
Magnipore

Differential single nucleotide changes of Oxford Nanopore Technologies sequencing signals for detection of RNA modification positions at the example of SARS-CoV-2

Compares two related ONT samples on a signal level

Jannes Spangenberg
Dr. Christian Höner zu Siederdisen
Dr. med. Andi Krumbholz
Dr. Manja Marz

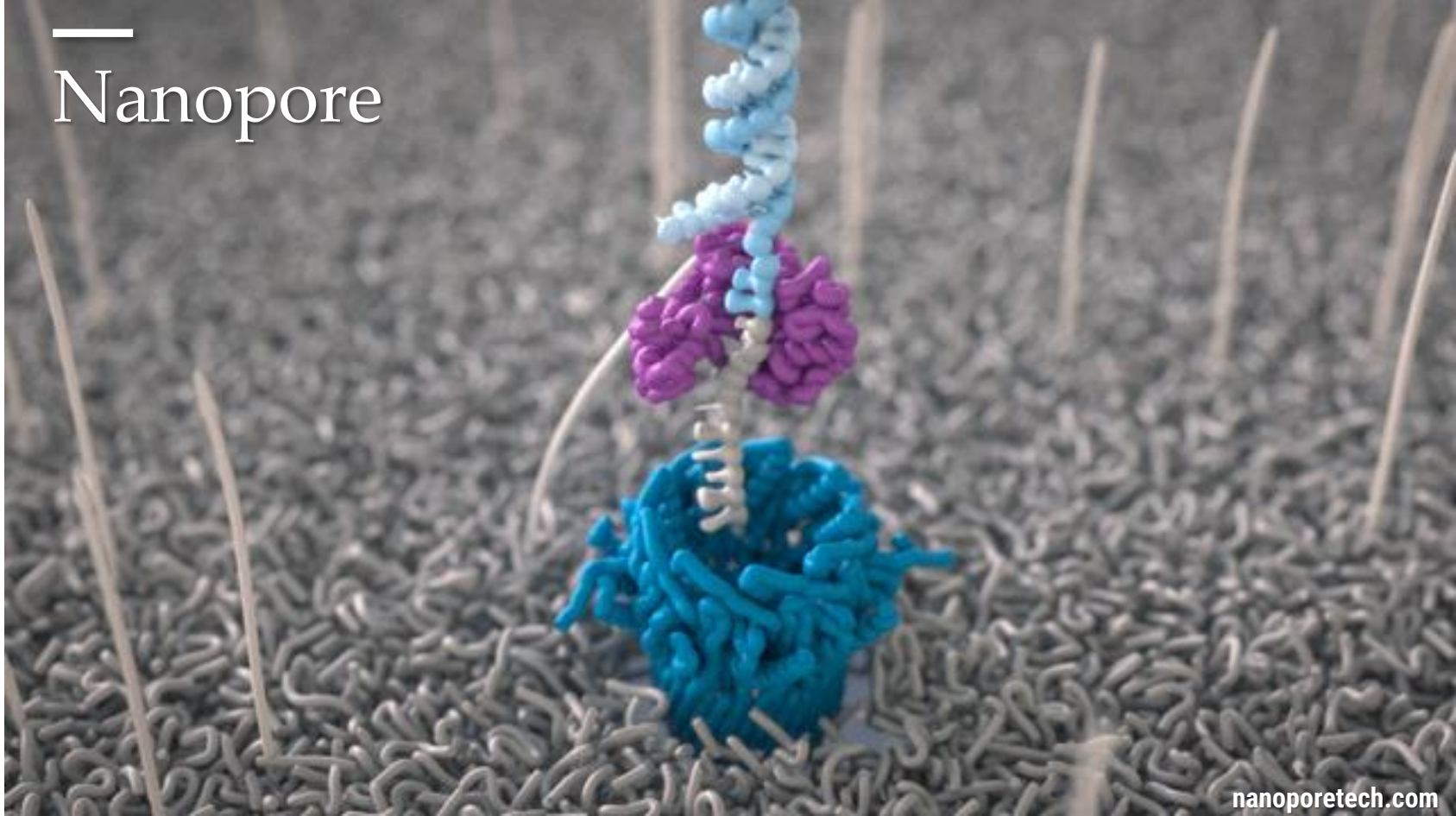
38th TBI Winterseminar Bled 2023

— Outline

- Introduction into Oxford Nanopore Technologies sequencing
- Magnipore pipeline

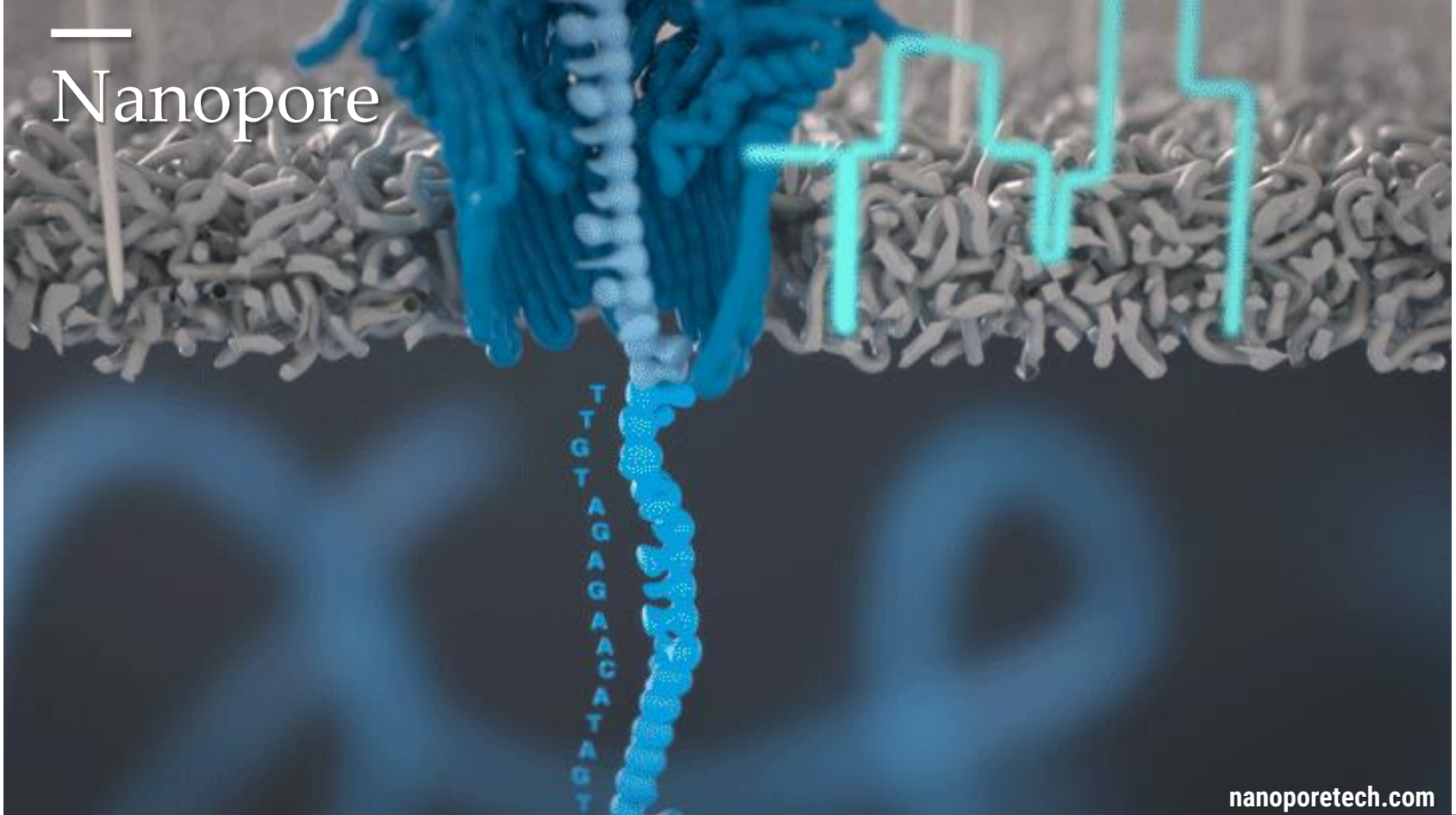


Nanopore



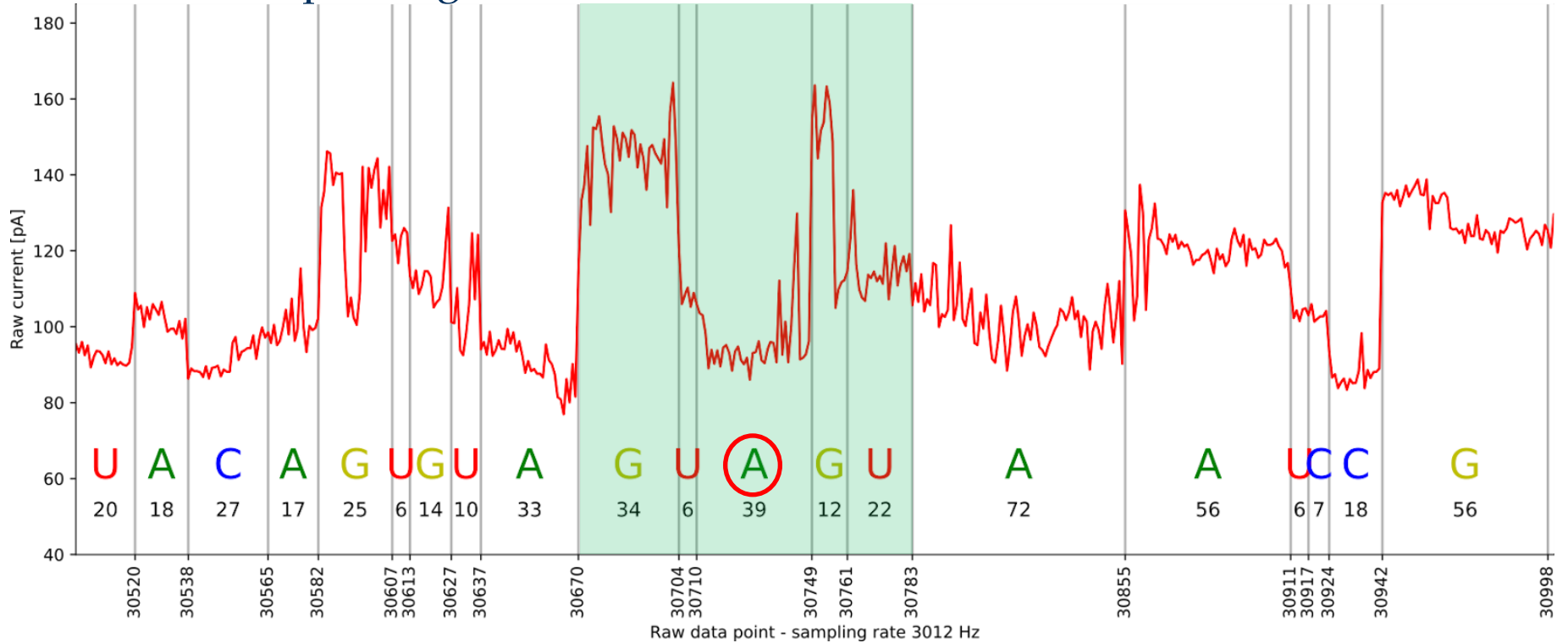
nanoporetech.com

Nanopore

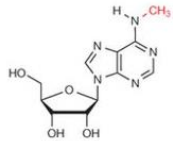
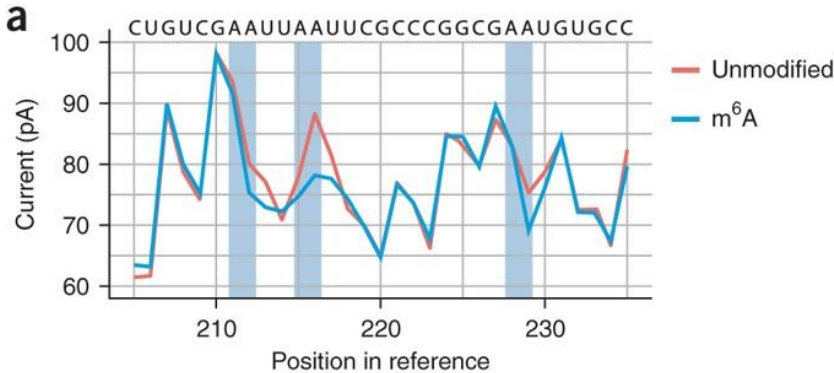


nanoporetech.com

Nanopore signal



Magnipore



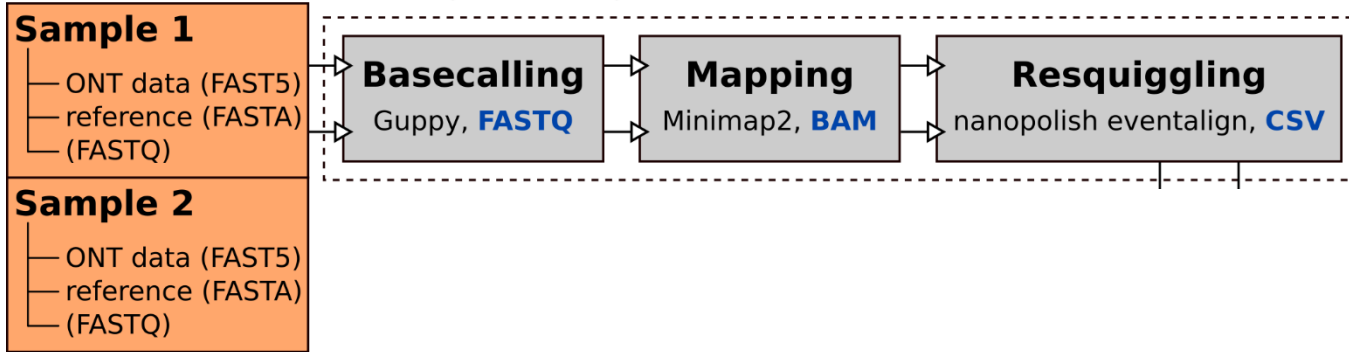
N⁶-methyladenosine (m⁶A)

Garalde, D., Snell, E., Jachimowicz, D. *et al.* Highly parallel direct RNA sequencing on an array of nanopores. *Nat Methods* 15, 201–206 (2018)

- Aim:
 - Find genomewide signal changes between two related samples (like virus variants)
- Task:
 - Samples should share alignable comparable regions
 - Compare normalised ONT signal, aggregated from multiple reads
 - Find significant signal changes between two samples
- Sources for signal changes
 - Bad signal segmentation
 - Molecular changes
 - Mutations
 - Modifications

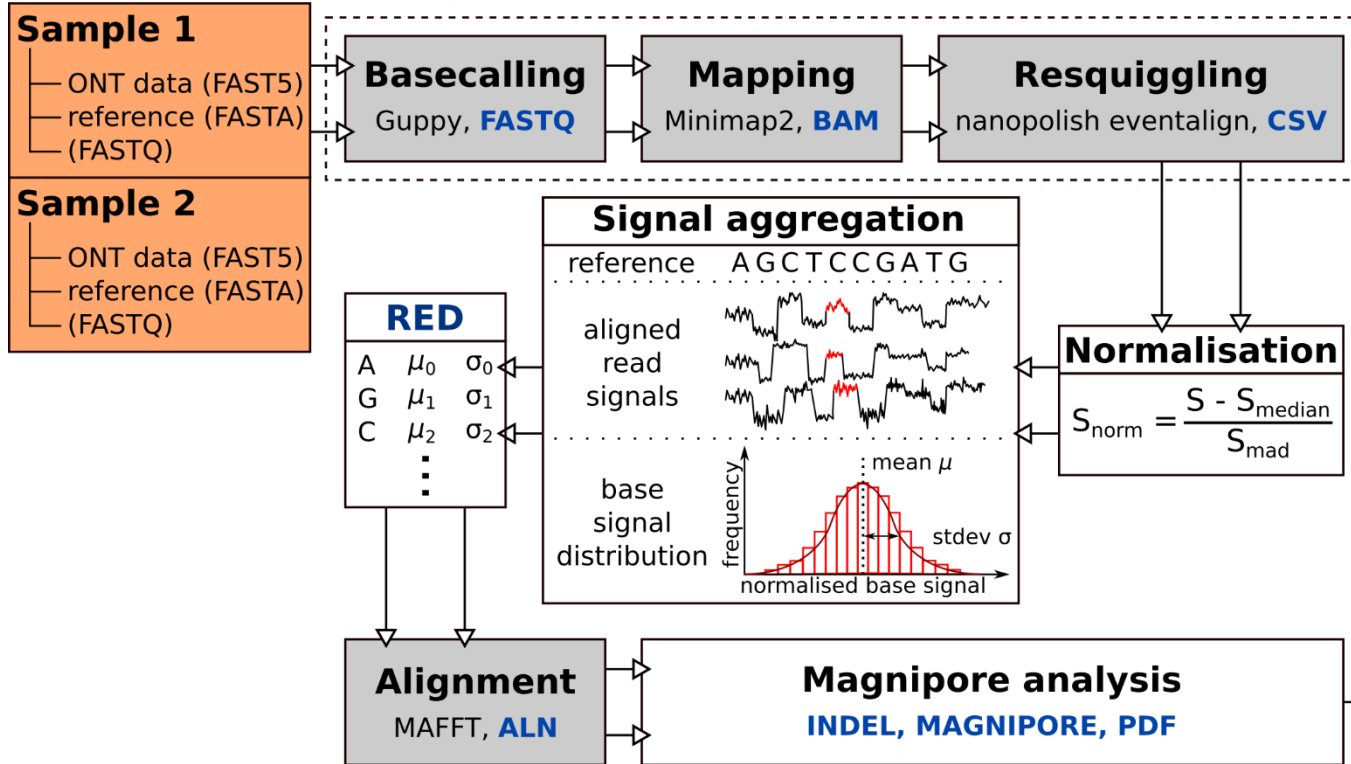
Input

Preprocessing signal segmentation & error correction



Input

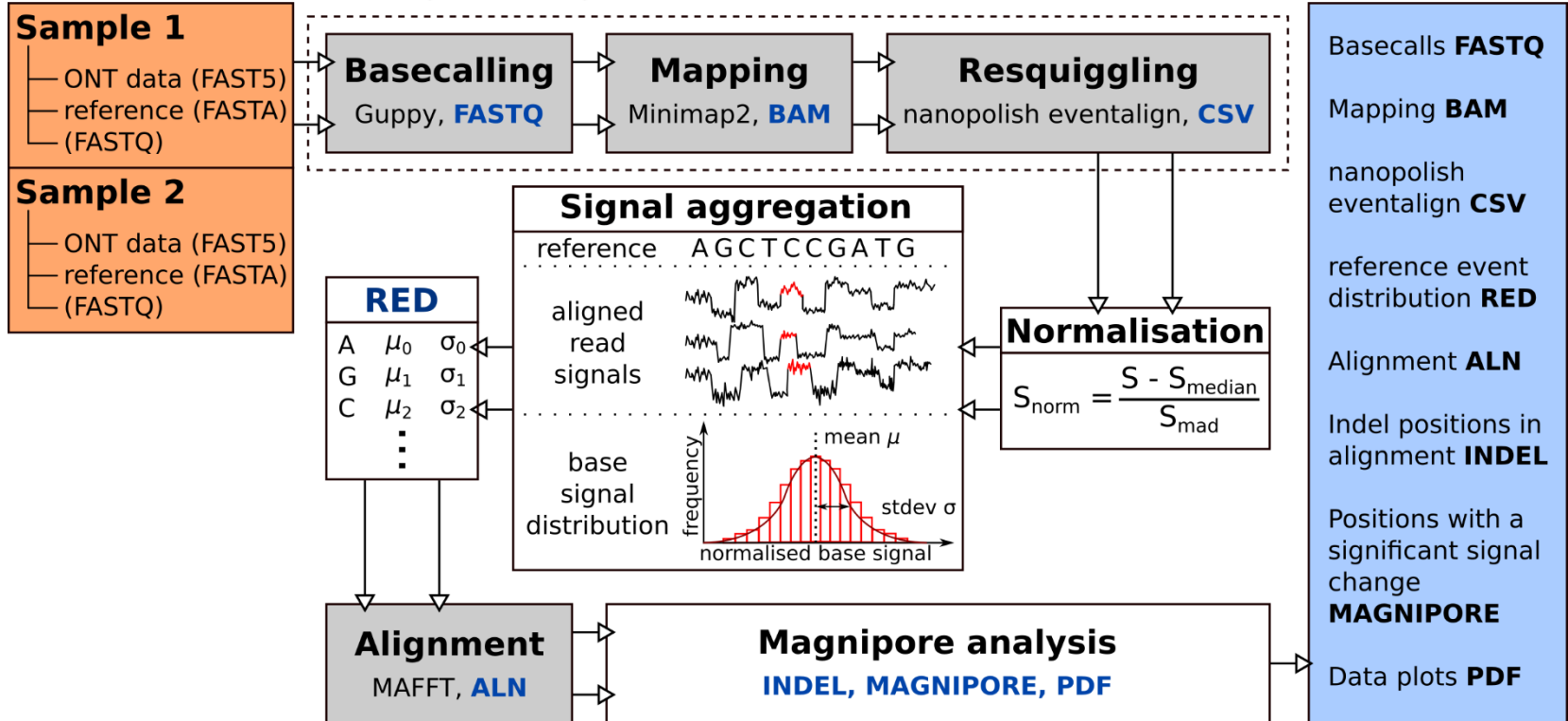
Preprocessing signal segmentation & error correction



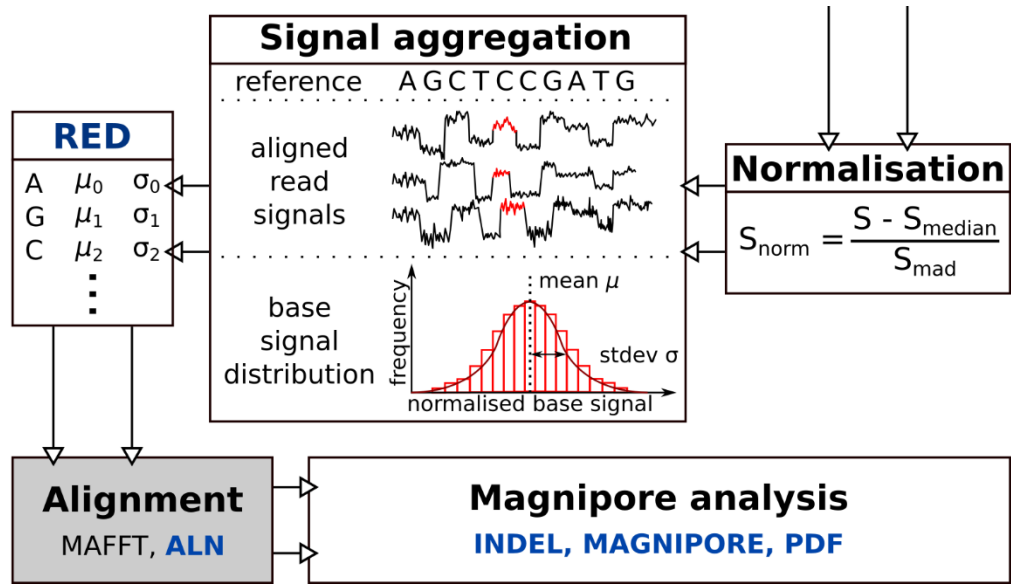
Input

Preprocessing signal segmentation & error correction

Output



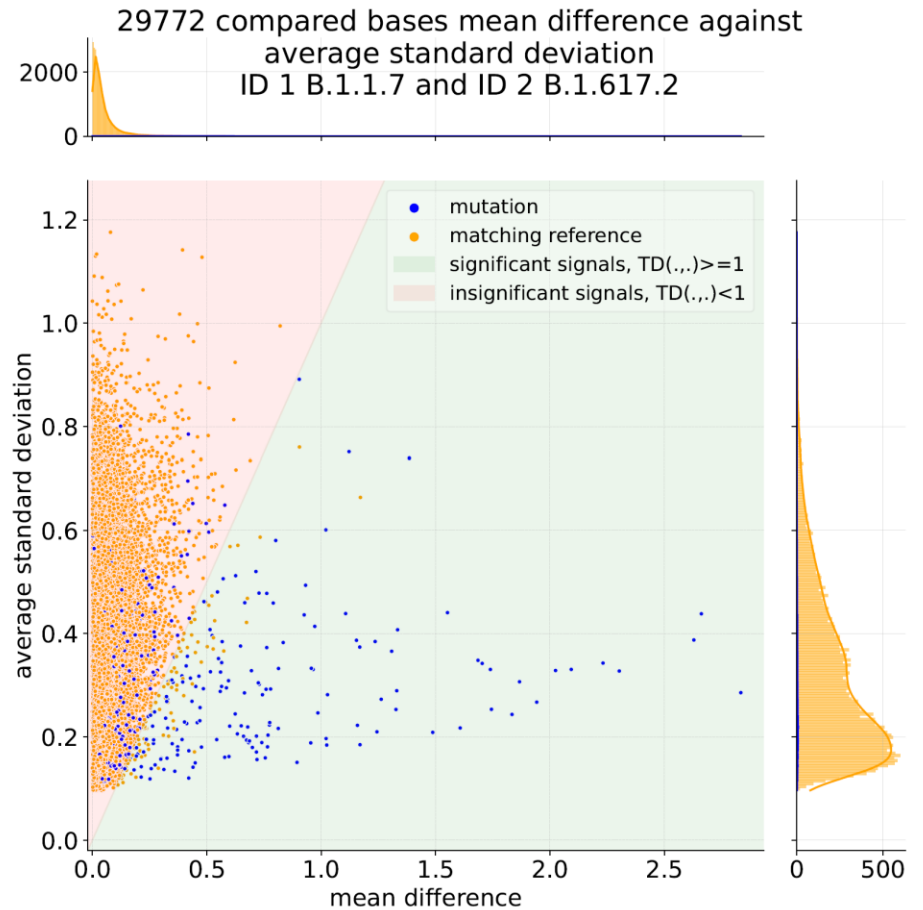
Magnipore pipeline



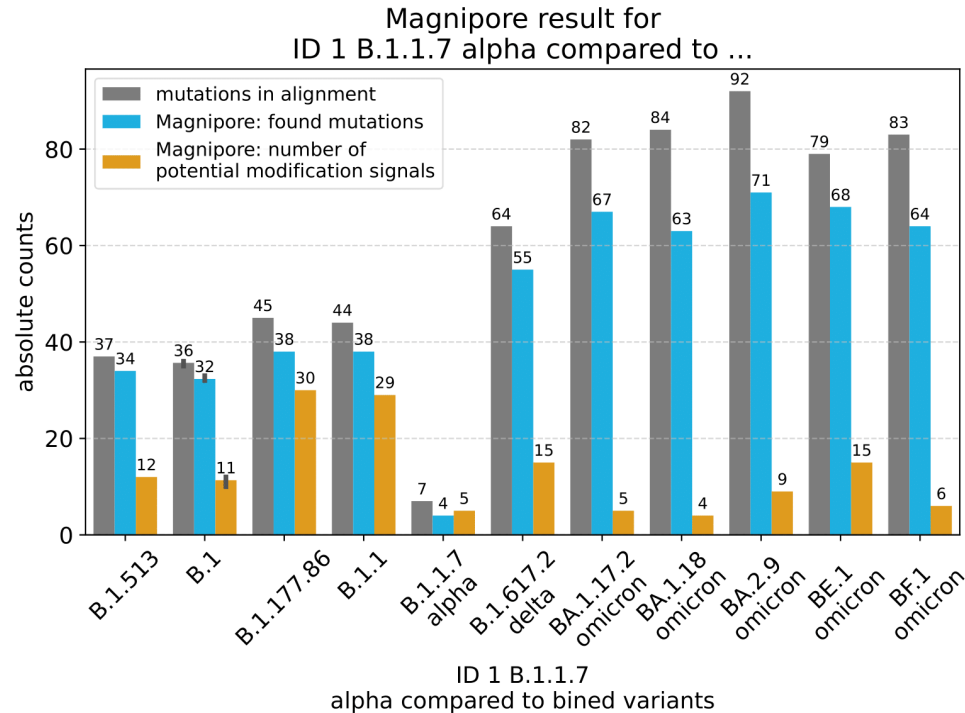
- Compare distributions positionwise for
 - aligned position i from sample 1
 - and position j from sample 2
 - Significance threshold:
 - $|\mu_{1,i} - \mu_{2,j}| > \frac{(\sigma_{1,i} + \sigma_{2,j})}{2}$
- Classify significant signal shifts
 - Mutation: reference mismatch around the significant position (± 3 bases)
 - Potential modification: references completely match around the significant position

Magnipore

- Mean difference: $|\mu_{1,i} - \mu_{2,j}|$
- Average stdev: $\frac{(\sigma_{1,i} + \sigma_{2,j})}{2}$
- Mutations show a strong shifts
- Potential modifications show smaller shifts compared to mutations



Looking at multiple comparisons

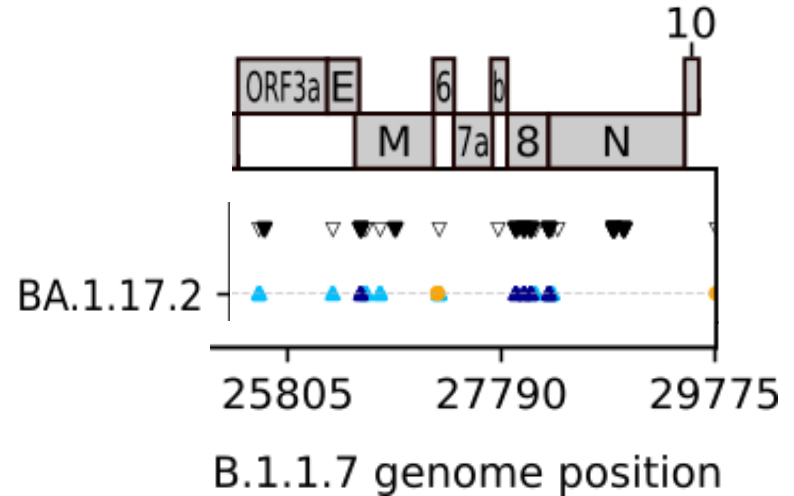


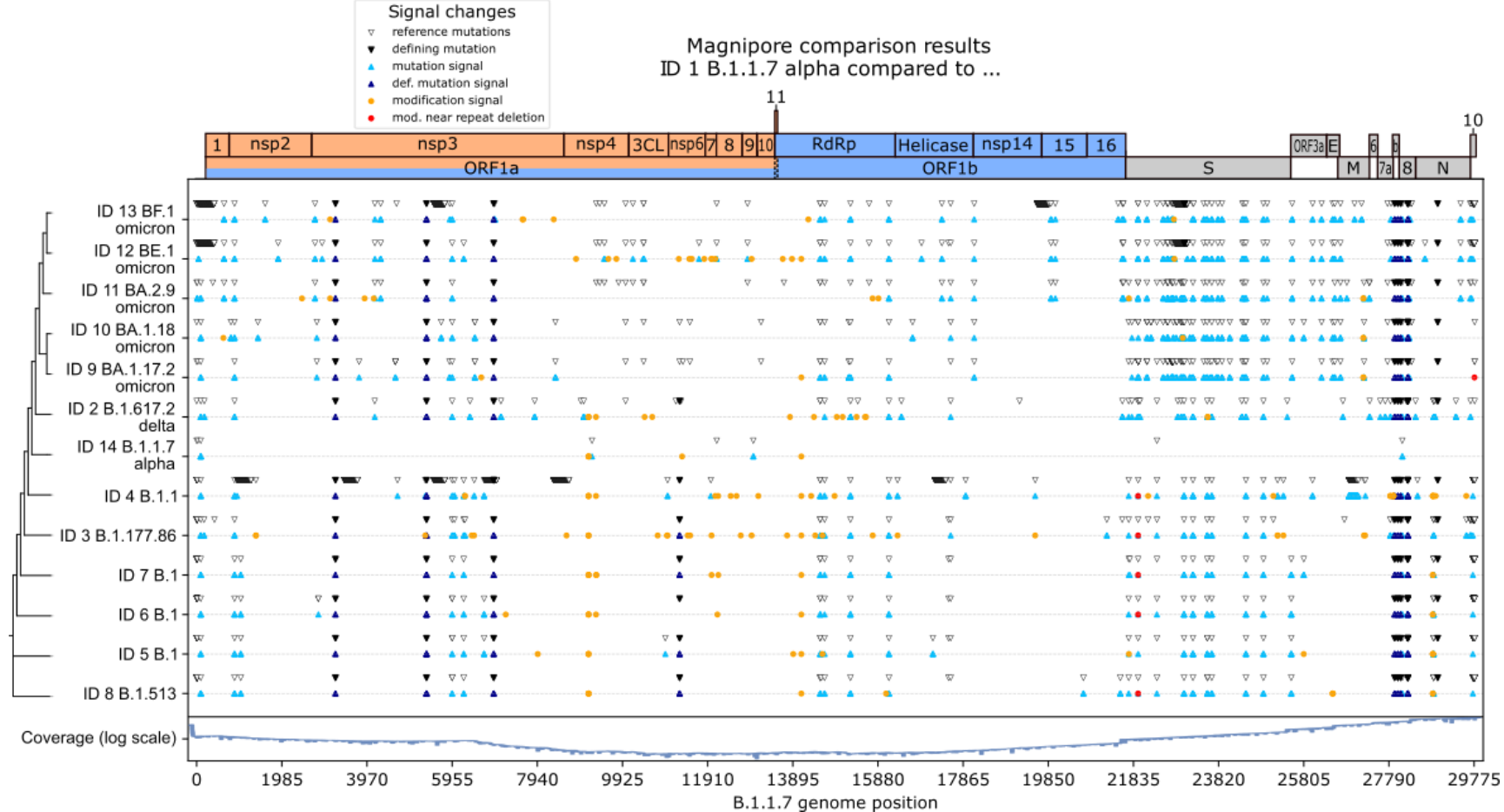
Looking at multiple comparisons

Magnipore comparison results
B.1.1.7 all c10 compared to ...

Signal changes

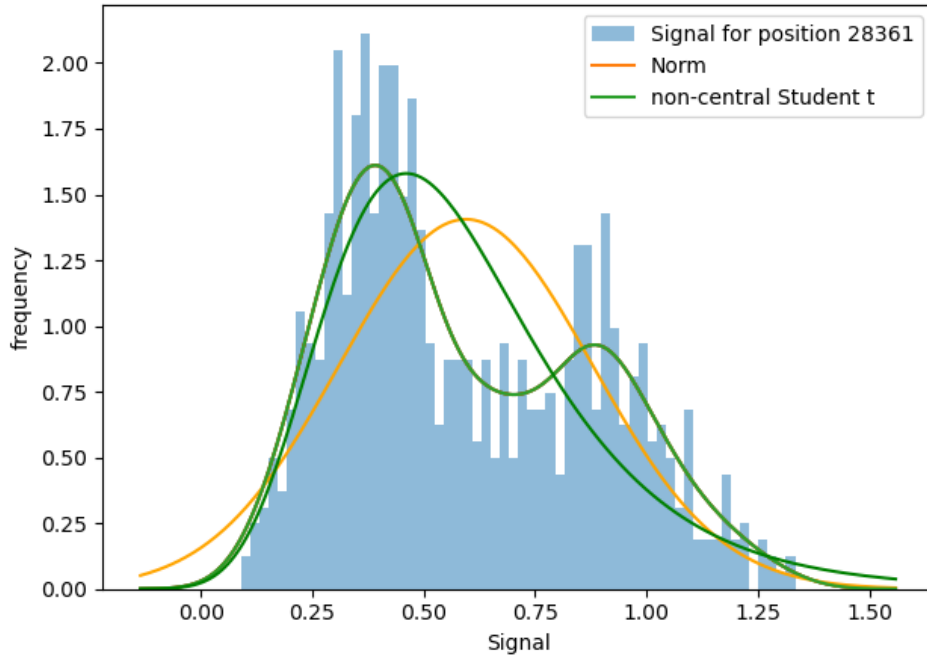
- ▽ reference mutations
- ▼ defining mutation
- ▲ mutation signal
- ▲ def. mutation signal
- modification signal
- mod. near repeat deletion



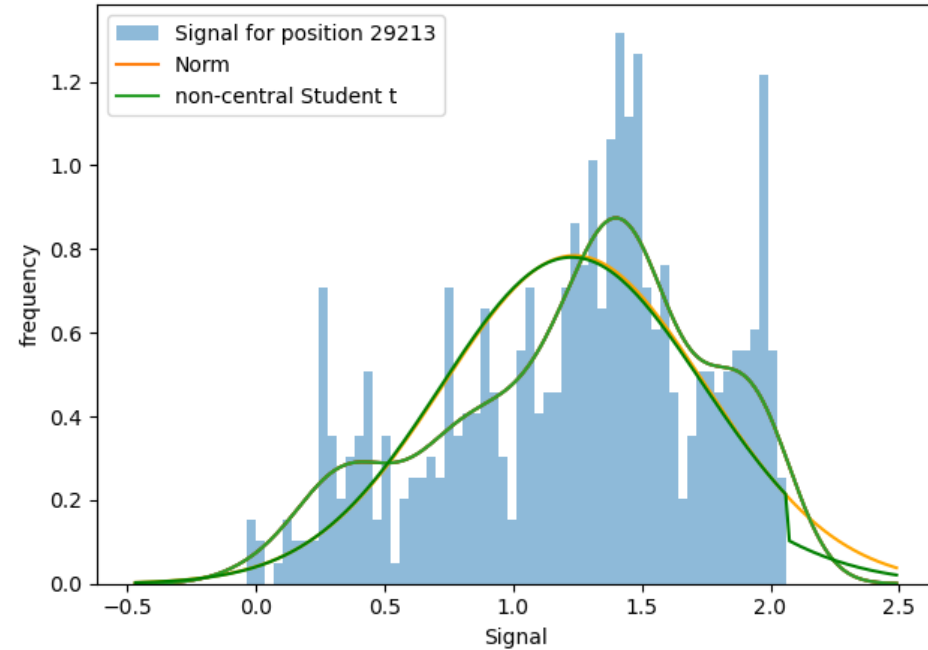


Base signal distribution not always normal

Sampled signal from 30 segments

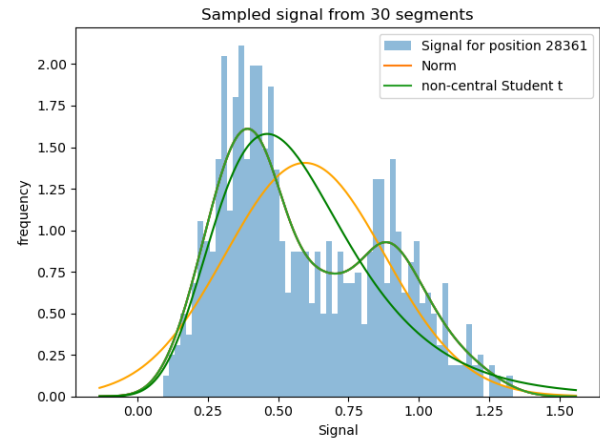
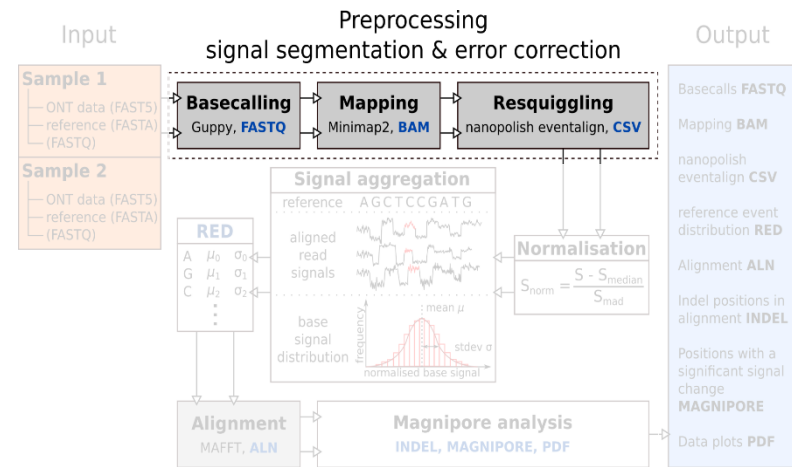


Sampled signal from 30 segments



Upcoming problems to solve

- Resquigling cannot distinguish between mutation and basecalling error
- Improve signal segmentation
- Find better distribution models (mixture)
 - Find differentiating signals within a single sample



Thanks to:

- Dr. Manja Marz
 - Dr. Christian Höner zu Siederdisen
 - Sebastian Krautwurst
 - Dr. Andi Krumbholz
-
- Wetlab:
 - Akash Srivastava
 - Milena Žarković
- and you!

Website: www.rna.uni-jena.de

Thanks for your attention!

Funded by
DFG Deutsche
Forschungsgemeinschaft
German Research Foundation



Thüringer Zentrum für
Lernende Systeme und Robotik



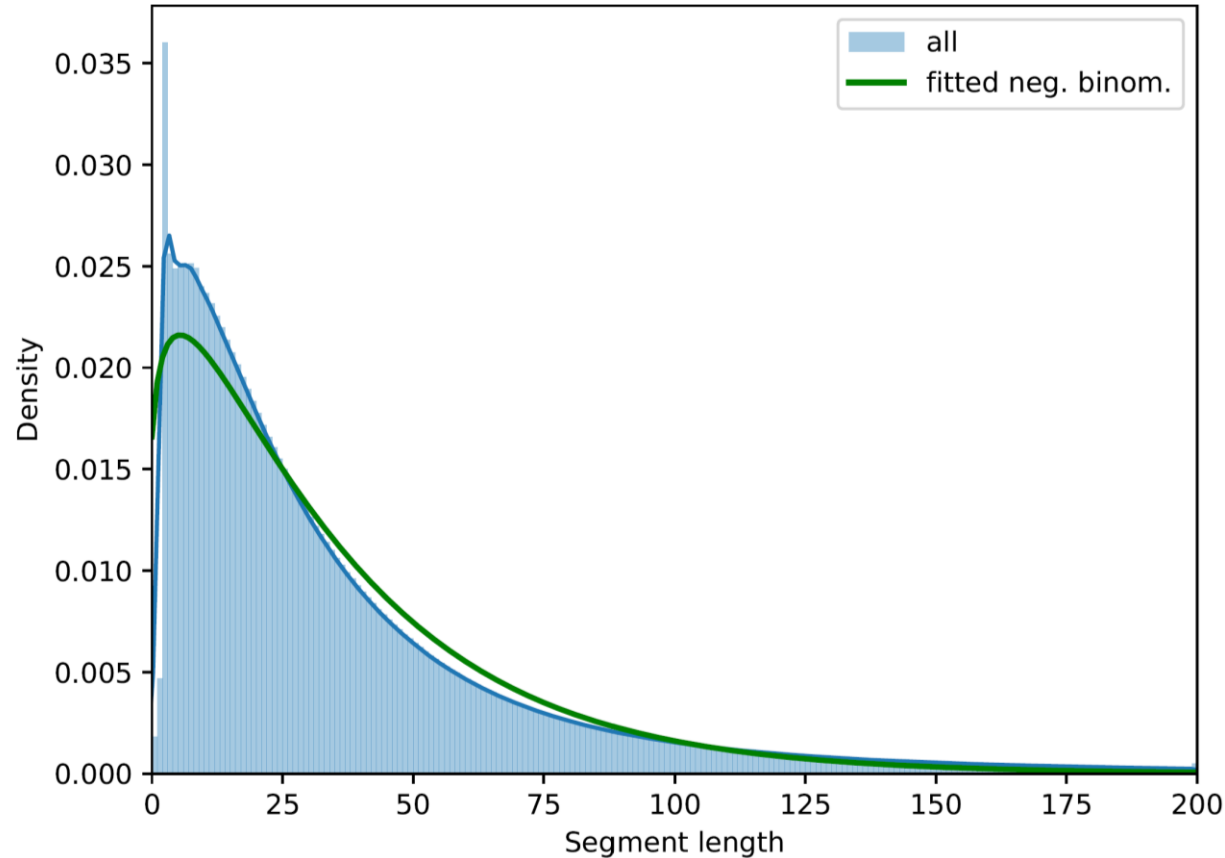
Funding (in part):

- German DFG Collaborative Research Centre AquaDiva (CRC 1076 AquaDiva)
- The German state of Thuringia via the Thüringer Aufbaubank (2021 FGI 0009)
- Carl-Zeiss-Stiftung within the program Scientific Breakthroughs in Artificial Intelligence (project “Interactive Inference”)
- Ministry for Economics, Sciences and Digital Society of Thuringia (Landesprogramm ProDigital, DigLeben-5575/10-9)

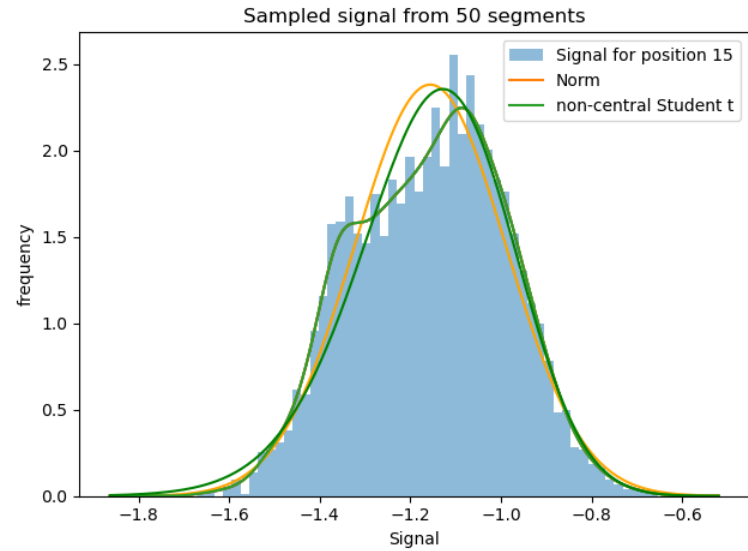
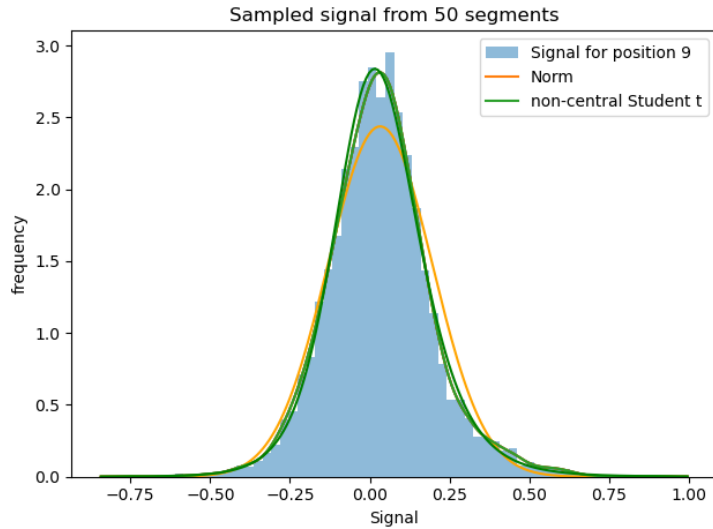
Supported by Oxford Nanopore Technologies

Signal segmentation

- IVT RNA data (not modified)
- 5-mer segment lengths tend to follow a negative binomial distribution?



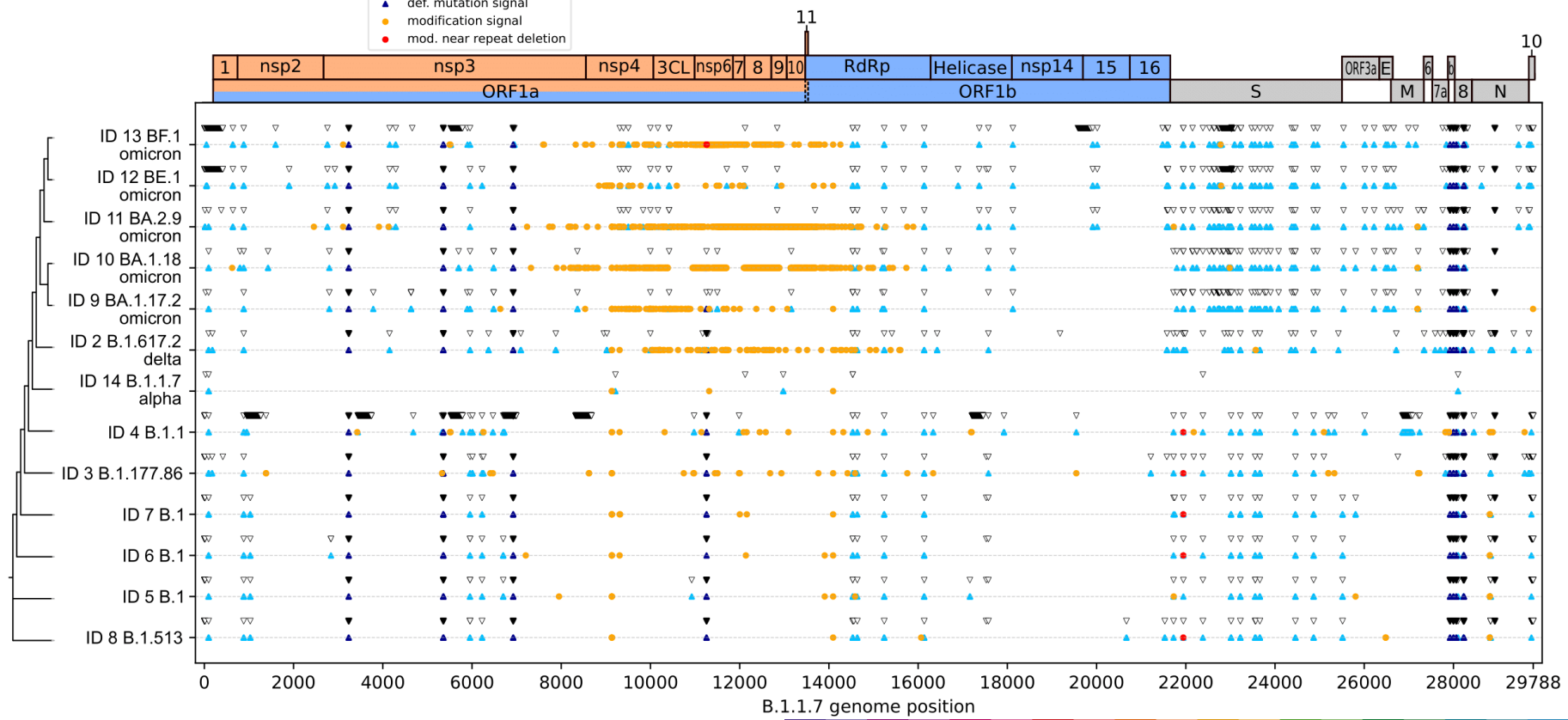
Base signal distribution normal?



Signal changes

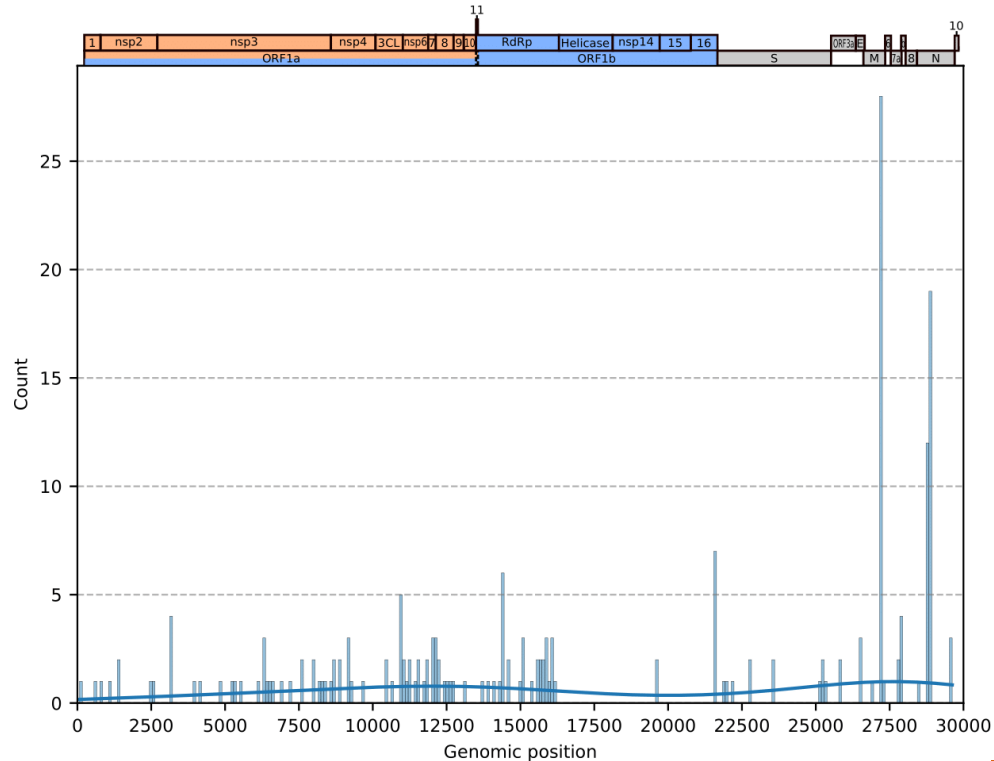
- ▽ reference mutations
- ▼ defining mutation
- ▲ mutation signal
- ▲ def. mutation signal
- modification signal
- mod. near repeat deletion

Magnipore comparison results ID 1 B.1.1.7 alpha compared to ...



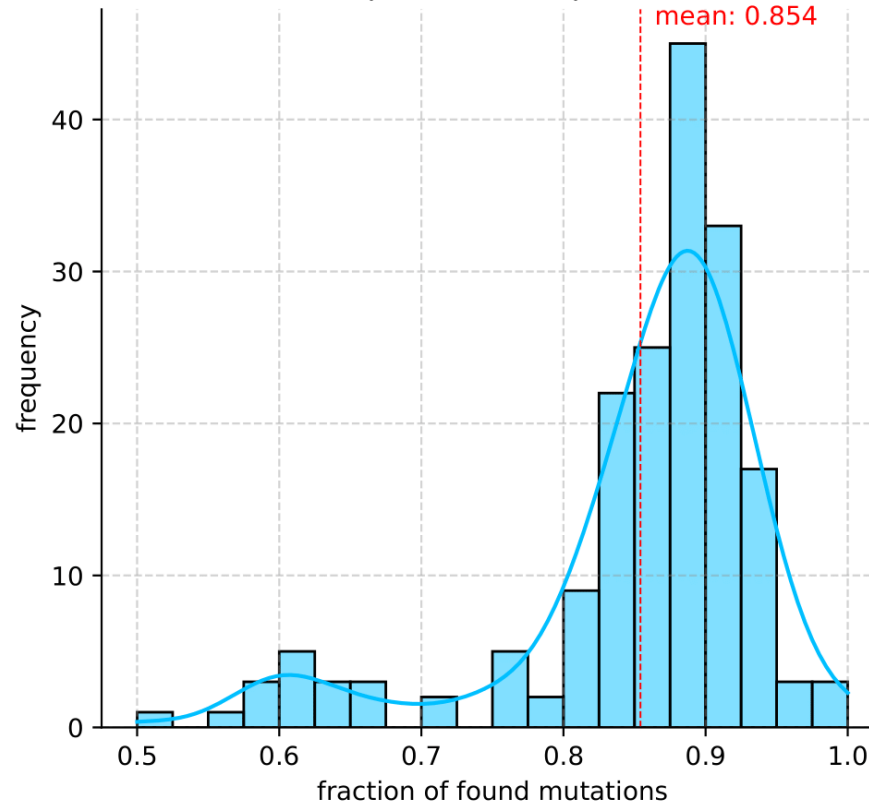
Where are differential signals with matching reference sequences?

Distribution of consistently reappearing potential modifications across SARS-CoV-2



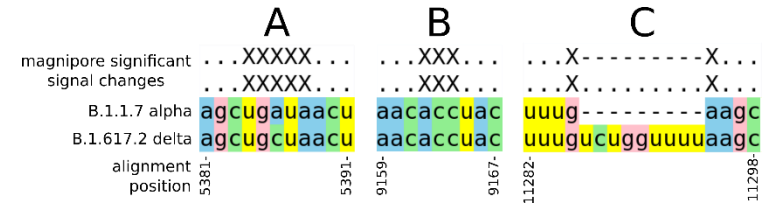
How many mutations are found by Magnipore

Distribution of fractions of found mutations in all pairwise comparisons

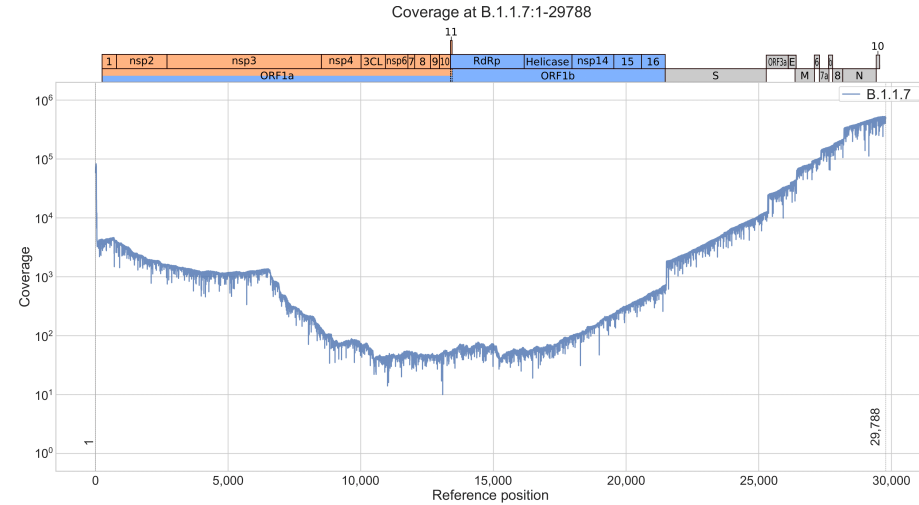
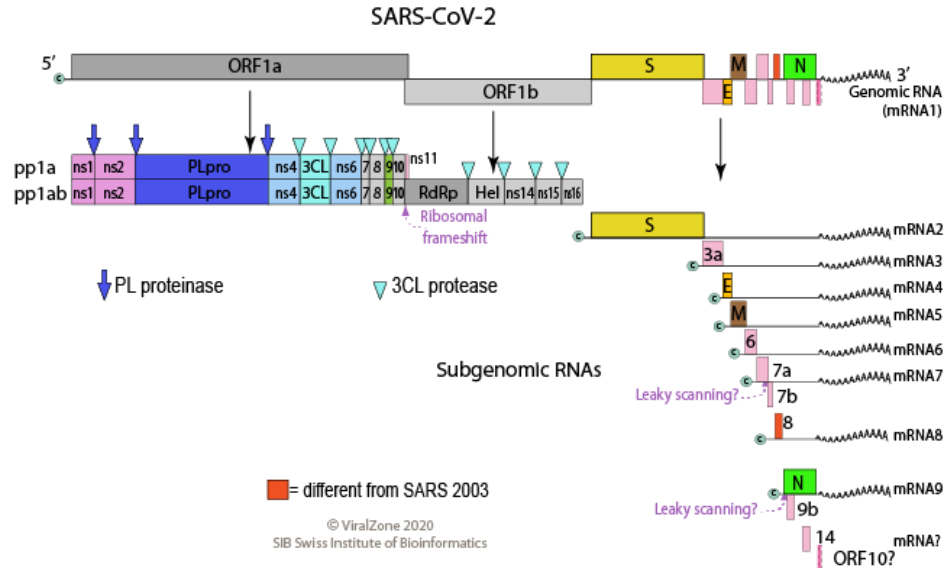


Further classification of significant differential signals

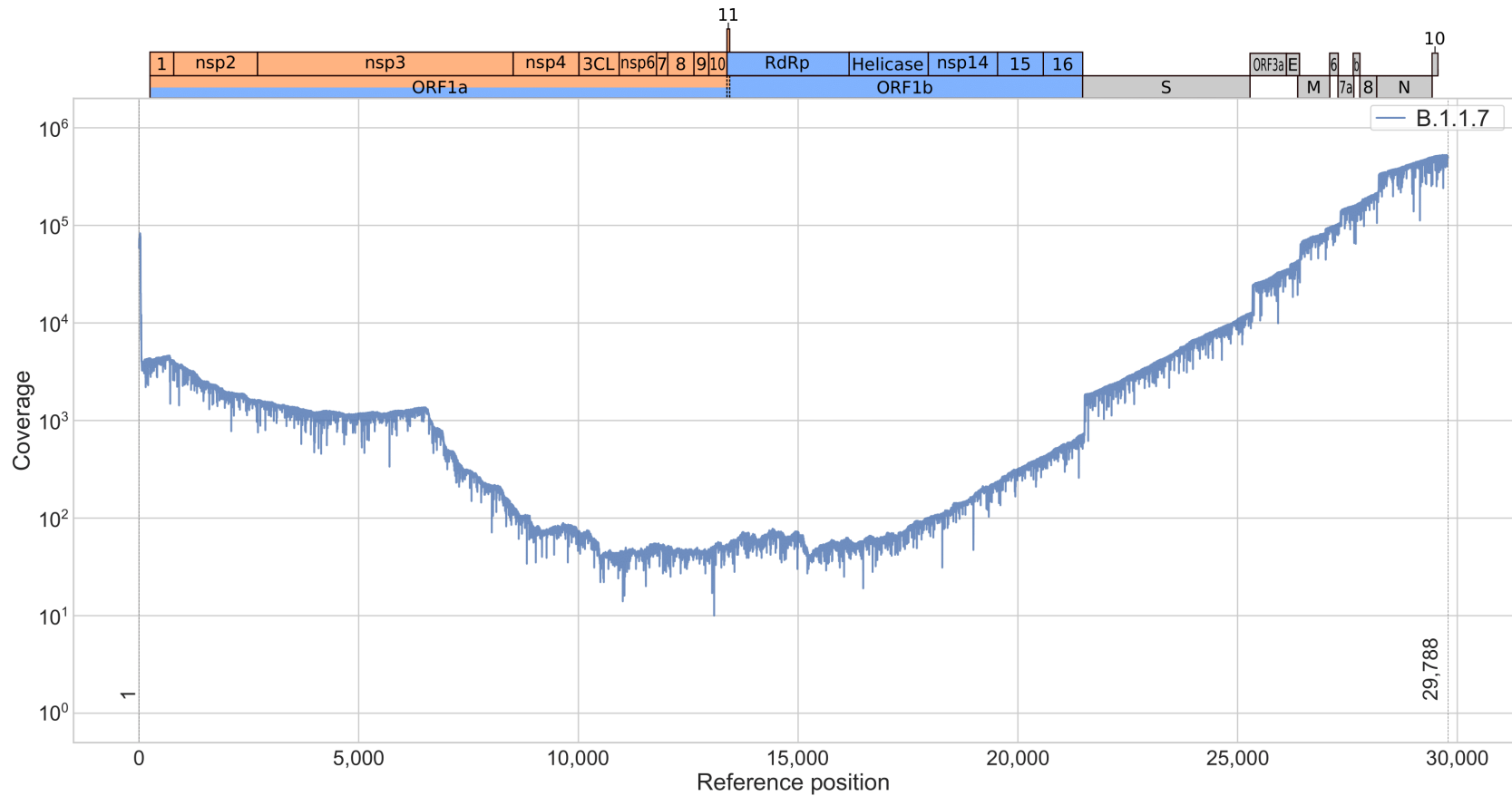
Reference alignment	Magnipore signal diff.	Designation	Symbol
mismatch	✓	base & signal mismatch mutation	▲
GGCCUA GGCACUA	X	undetected mutation or error in reference	should not occur
mismatch with gap	✓	indel mutation	▲
GGCACUA CUA GGCA---UUA	X	undetected mutation or error in reference	should not occur
match	✓	potential modification	●
GGCACUA GGCACUA	X	bases & signals match	standard
match with gap	✓	potential modification near repeat deletion	●
GGCACUA CUA GGCA---CUA	X	bases & signals match	standard



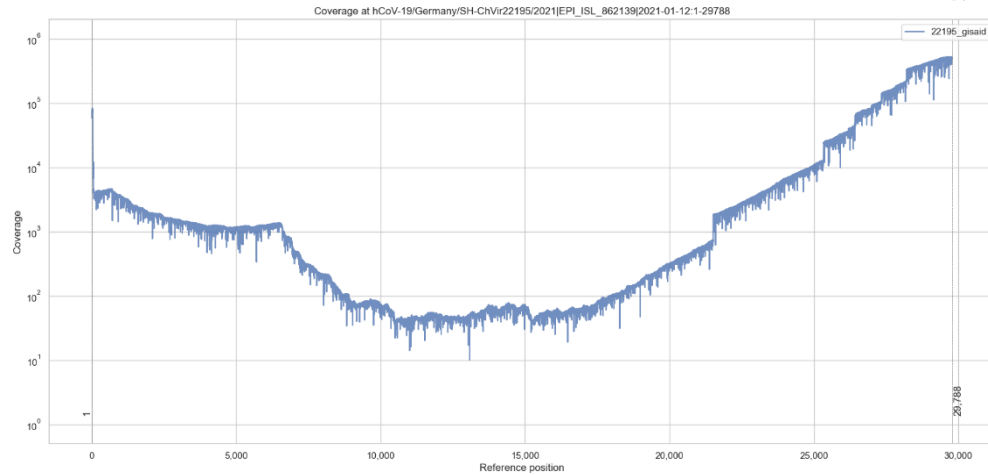
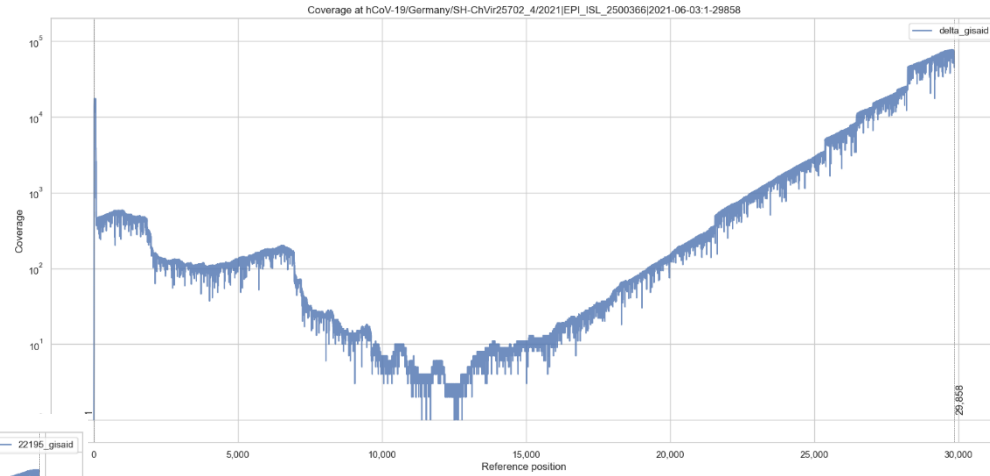
Low coverage in ORF1ab



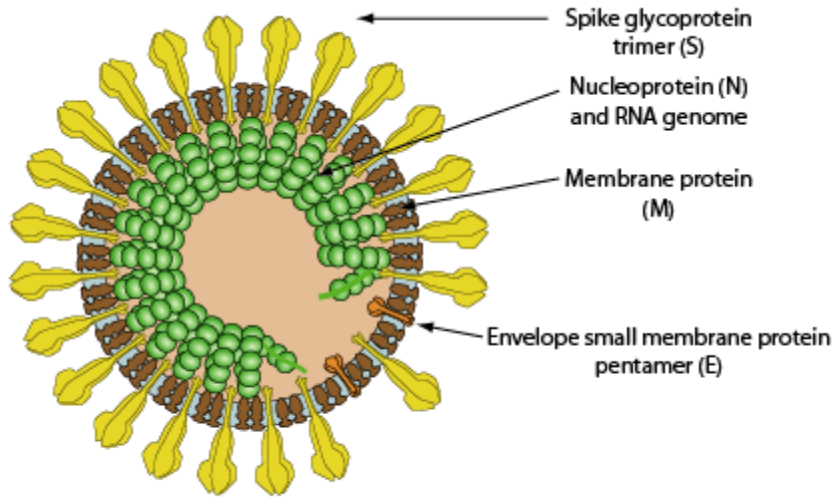
Coverage at B.1.1.7:1-29788



SARS-CoV-2 coverage



SARS-CoV-2



© ViralZone 2020
SIB Swiss Institute of Bioinformatics

