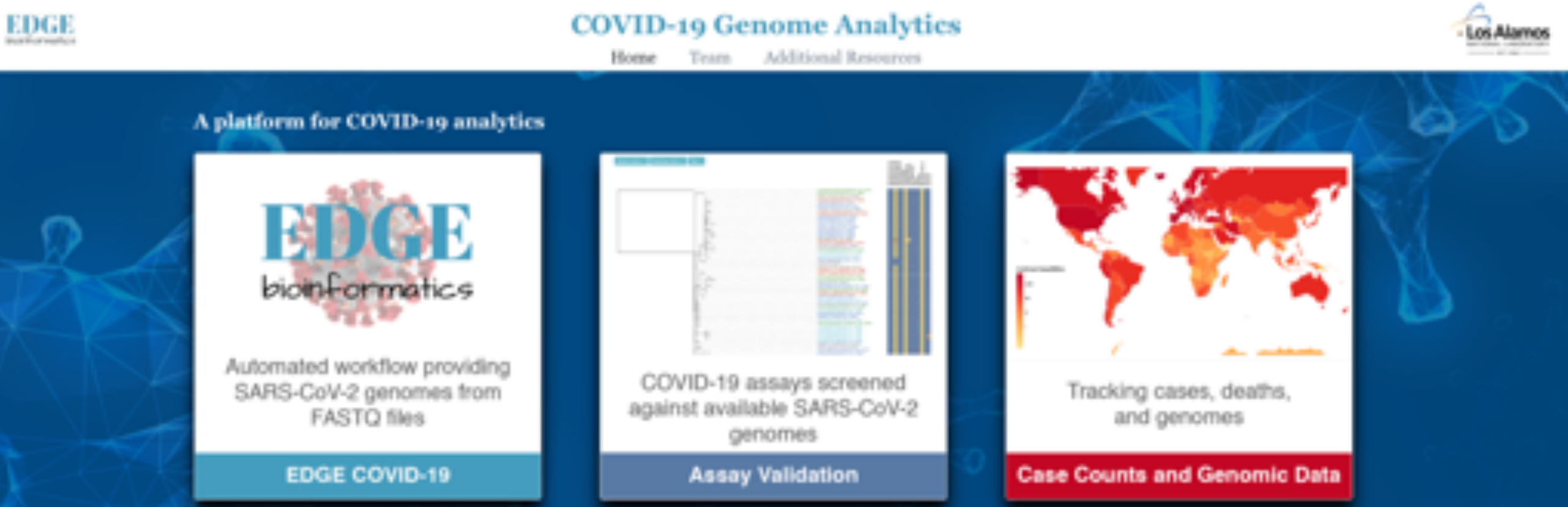


- As genomics is used for biosurveillance of outbreak pathogens, it can help reveal where diagnostic assays may fail – thus we advocate for:
- 1) robust genomic data to be continually generated (even prior to outbreaks) to inform us of pathogen presence and diversity/evolution; 2) continuous tracking of mutations that may affect diagnostic assays and therapeutic targets; 3) automated re-design of assays and suggestion of alternative targets for therapeutic design

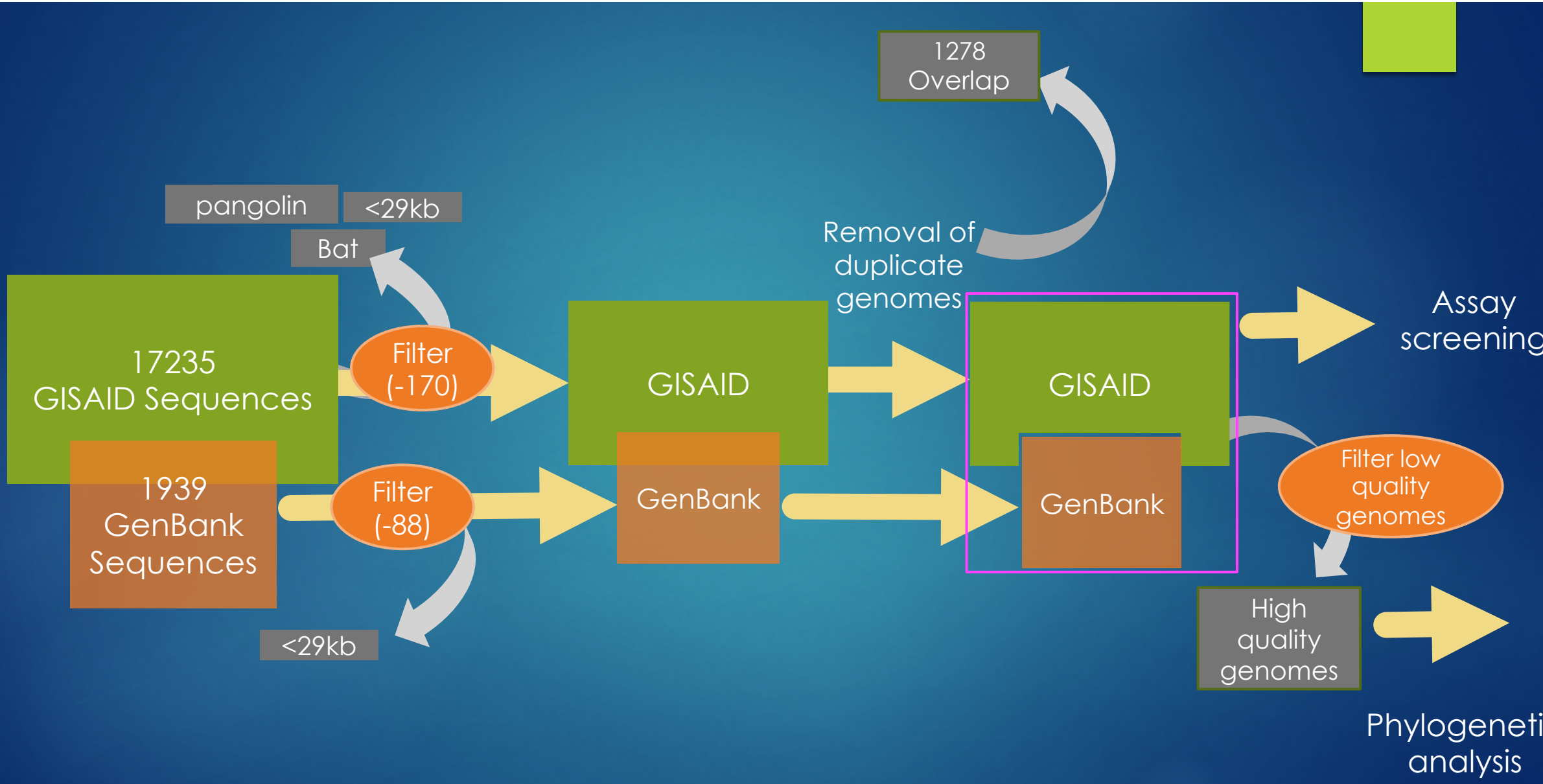
Assay Validation Method

- Our team has developed a web-based application to monitor existing PCR assays that are currently in use around the world
  - Assays from US CDC, China CDC, Charité, HKU, Japan NIID, others
  - SARS-CoV-2 genomes downloaded from GenBank and GISAID daily
- ThermonucleotideBLAST: assesses likelihood of detection success using known thermodynamic parameters (delta G and Tm)
- Visualization
  - Phylogenetic tree created with PhaME and rendered with PhyD3
  - Heatmap of thermodynamic mismatches
  - Table of mismatches and recall
  - Top ranked assays by recall

Workflow

- Genome sequences downloaded daily from GenBank and GISAID
  - GenBank: 1624, GISAID: 15498 (as of 2020-04-30 08:36)
  - GenBank: 1939, GISAID: 17235 (as of May 7th 3PM)
- Filter out bat, pangolin, and sequences shorter than 29 kb
  - GenBank: 86 (+2) removed. GISAID: 170 removed
- Remove overlapping sequences
  - Overlap: 1227 (+51). Final database total: 15639 (+1999)
- Each assay validated against every sequence, producing thermodynamic mismatches and recall value (True positive rate, sensitivity)
  - $Recall = \frac{True\ Positives}{True\ Positives + False\ Negatives}$
- Additional low quality genome removal for phylogenetic analysis
- Render visualizations

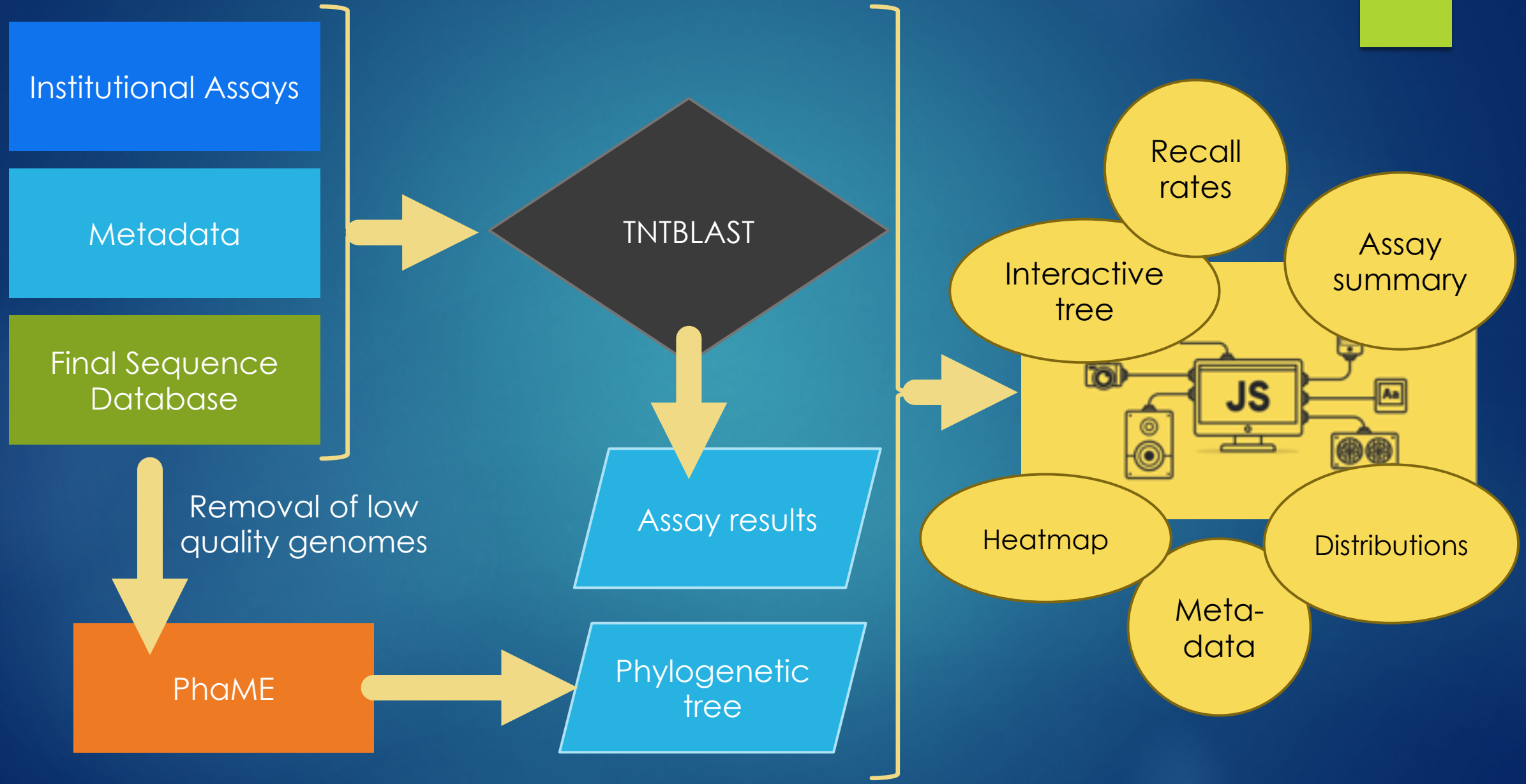
COVID-19 Genome Analytics



EDGE bioinformatics

COVID-19 Genome Analytics

Los Alamos NATIONAL LABORATORY EST. 1943

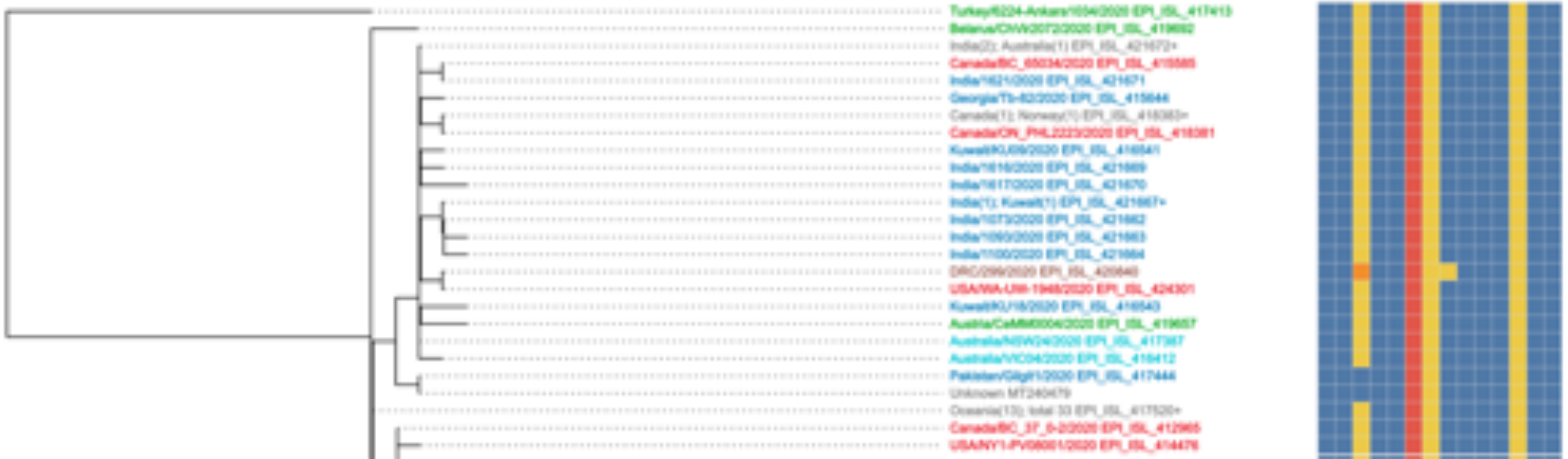


Assay Validation Dashboard

- Top ranked assays by recall
  - Institut Pasteur: Two assays against RdRP (nCoV_IP2, nCoV_IP4)
  - Charité, Berlin: Two RdRP, one Sarbeco screening assay against Envelope
  - Hong Kong University: Assay targeting ORF1b

Tree/heatmap view of assay results

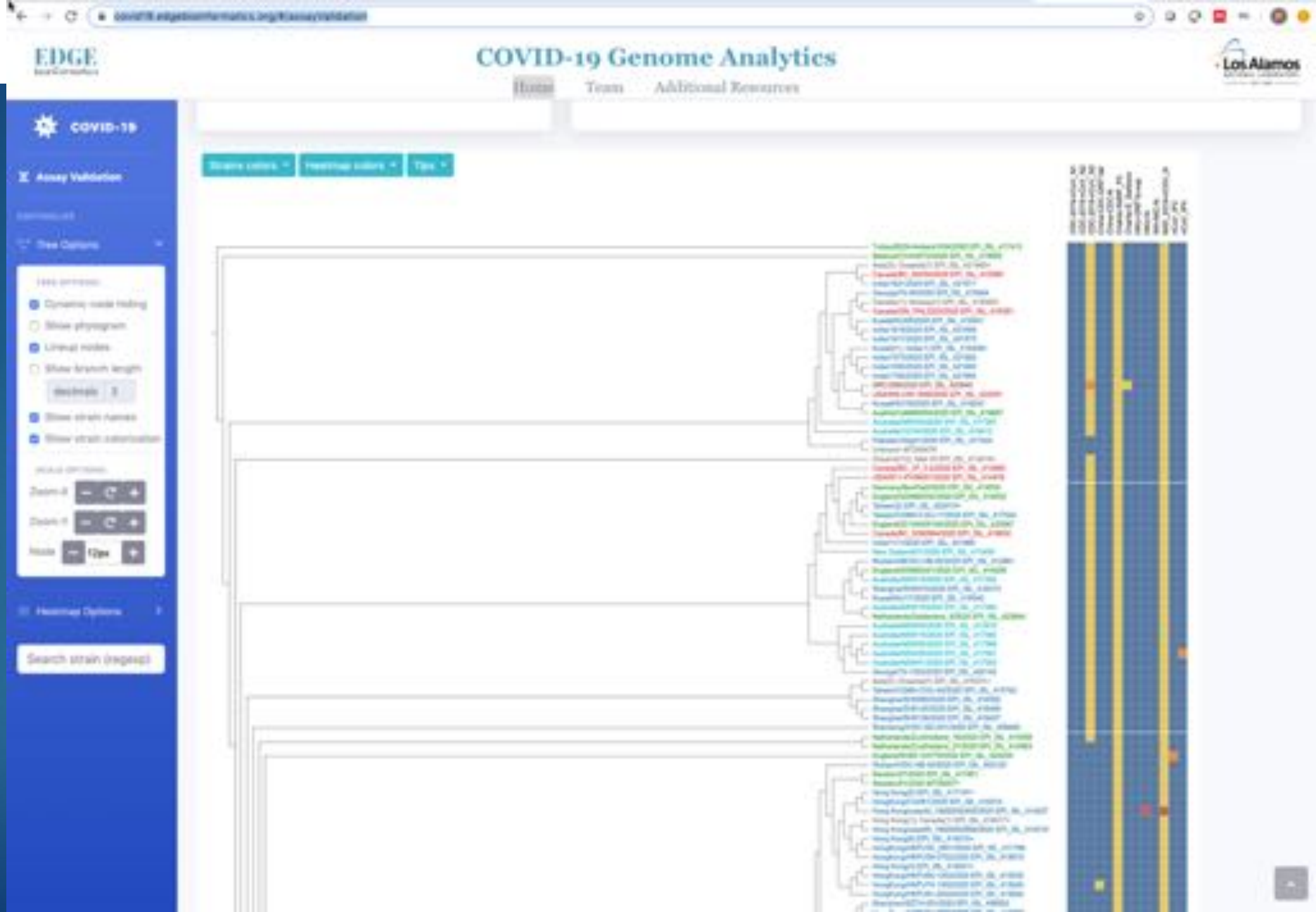
- Heatmap with tree can show evolutionary patterns of mismatches
  - Charité: probe with two mms (P1), reverse primer with one mm (P2)
    - P2 designed originally for SARS and bat-SARS coronaviruses
  - USA CDC: Mismatch in forward primer of N3



EDGE bioinformatics

Known issues with CDC N3 assay

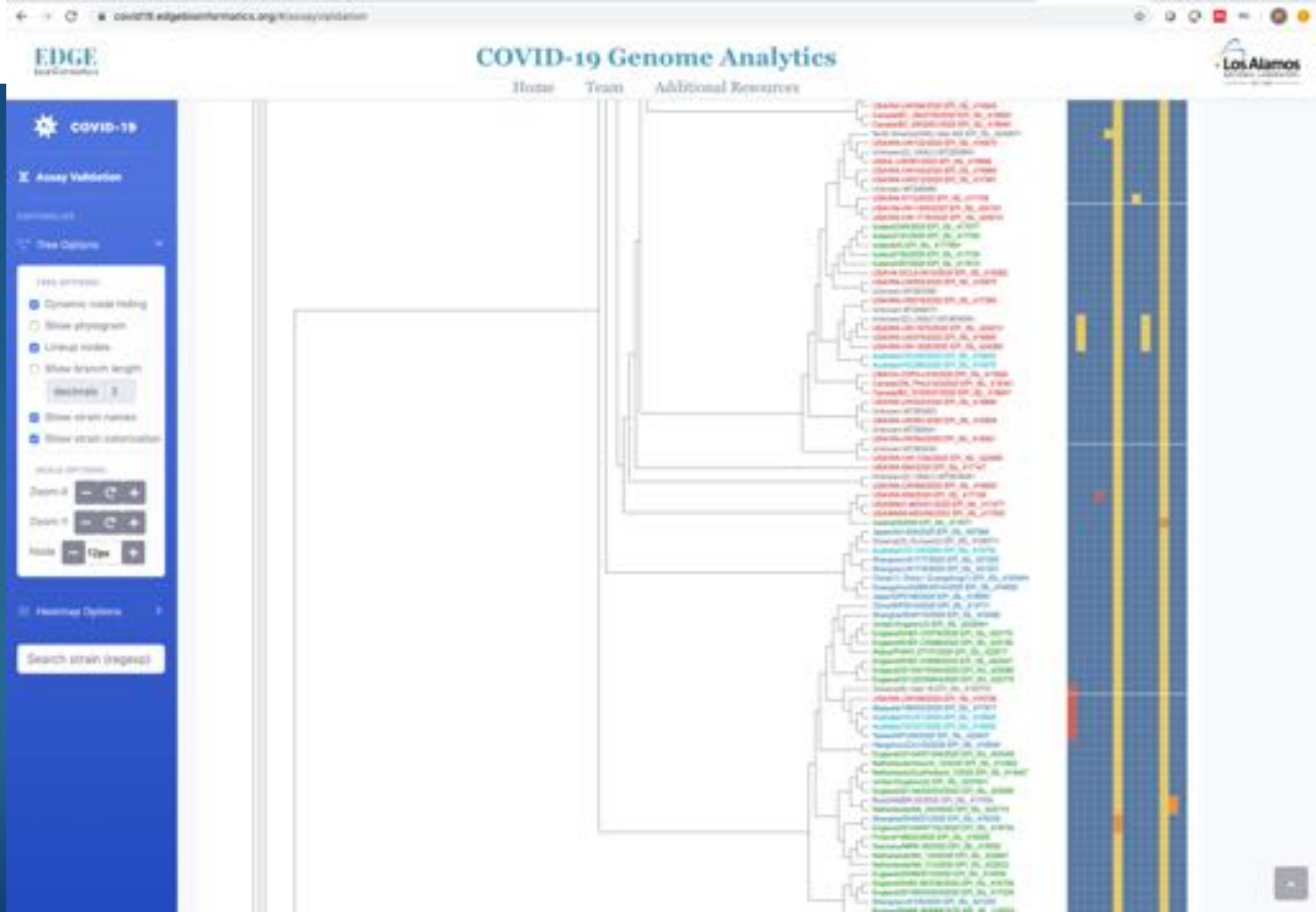
Several clades show potential issues



EDGE bioinformatics

Potential issues for some newer clades with CDC N2 and N1 assays

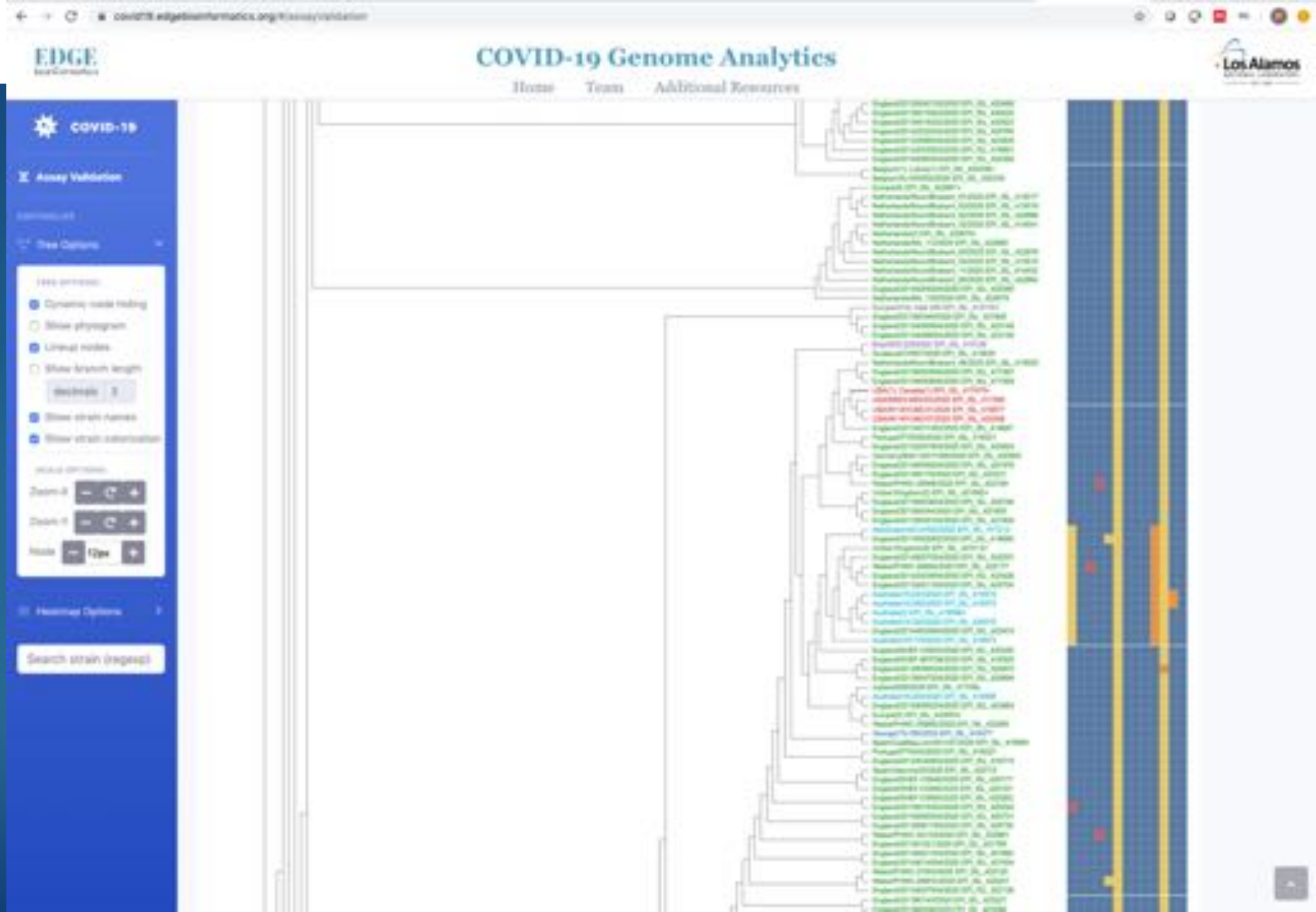
Including some US isolates



EDGE bioinformatics

An additional clade shows potential issues with CDC N1

Includes strains from Australia/ England/ New Zealand



Remaining issues

- Display/interaction/presentation
- Genome quality assessment
  - *identified several strains already with errors