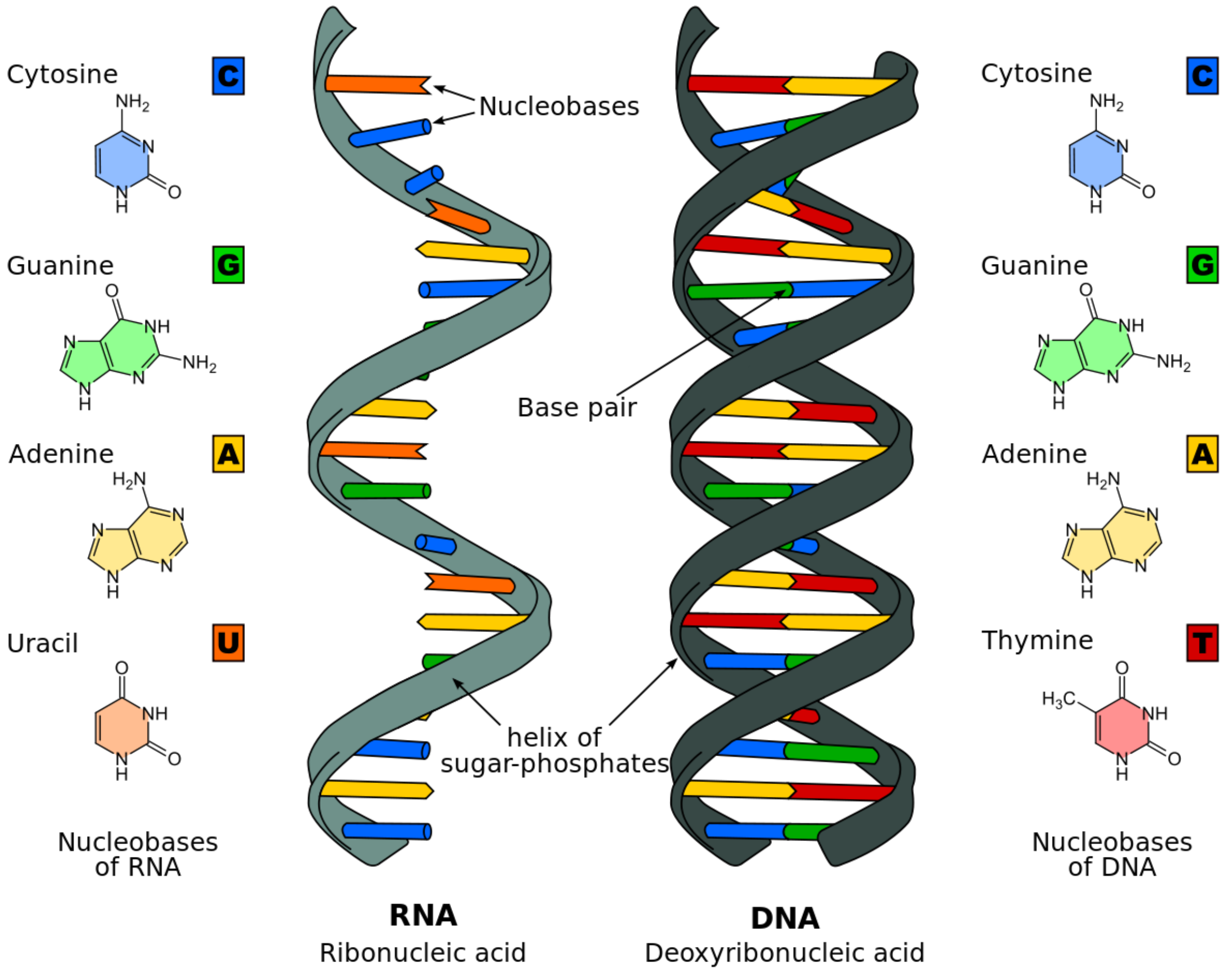
SARS-CoV-2 Variants in Michigan

Dr Heather Blankenship
Bioinformatics and Sequencing Section Manager
Michigan Department of Health and Human Services

Allen Stout
COVID-19 Surveillance and Investigation Unit Manager
Michigan Department of Health and Human Services

Nucleic Acid – DNA v RNA

- Nucleic acid of an organism can either be RNA or DNA
- RNA is more unstable and more likely to accumulate mutations
- Molecular tests can identify the presence of nucleic acid even if a viable organism is not present



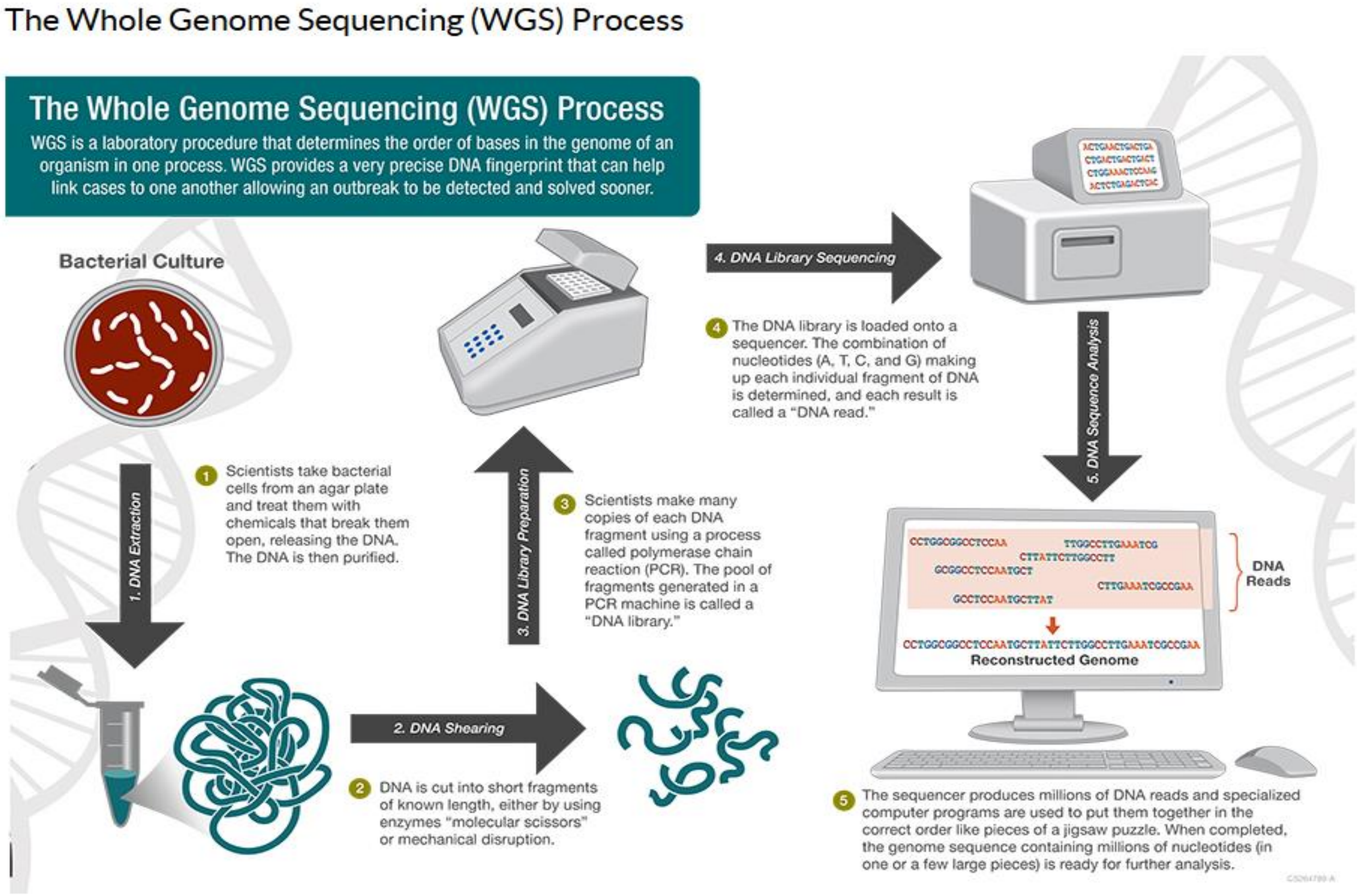
Mutations

The cat chased the ball around the house.

- Point Mutation (Change a single nucleotide to a different nucleotide)
  - Silent (no change) **The cat chased the ball around the house.**
  - Missense (changes the outcome) **The cat chased the ball around the mouse.**
  - Nonsense (will result in early stop) **The cat chased.**

- Frameshift Mutation (Addition or Deletion of Nucleotides)
  - Addition (adds nucleotides) **The cat of chased the ball around the house.**
  - Deletion (removes nucleotides) **The cat chased the_around the house.**

Whole Genome Sequencing Process – Pure Isolate



The Whole Genome Sequencing (WGS) Process

The Whole Genome Sequencing (WGS) Process

WGS is a laboratory procedure that determines the order of bases in the genome of an organism in one process. WGS provides a very precise DNA fingerprint that can help link cases to one another allowing an outbreak to be detected and solved sooner.

① Scientists take bacterial cells from an agar plate and treat them with chemicals that break them open, releasing the DNA. The DNA is then purified.

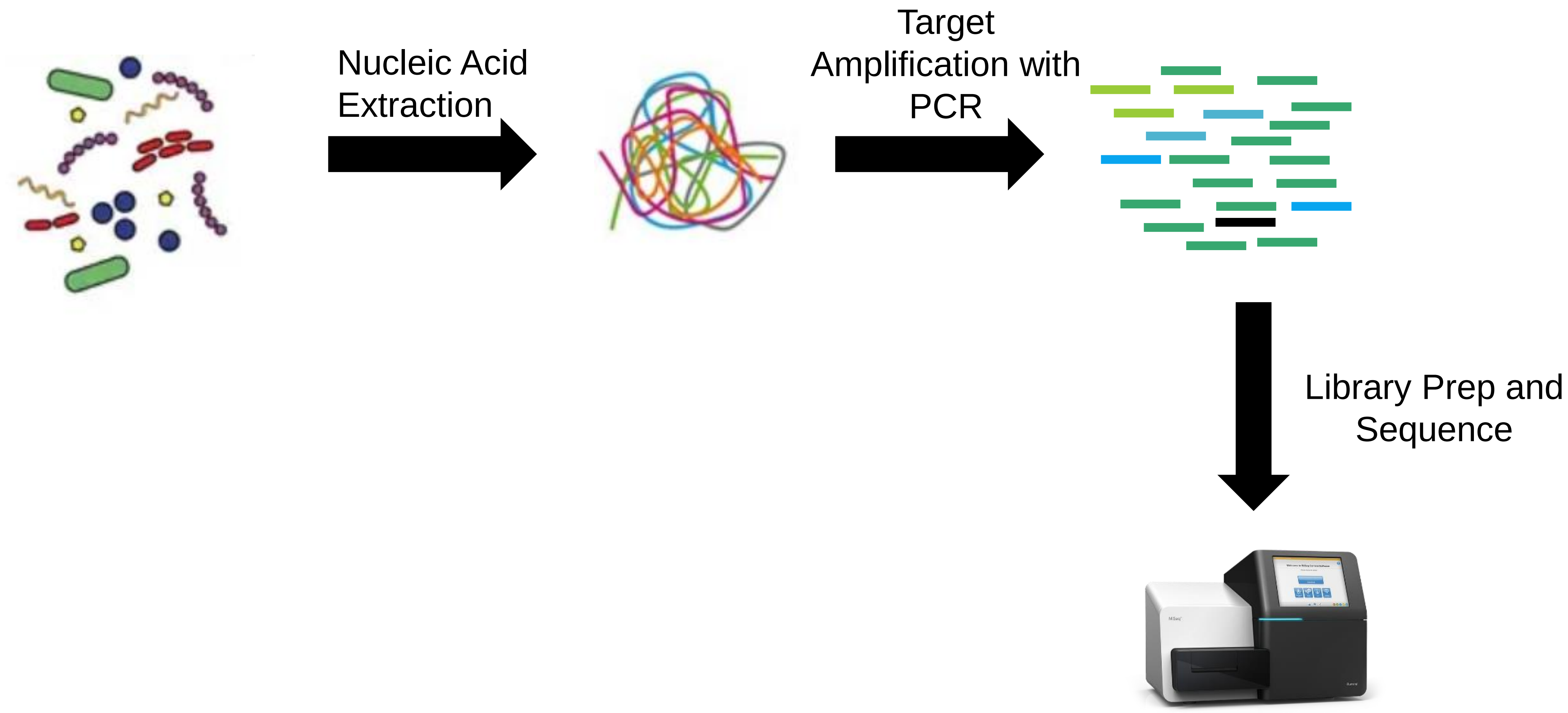
② DNA is cut into short fragments of known length, either by using enzymes "molecular scissors" or mechanical disruption.

③ Scientists make many copies of each DNA fragment using a process called polymerase chain reaction (PCR). The pool of fragments generated in a PCR machine is called a "DNA library."

④ The DNA library is loaded onto a sequencer. The combination of nucleotides (A, T, C, and G) making up each individual fragment of DNA is determined, and each result is called a "DNA read."

⑤ The sequencer produces millions of DNA reads and specialized computer programs are used to put them together in the correct order like pieces of a jigsaw puzzle. When completed, the genome sequence containing millions of nucleotides (in one or a few large pieces) is ready for further analysis.

Whole Genome Sequencing Process – Metagenomics

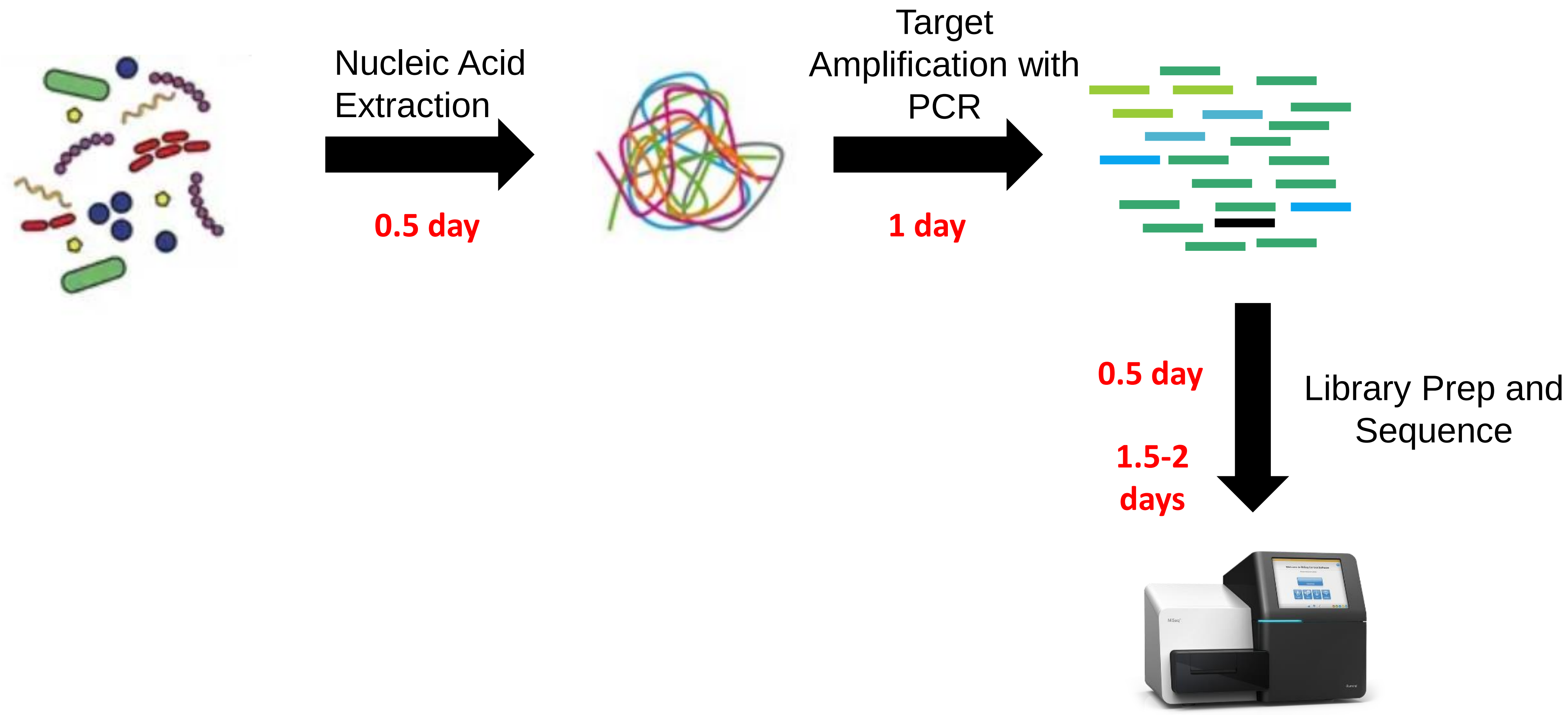


Nucleic Acid Extraction

Target Amplification with PCR

Library Prep and Sequence

Whole Genome Sequencing Process – SARS-CoV-2



Nucleic Acid Extraction

0.5 day

Target Amplification with PCR

1 day

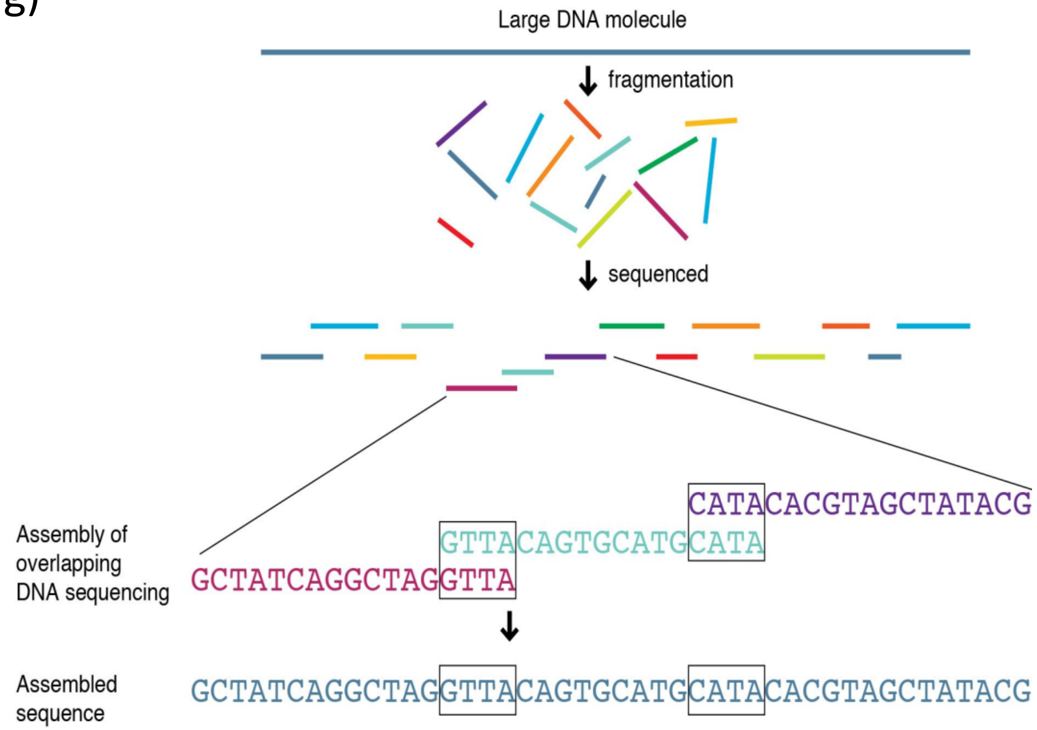
0.5 day

1.5-2 days

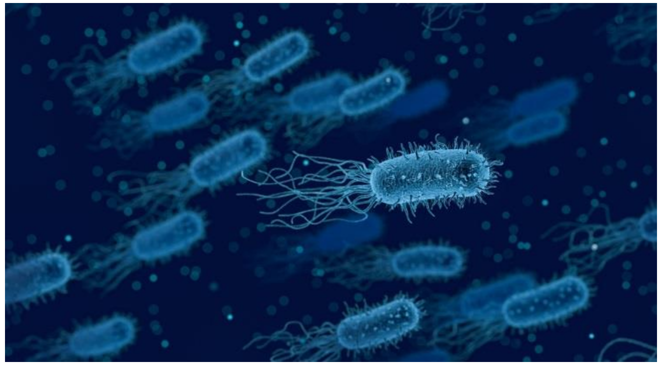
Library Prep and Sequence

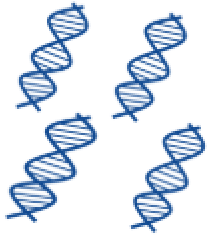
Analysis: 0.5 day

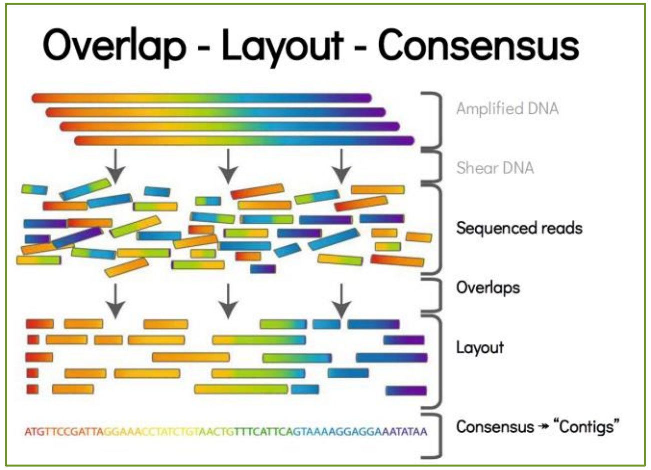
1 box set = 1,084,440 words
~ 5,442,200 letters (5 letters/wd avg)



WGS Pipeline







Torstern Seeman

Now what?

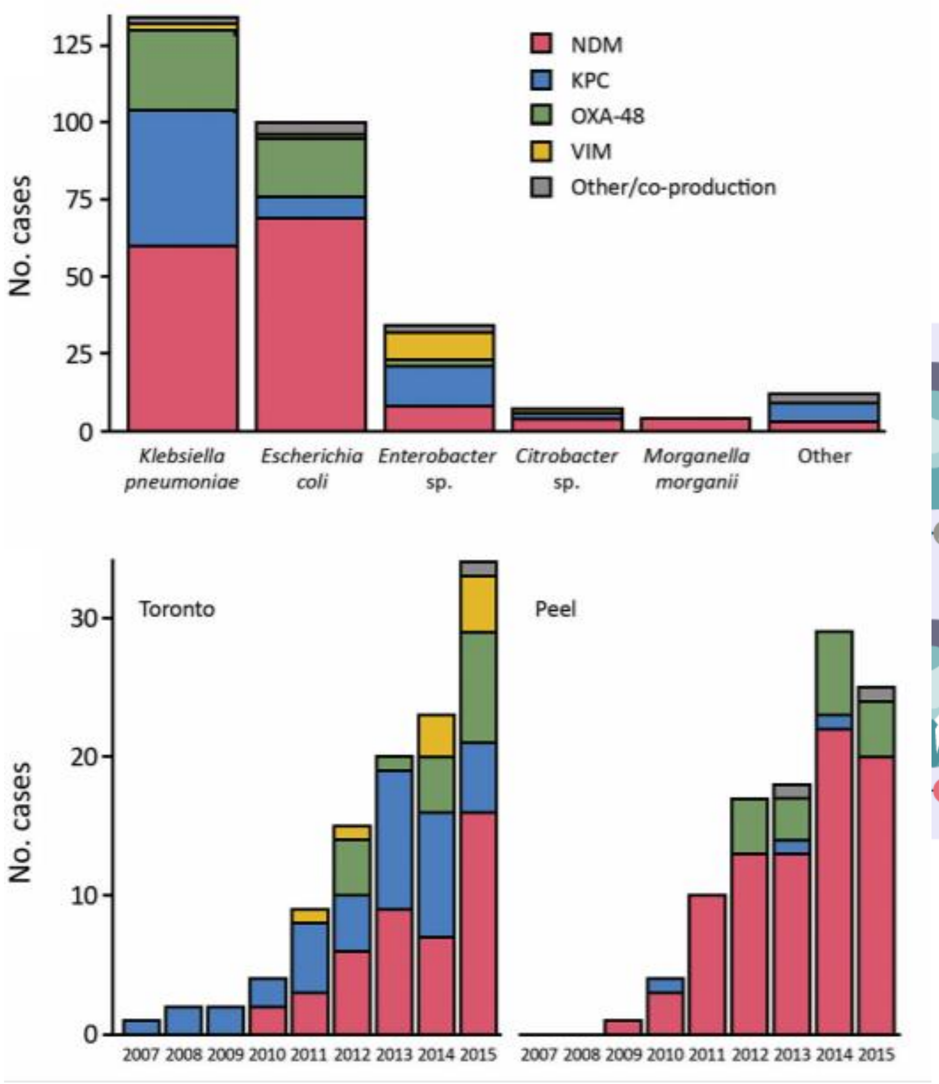
Use of Pathogen Genomics

Investigate Outbreaks

Transmission Patterns

Monitor Virulence Genes

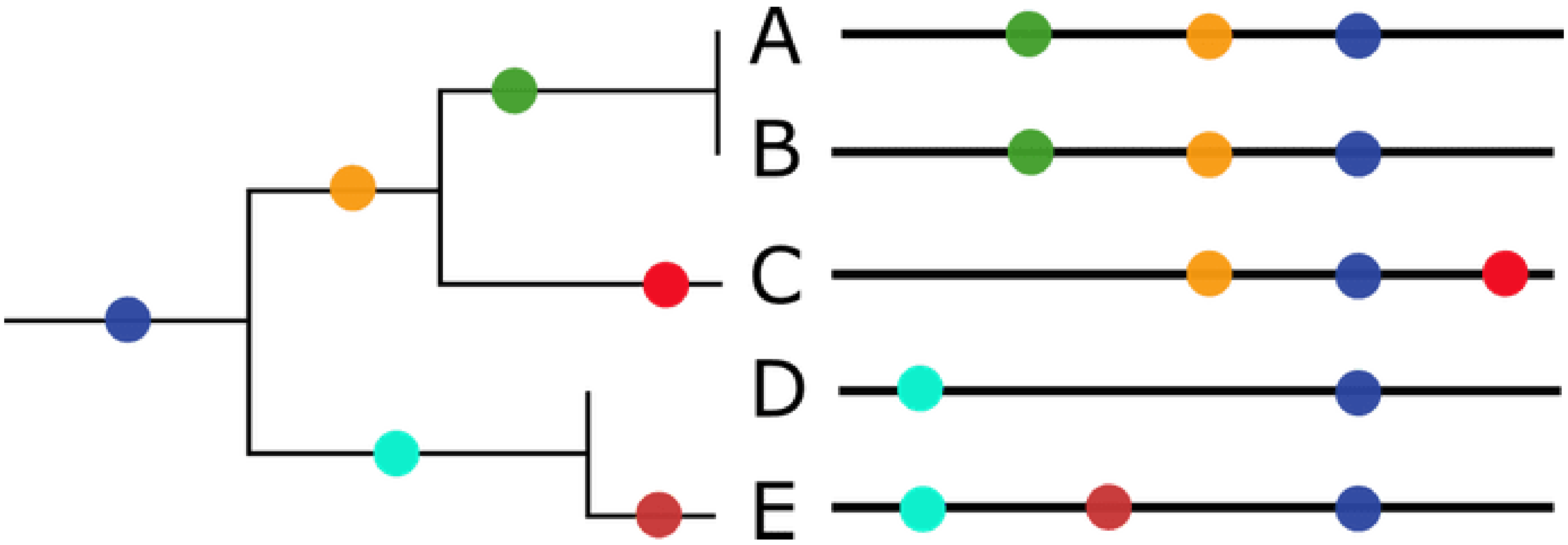
Develop Interventions



Healthline.org, JL Gardy & NJ Loman, Nature Review Genetics 2018, CDC

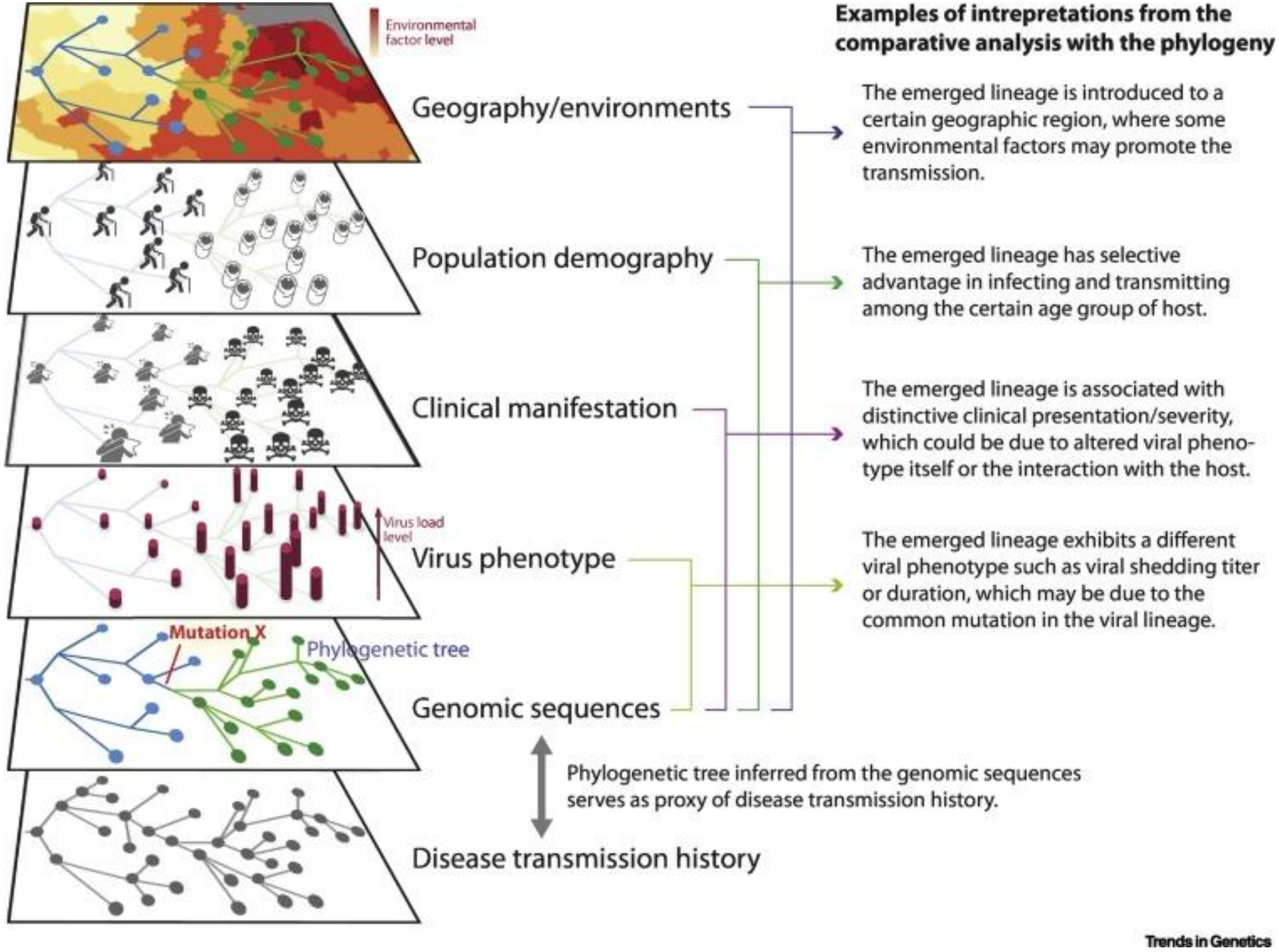
Using Mutations to Examine Relatedness

- We can use the mutations that are generated to build trees that allow us to examine how related to isolates may be
- The blue mutation is shared by all isolates and allows us to hypothesize that it was the first mutation to occur



nextstrain.org

Overlaying Epidemiological and Genomic Data



Examples of intrepretations from the comparative analysis with the phylogeny

The emerged lineage is introduced to a certain geographic region, where some environmental factors may promote the transmission.

The emerged lineage has selective advantage in infecting and transmitting among the certain age group of host.

The emerged lineage is associated with distinctive clinical presentation/severity, which could be due to altered viral phenotype itself or the interaction with the host.

The emerged lineage exhibits a different viral phenotype such as viral shedding titer or duration, which may be due to the common mutation in the viral lineage.

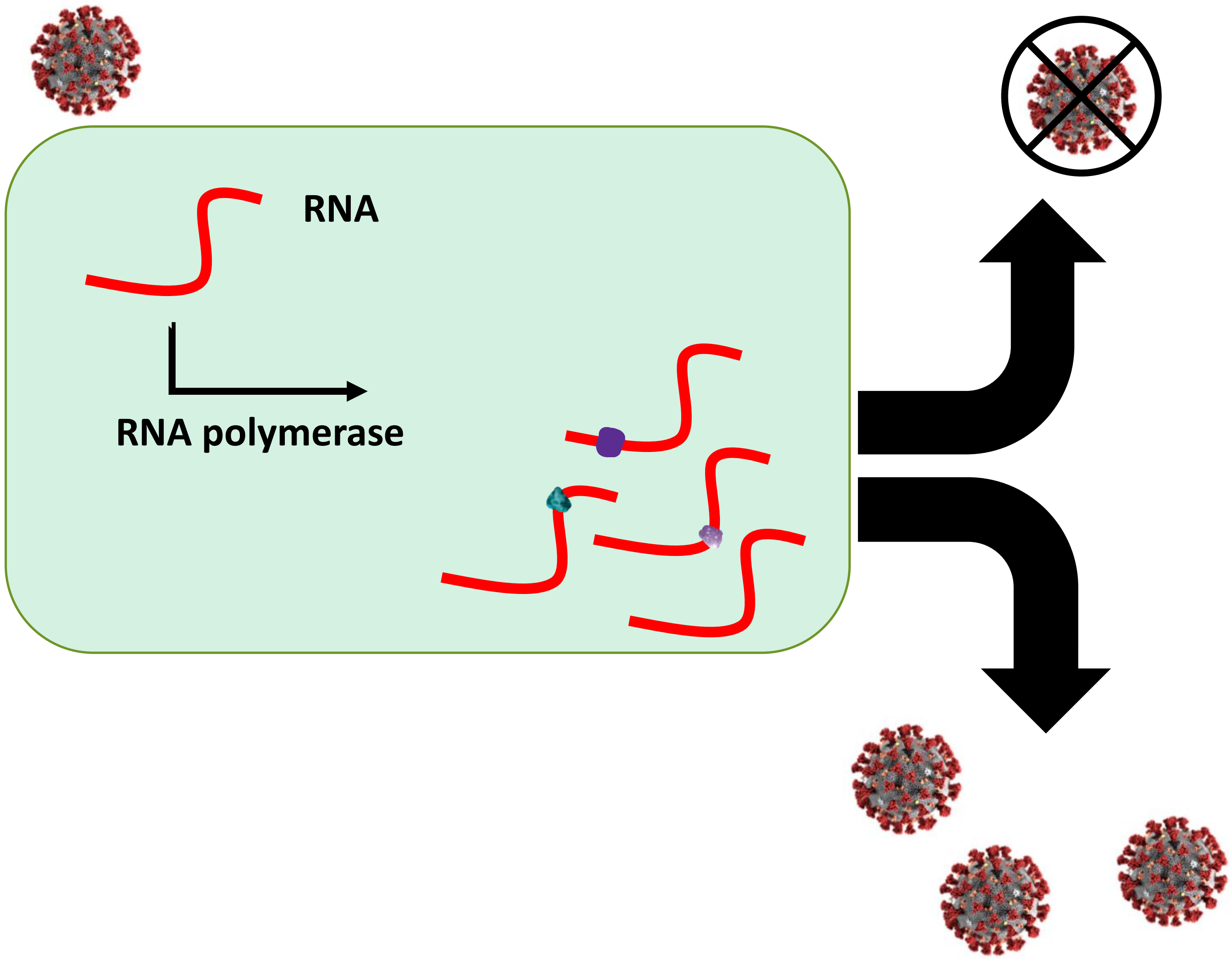
Phylogenetic tree inferred from the genomic sequences serves as proxy of disease transmission history.

Trends in Genetics

Lam et al. (2020)

SARS-CoV-2 Specifics

SARS-CoV-2



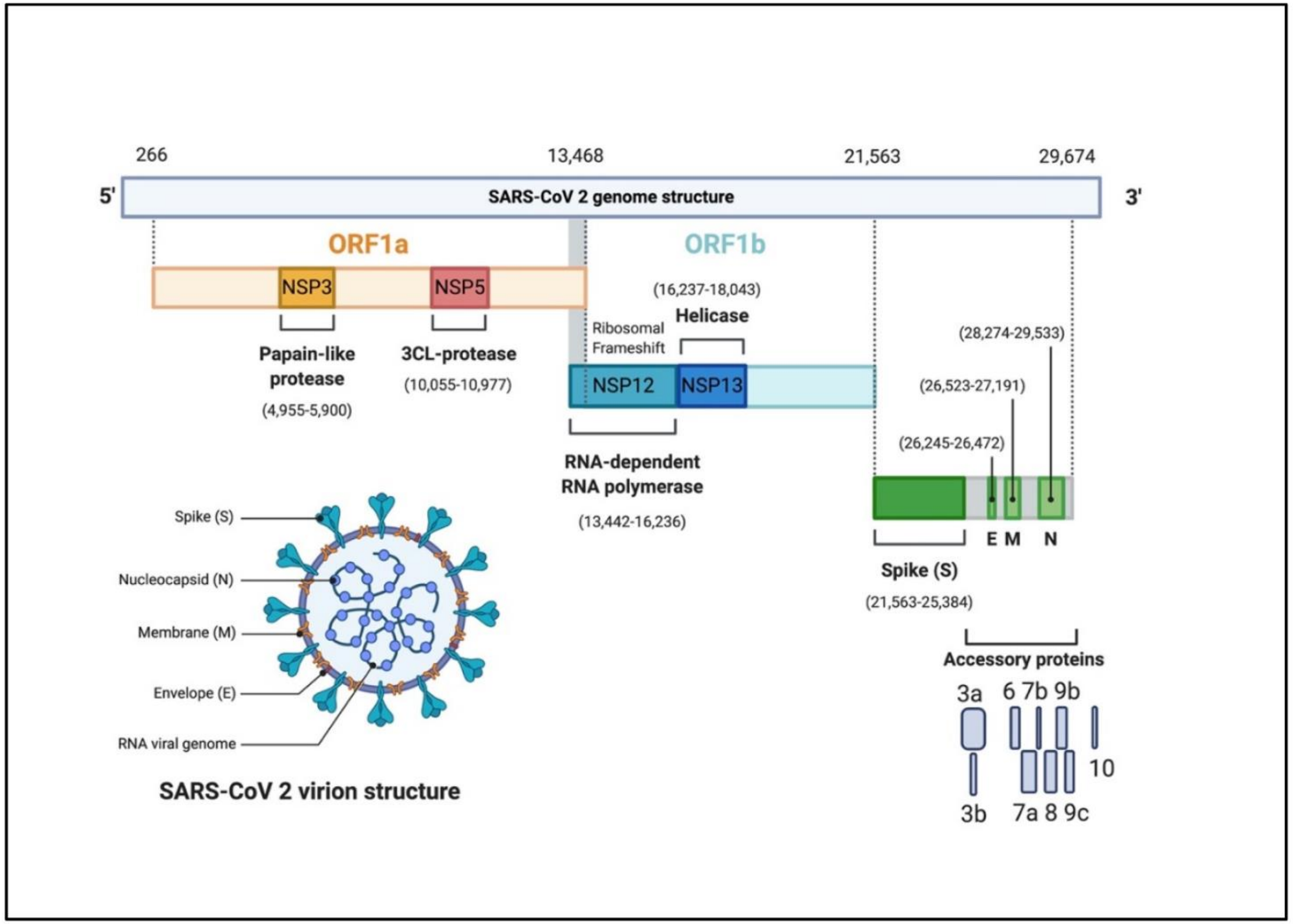
- RNA viruses infect human (and other eukaryotic) cells because they lack the ability to replicate themselves

- The RNA polymerase is highly error prone and can result in mutations in the genome

- Most mutations will not affect the virus due to:
  - Redundancy in genetic code
  - No longer viable virus
  - Slight change in protein

SARS-CoV-2 Genome

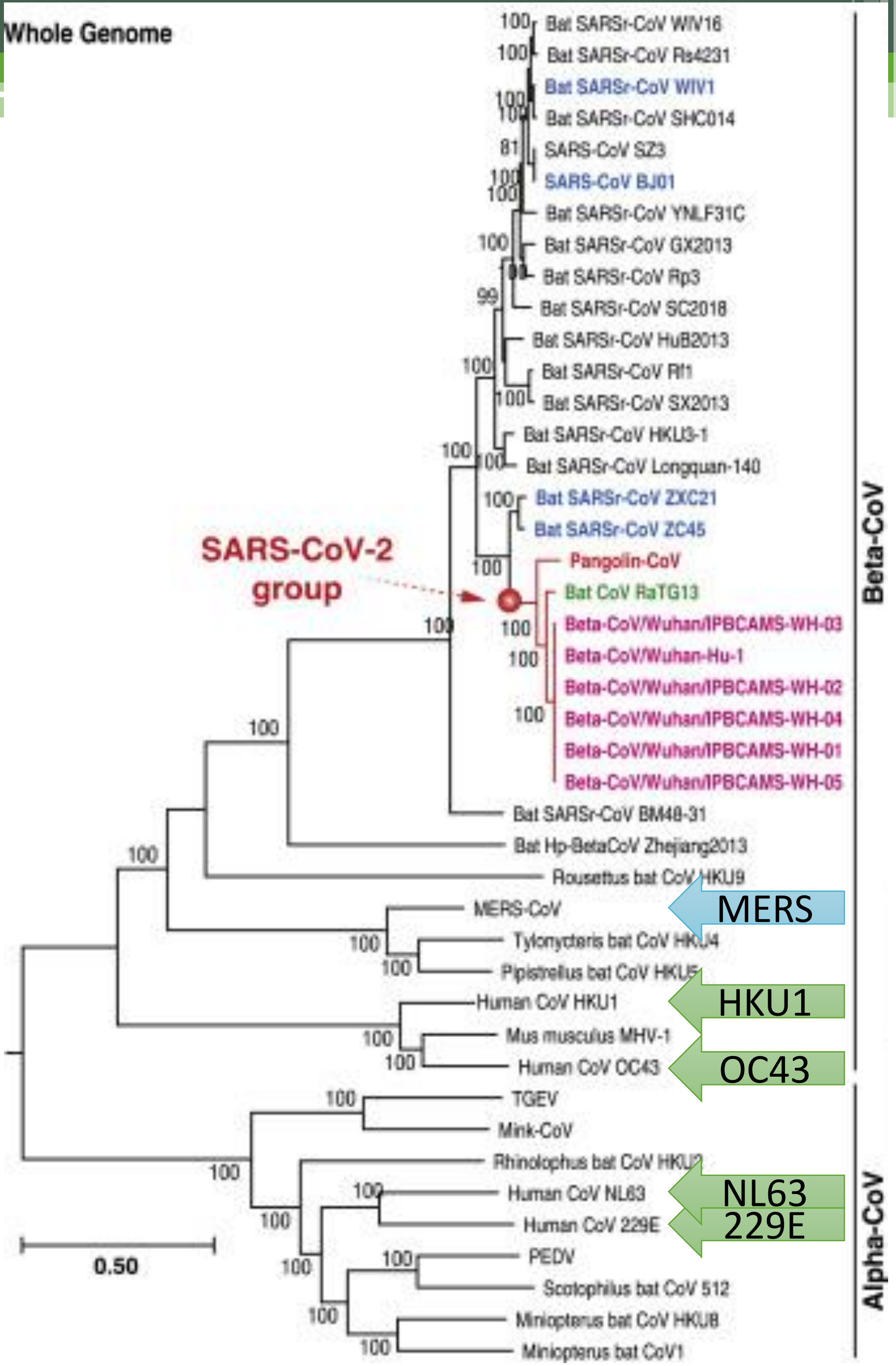
- 4 main structural proteins that are highly related based on sequence similarity to SARS-CoV-2 and MERS
  - S: spike
  - N: Nucleocapsid
  - E: Envelope
  - M: Membrane
- 11 protein coding genes and 12 expressed proteins
- ORF1a and ORF1ab – replication and protein modification



Alanagreh et al (2020) Pathogens

SARS-CoV-2

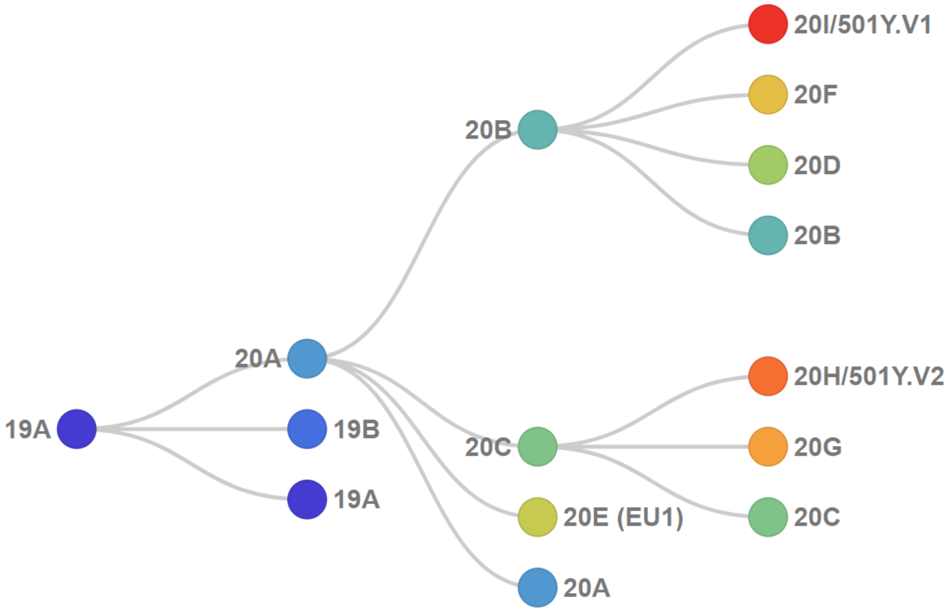
- Origin is hypothesized to be bat or pangolin coronavirus
- 89% sequence identity to other coronaviruses
- Common cold coronaviruses highlighted in red
- Uniqueness made it hard to identify genomic material for reference and positive control at the start of the pandemic



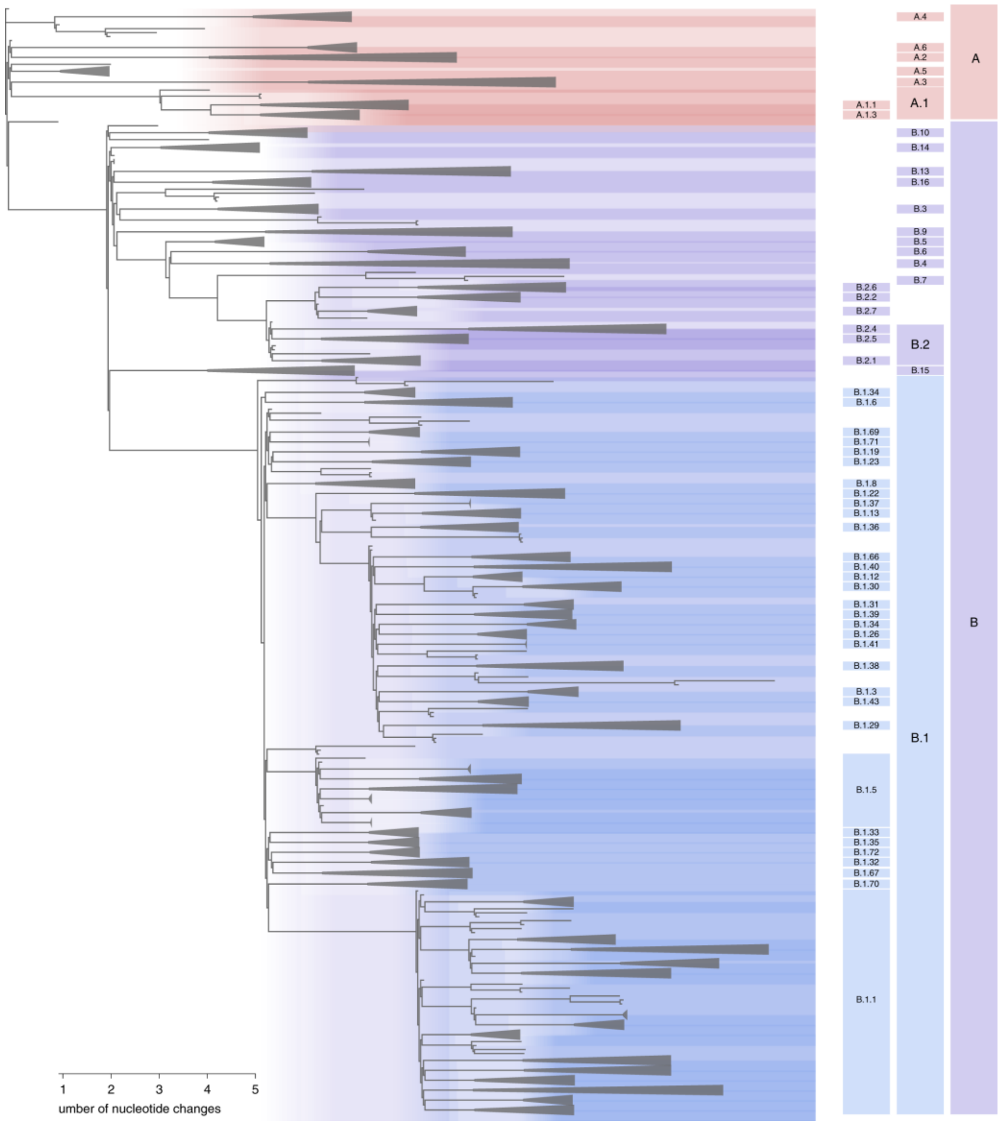
Adapted from Zhang et al (2020) Current Biology.

Clades (Typing)

- Differences in clade classification may have varying levels of discrimination among isolates
- Each clade is characterized by specific mutations

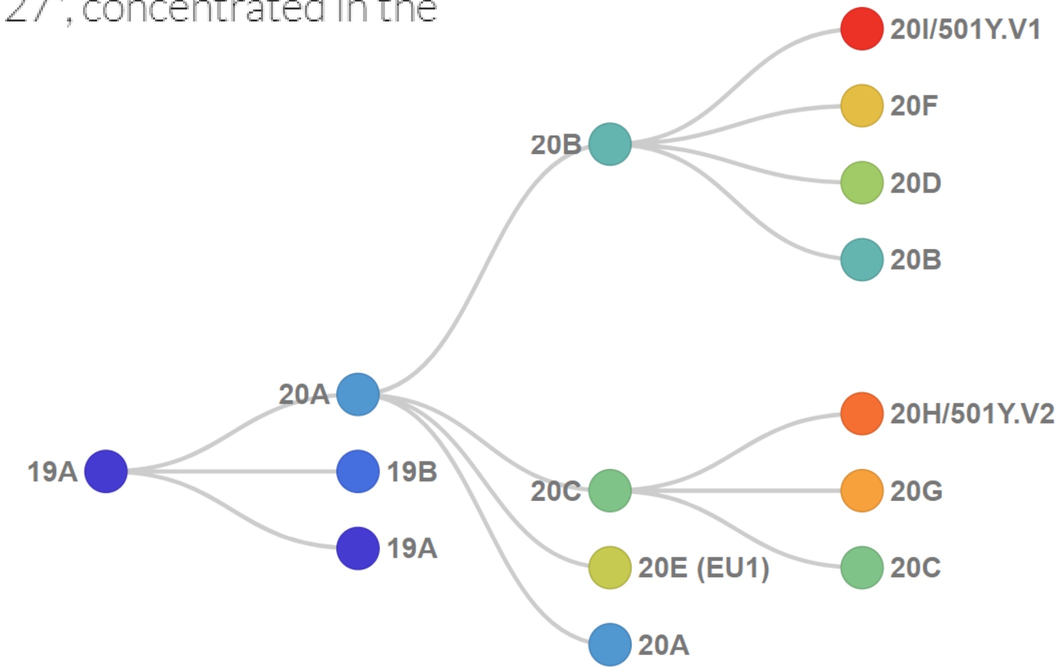


Nextclade/Nextstrain



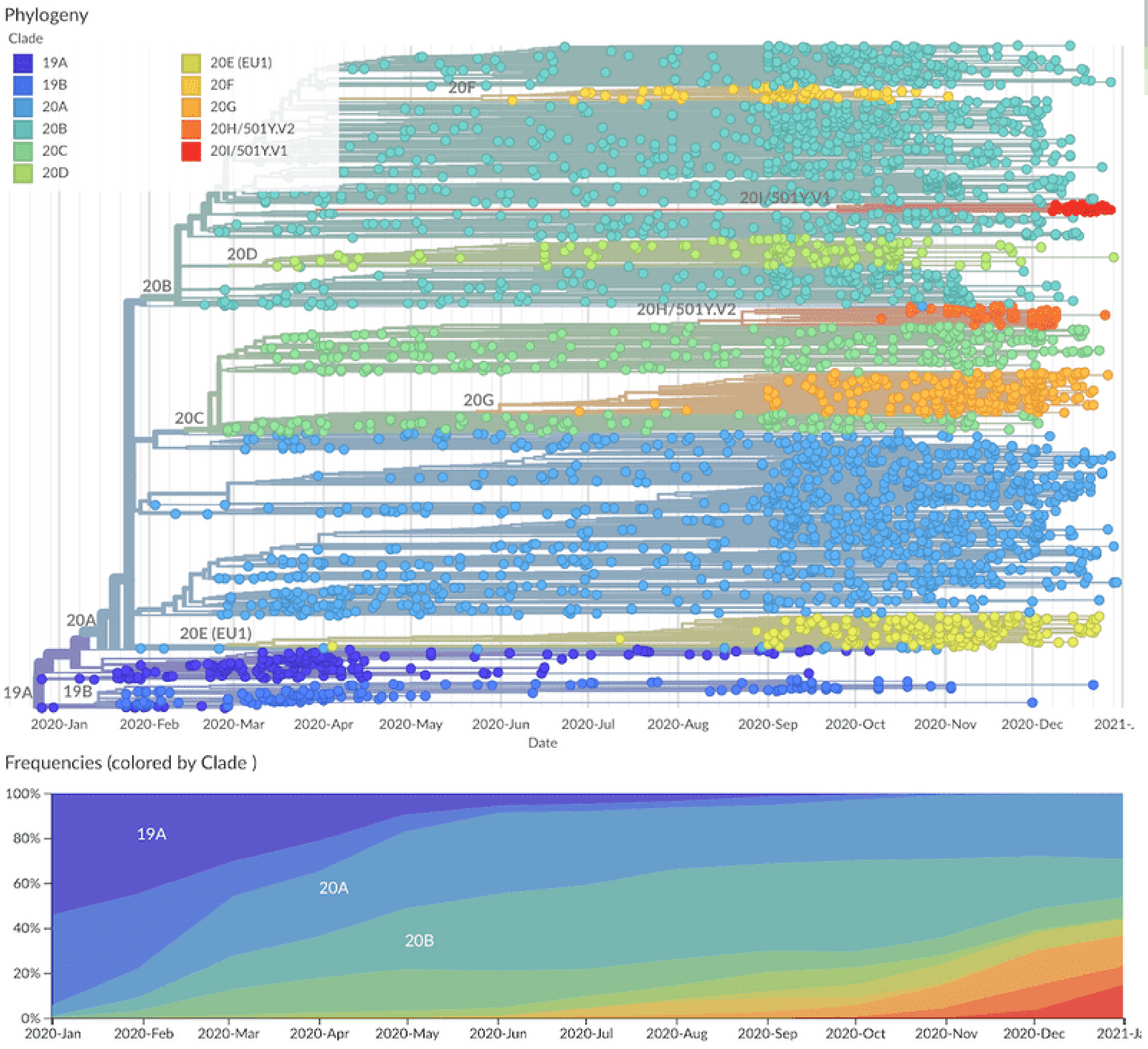
Pangolin

- `20A`: basal pandemic lineage bearing S 614G that's globally distributed
- `20B`: derived from 20A bearing N 203K, N204R and ORF14 50N, also globally distributed
- `20C`: derived from 20A bearing ORF3a 57H and ORF1a 265I, also globally distributed
- `20D`: derived from 20B bearing ORF1a 1246I and ORF1a 3278S, concentrated in South America, southern Europe and South Africa
- `20E`: derived from 20A bearing N 220V, ORF10 30L, ORF14 67F and S 222V, concentrated in Europe
- `20F`: derived from 20B bearing ORF1a 300F and S 477N, concentrated in Australia
- `20G`: derived from 20C bearing ORF1b 1653D, ORF3a 172V, N 67S and N 199L, concentrated in the United States
- `20H/501Y.V2`: derived from 20C bearing S 80A, S 215G, S 484K, S 501Y, S 701V, concentrated in South Africa
- `20I/501Y.V1`: derived from 20B bearing S 501Y, S 570D, S 681H, ORF8 27*, concentrated in the United Kingdom

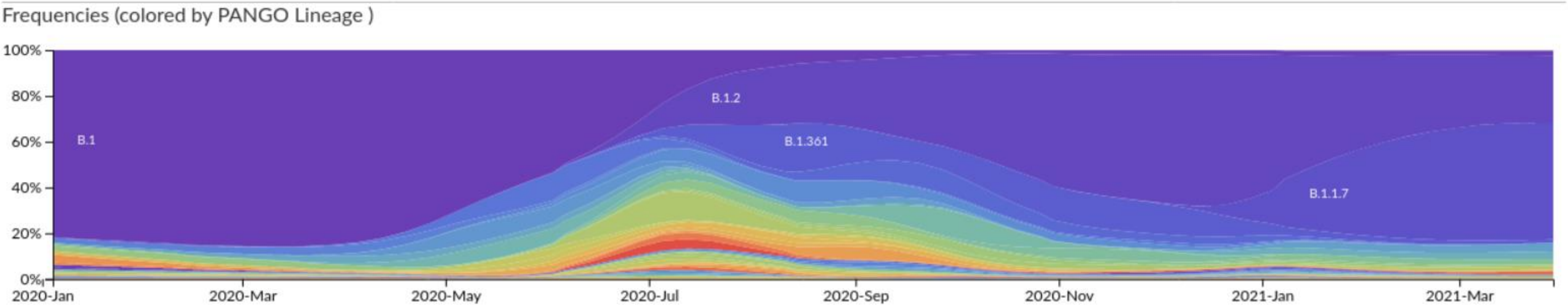


Clades

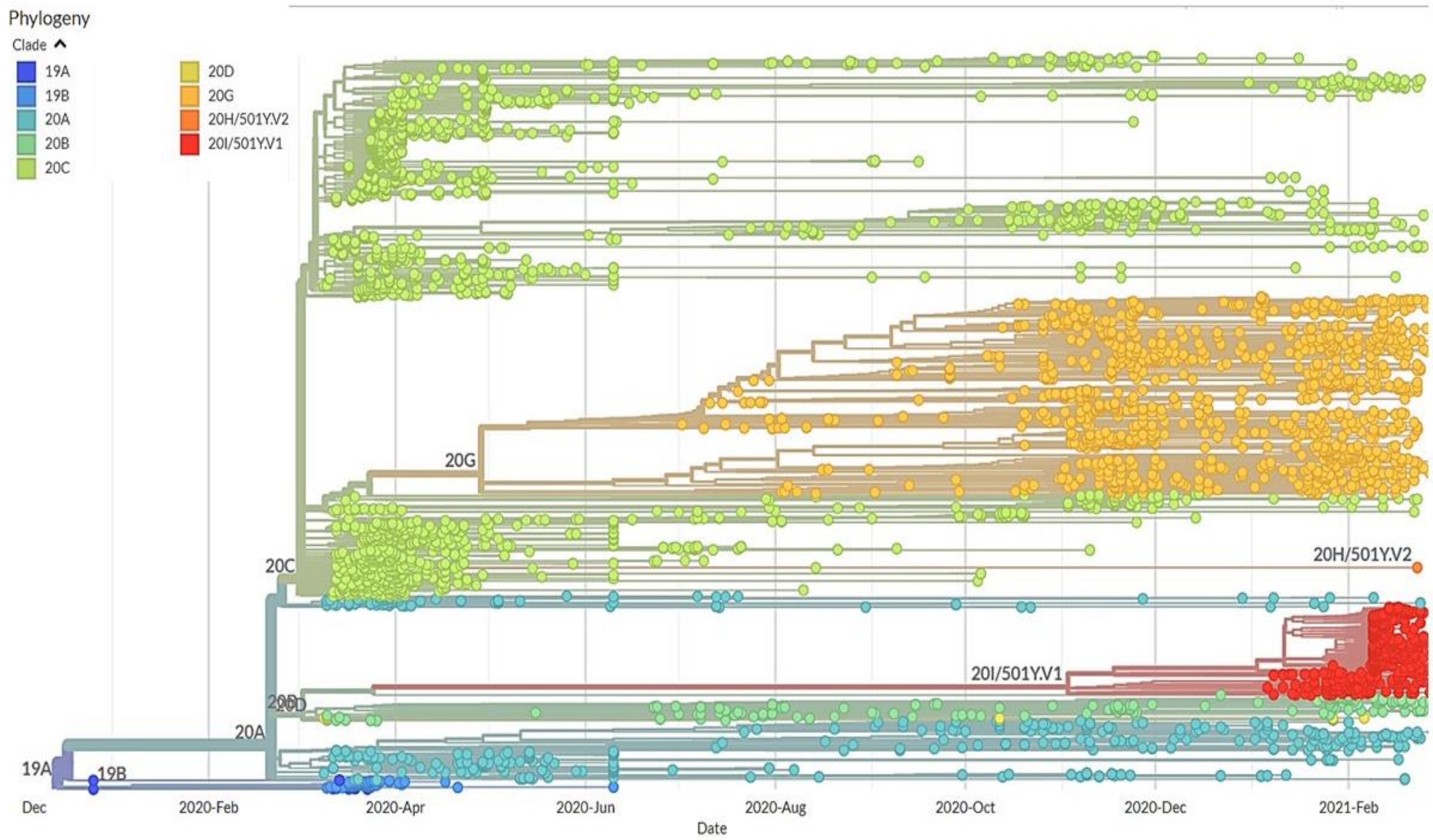
1. A clade reaches >20% global frequency for 2 or more months
2. A clade reaches >30% regional frequency for 2 or more months
3. A VOC ('variant of concern') is recognized (applies currently to 501Y.V1 and 501Y.V2)



Clades in Michigan



10,000 ft view of Michigan



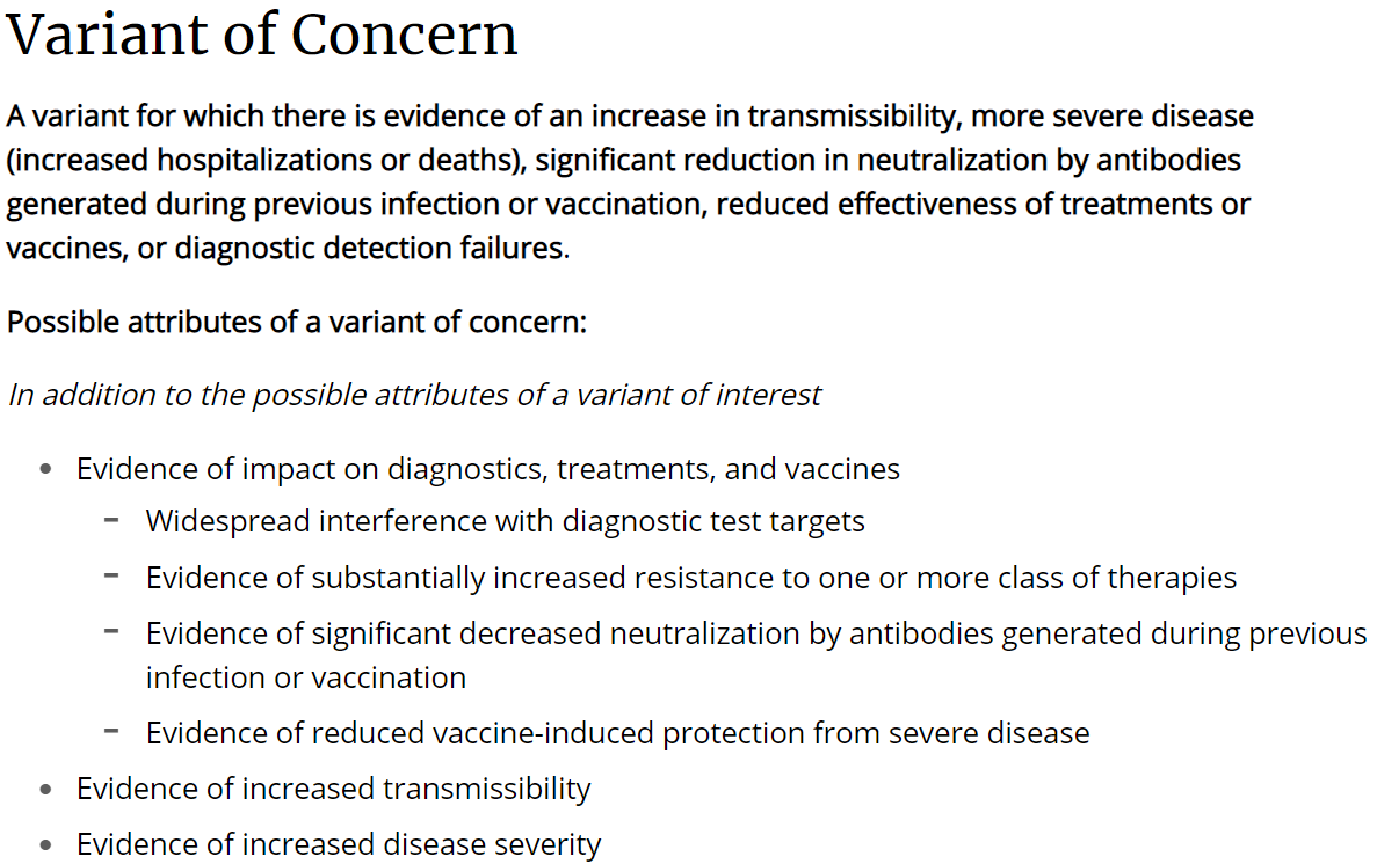
Variant of Interest

A variant with specific genetic markers that have been associated with changes to receptor binding, reduced neutralization by antibodies generated against previous infection or vaccination, reduced efficacy of treatments, potential diagnostic impact, or predicted increase in transmissibility or disease severity.

Possible attributes of a variant of interest:

- Specific genetic markers that are predicted to affect transmission, diagnostics, therapeutics, or immune escape
- Evidence that demonstrates it is the cause of an increased proportion of cases or unique outbreak clusters
- Limited prevalence or expansion in the US or in other countries

CDC



Variant of Concern

A variant for which there is evidence of an increase in transmissibility, more severe disease (increased hospitalizations or deaths), significant reduction in neutralization by antibodies generated during previous infection or vaccination, reduced effectiveness of treatments or vaccines, or diagnostic detection failures.

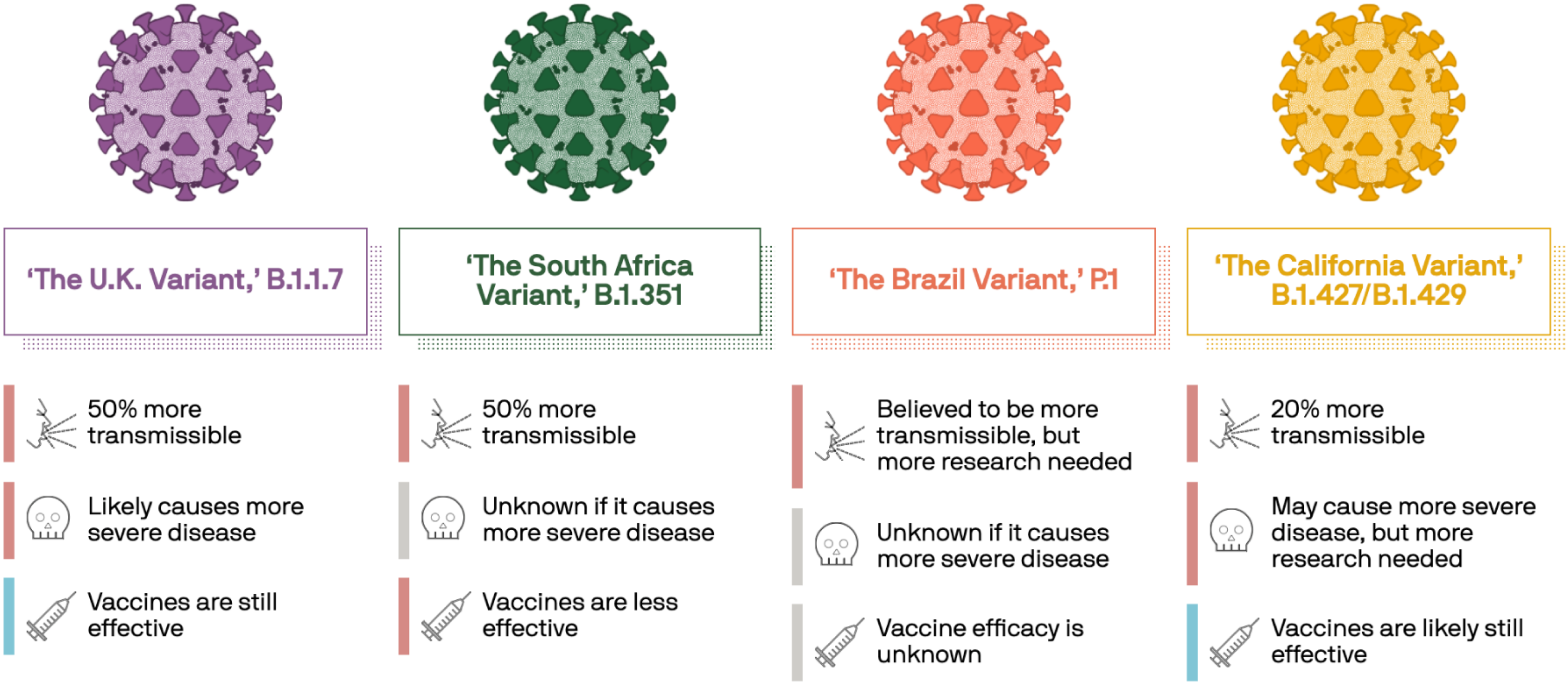
Possible attributes of a variant of concern:

In addition to the possible attributes of a variant of interest

- Evidence of impact on diagnostics, treatments, and vaccines
    - Widespread interference with diagnostic test targets
    - Evidence of substantially increased resistance to one or more class of therapies
    - Evidence of significant decreased neutralization by antibodies generated during previous infection or vaccination
    - Evidence of reduced vaccine-induced protection from severe disease
- Evidence of increased transmissibility
- Evidence of increased disease severity

CDC

Variants of Concern



'The U.K. Variant,' B.1.1.7	'The South Africa Variant,' B.1.351	'The Brazil Variant,' P.1	'The California Variant,' B.1.427/B.1.429
🟥 50% more transmissible	🟥 50% more transmissible	🟥 Believed to be more transmissible, but more research needed	🟥 20% more transmissible
🟥 Likely causes more severe disease	⬜ Unknown if it causes more severe disease	⬜ Unknown if it causes more severe disease	🟥 May cause more severe disease, but more research needed
🟦 Vaccines are still effective	🟥 Vaccines are less effective	⬜ Vaccine efficacy is unknown	🟦 Vaccines are likely still effective

Variant of High Consequence

A variant of high consequence has clear evidence that prevention measures or medical countermeasures (MCMs) have significantly reduced effectiveness relative to previously circulating variants.

Possible attributes of a variant of high consequence:

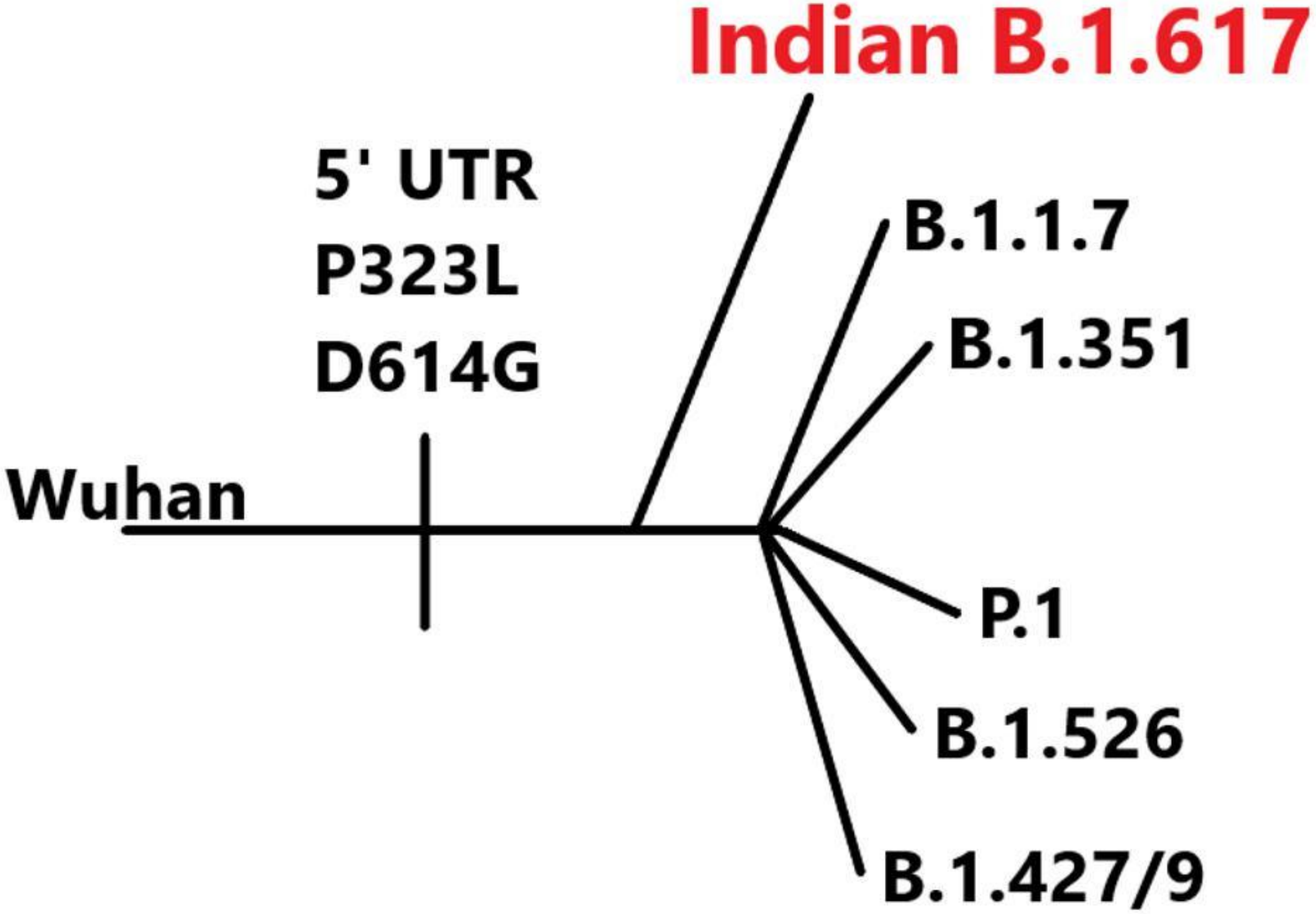
In addition to the possible attributes of a variant of concern

- Impact on Medical Countermeasures (MCM)
  - Demonstrated failure of diagnostics
  - Evidence to suggest a significant reduction in vaccine effectiveness, a disproportionately high number of vaccine breakthrough cases, or very low vaccine-induced protection against severe disease
  - Significantly reduced susceptibility to multiple Emergency Use Authorization (EUA) or approved therapeutics
  - More severe clinical disease and increased hospitalizations

CDC

Other Variants

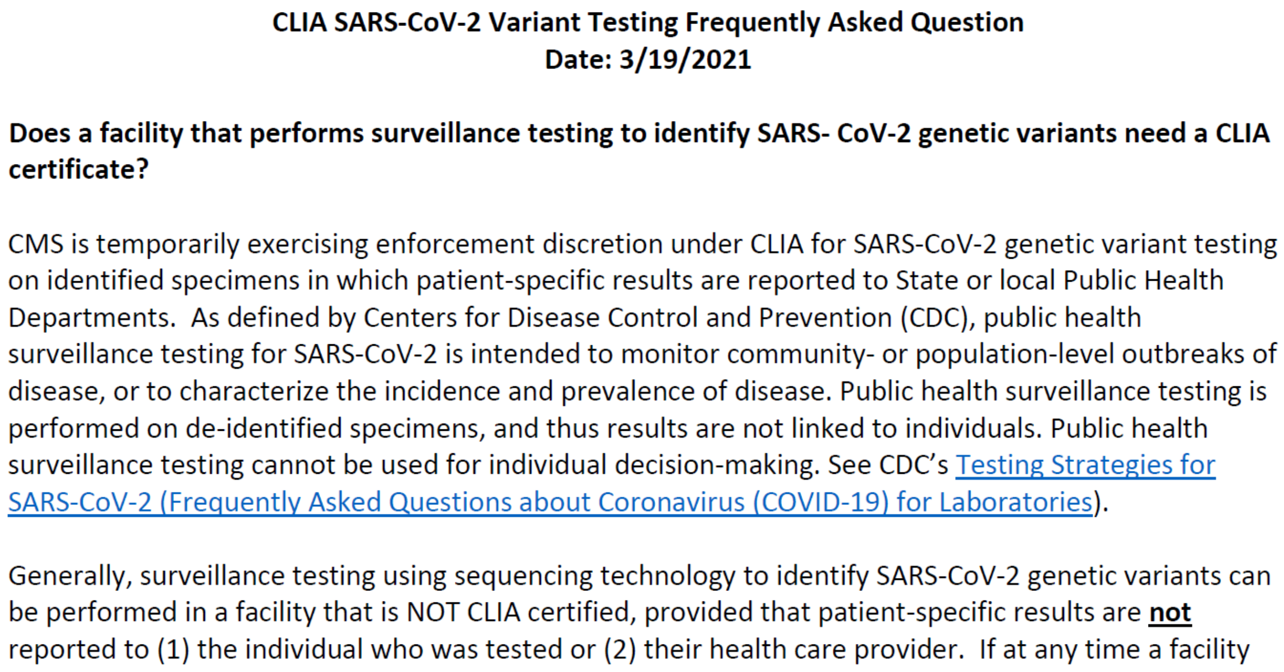
- Spike 614G – March 2020 [Nextstrain]



CMS Guidance for SARS-CoV-2 Sequencing

- De-identified samples
- Report to Public Health only
- Validated test

- Reports do not go back to patients or providers for clinical care use



CLIA SARS-CoV-2 Variant Testing Frequently Asked Question
Date: 3/19/2021

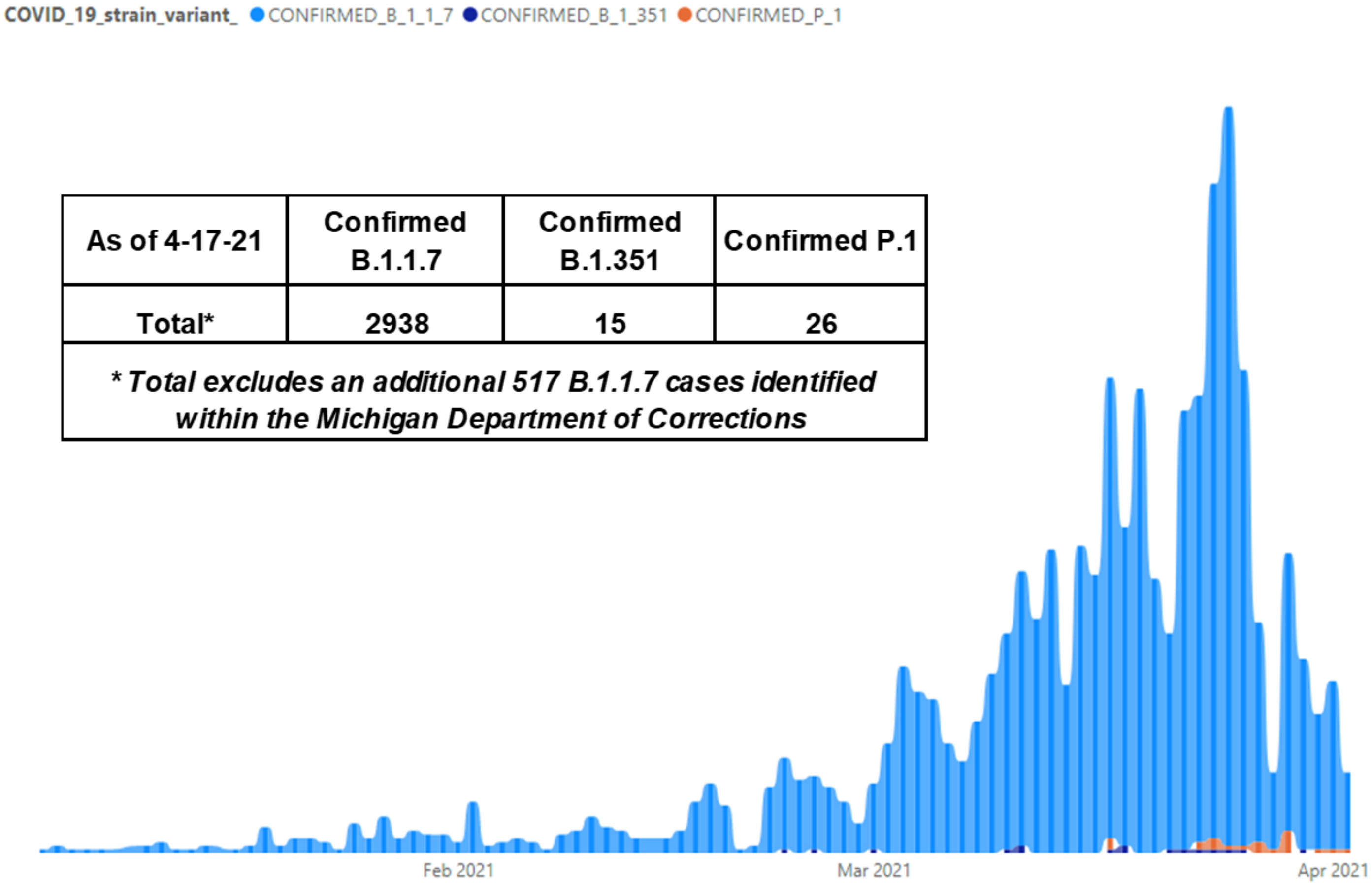
Does a facility that performs surveillance testing to identify SARS- CoV-2 genetic variants need a CLIA certificate?

CMS is temporarily exercising enforcement discretion under CLIA for SARS-CoV-2 genetic variant testing on identified specimens in which patient-specific results are reported to State or local Public Health Departments. As defined by Centers for Disease Control and Prevention (CDC), public health surveillance testing for SARS-CoV-2 is intended to monitor community- or population-level outbreaks of disease, or to characterize the incidence and prevalence of disease. Public health surveillance testing is performed on de-identified specimens, and thus results are not linked to individuals. Public health surveillance testing cannot be used for individual decision-making. See CDC's Testing Strategies for SARS-CoV-2 (Frequently Asked Questions about Coronavirus (COVID-19) for Laboratories).

Generally, surveillance testing using sequencing technology to identify SARS-CoV-2 genetic variants can be performed in a facility that is NOT CLIA certified, provided that patient-specific results are **not** reported to (1) the individual who was tested or (2) their health care provider. If at any time a facility

SARS-CoV-2 Variants of Concern Community Spread

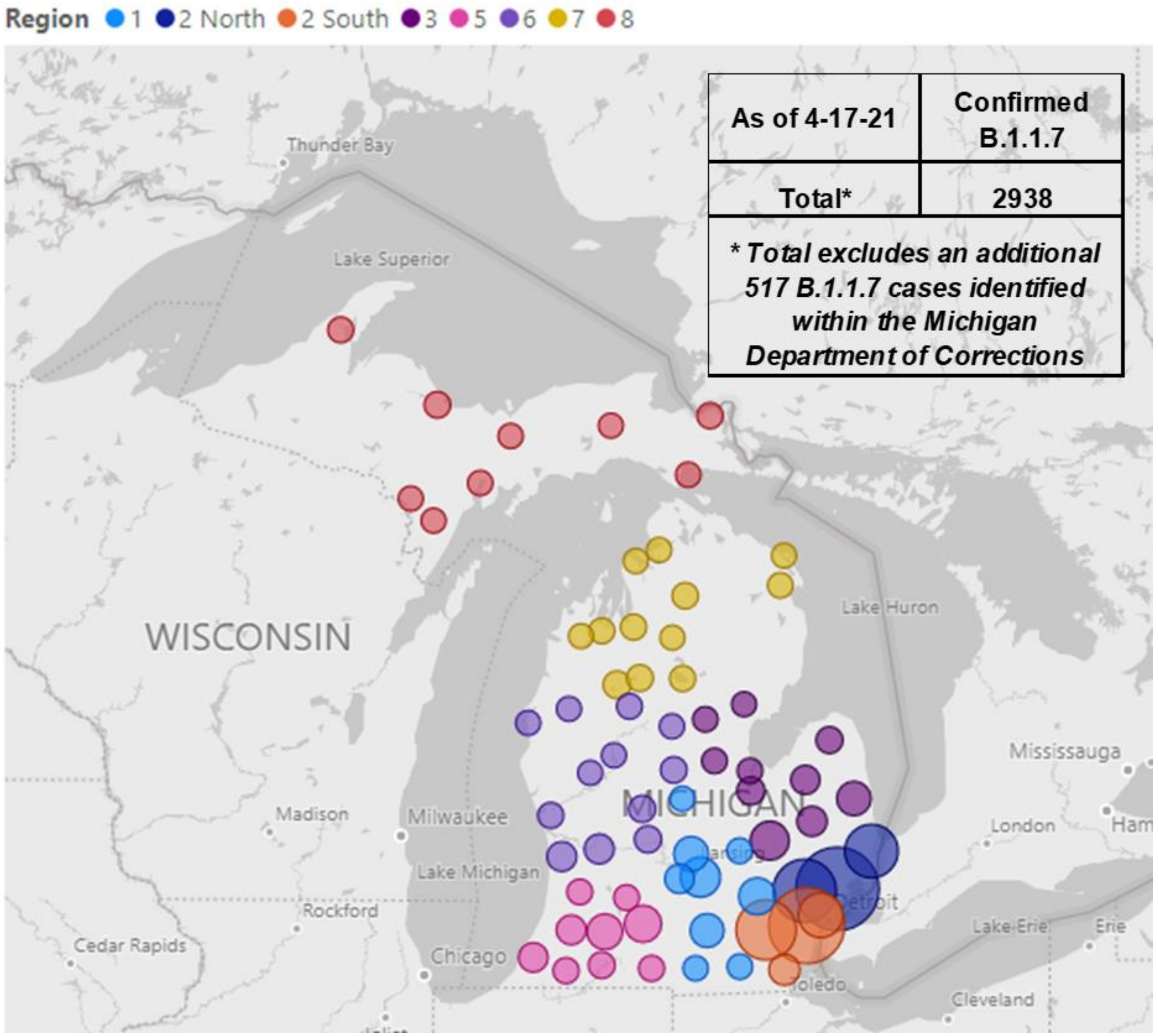
MDHHS identified the first case of B.1.1.7 in Washtenaw County on January 16th | B.1.351 in Jackson County on March 8th | P.1 in Bay County on March 30th



As of 4-17-21	Confirmed B.1.1.7	Confirmed B.1.351	Confirmed P.1
Total*	**2938**	**15**	**26**

Total excludes an additional 517 B.1.1.7 cases identified within the Michigan Department of Corrections

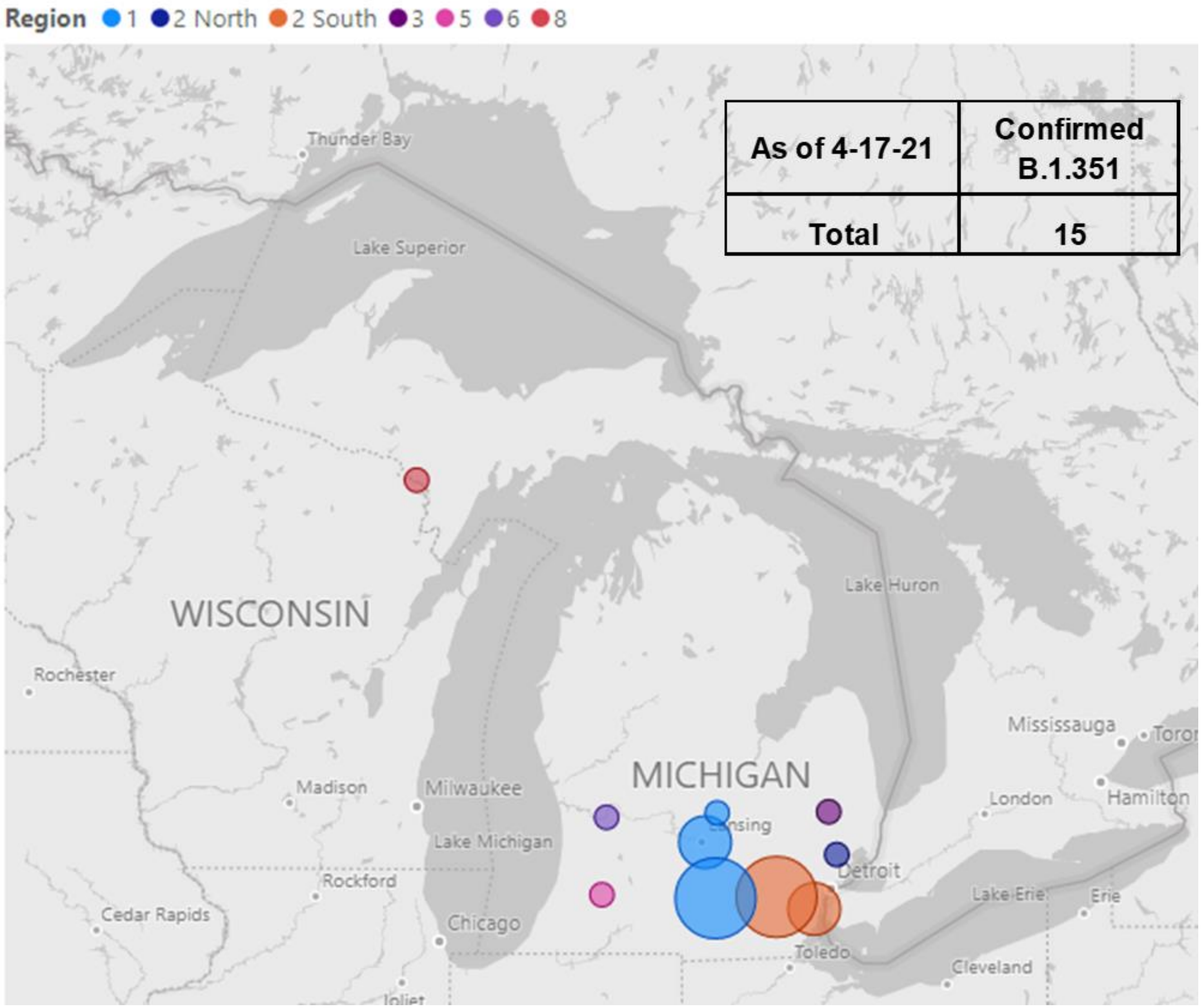
B.1.1.7 Distribution by Public Health Preparedness Region



As of 4-17-21	Confirmed B.1.1.7
Total*	2938

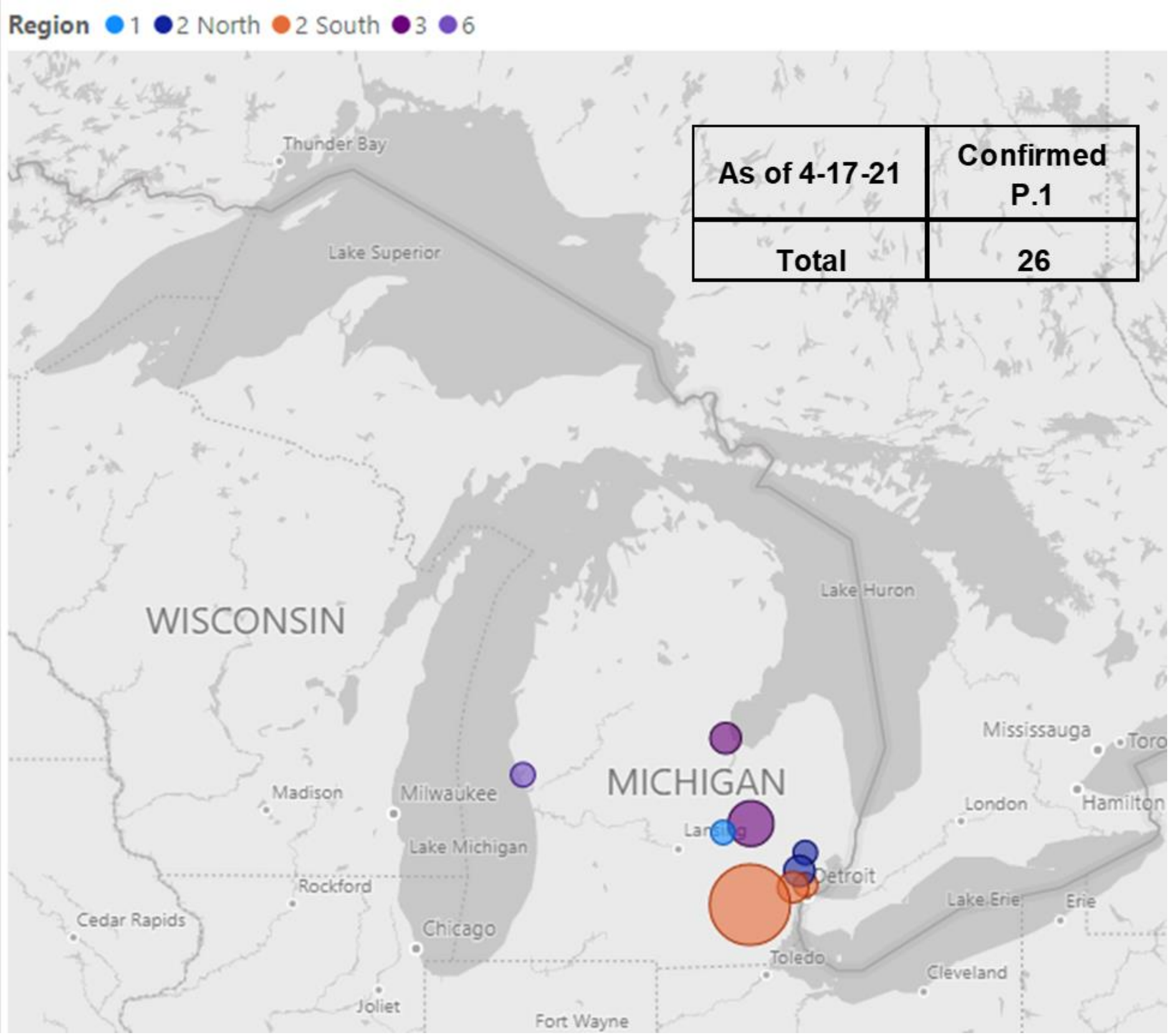
** Total excludes an additional 517 B.1.1.7 cases identified within the Michigan Department of Corrections*

B.1.351 Distribution by Public Health Preparedness Region



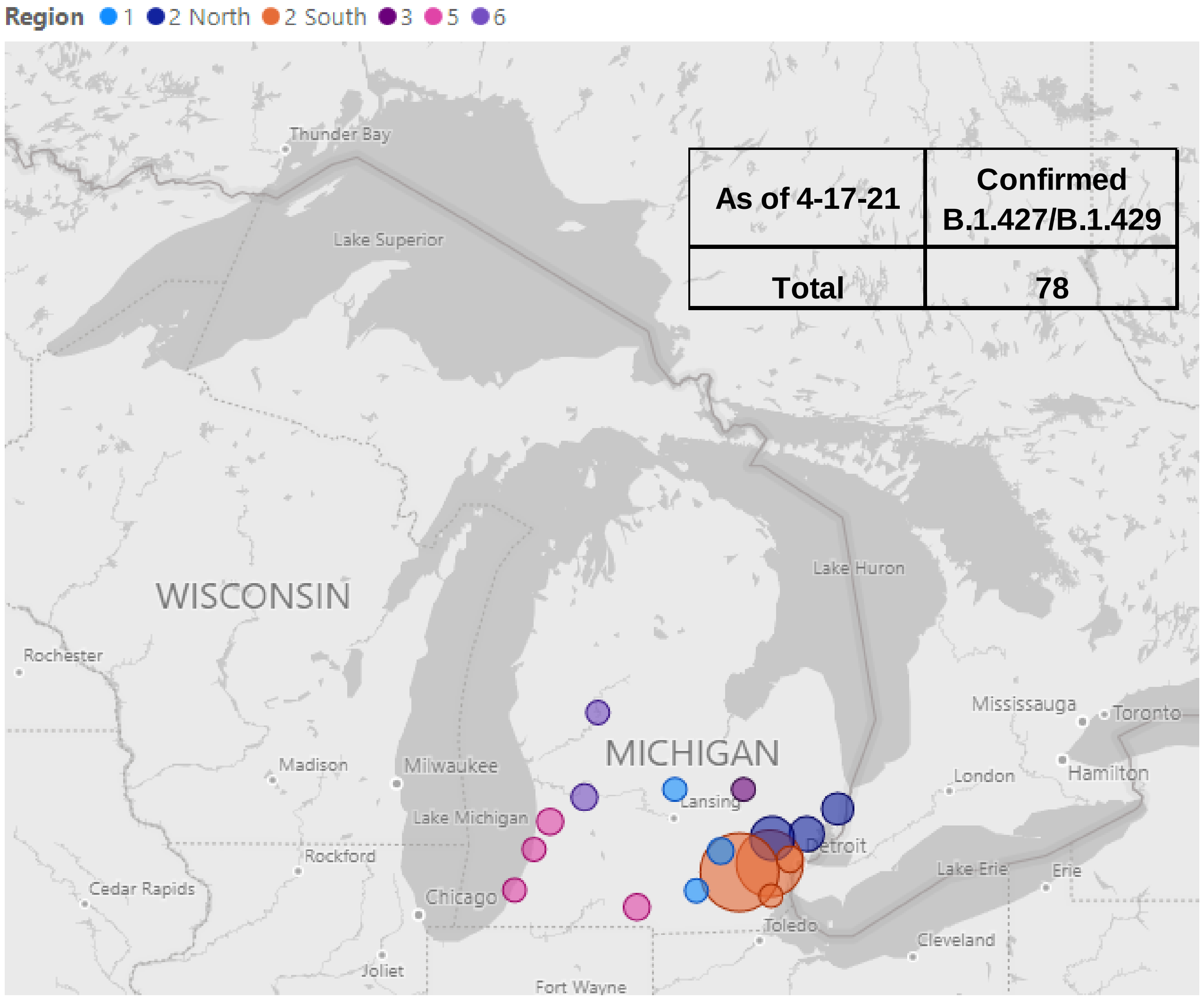
As of 4-17-21	Confirmed B.1.351
Total	15

P.1 Distribution by Public Health Preparedness Region



As of 4-17-21	Confirmed P.1
Total	26

B.1.427/B.1.429 Distribution by Public Health Preparedness Region

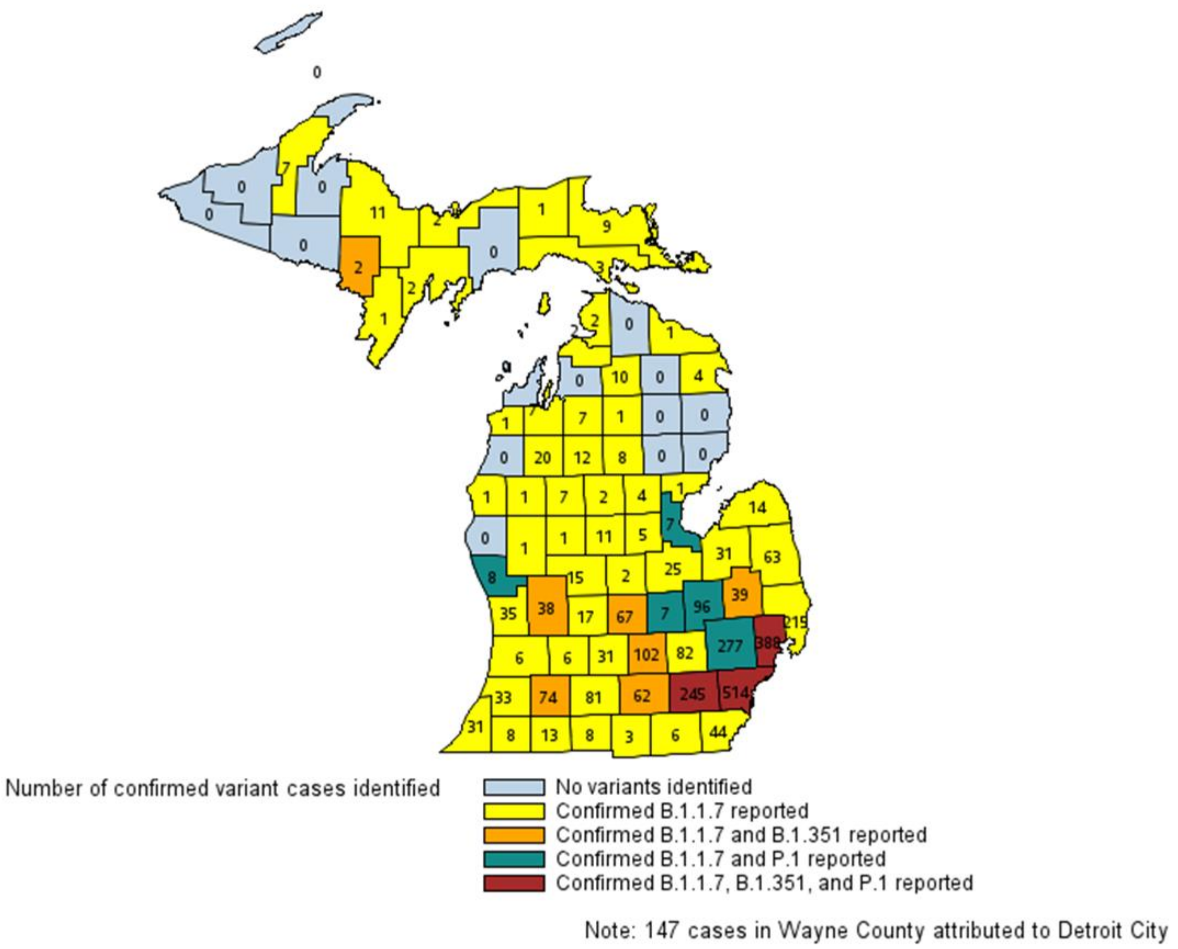


As of 4-17-21	Confirmed B.1.427/B.1.429
Total	78

SARS-CoV-2 Variants of Concern Current Status

Variants of Concern by County

April 17 2021



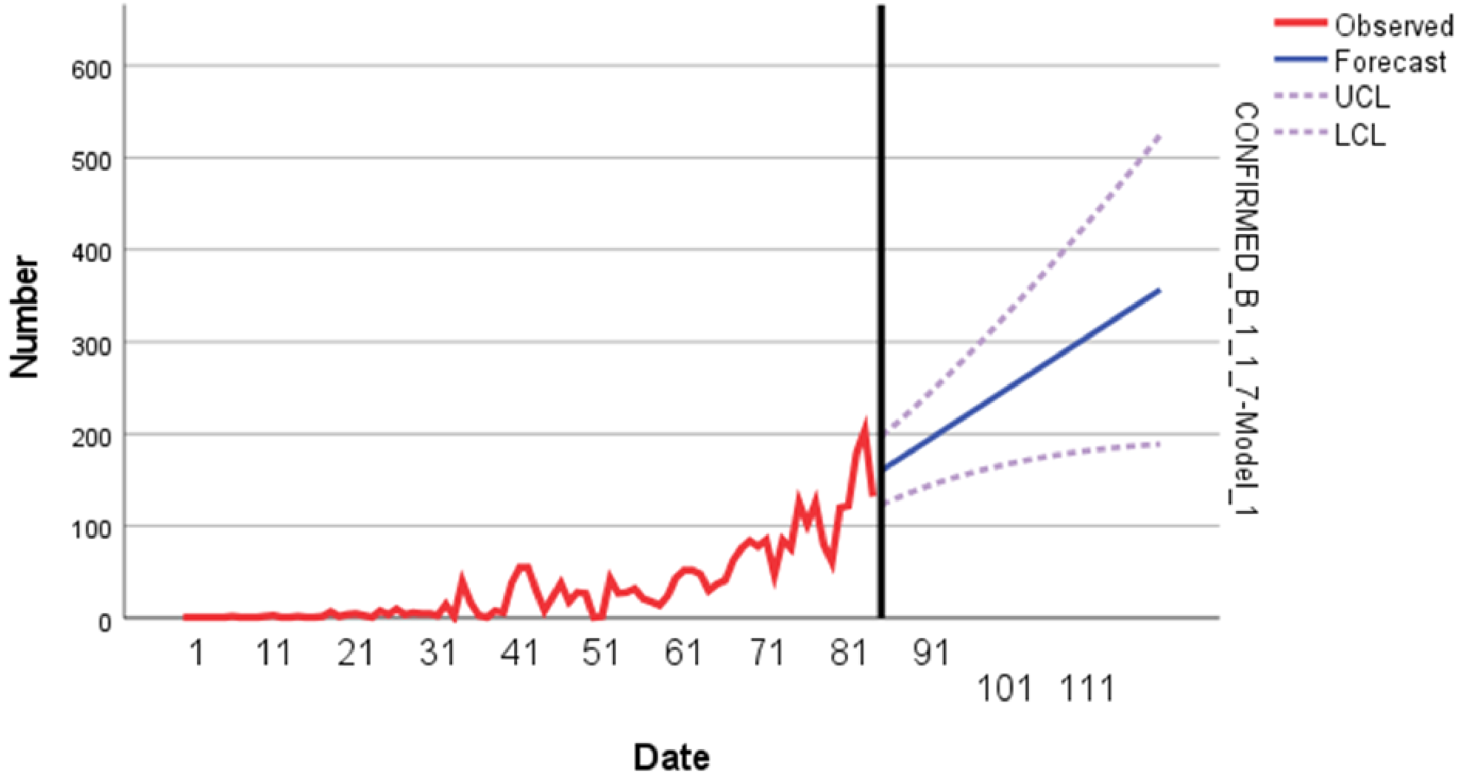
Number of confirmed variant cases identified

- 🟦 No variants identified
- 🟨 Confirmed B.1.1.7 reported
- 🟧 Confirmed B.1.1.7 and B.1.351 reported
- 🟩 Confirmed B.1.1.7 and P.1 reported
- 🟥 Confirmed B.1.1.7, B.1.351, and P.1 reported

Note: 147 cases in Wayne County attributed to Detroit City

SARS-CoV-2 Variants of Concern Forecast

Data modeling indicates that B.1.1.7 variant cases will likely continue to increase in Michigan for the near future, while other variants of concern (B.1.351, P.1, B.1.427/B.1.429) have insufficient data to enable forecasting at this time



Questions?

Heather M Blankenship, PhD

Michigan Department of Health and Human Services
Bureau of Laboratories
Bioinformatics and Sequencing Section Manager

Email:
blankenshiph@michigan.gov

Allen W Stout, MS

Michigan Department of Health and Human Services
Division of Communicable Disease
COVID-19 Surveillance and Investigation Unit Manager

Email:
stouta4@michigan.gov