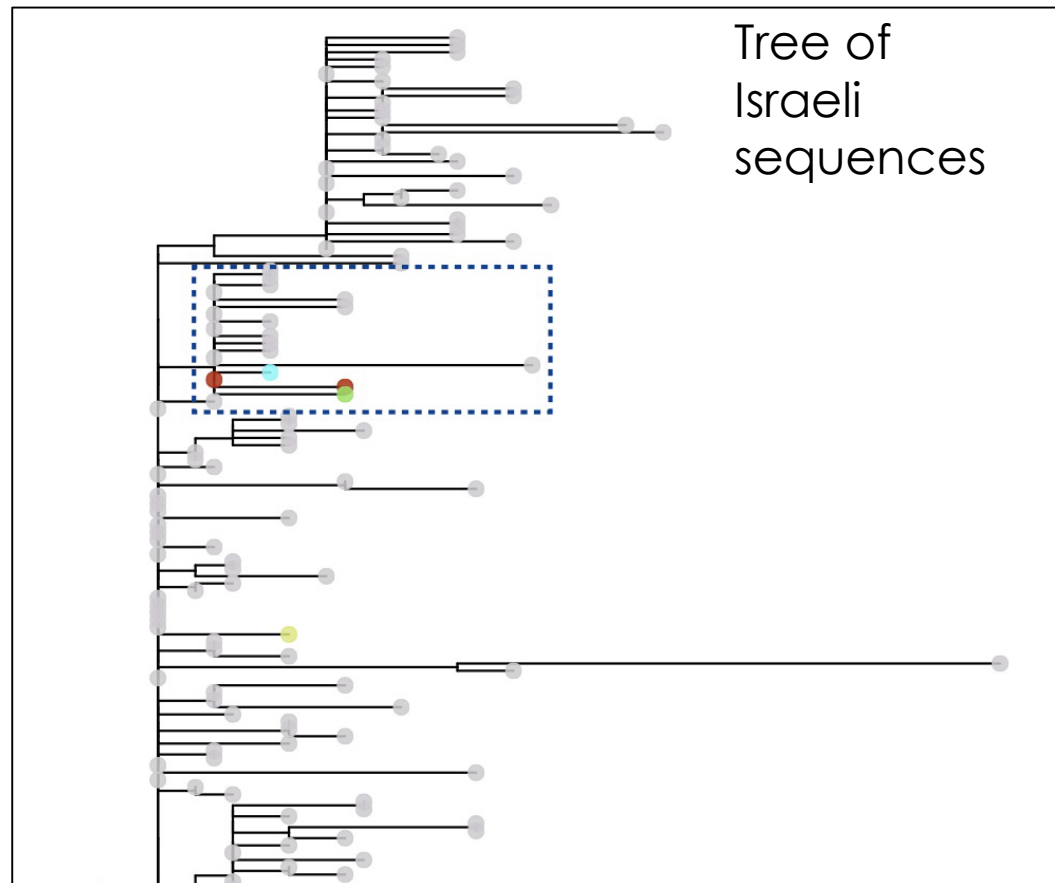
#	Genome coordinates	Length	ORF/Genomic location	Suggested effect	Number of samples found in	Sample IDs	Number of reads supporting deletion
1	686-694	9nt	ORF1ab polyprotein	Deletion of 3 amino acids	2	2086008, 130710157	3575, 1852
2	3882-3899	18nt	ORF1ab polyprotein	Deletion of 6 amino acids and an additional single amino acid mutation	2	2089839, 2089852	427, 605
3	27387-27396	10nt	End of ORF6 and start of ORF7a	Stop codon of ORF6 is recreated. Start codon of ORF7a is deleted with no in-frame replacement	1	13077726	3801
4	28254	1nt	End of ORF8	Last amino acid is replaced by a 5 amino acid addition	1	2086033	2849
5	29746-29748	3nt	3' UTR	Non-coding, unknown	2	51137844, 51141225	42,147

- Three of five deletions are in-frame
- Two are not in frame and affect ORF7a and ORF8

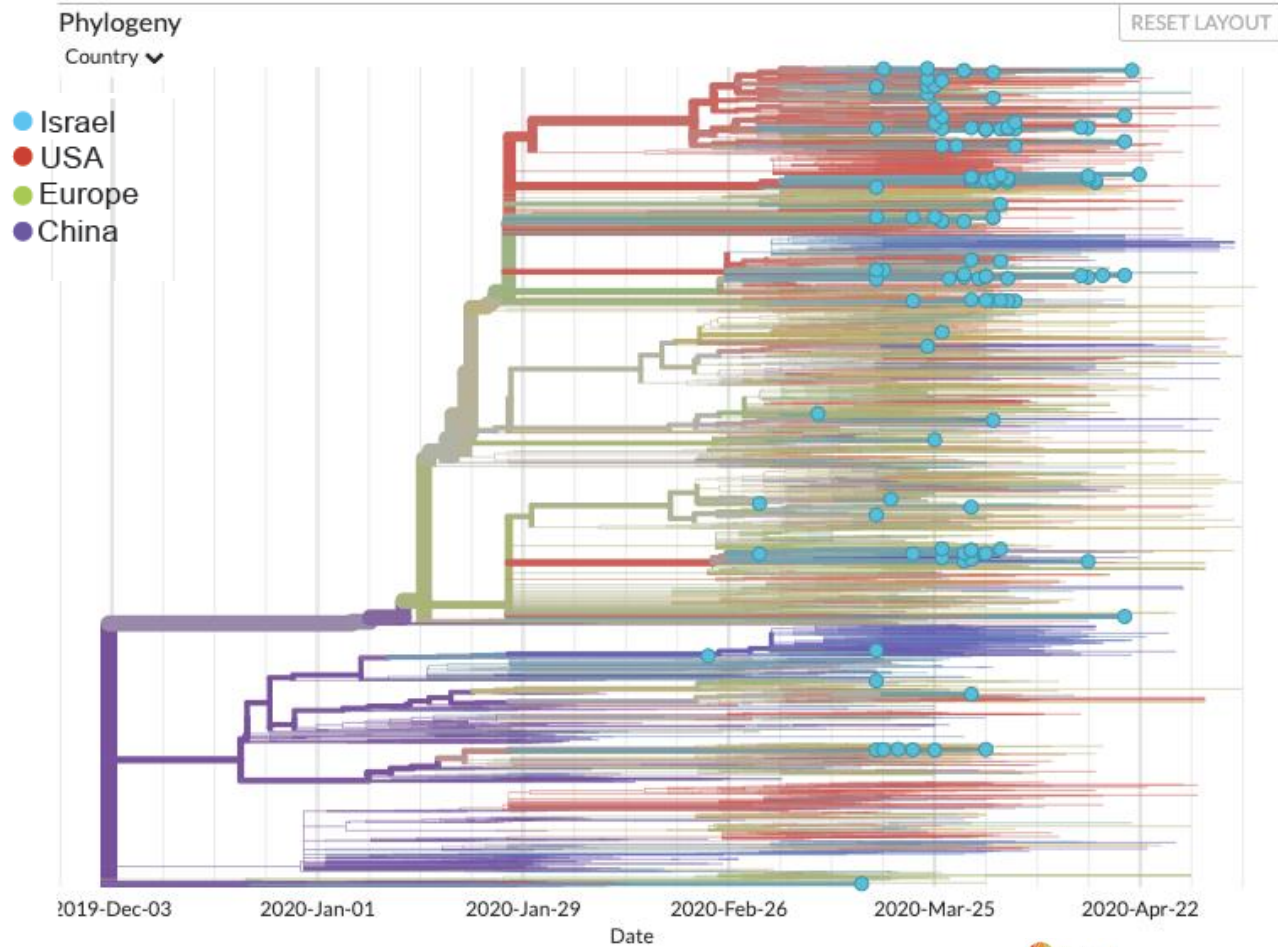
ATGGCATA~~TAC~~AGGTAACGTATC

Some of the deleted sequences cluster together

- Four of the deleted samples in the same clade
- This clade is defined by a non-synonymous mutation in NSP16: could this mutation potentiate deletions...?



Comparison of Israeli diversity & global diversity



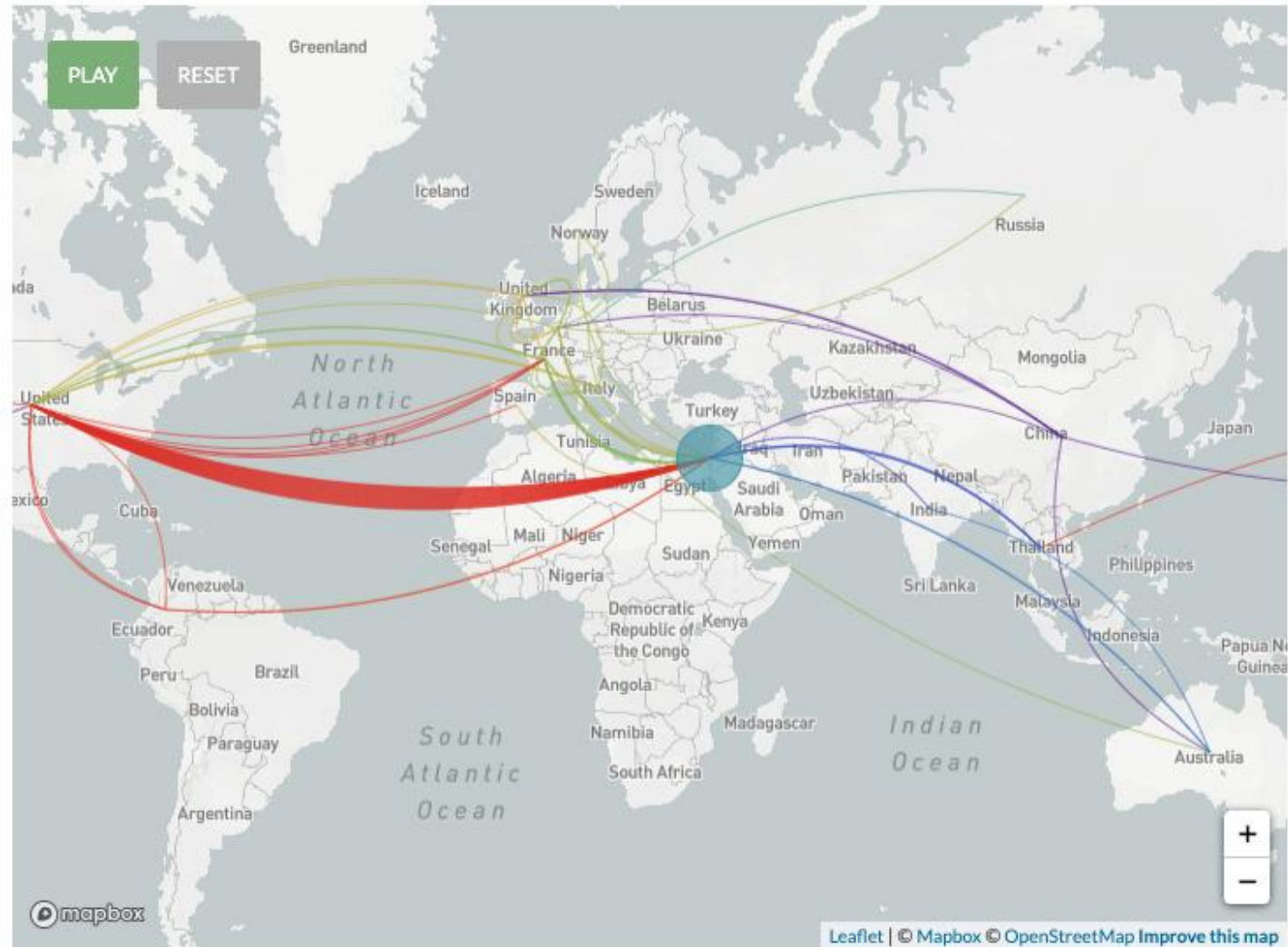
nextstrain.org

Nextstrain

Real-time tracking of pathogen evolution

Comparison of Israeli diversity & global diversity

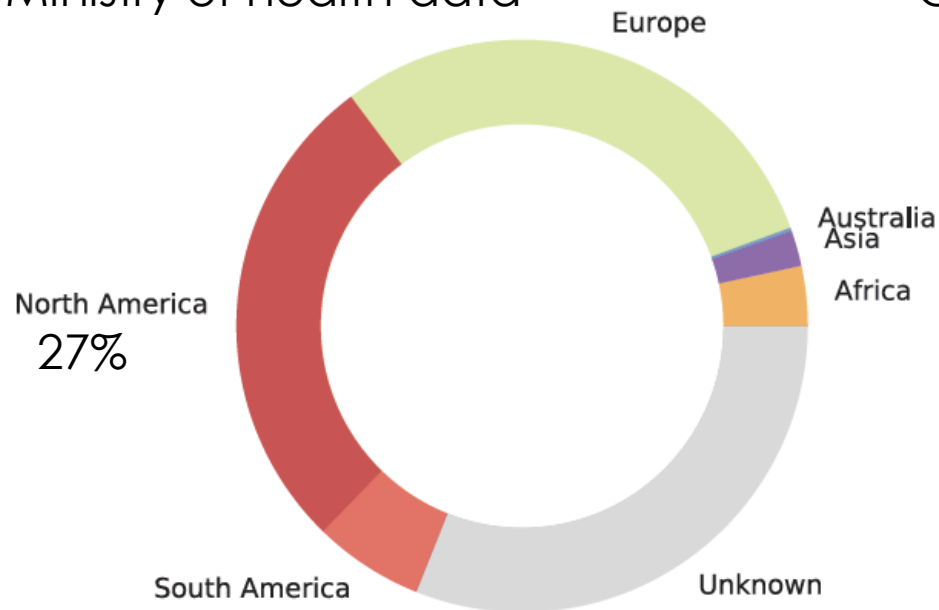
- **70%** of transmission chains in Israel are from the U.S.
- The rest from Europe
- Almost none from south east Asia



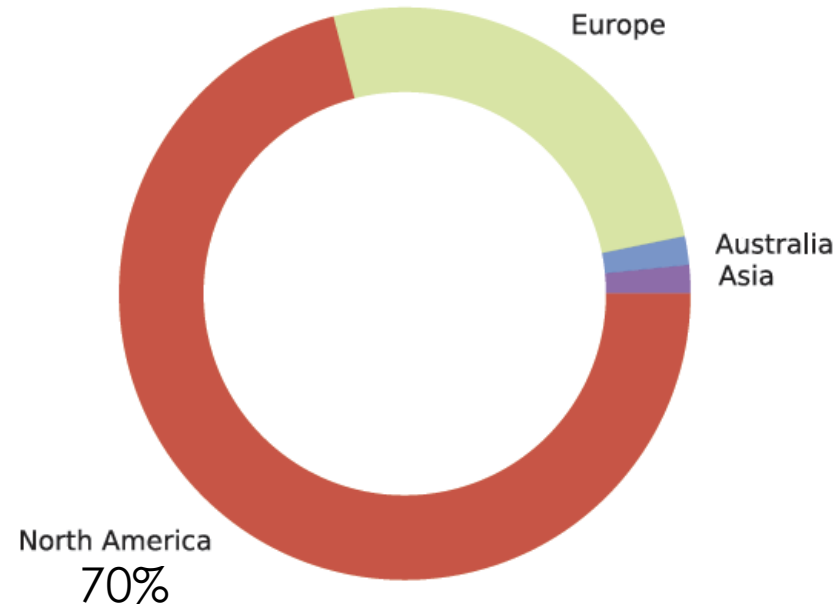
Transmission chains vs reported travelers

- Travelers returning from US contributed dramatically to spread in Israel

Infected returning travelers
Ministry of health data

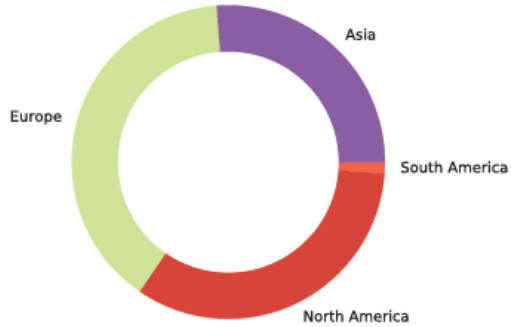


Inferred clade importations
Our sequencing data

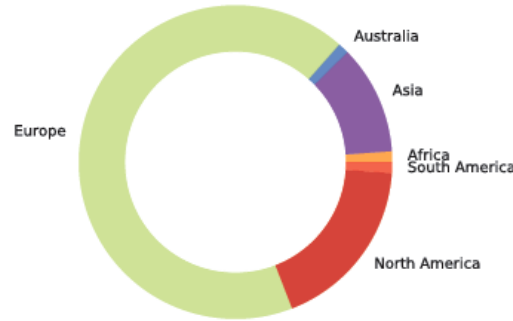


We do not see this in other countries

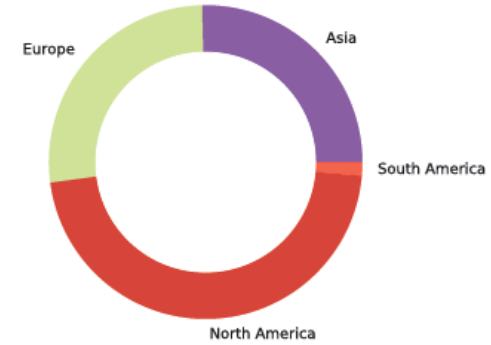
Australia



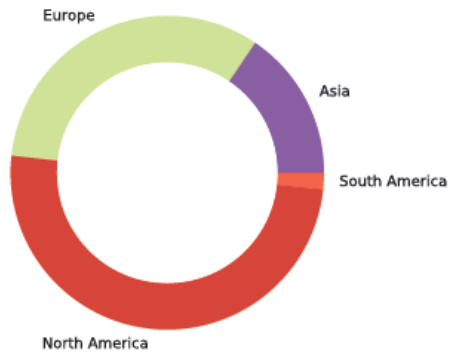
Belgium



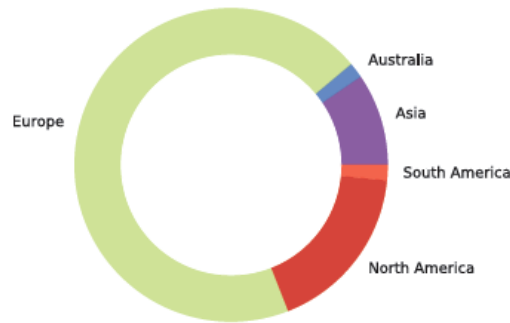
Canada



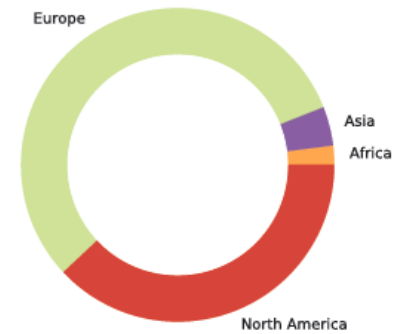
France



Netherlands

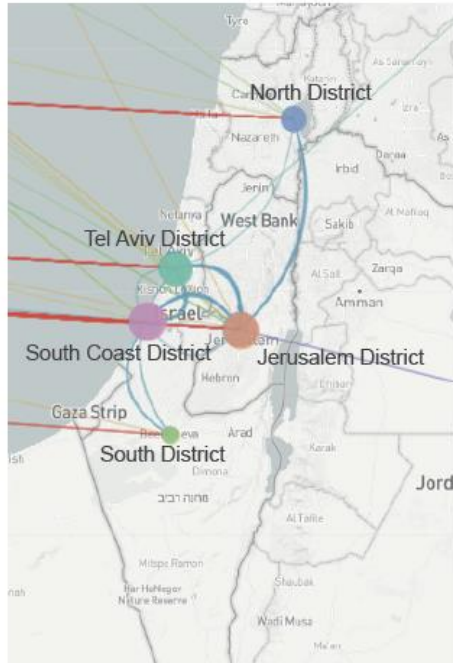


Sweden

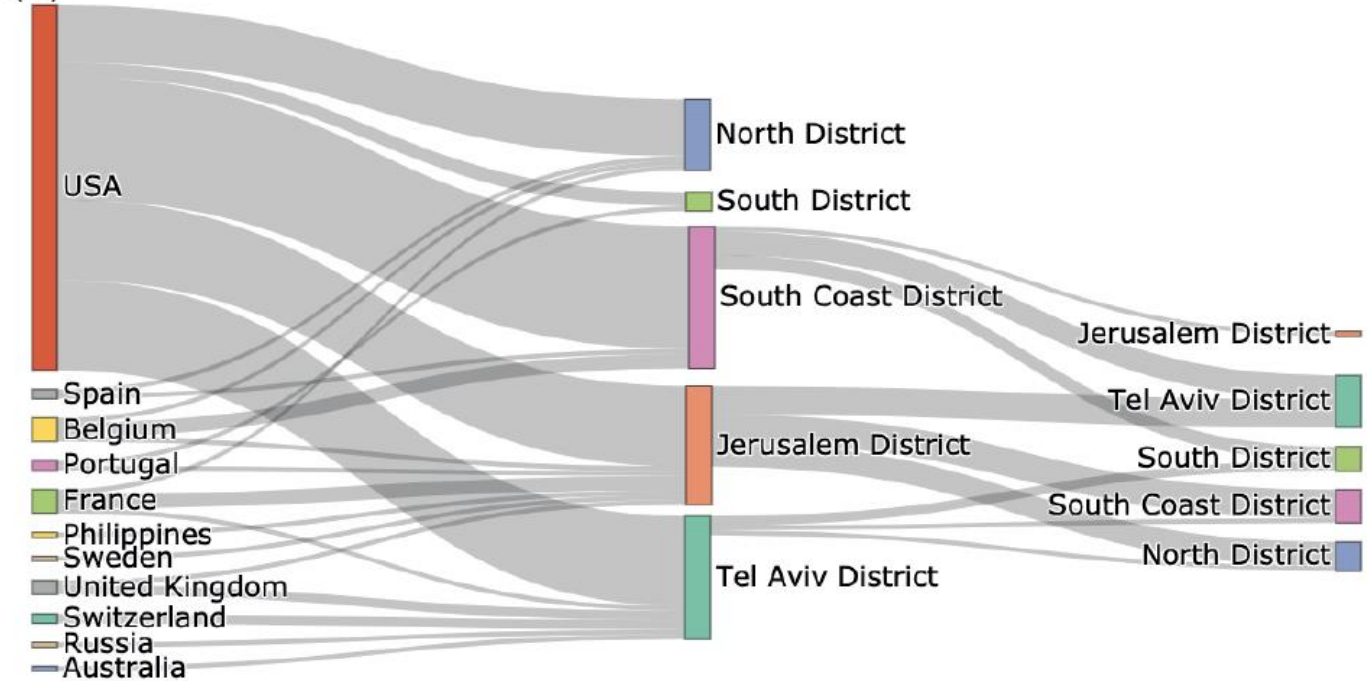


Spread inside Israel

(A)

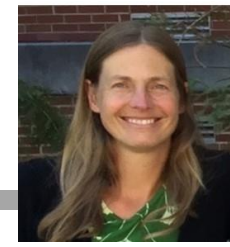


(B)



- USA introductions to all regions of Israel
- The virus travelled inside Israel

Phylodynamic model

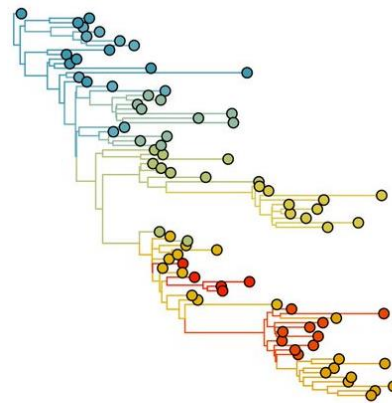


Katia Koelle

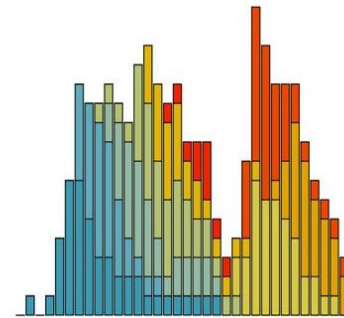


Michael Martin

- Fit a Susceptible-Exposed-Infected-Recovered (SEIR) compartmental model to the phylogeny
- Combine epidemiology and evolution



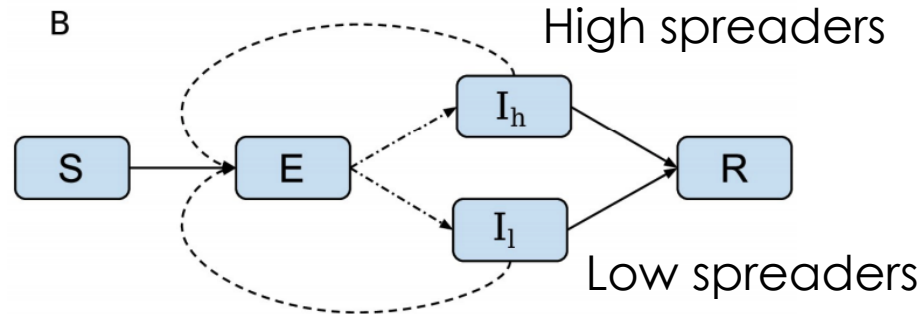
Time-scaled phylogeny



Infections through time

Details

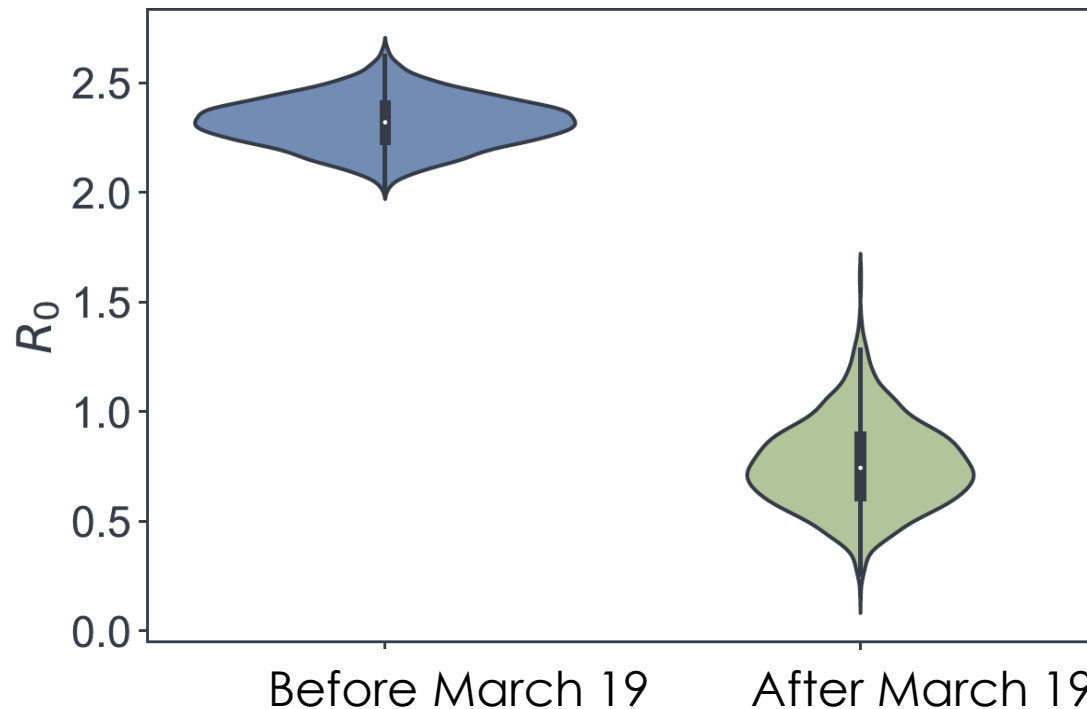
- Based on BEAST + PhyDyn packages



- Some people might spread more than others (**transmission heterogeneity**)
- Non-infectious period following exposure (E)
- Transmission dynamics (R_0) might change following social distancing

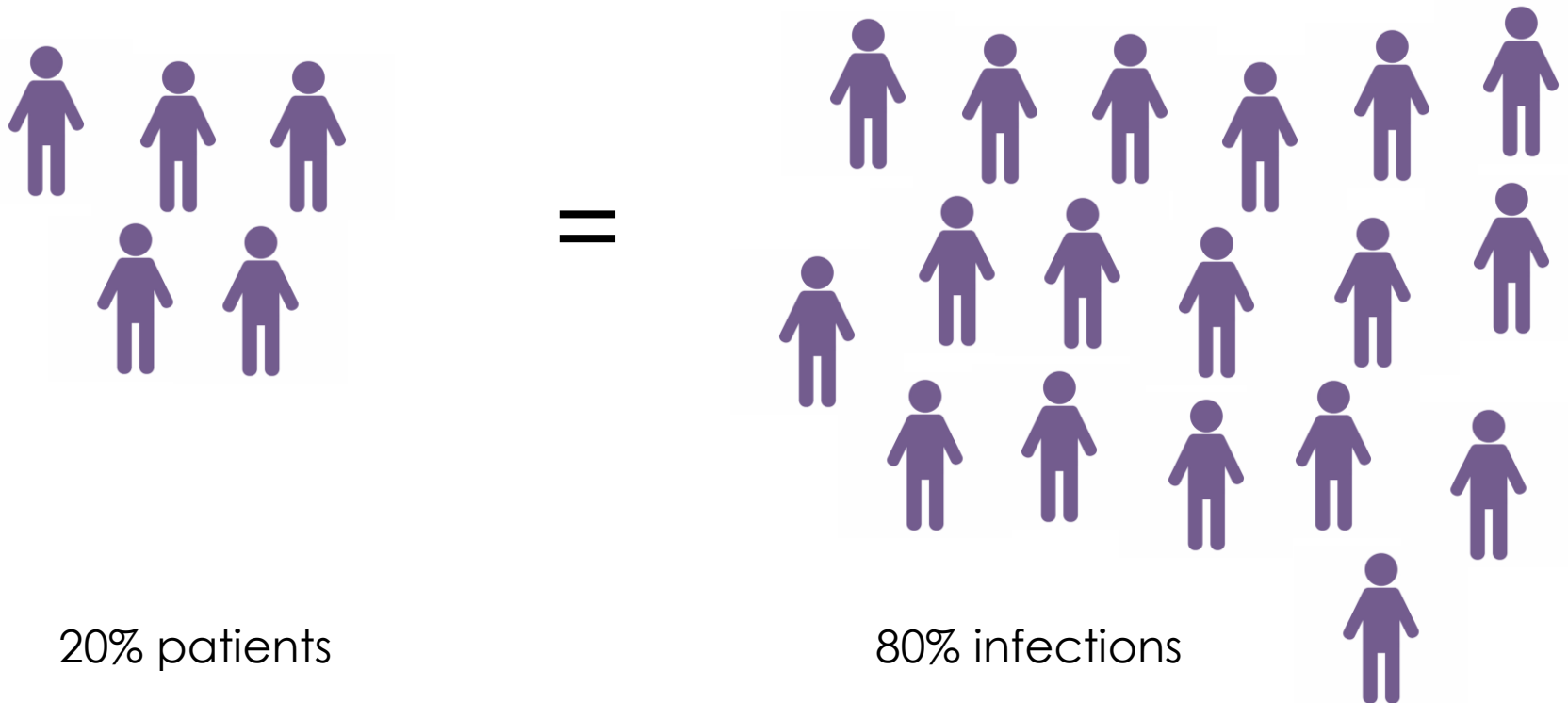
R_0 (basic reproductive number) across time

- R_0 : mean number of individuals each person infects



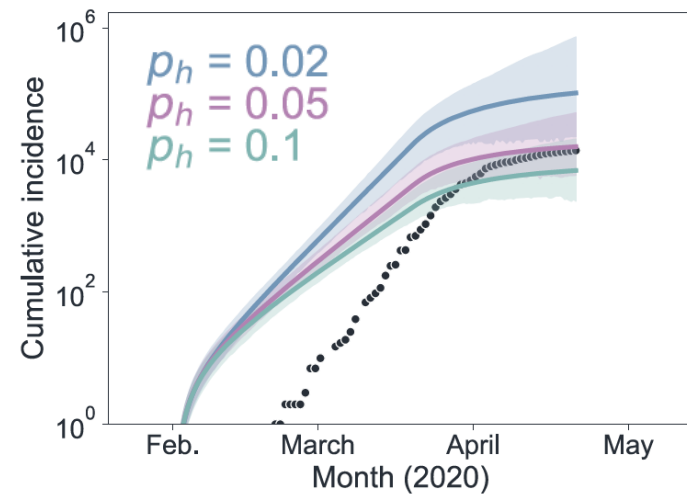
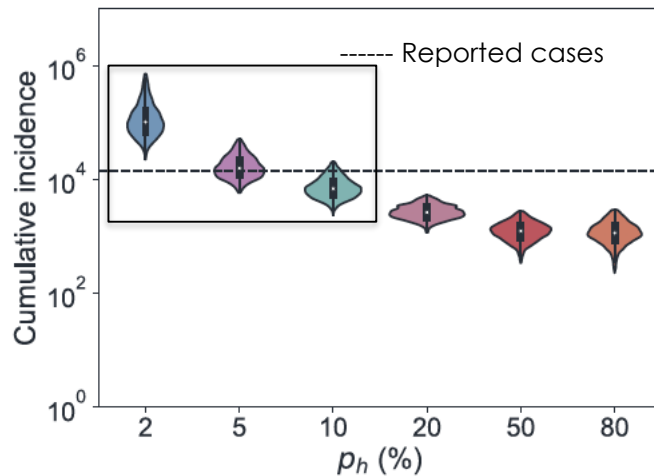
Transmission heterogeneity

- R_0 represents an average
- 20/80 rule: for many infectious diseases, 20% of infected individuals responsible for 80% of cases



Extreme superspreading dynamics

- p_h – proportion of “high-spread” patients responsible for 80% of infections



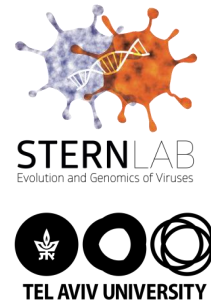
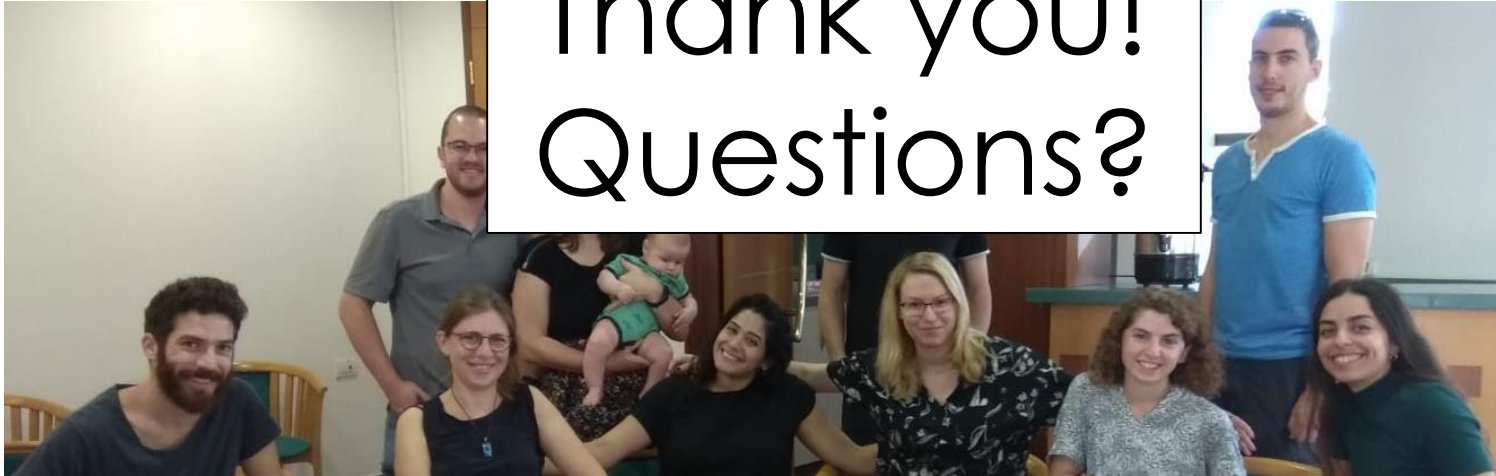
**Extreme super-spreading
dynamics: 2-10% of patients
responsible for 80% of infections**

Summary

- Transmission from US travelers: gap in policy (European travelers quarantined, US not)
- Superspreaders: “biological” (more virus?) and/or “social” (large gatherings?)
- Corollary: most individuals do not contribute to spread at all.
- Sequencing data alone shows effectiveness of shelter-in-home measures implemented March 19

Thanks

Thank you!
Questions?



Donations from
Millner
foundation
and AppFlyer