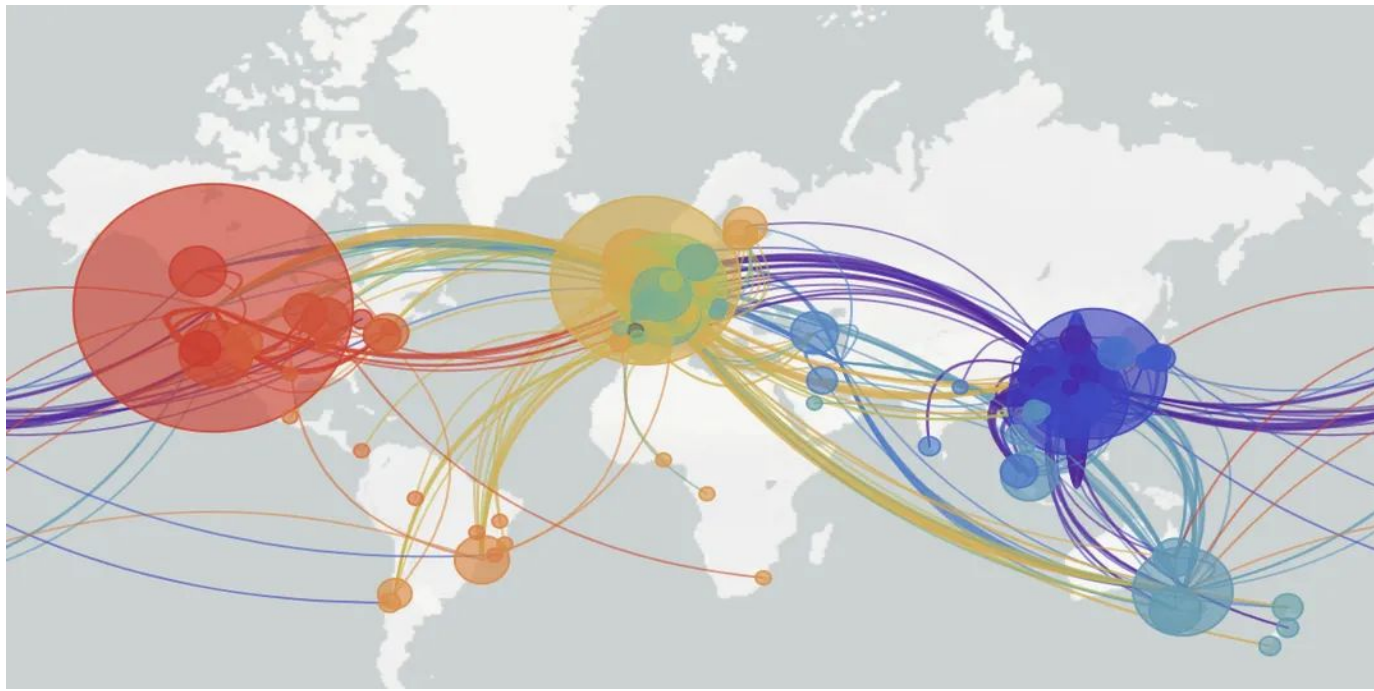# Collaboration in the Time of COVID

Allison Lai

Dr. Bryan Lewis

Background

GISAID
- "Global Initiative on Sharing All Influenza Data"
- Public genome repository for viruses
- As of July 26th, _12,141,074_ COVID sequences have been uploaded to the database

Why is it Significant?
- Has been crucial to identifying and tracking spread of new variants as well as development of vaccines
- In the past, there has been reluctance to share sequences like these, but GISAID has helped put "contributors from higher and lower resource environments on more equal footing" (source)

What I've Done

Comparison of a country's GISAID contributions vs. how they've fared during COVID
- Measures used:
  - GHSI - measures the capacities of 195 countries to prepare for epidemics and pandemics on a scale of 0-100
  - % of cumulative cases sequenced and submitted to GISAID
- Identify outlying countries
  - look for "low hanging fruit"

Look at submissions where the region of origin is different from the region it was sequenced in
- Questions:
  - Which countries are the origins the most? Which ones sequenced most?
  - Did this happen more during different phases of the pandemic?
  - Of countries with low health security index did they provide more sequences than those with higher indices?
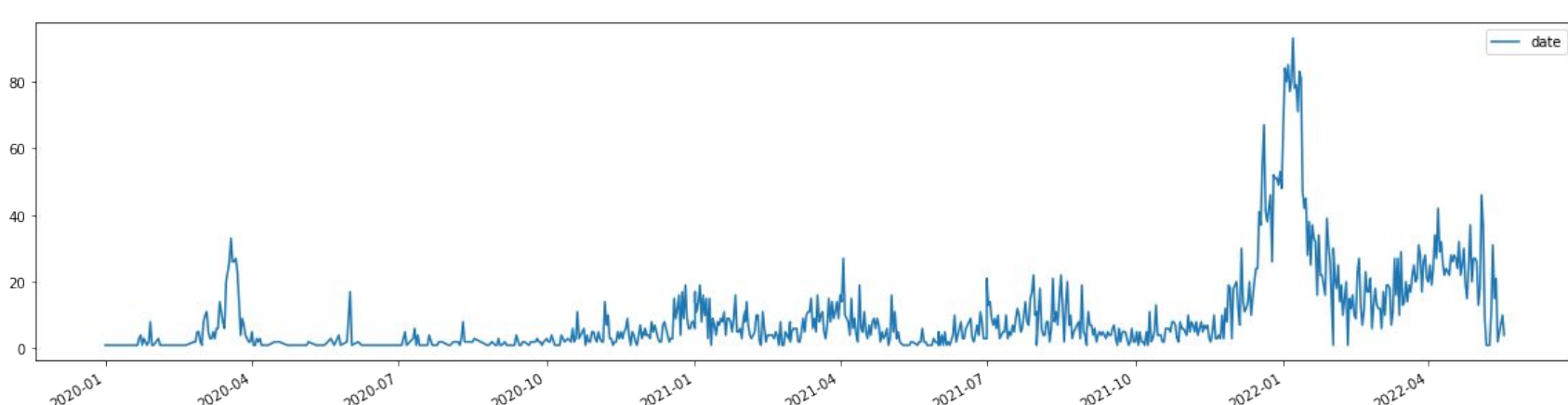

Figure #2. Inter-regional sequencing 2020-2022

References
- "Coronavirus Vaccination Rates." Johns Hopkins Coronavirus Resource Center, https://coronavirus.jhu.edu/.
- "The 2021 Global Health Security Index." GHS Index, 8 Dec. 2021, https://www.ghsindex.org/.
- Predictive of Coronavirus Pandemic Responses among Organization for Economic Cooperation and Development Countries." PloS One, Public Library of Science, 7 Oct. 2020, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7540886/.

Project Goals

1. Assess equity of research and resources in the global sequencing community
- Contrast level of GISAID contributions (sequencing) to how the country has fared during the pandemic
- Main question:
  - Are there certain countries contributing to the global initiative who may not have enough capacity to combat Covid themselves?

. 2. Find inequalities
- (i.e a country that has contributed a lot of sequences but does not have enough vaccines)
- This can help us understand what barriers might prevent these countries from getting help or sharing genomic data on SARS-CoV-2.
- Main question(s):
  - What is the story behind these inequalities?
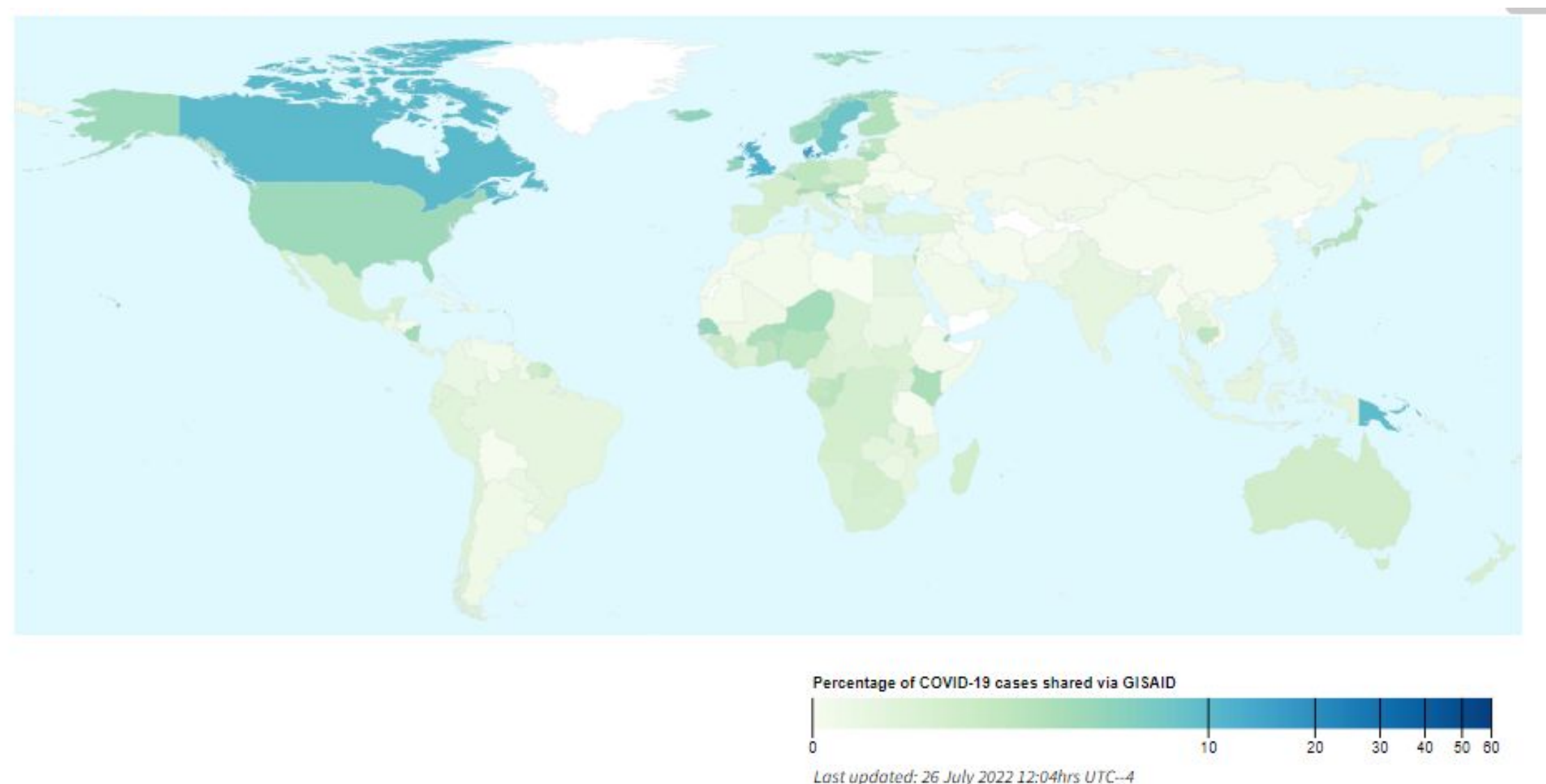 - Are these countries getting the recognition they deserve?

Figure #1. Percentage of COVID-19 cases shared to GISAID by each country since January 2020 (source)

Current/Future Work

Narrow in on countries that have a lot of "outside sequencing"
- Look at fraction of sequences "in country" vs "outside country"
- Issue: scientific imperialism
  - Nations with more resources can sample other countries' samples and publish results without building capacity or credit
- Is there a relationship between a country's capacity (measured by GHSI) and how many outside sequences they do?

Tracking where new variants are first sequenced
- Fraction of submissions that are sequenced in vs out of country of origin when the variant was relatively new (first 60 days)
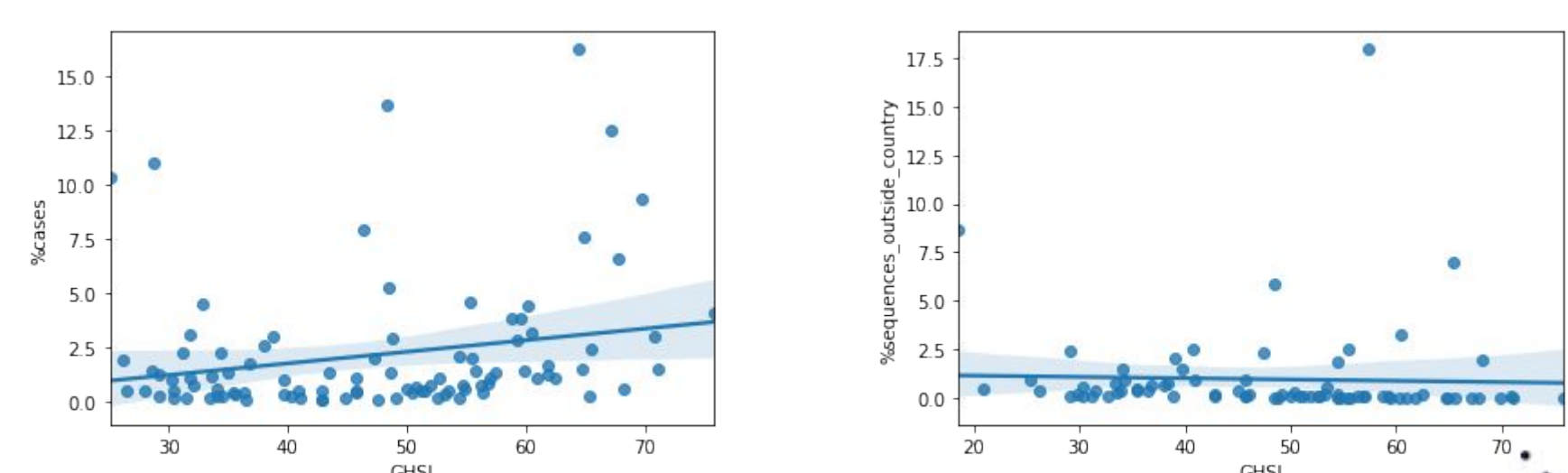- How often was the first detection sequenced in another country?

index country	
Japan	4232
South Korea	1212
Singapore	978
Argentina	207
Thailand	185
Estonia	113
Brazil	65
India	61
Philippines	61
Italy	52
USA	48

Figure #3. left: GHSI vs. % cumulative cases sequenced right: GHSI vs. %sequences from outside country