

T W I S T

BIOSCIENCE

Detecting known and novel mutations in a host's SARS-CoV-2 population by ultra-deep sequencing with unique molecular barcodes

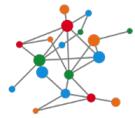
Presented by:

Prof. Dr. Stephan Ossowski

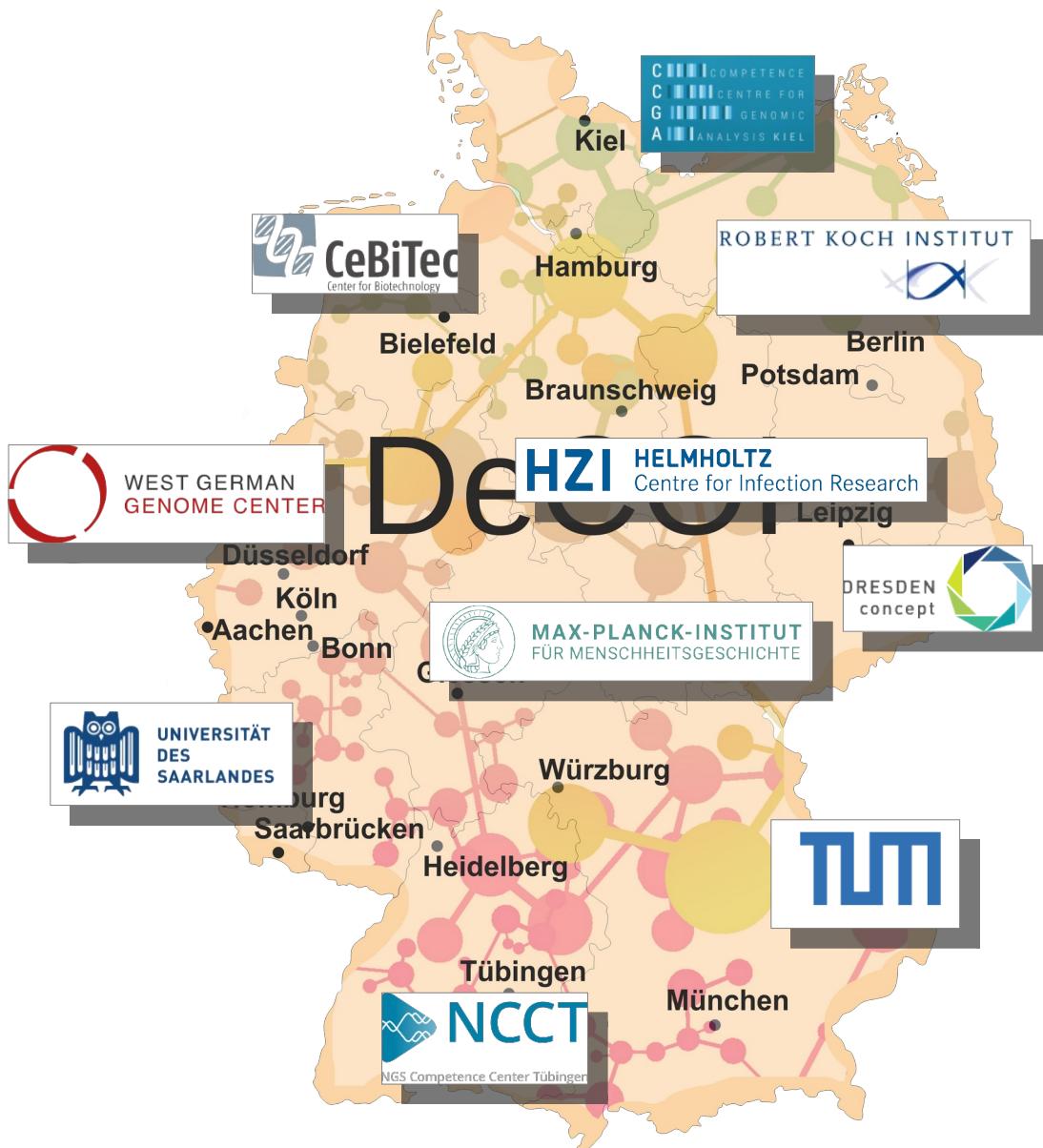
Institute for Medical Genetics and Applied Genomics at the University of Tübingen

DeCOI

German COVID-19 OMICS Initiative



- Viral Genome Sequencing Group



Decentralized / federated

- Since March 2020
- Heterogeneous sequencing landscape
- Various established protocols
- Exchange of protocols, best practices, QC materials and tools

Open

- Fast data sharing with community
- Genomes and metadata submission: GISAID & RKI DESH

Funding sources:

- DFG
- NUM
- Illumina
- Central government & states

Lead: Alexander Dilthey (Düsseldorf)

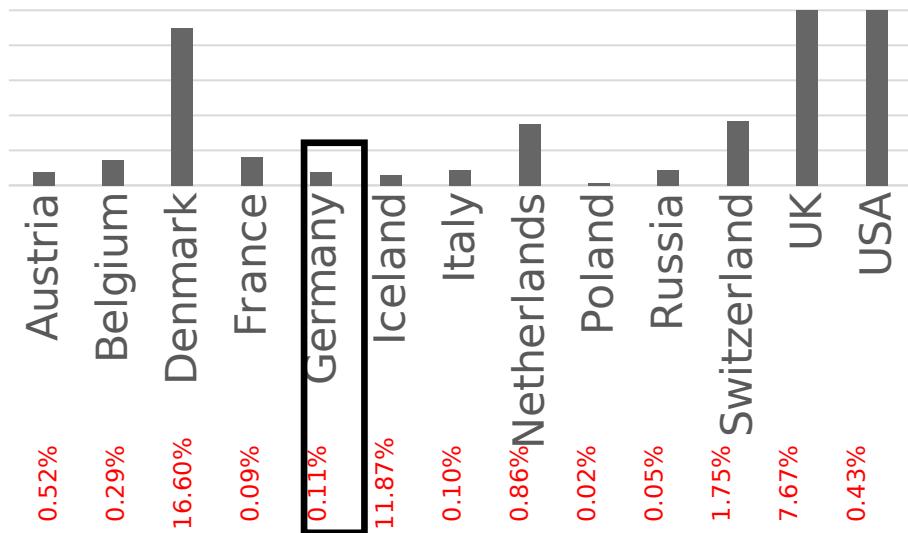
SARS-CoV-2 genome sequencing in Germany: end of last year

GISAID: 753 viral genomes from Germany – 473 from DeCOI members (14 Nov 2020)

DeCOI current numbers:

| | Total | On GISAID |
|------------|-------|-----------|
| Bielefeld | 135 | 50 |
| Dresden | 30 | 0 |
| Düsseldorf | 477 | 322 |
| Munich | 260 | 1 |
| RKI | NA | 23 |
| Saarbrücke | | |
| n | 65 | 25 |
| Tübingen | 76 | 52 |
| | 1043 | 473 |

GISAID genomes by country



Proportion cases sequenced (WHO/ GISAID)

SARS-CoV-2 genome sequencing in Germany: January 2021

Sequencing of viral genomes for 5-10% of positive cases:

HEALTHCARE & PHARMA JANUARY 14, 2021 / 1:39 PM / UPDATED 19 DAYS AGO

Germany plays catch-up in bid to monitor coronavirus mutations

By Ludwig Burger, Douglas Busvine

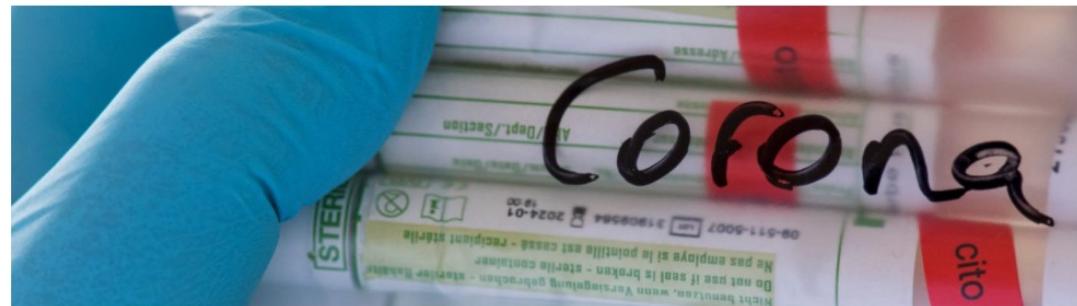
4 MIN READ



FRANKFURT/BERLIN (Reuters) - Germany wants to jump-start gene sequencing efforts to closely track coronavirus mutations and catch up with European nations such as Britain and Denmark which have taken the lead on decoding viral genomes.



SWR > SWR Aktuell > Baden-Württemberg



31,5 MILLIONEN EURO FÜR SEQUENZIERUNG

Baden-Württemberg untersucht alle positiven Corona-Tests auf Mutationen

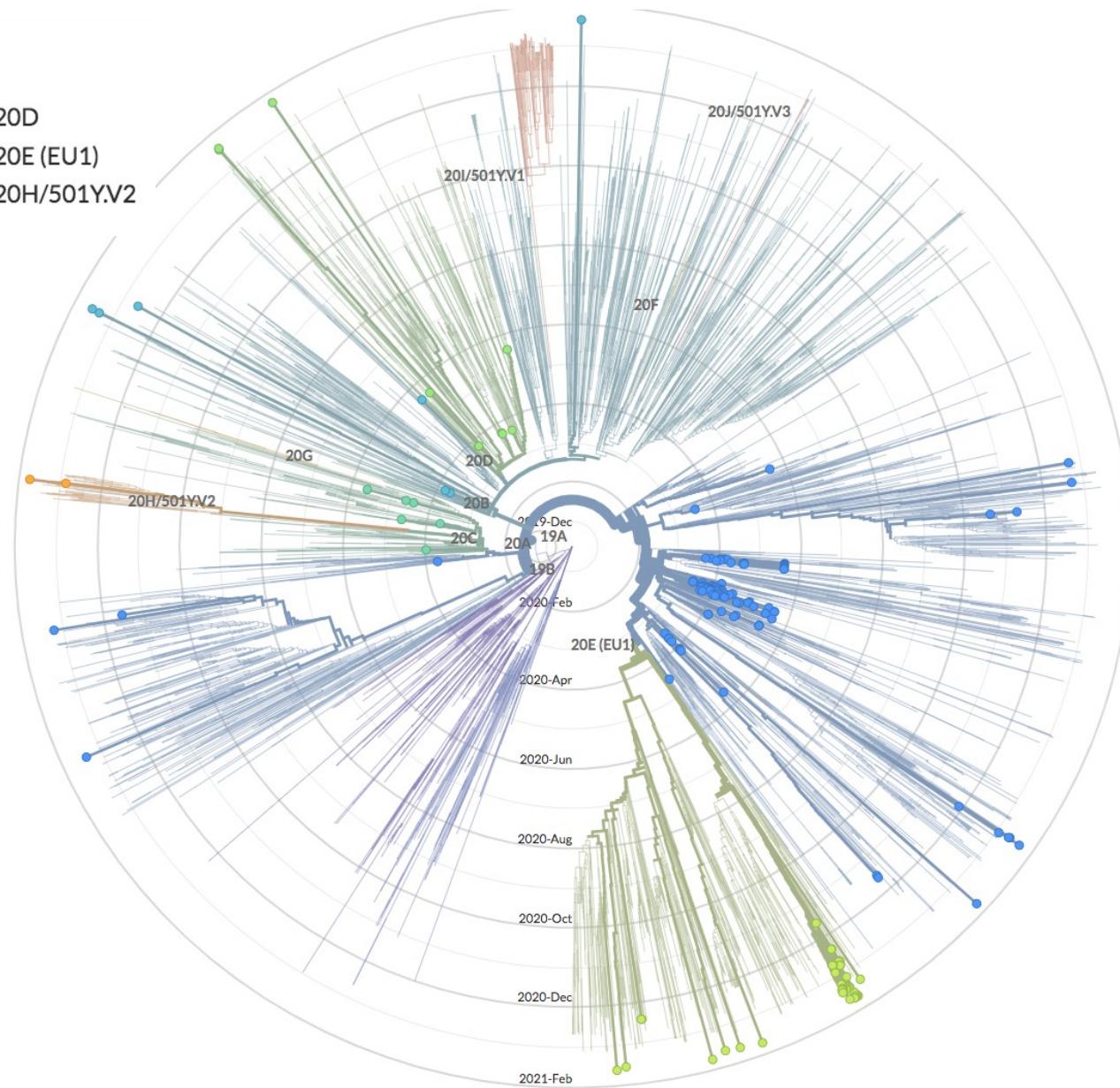
NextStrain: European Clades

Phylogeny

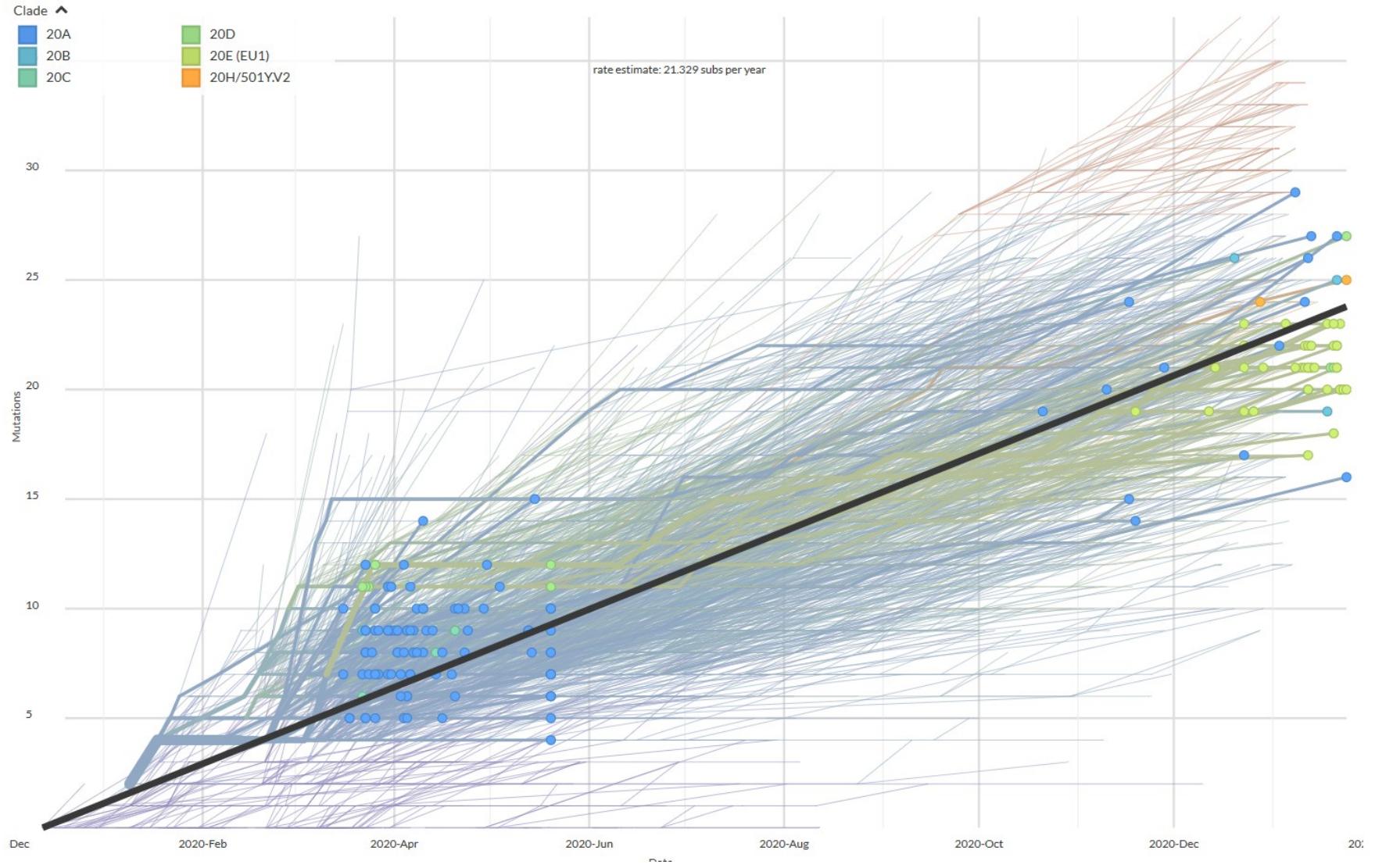
Clade ▾

- 20A
- 20B
- 20C

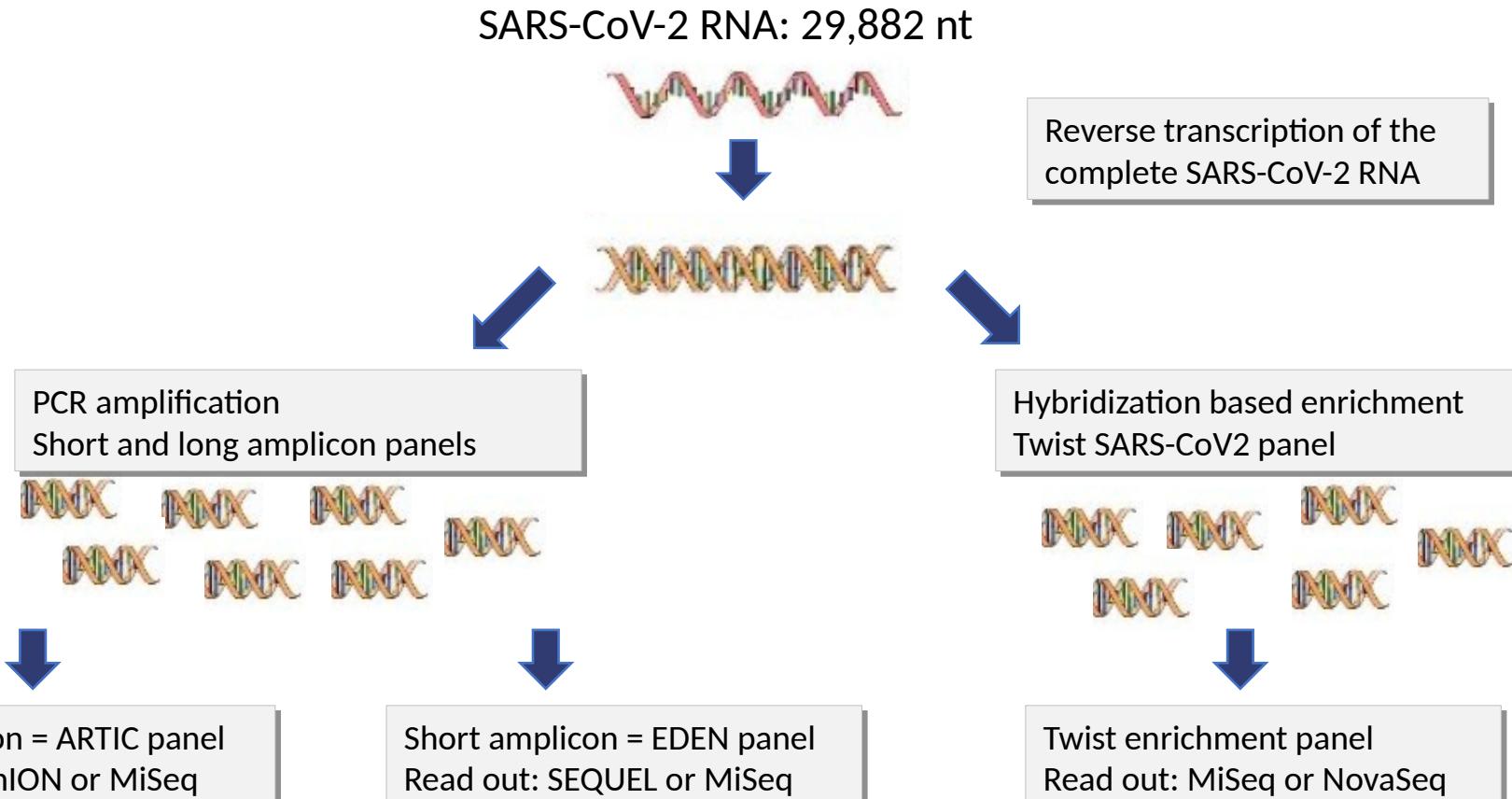
- 20D
- 20E (EU1)
- 20H/501Y.V2



NextStrain: Linearly Increasing Divergence of Strains

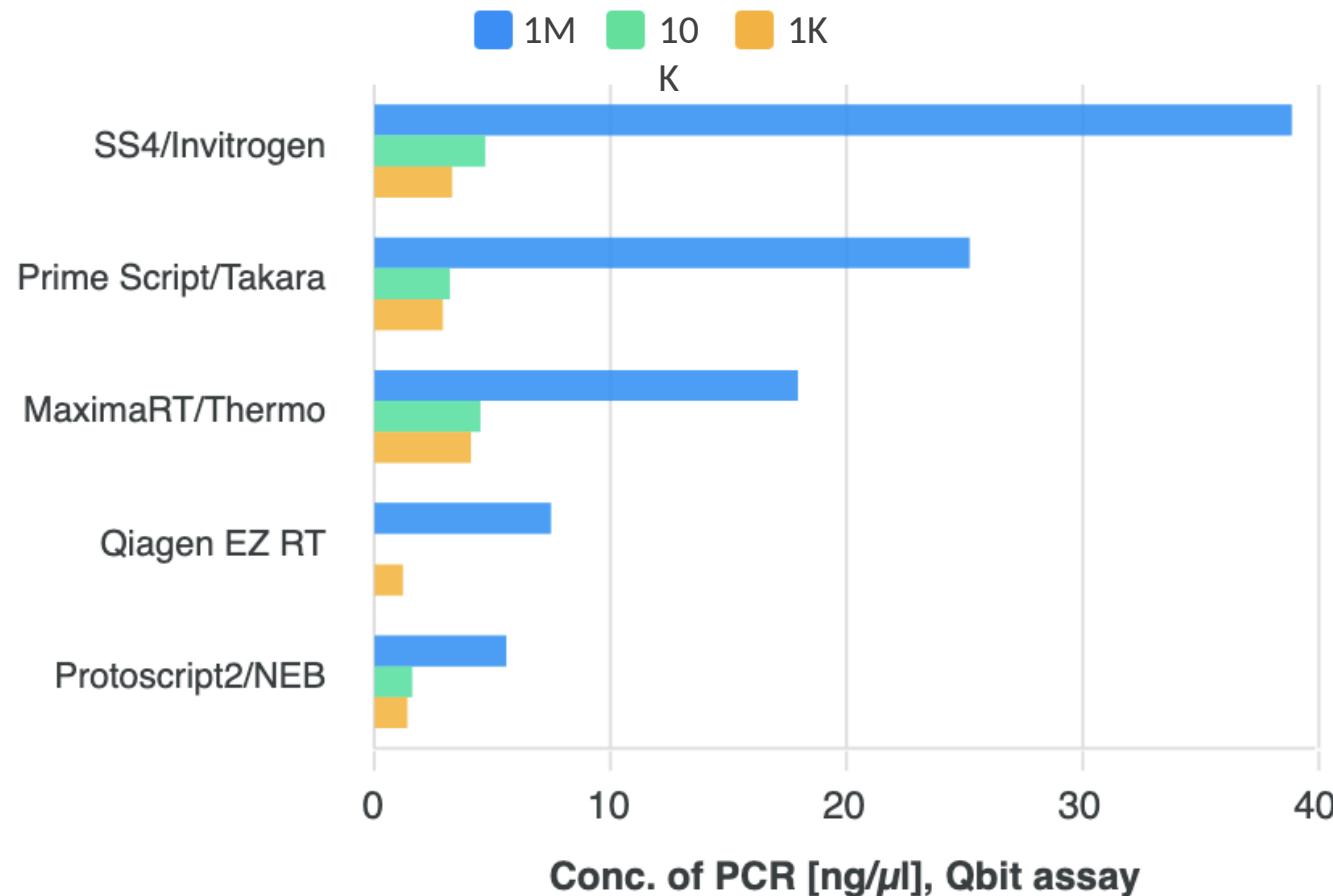


Sequencing Strategies for SARS-CoV2 Genomes





All RT reactions performed with 130 ng UHR + Twist positive control *



*

| CATALOG # | NAME |
|-----------|---|
| 102019 | Twist Synthetic SARS-CoV-2 RNA Control 1 (MT007544.1) |

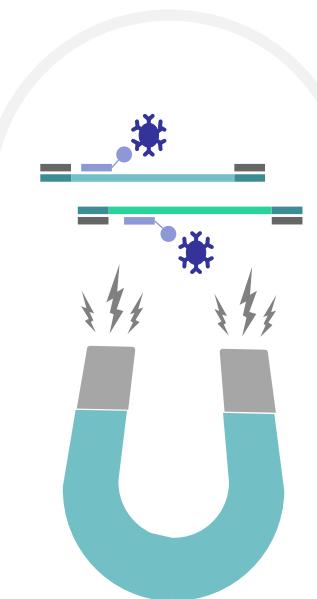
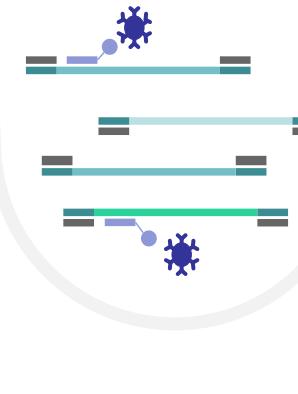
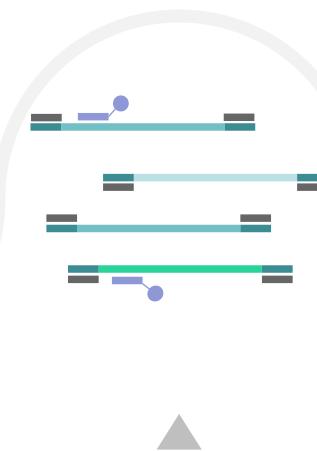
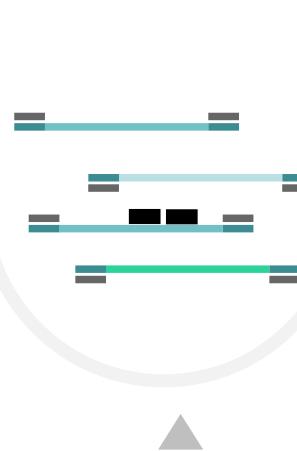
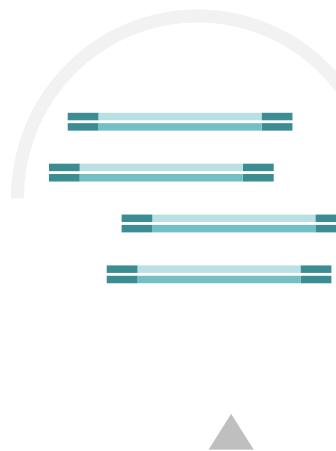
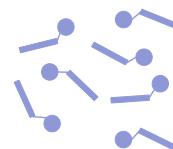
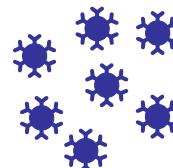


Prepped Library Fragments

Blocked Library Fragments

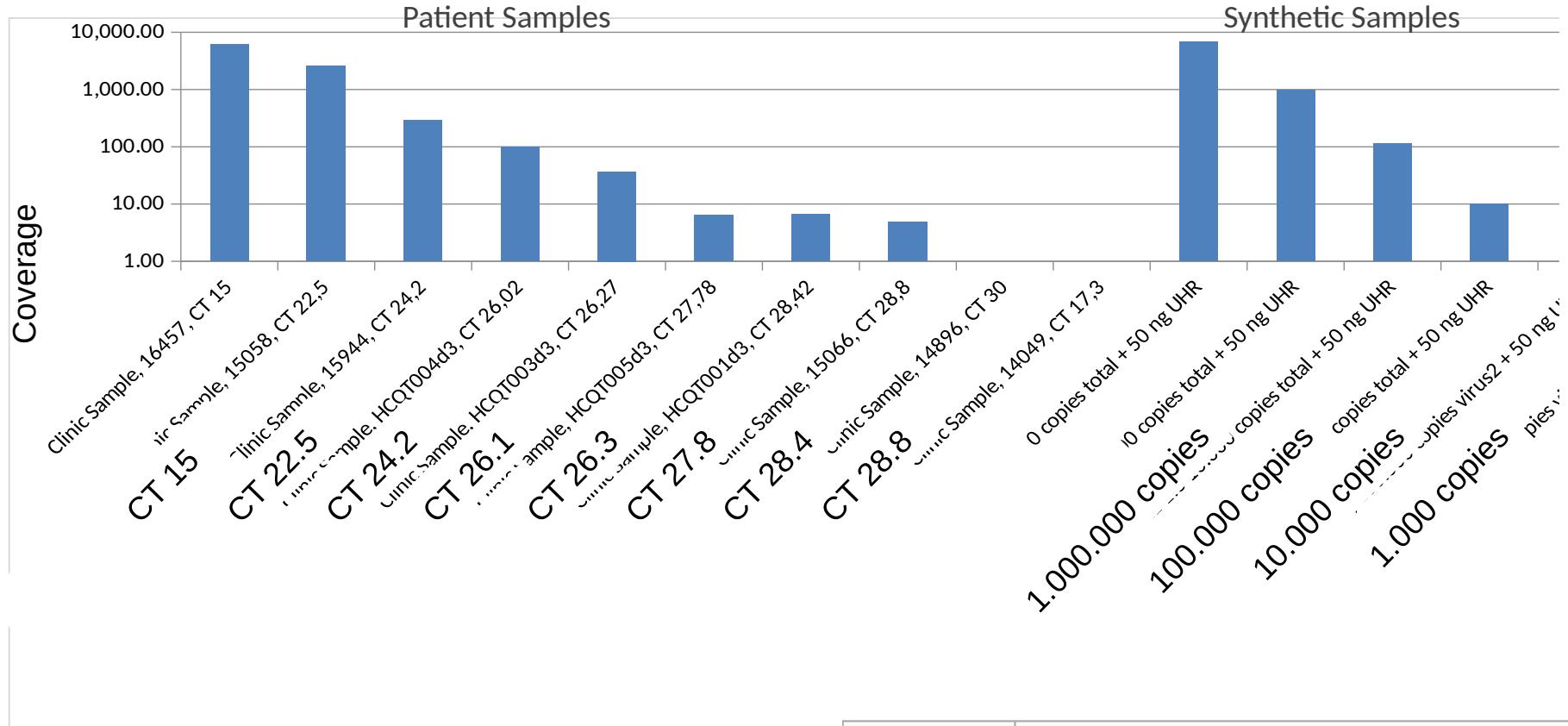
Hybridized Library Fragments

Bead-Bound Target Sequences

Universal Blockers
and COT-1TWIST
DNA ProbesStreptavidin
Beads

- Sense Strand
- Anti-Sense Strand
- Dual Adapters
- Biotin

<https://www.twistbioscience.com/resources/product-sheet/twist-sars-cov-2-research-panel>



Experiments: Michaela Pogoda

Clinical samples collected & processed by:

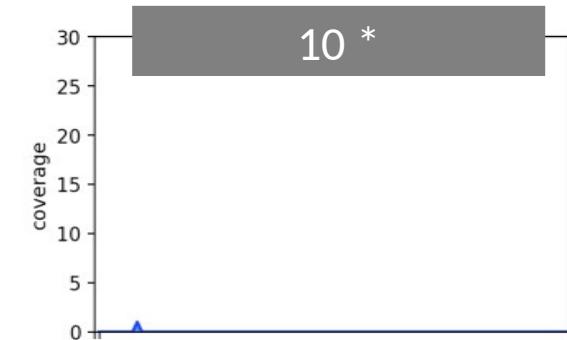
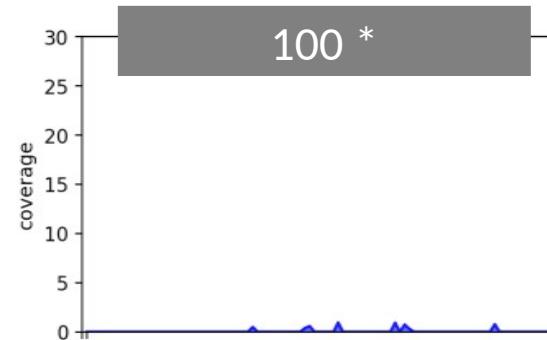
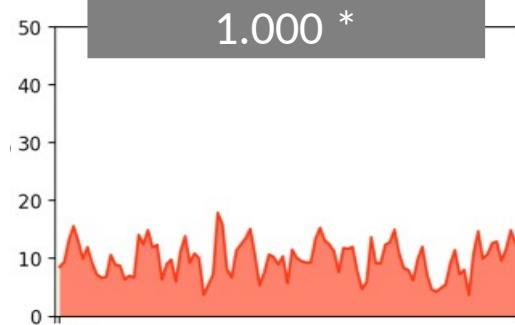
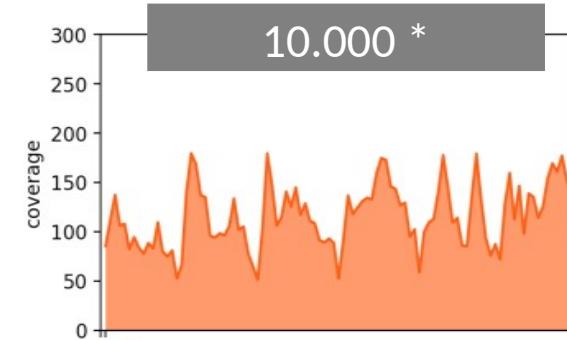
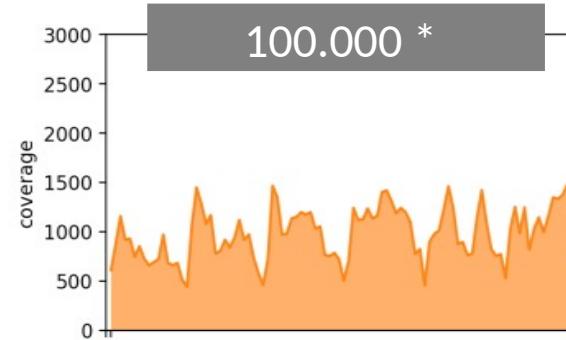
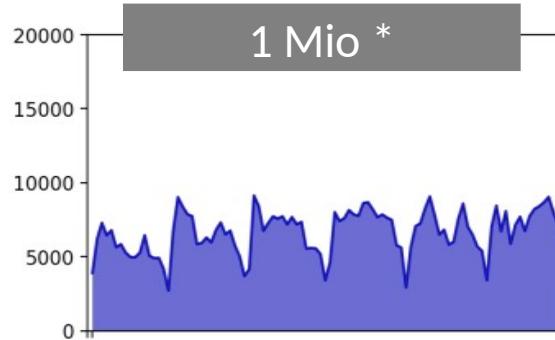
Michael Bitzer, Siri Göpel, Tina Ganzenmüller,
Michael Sonnabend, Angel Angelov, Silke Peter,
Daniela Bezdan, Angelika Iftner, Thomas Iftner

| CATALOG # | NAME |
|-----------|---|
| 102019 | Twist Synthetic SARS-CoV-2 RNA Control 1 (MT007544.1) |
| 102024 | Twist Synthetic SARS-CoV-2 RNA Control 2 (MN908947.3) |



* Dilution: Synthetic SARS-CoV2 copies per ul

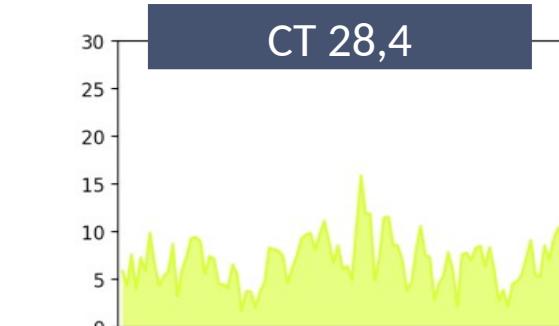
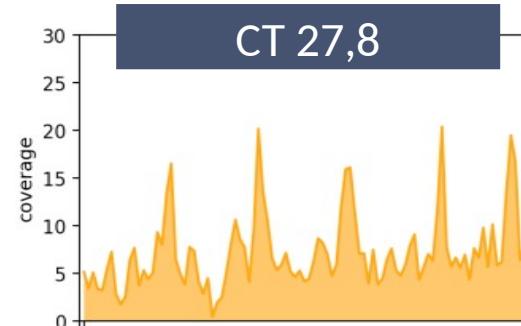
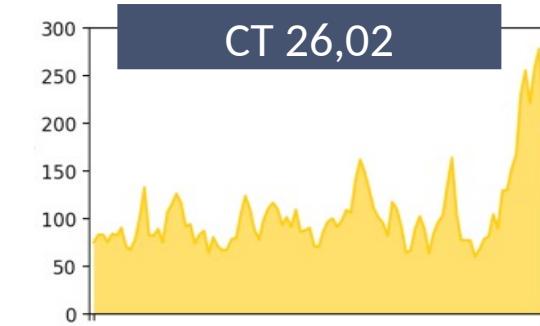
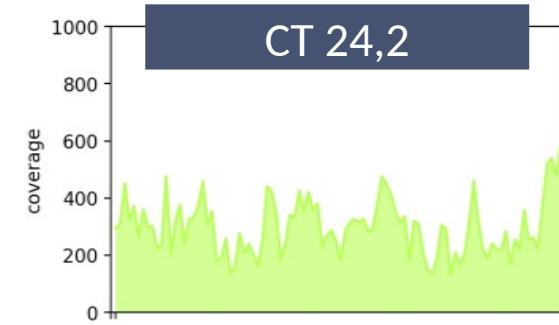
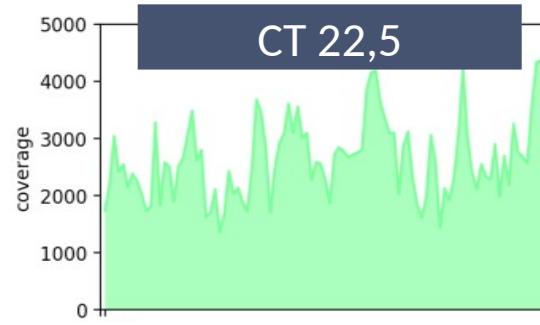
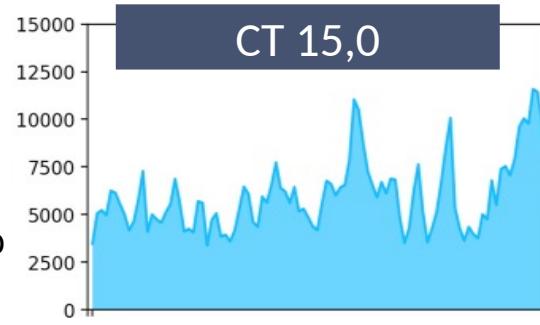
Coverage



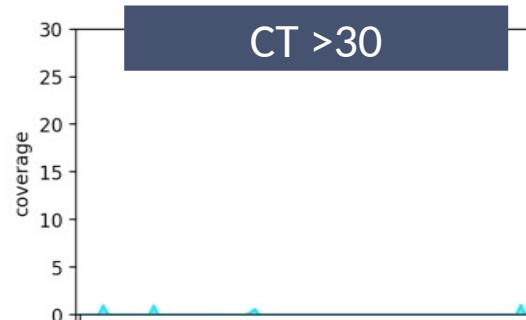
Genomic position



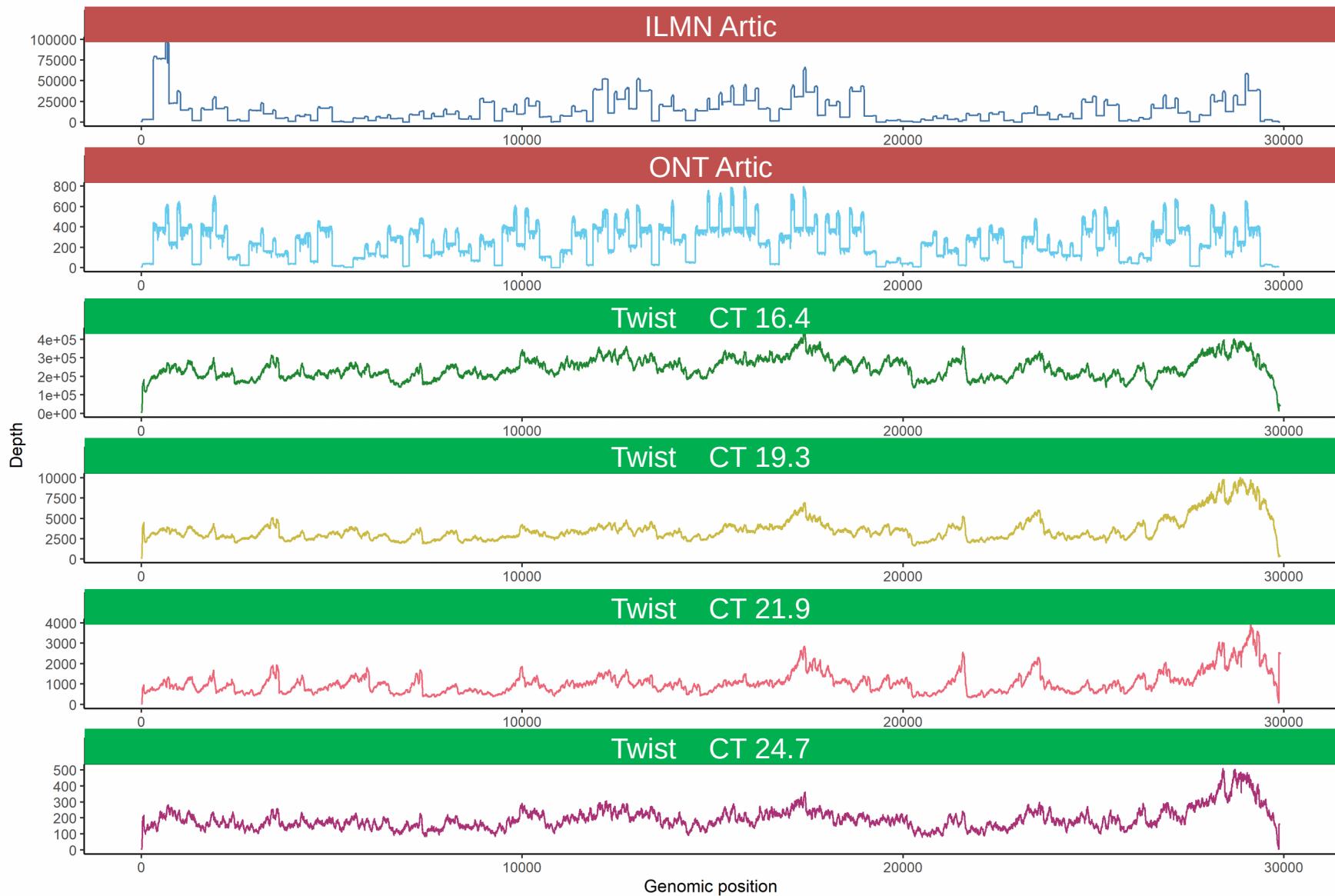
Coverage

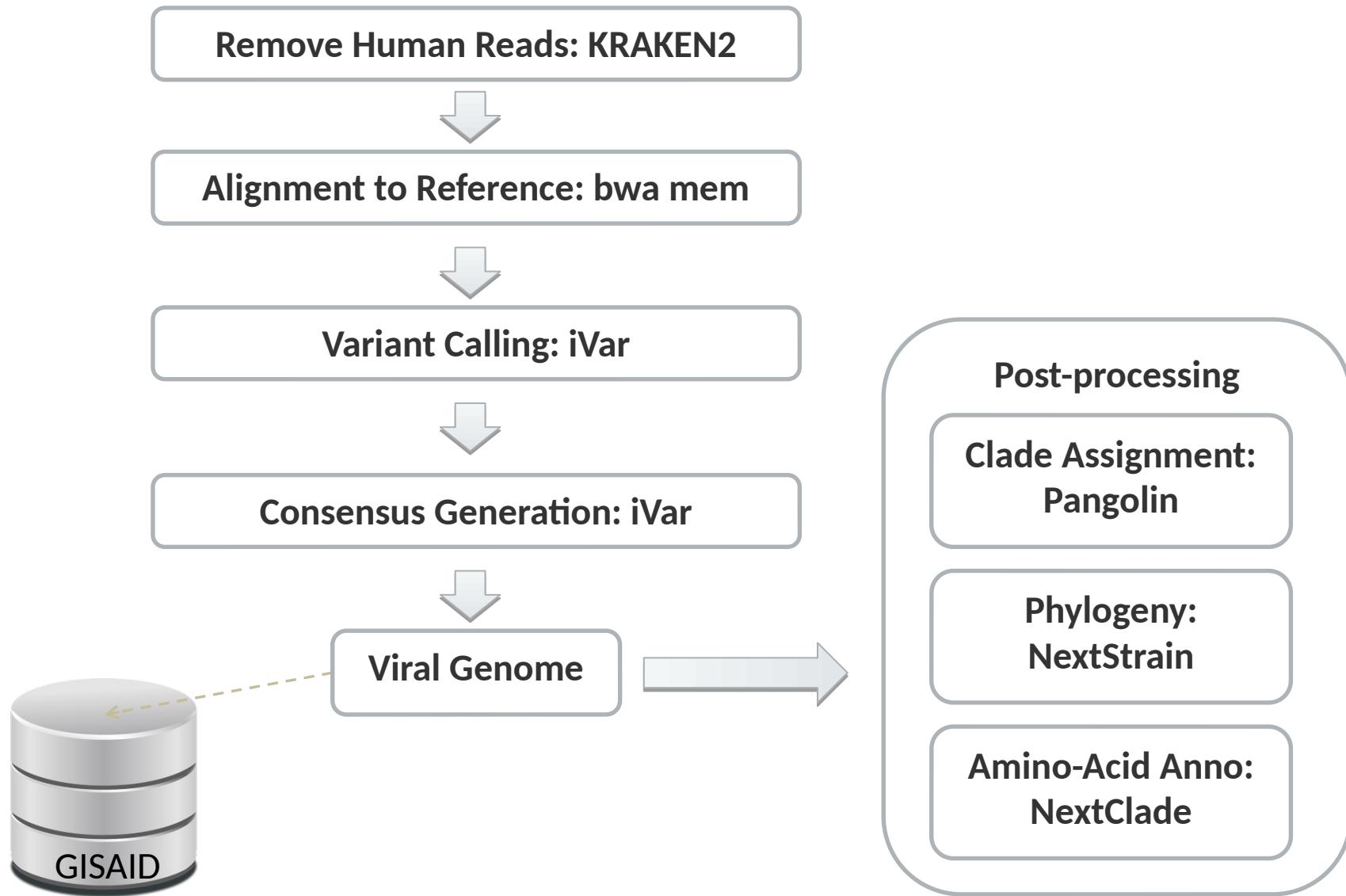


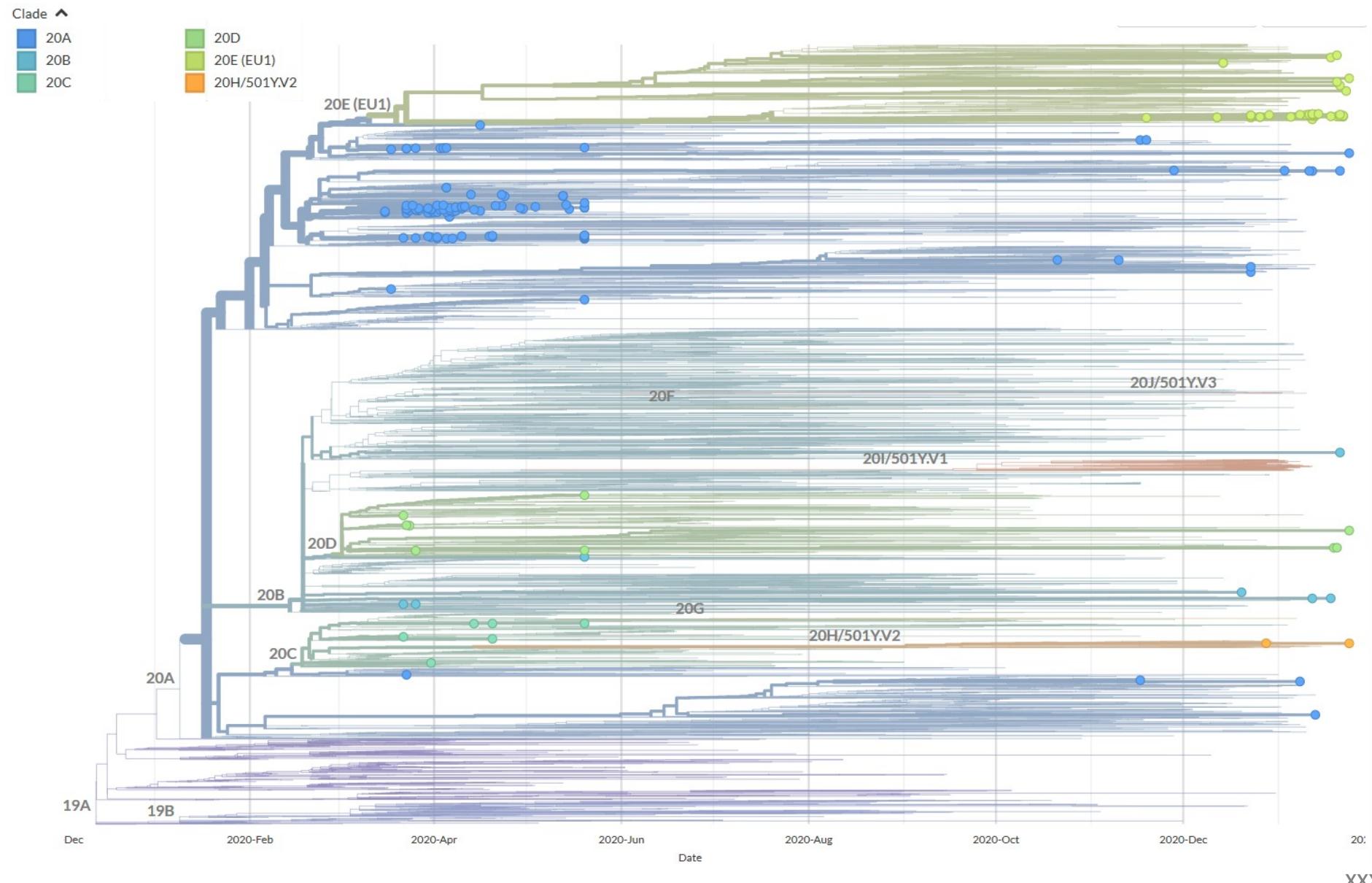
Genomic position



- We do not get continuous coverage for samples with Ct > 28
- We get hardly any reads with Ct > 30

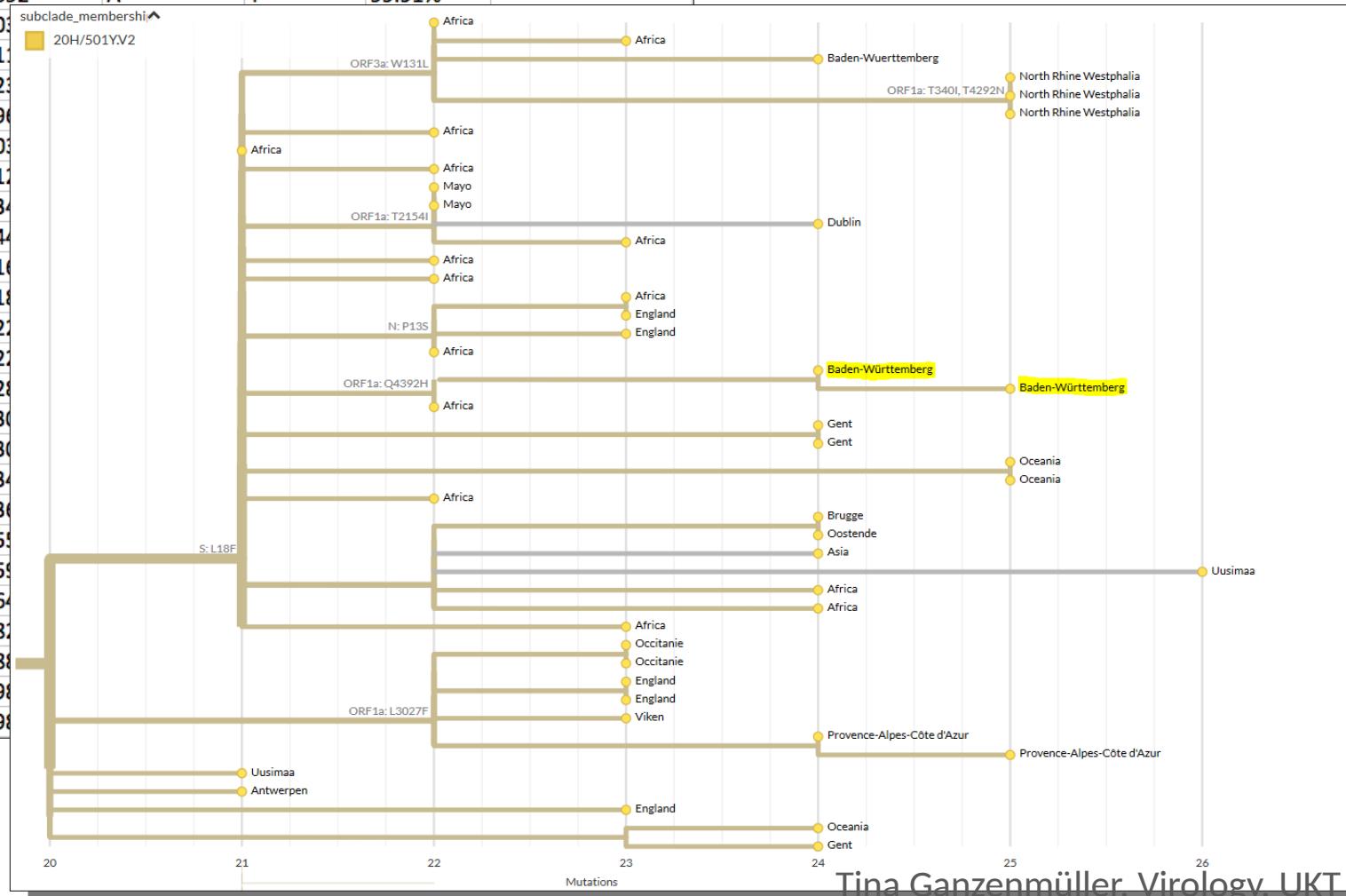


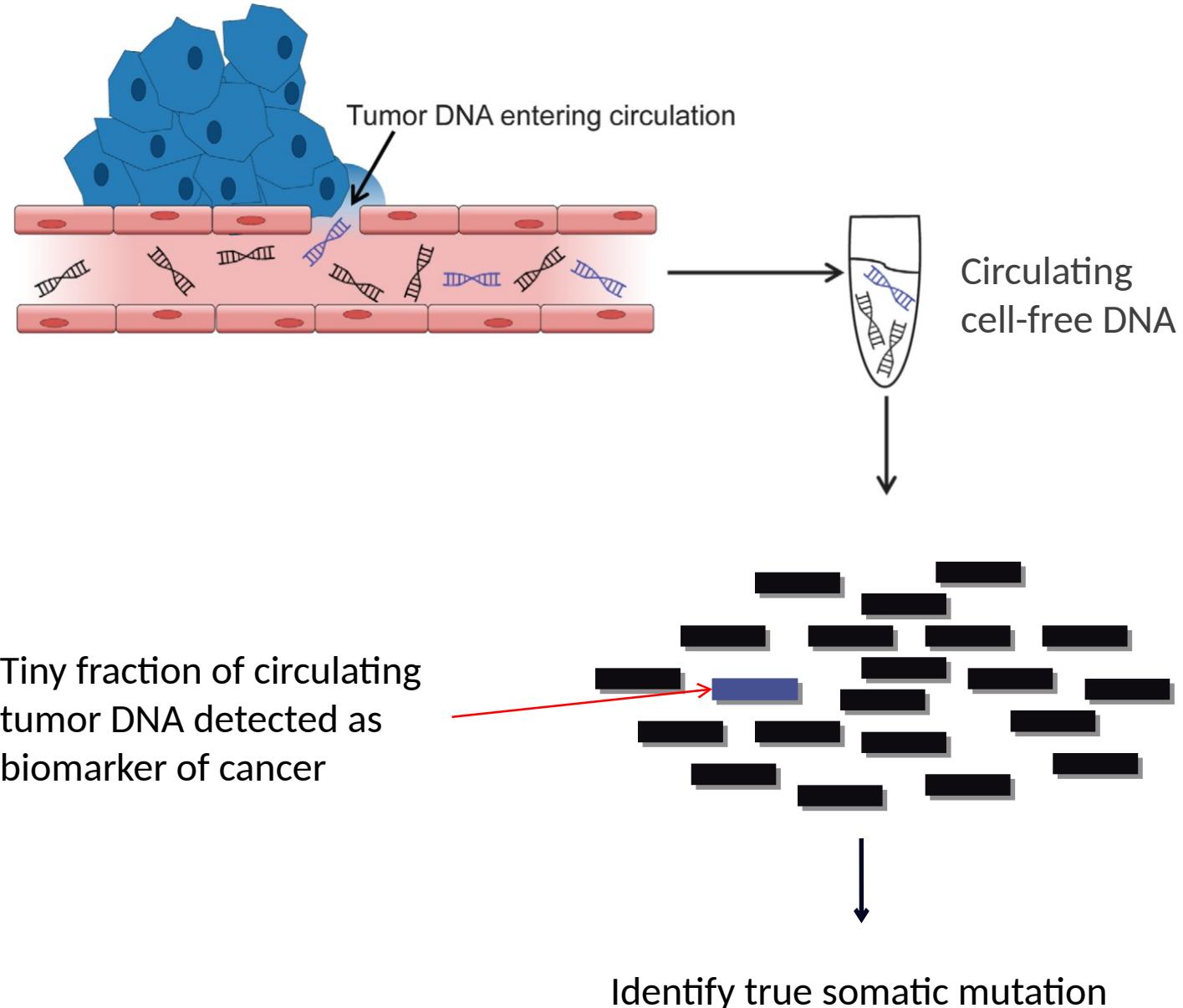






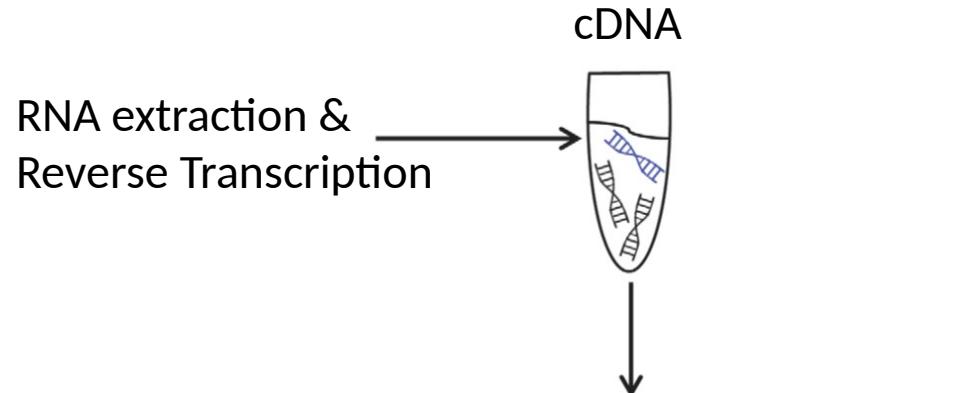
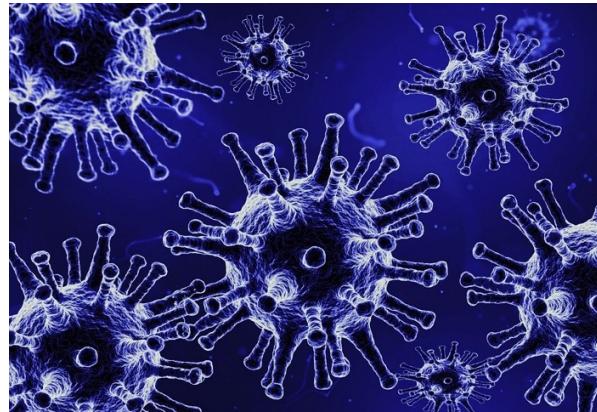
| | A | B | C | D | E | F |
|---|------------------|----------|-----------|-------------|--------|-------------------|
| 1 | Reference genome | Position | Reference | Alternative | AF | Amino-Acid change |
| 2 | MN908947.3 | 174 | G | T | 99.87% | |
| 3 | MN908947.3 | 241 | C | T | 99.98% | |
| 4 | MN908947.3 | 1059 | C | T | 99.96% | |
| 5 | MN908947.3 | 2692 | A | T | 99.91% | |



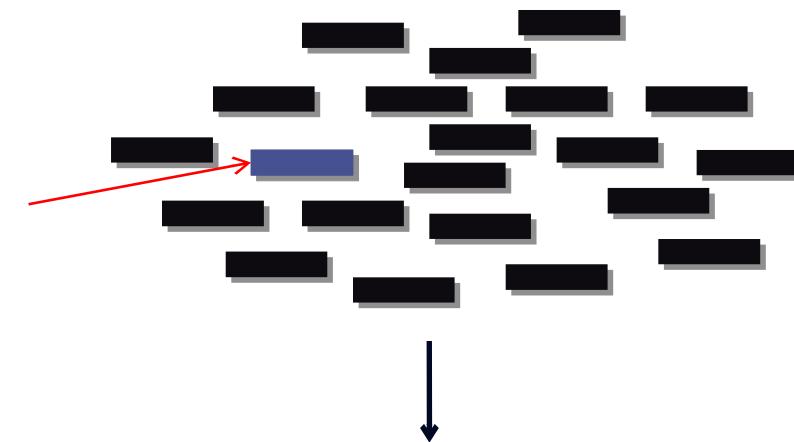




Virus population in the host



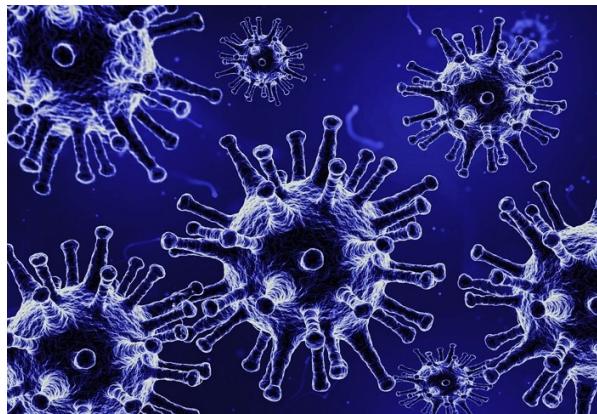
Tiny fraction of the SARS-CoV-2
viruses in a sample carry a novel
mutation



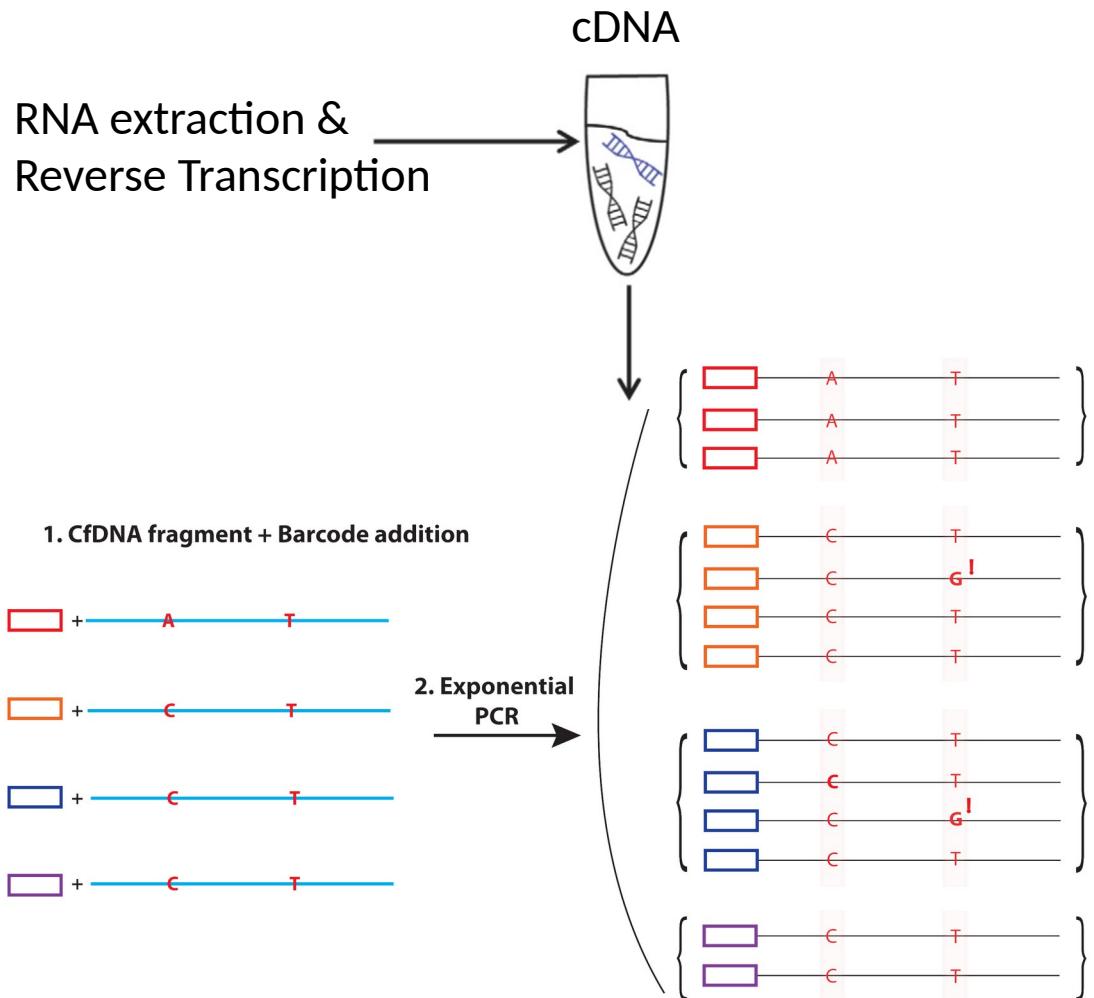
Ultra-Deep Sequencing: 1,000,000x Coverage



Virus population in the host

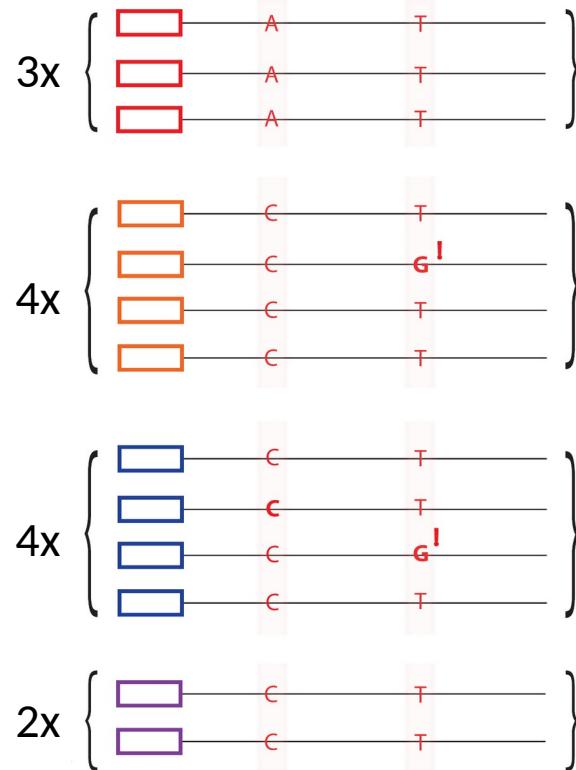


Unique Molecular Barcodes (UMI)

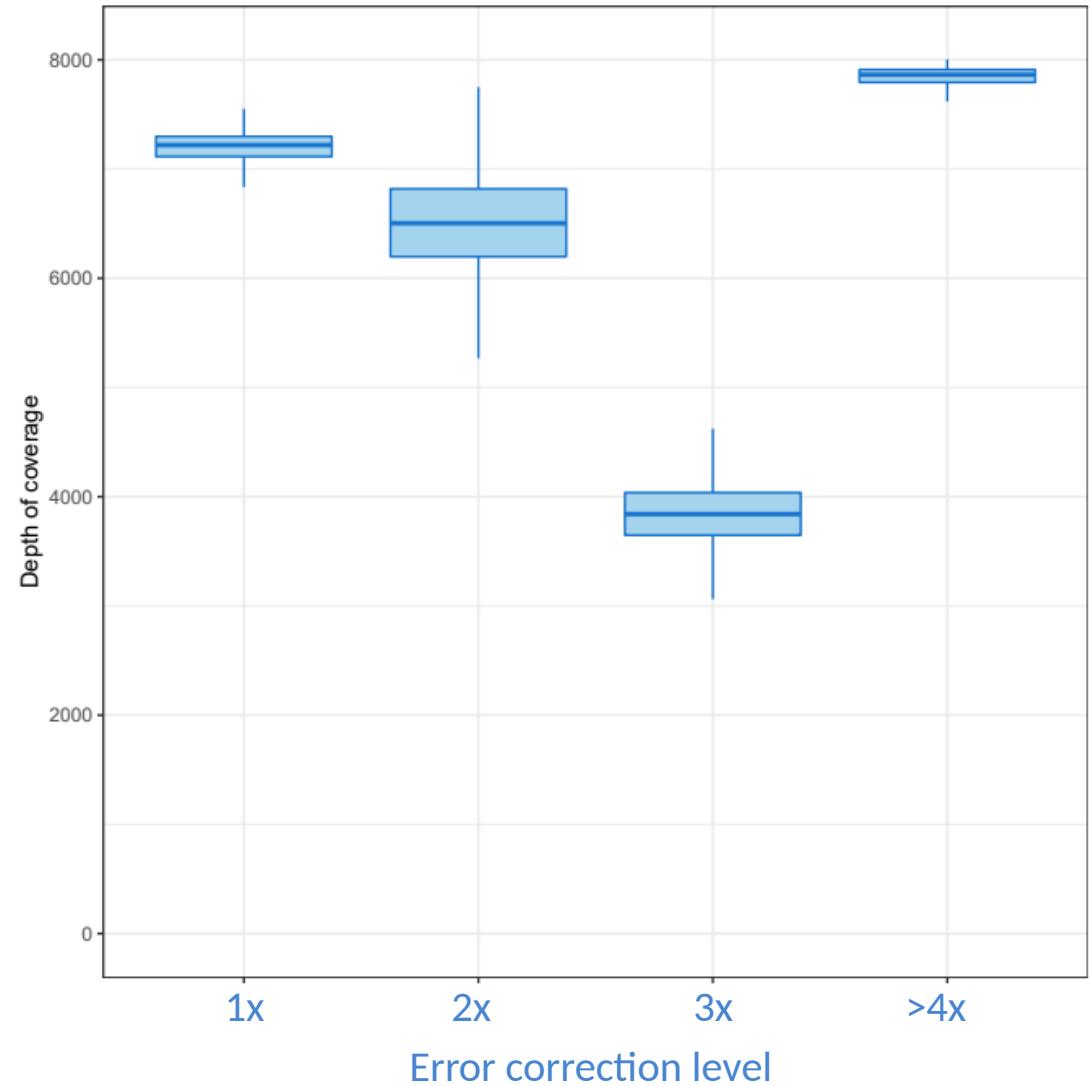




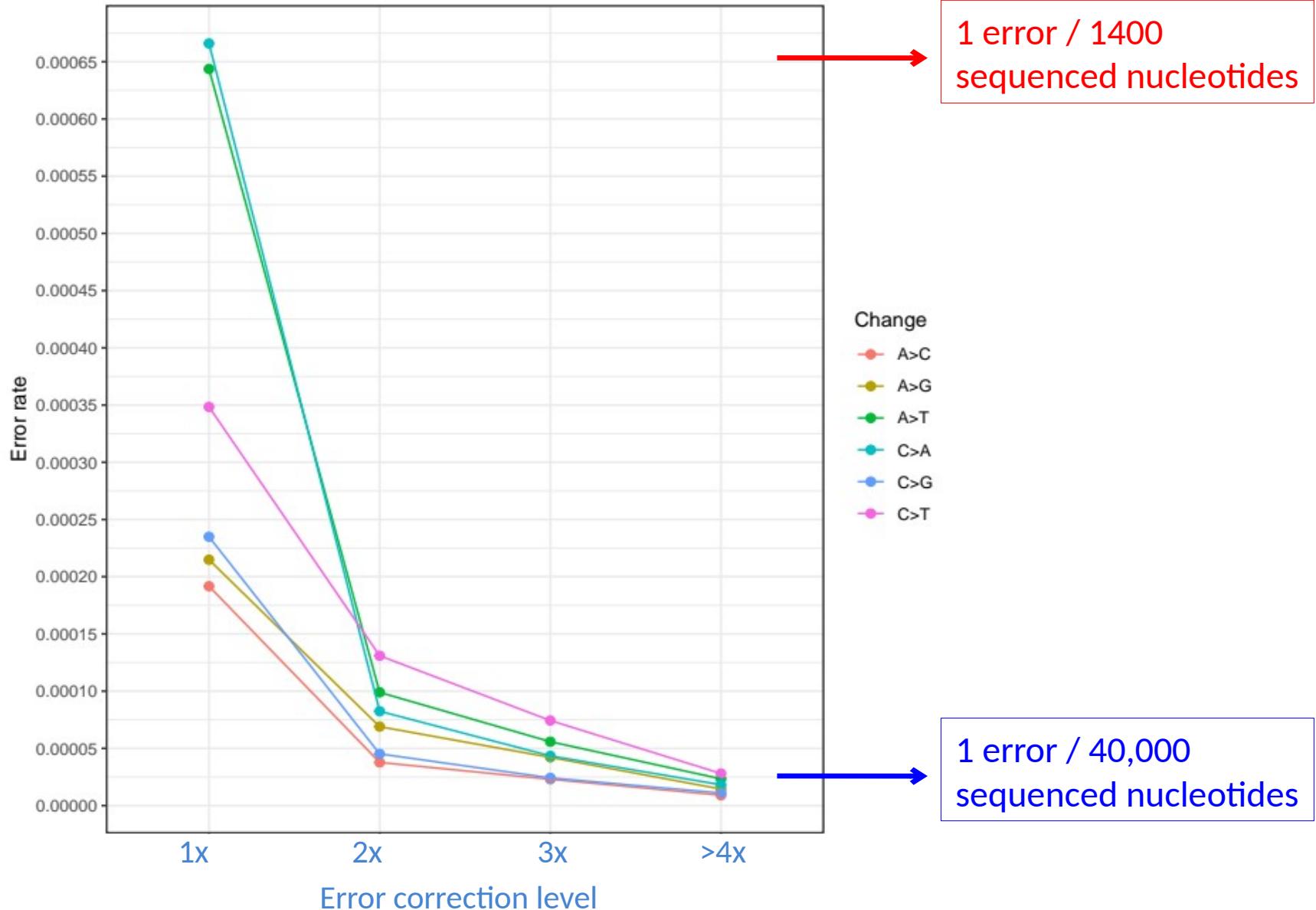
Error correction level:

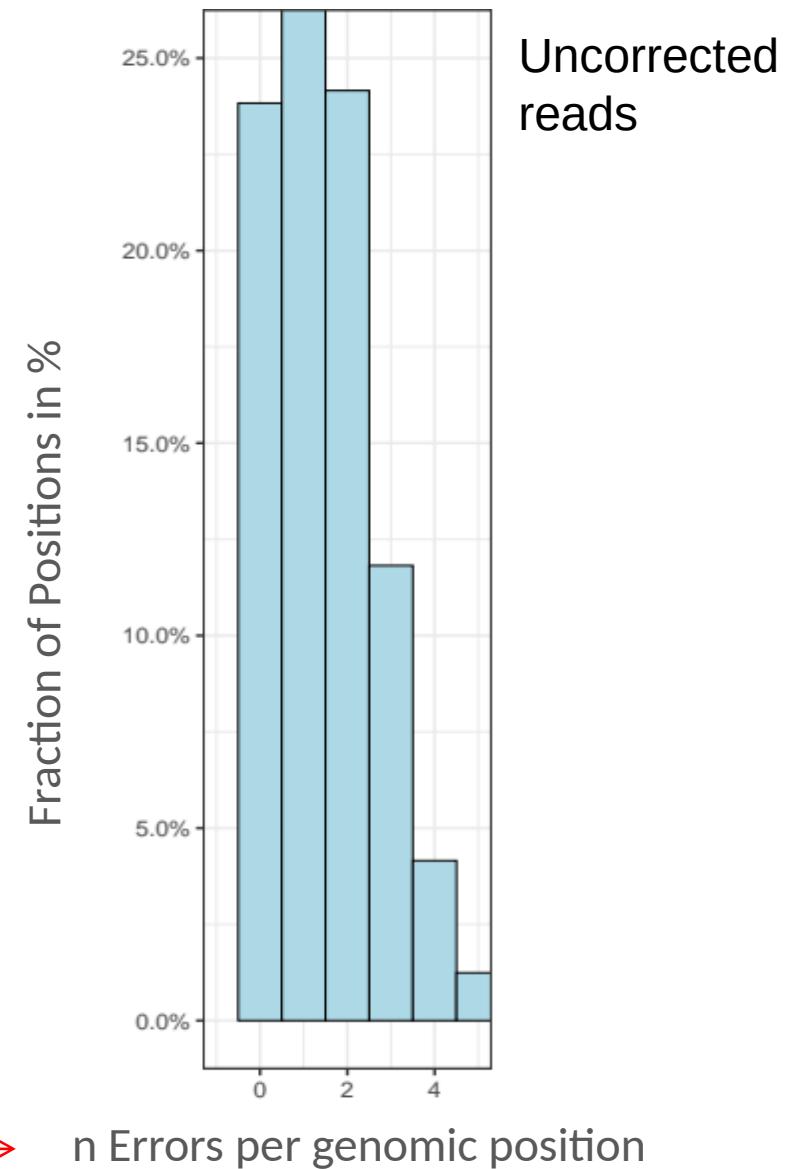
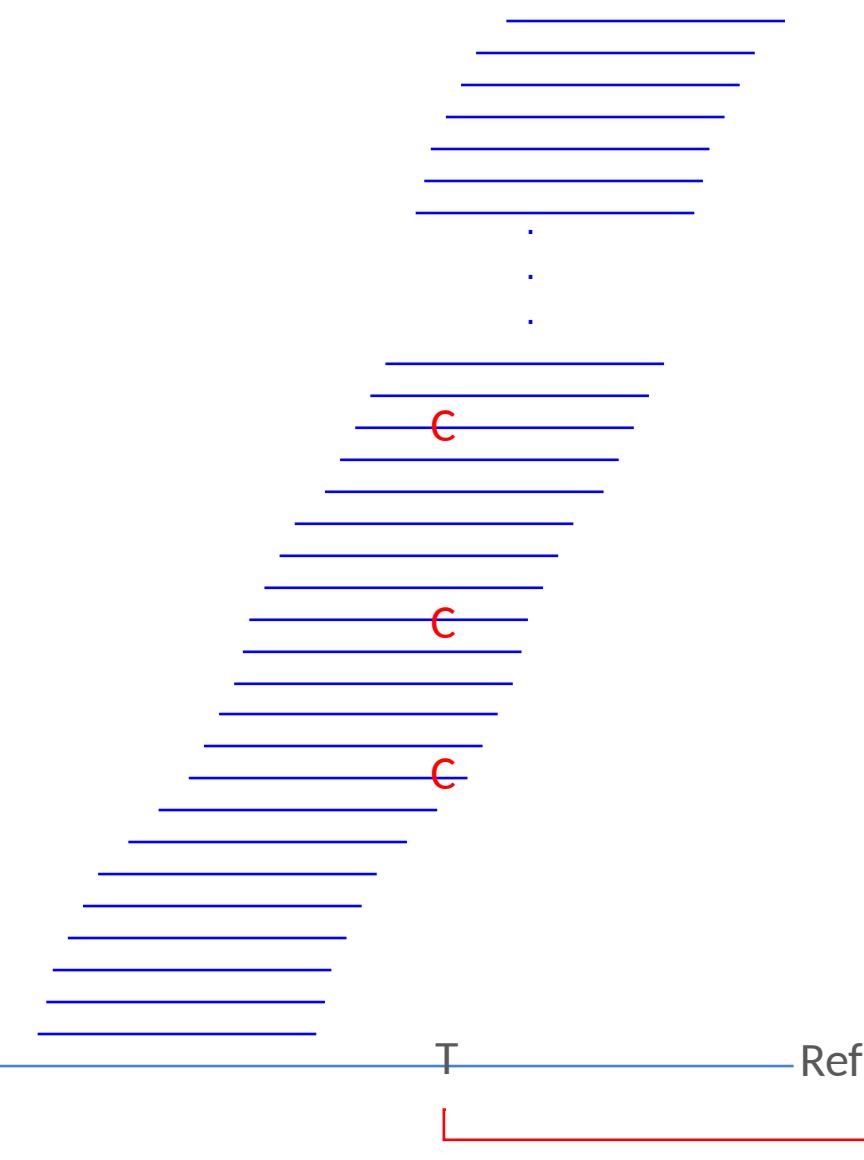


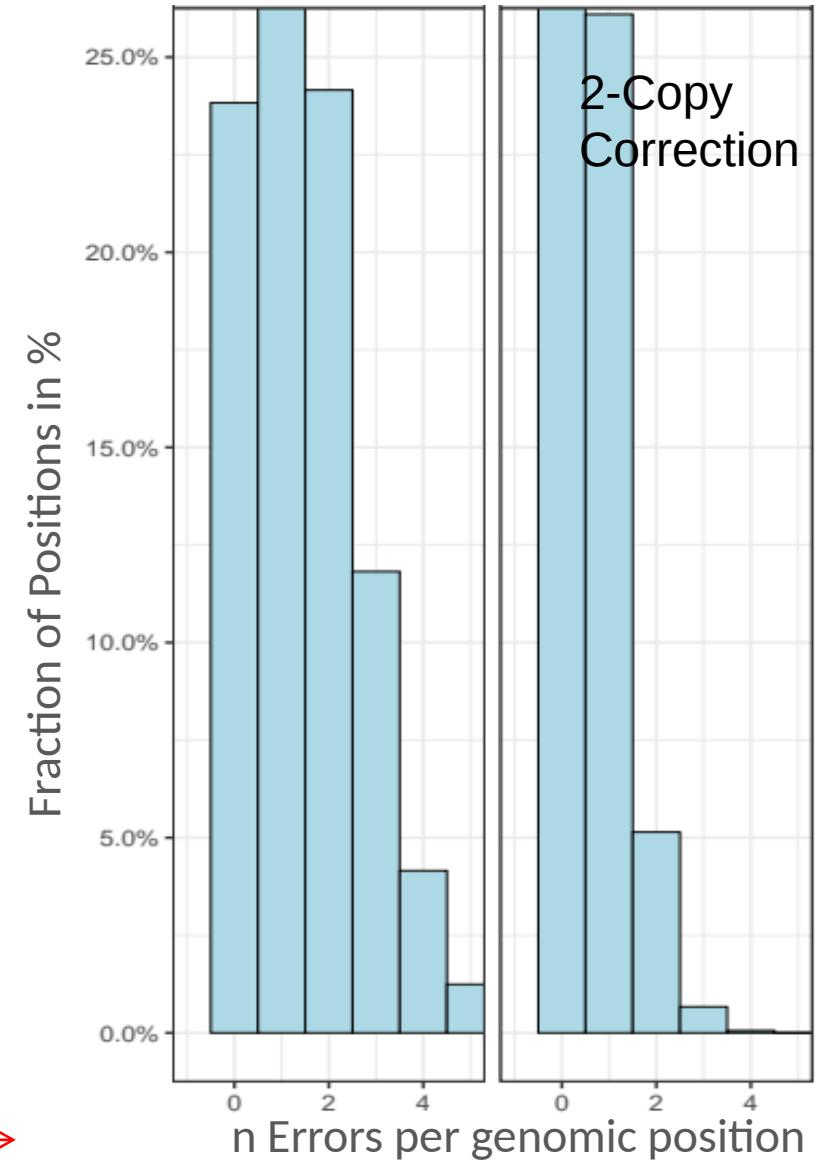
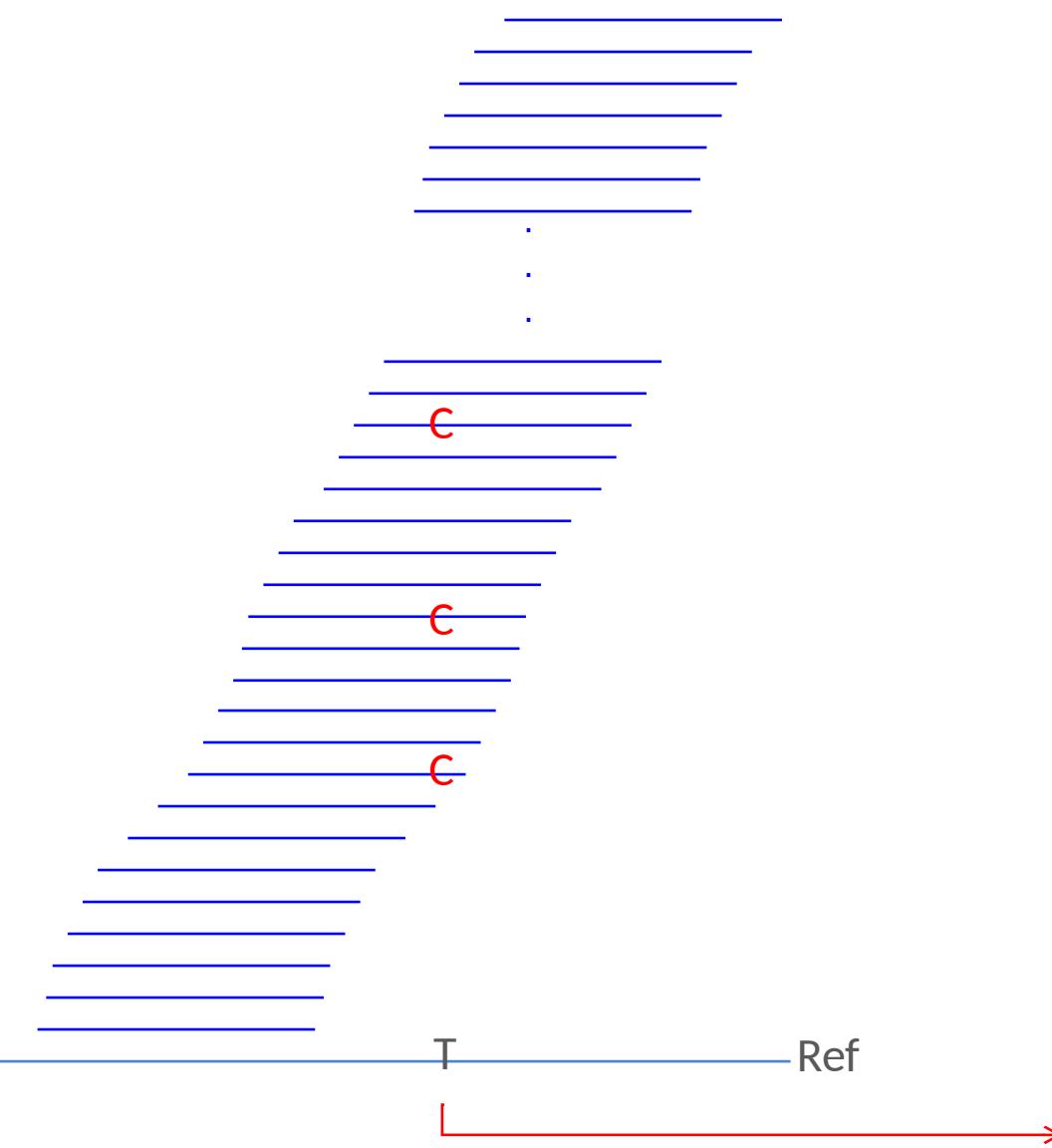
Depth of coverage per correction level:

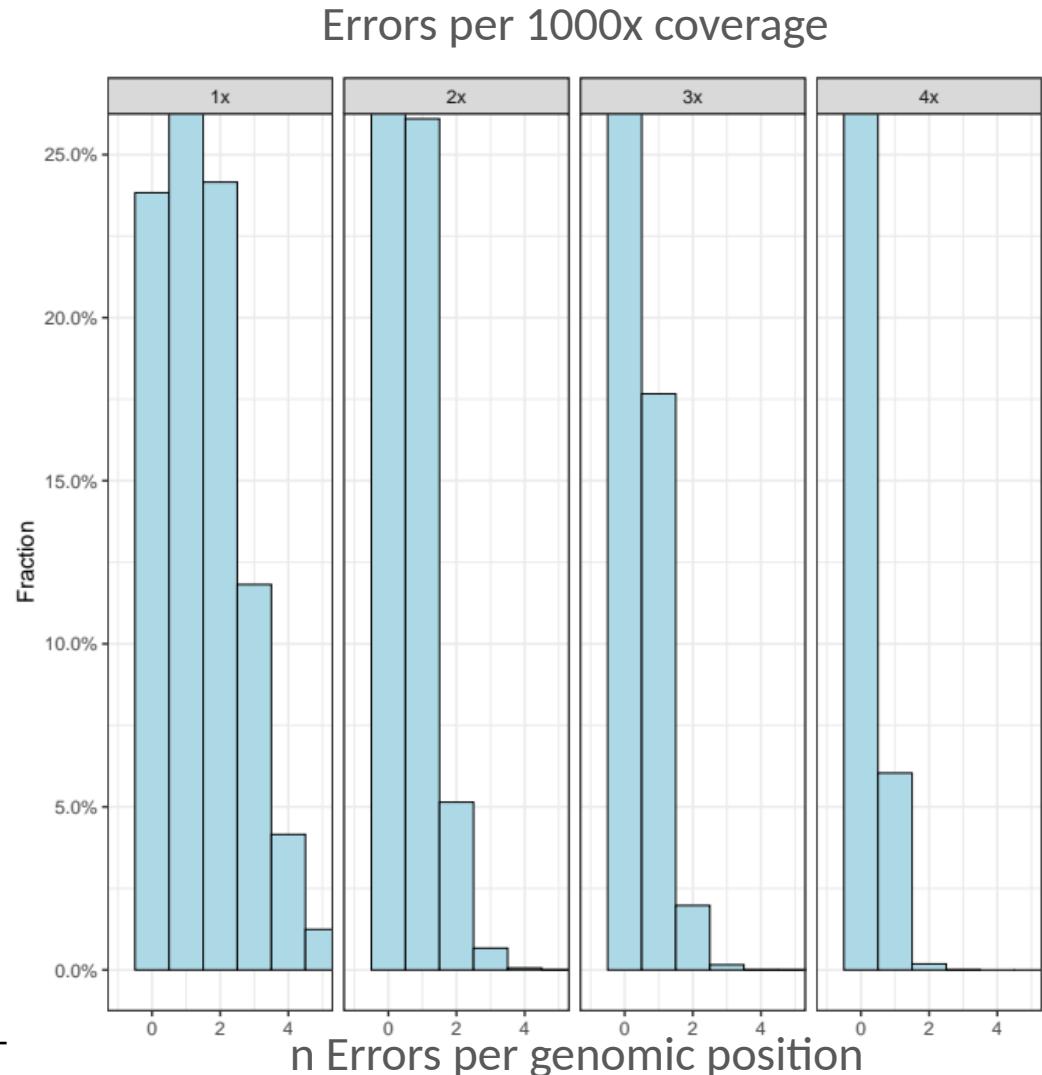
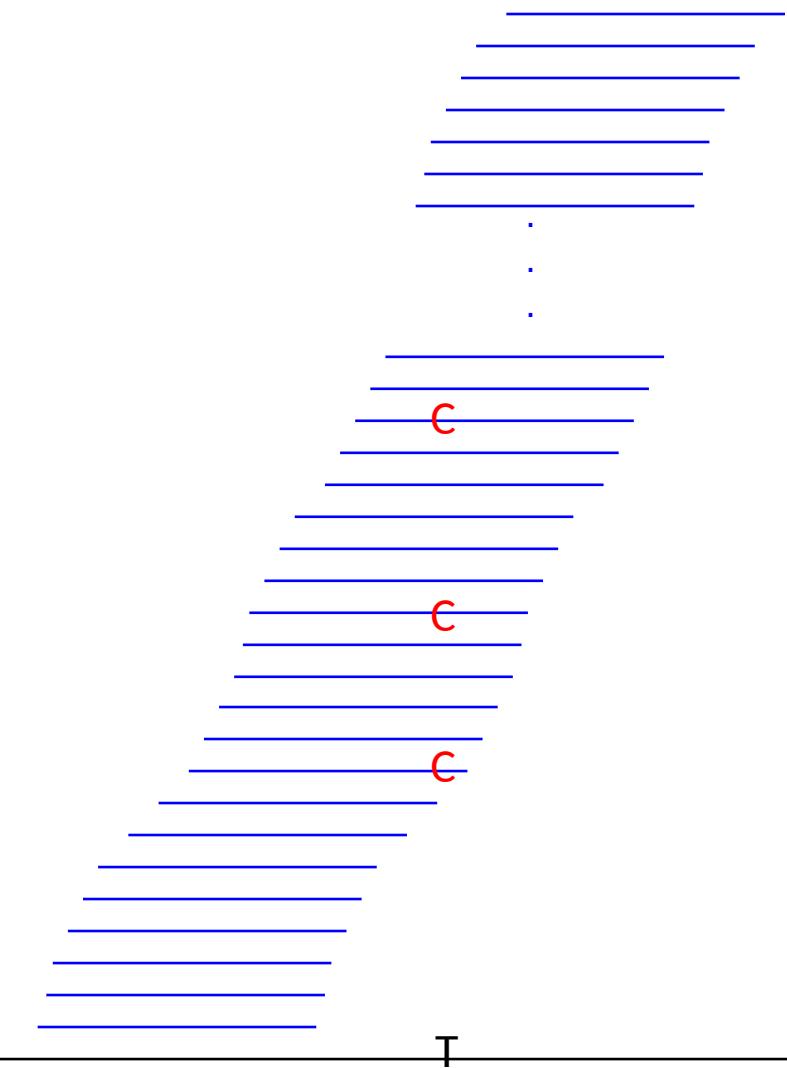


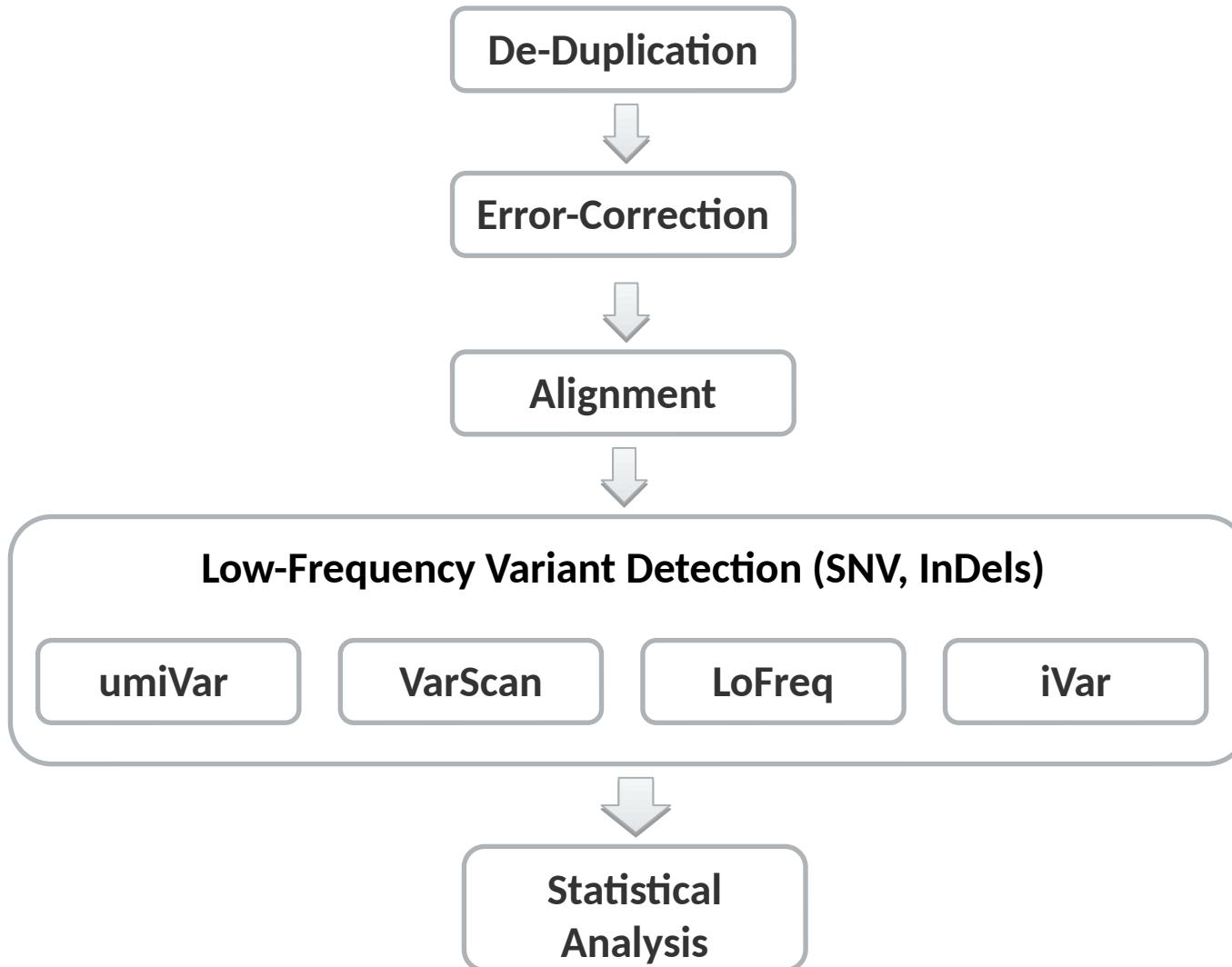
Error Rates per Nucleotide





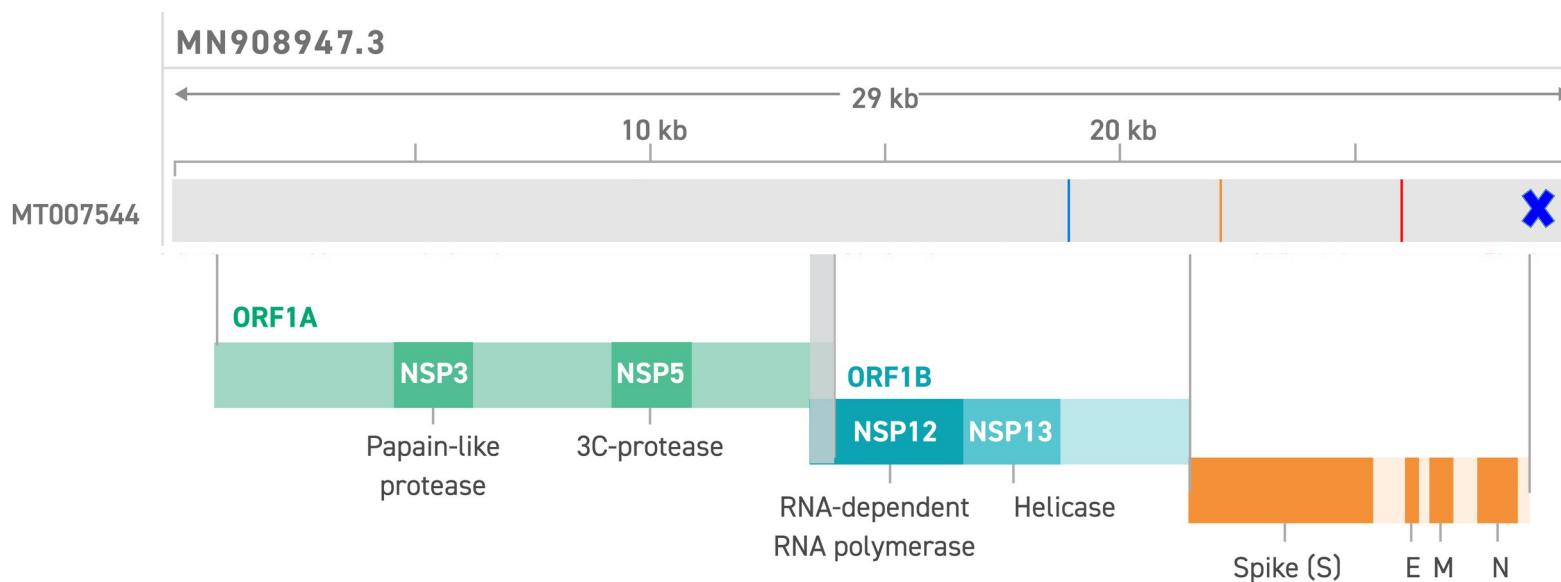
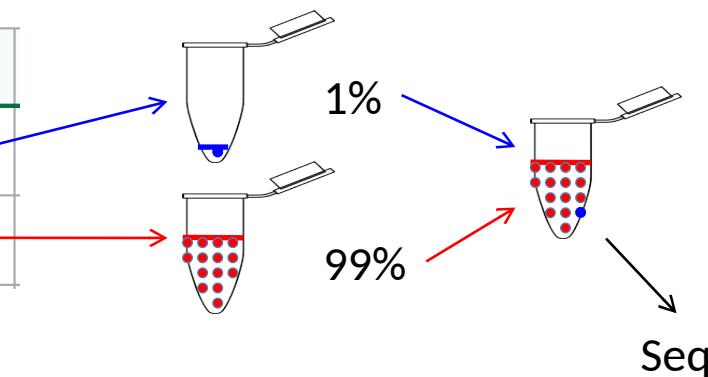


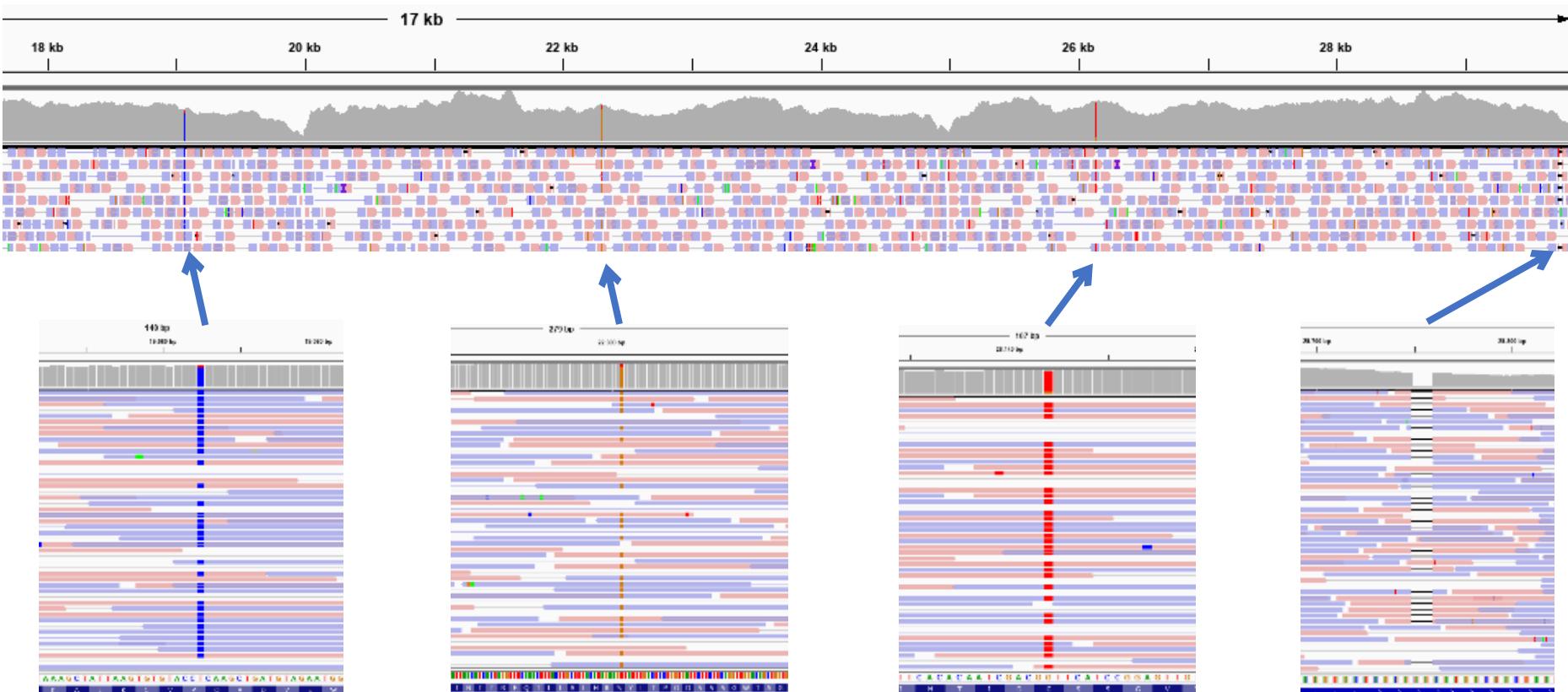






| CATALOG # | NAME |
|-----------|---|
| 102019 | Twist Synthetic SARS-CoV-2 RNA Control 1 (MT007544.1) |
| 102024 | Twist Synthetic SARS-CoV-2 RNA Control 2 (MN908947.3) |





| Position | 19065 | 22303 | 26144 | 29754 |
|----------|-------------------|---------------------|-------------------|-------------------|
| Reads | 6087 | 6566 | 7025 | 38 |
| A | 11 | 3 | 2 | 0 |
| C | 5994 (99%) | 1 | 1 | 38 (1%) |
| G | 2 | 6462 (98.5%) | 120 (2%) | 0 |
| T | 77 (1%) | 98 (1.5%) | 6901 (98%) | 0 |
| Dels | | | | 3769 (99%) |

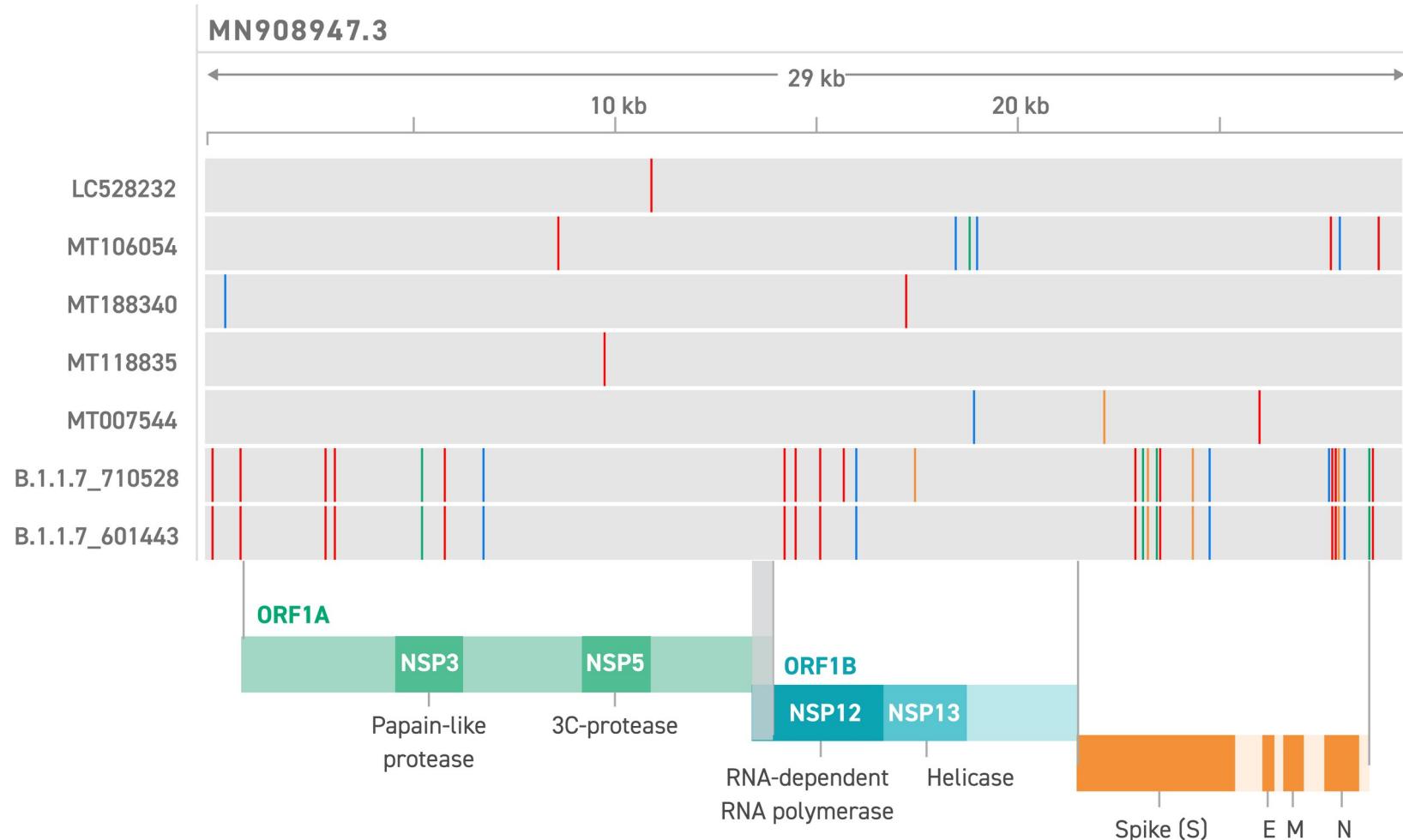


| Position | Change | %1 Mix (reads / AF) | %0.5 Mix (reads / AF) | %0.1 Mix (reads / AF) |
|----------|----------|---------------------|-----------------------|-----------------------|
| 19065 | T -> C | 348/38498 (0.9%) | 204/41687 (0.49%) | 78/40984 (0.19%) |
| 22303 | T -> G | 385/45576 (0.85%) | 229/50939 (0.45%) | 64/49169 (0.13%) |
| 26144 | G -> T | 507/47843 (1.05%) | 349/51955 (0.67%) | 113/49886 (0.23%) |
| 29754 | C -> DEL | 180/20379 (0.89%) | 66/22176 (0.30%) | 19/21622 (0.09%) |

- SNVs and indels can be detected even at 0.1% fraction
- About 50% drop in coverage around the indel, but still detectable
- Test with synthetic SARS-CoV-2 strains with more SNVs and indels would be helpful to better estimate sensitivity



Twist Synthetic Viral Controls:



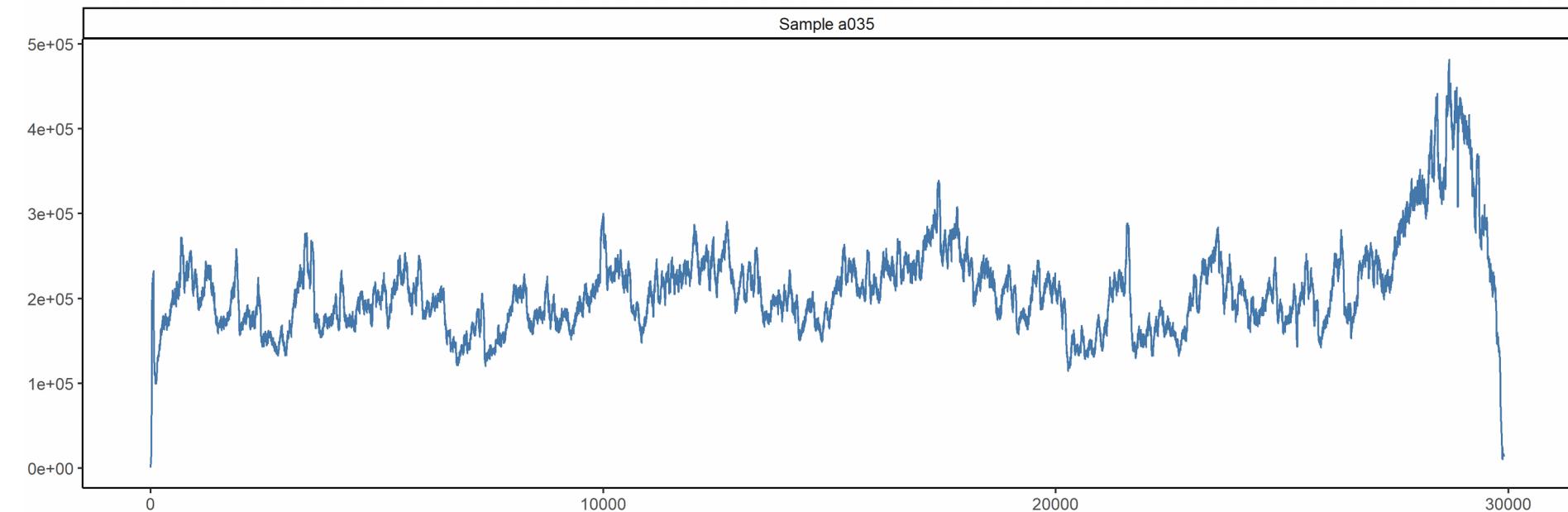
Ultra-deep Sequencing of Clinical Samples



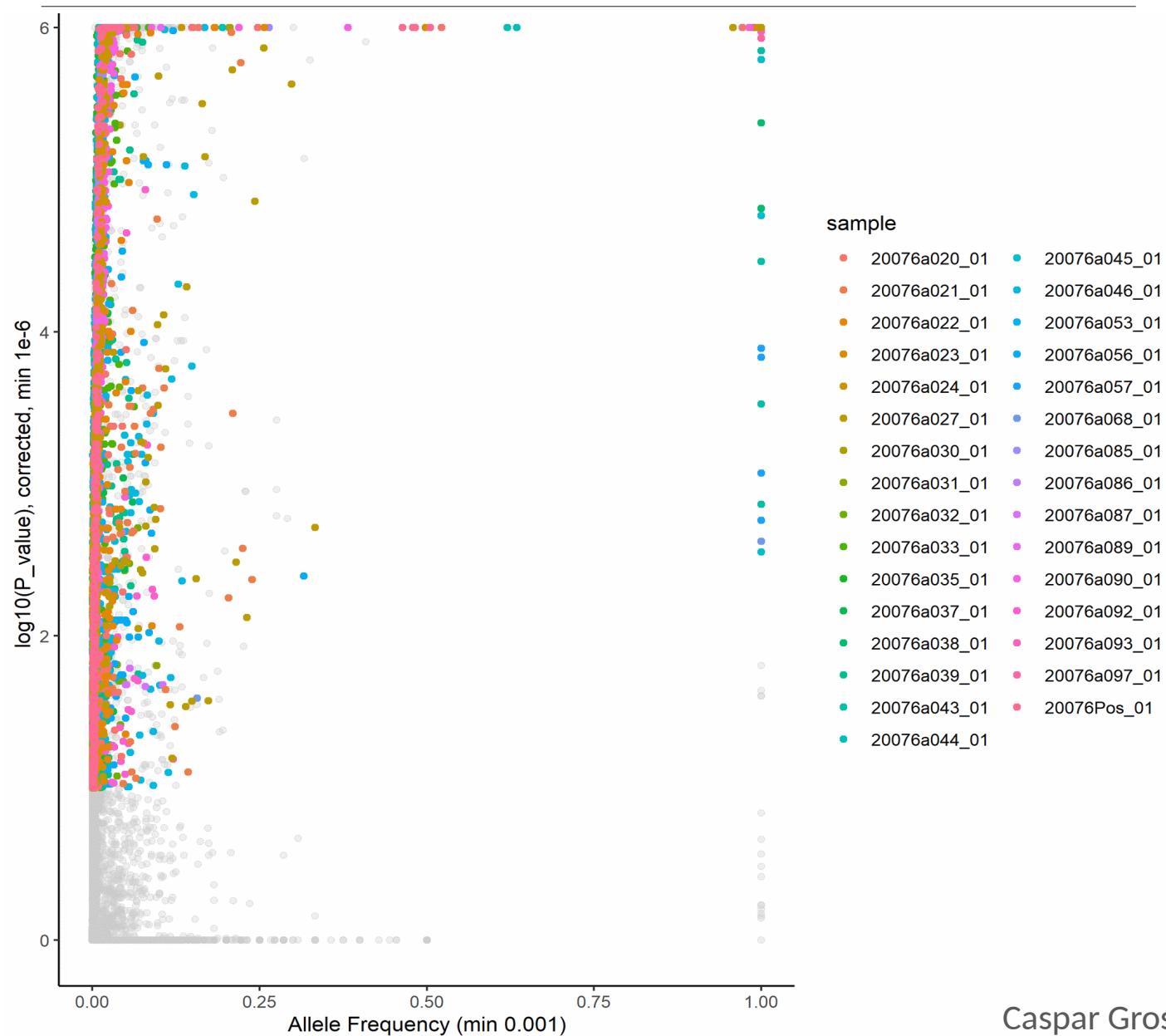
| Sample Name | on-target read % | target region read depth ▾ |
|--------------|------------------|----------------------------|
| 20076a044_01 | 98.53% | 259 497.83x |
| 20076a035_01 | 98.56% | 226 347.07x |
| 20076a130_01 | 99.45% | 172 681.32x |
| 20076a089_01 | 99.57% | 145 399.58x |
| 20076Pos_01 | 99.33% | 111 116.48x |
| 20076a128_01 | 99.66% | 109 956.77x |
| 20076a129_01 | 99.65% | 106 973.06x |
| 20076a126_01 | 99.61% | 89 018.13x |
| 20076a085_01 | 99.69% | 45 962.27x |
| 20076a125_01 | 99.59% | 44 879.46x |
| 20076a117_01 | 99.81% | 38 771.14x |
| 20076a006_01 | 99.80% | 33 153.88x |
| 20076a090_01 | 99.62% | 31 389.30x |
| 20076a149_01 | 89.88% | 27 506.71x |
| 20076a024_01 | 99.57% | 24 323.87x |
| 20076a030_01 | 99.41% | 24 232.84x |
| 20076a137_01 | 82.64% | 23 460.78x |
| 20076a023_01 | 97.34% | 22 914.47x |
| 20076a132_01 | 90.19% | 22 462.99x |
| 20076a153_01 | 89.93% | 15 859.50x |

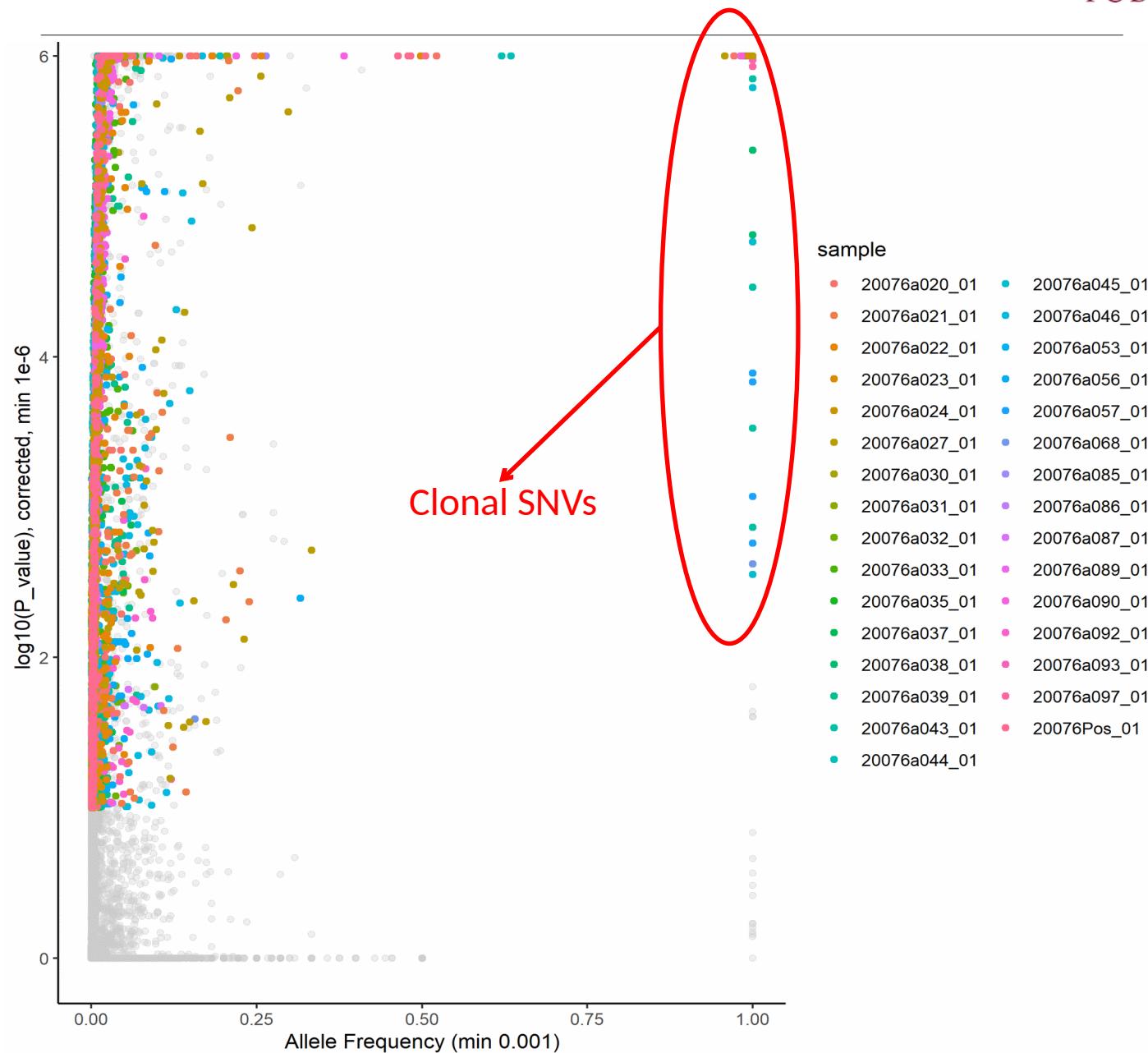
... total of 80 clinical samples sequenced ultra-deep with UMI

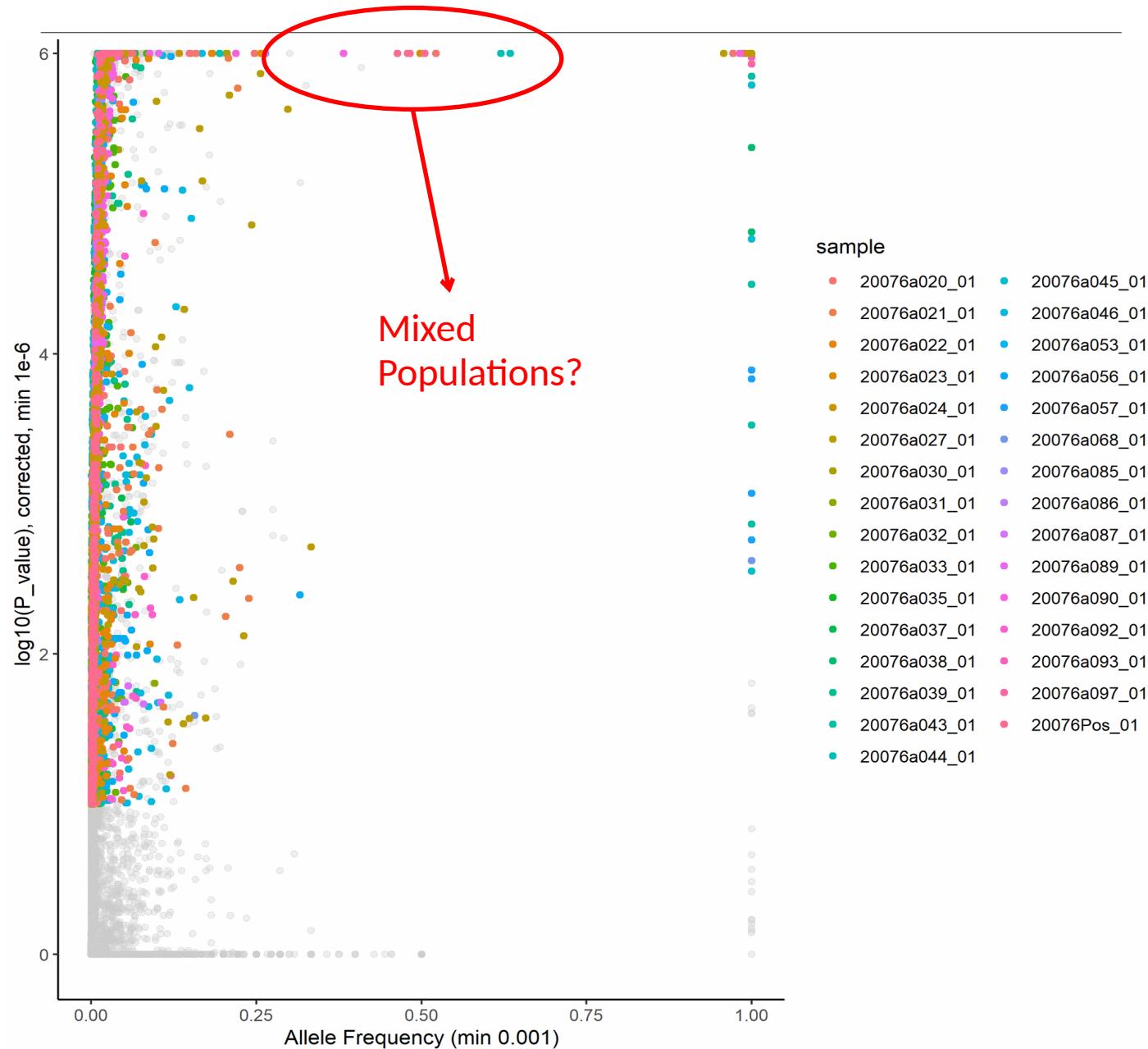
Michaela Pogoda

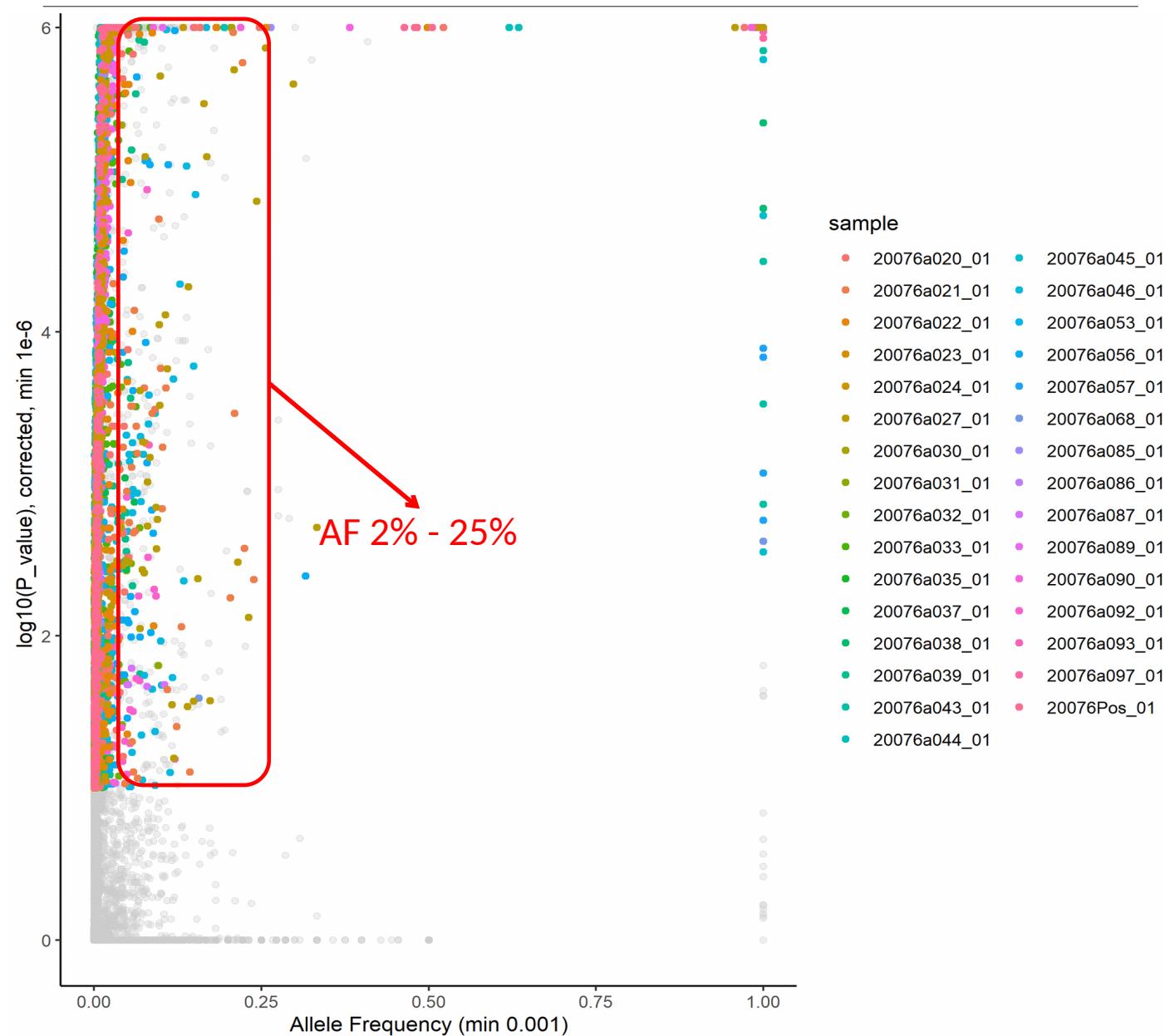


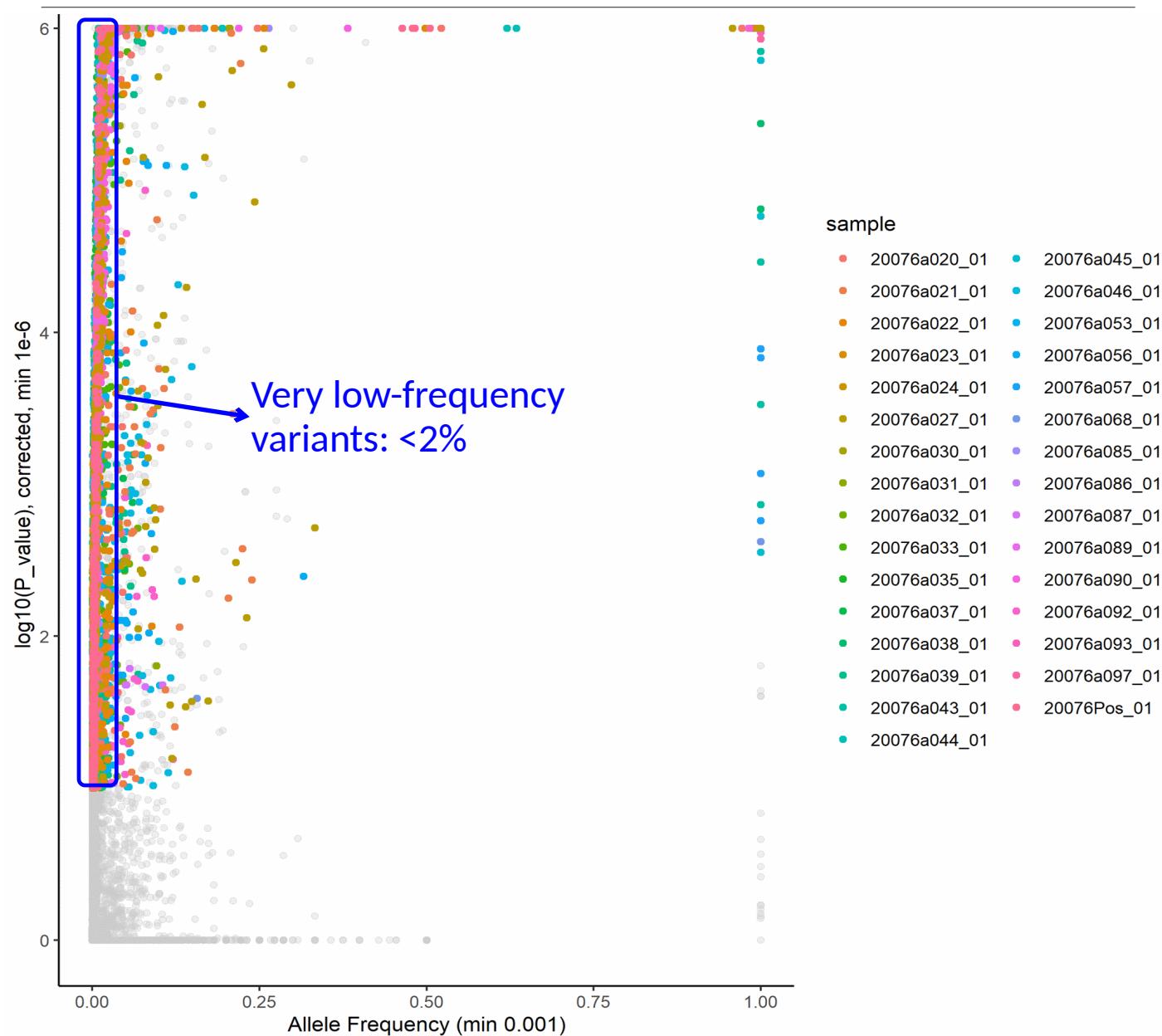
- De-duplication = merging of PCR duplicates using UMIs
- Leads to ‘true coverage’ representing unique RNA molecules
- Errors are corrected during de-duplication
- Sample 35: **226,000x => 25,000x coverage**







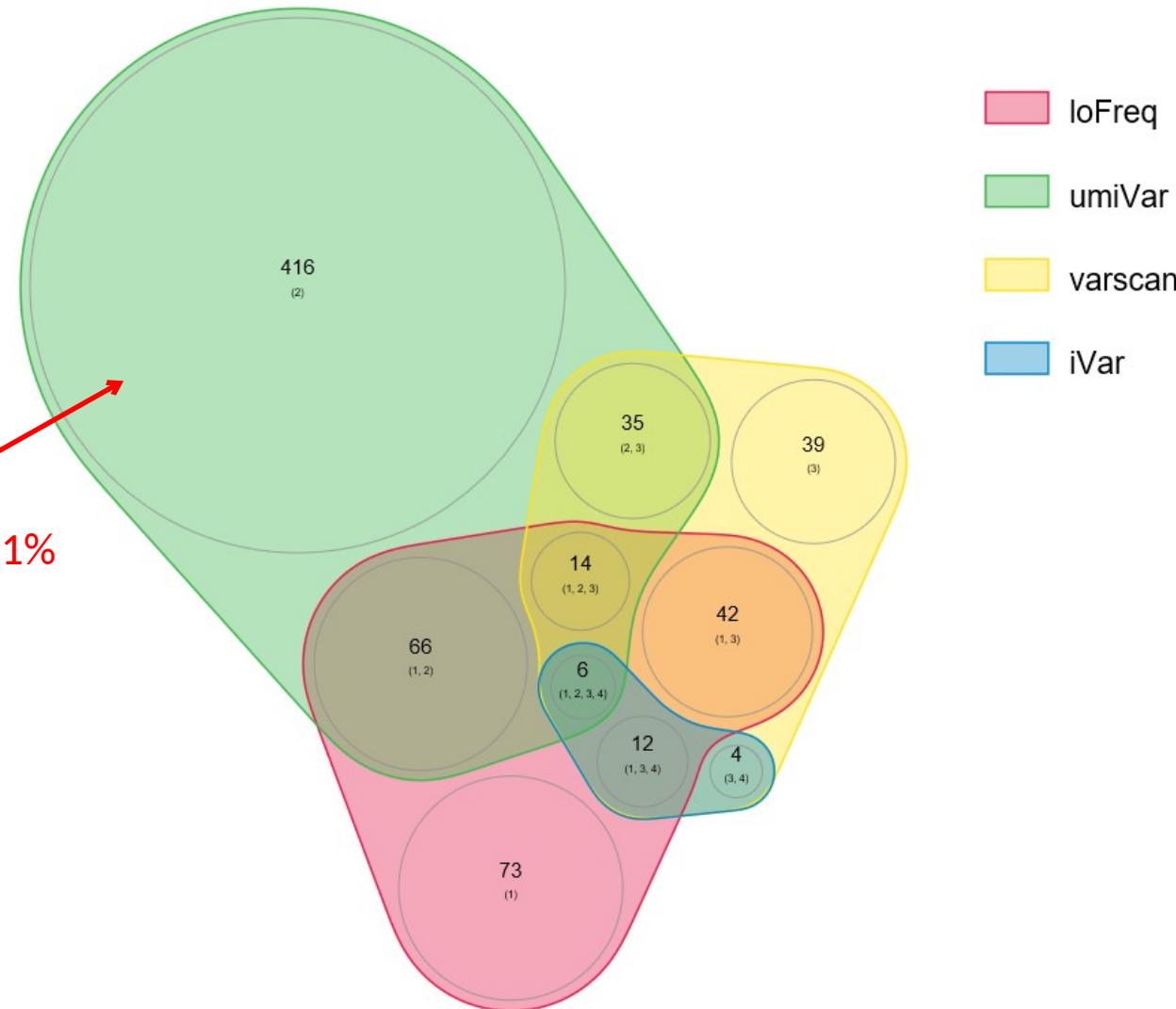


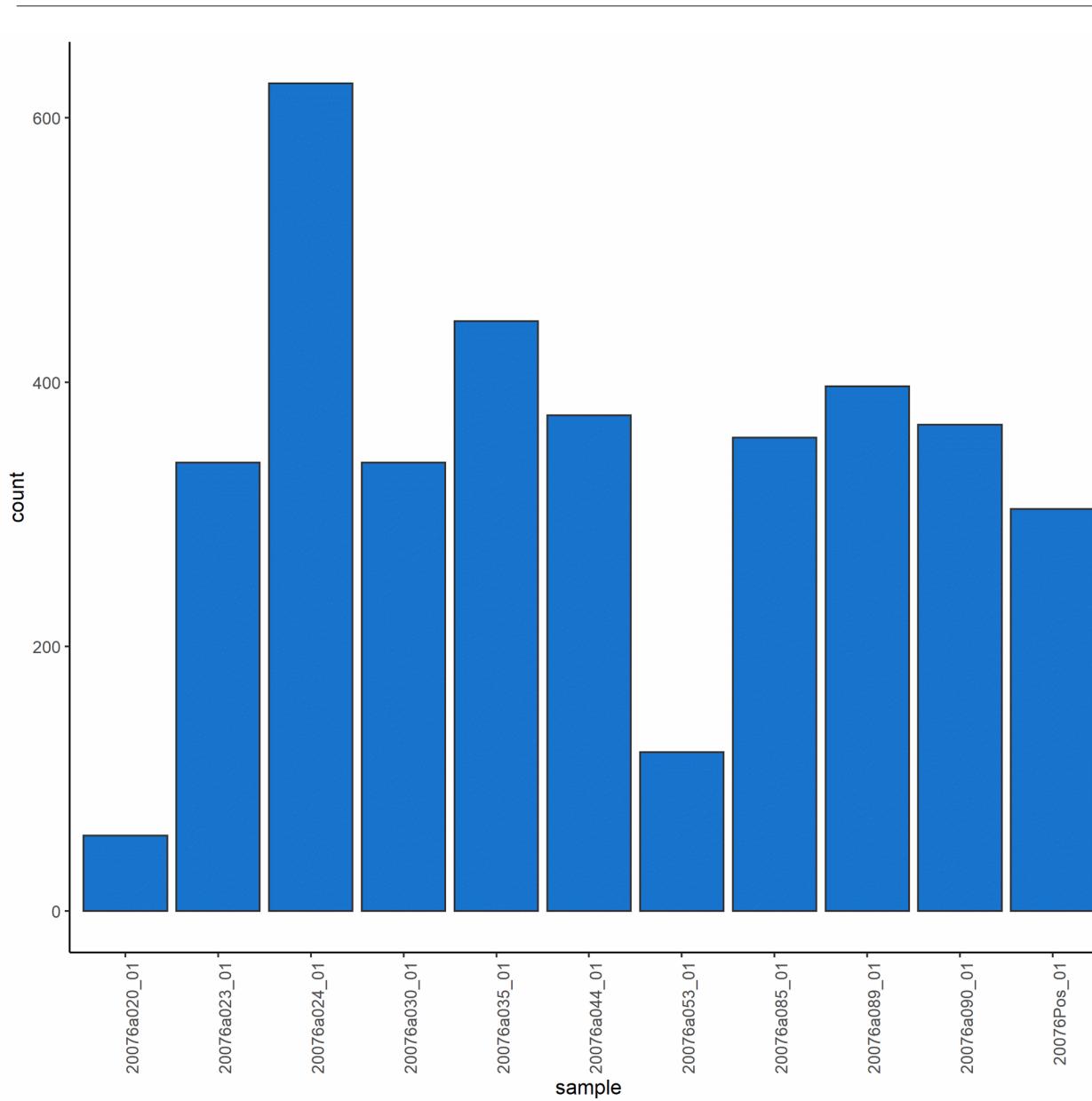




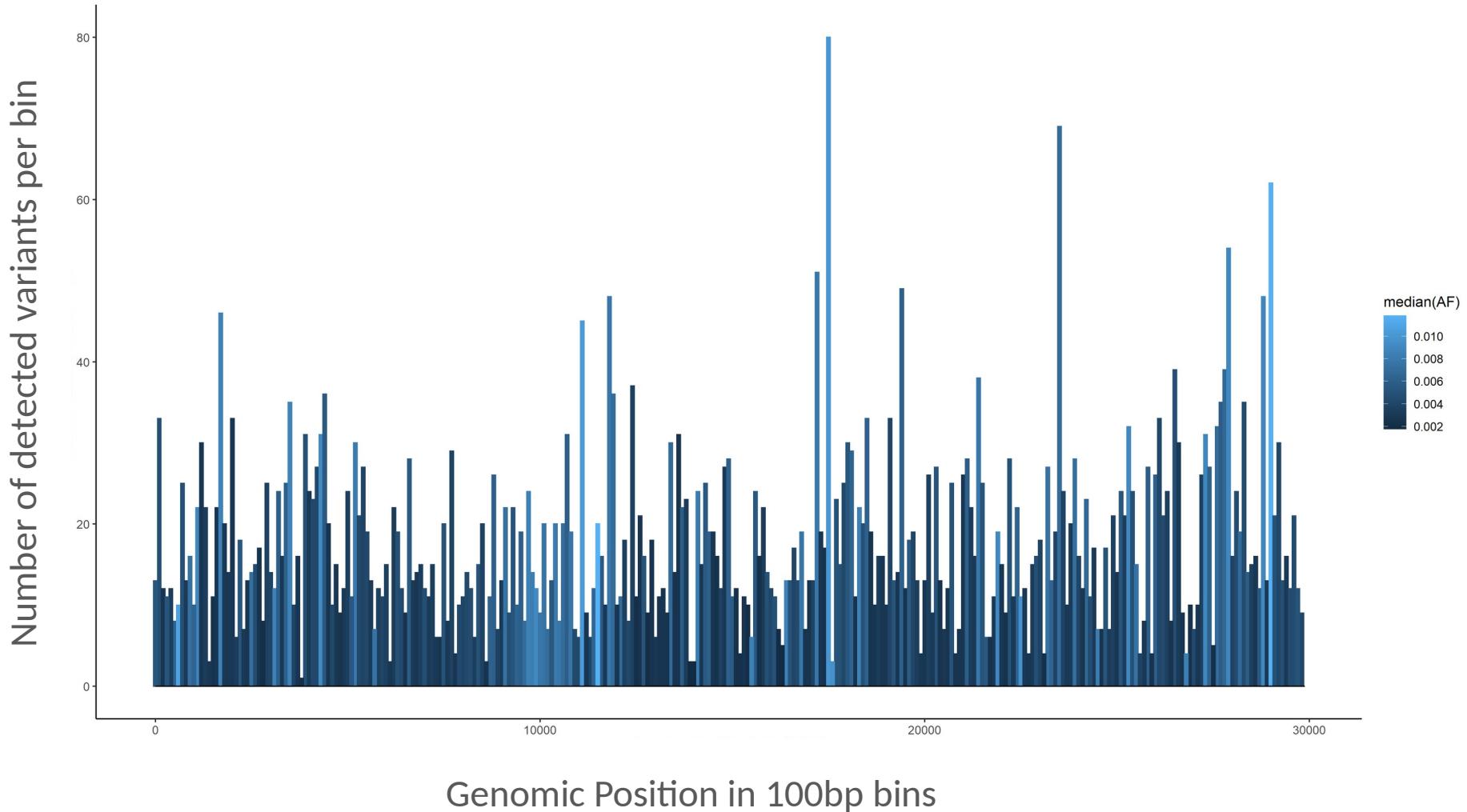
Sample 35

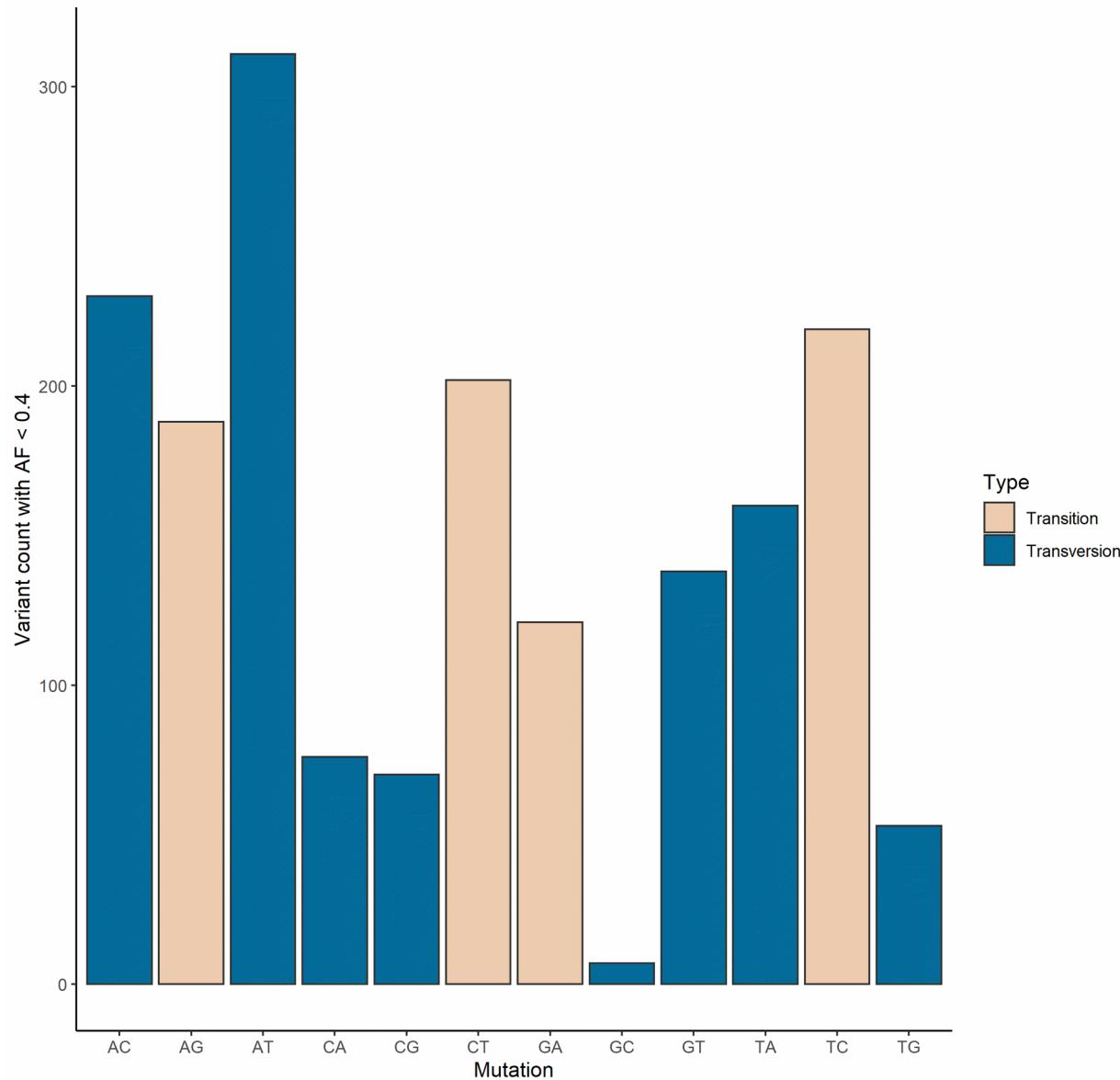
Allele Fraction < 1%



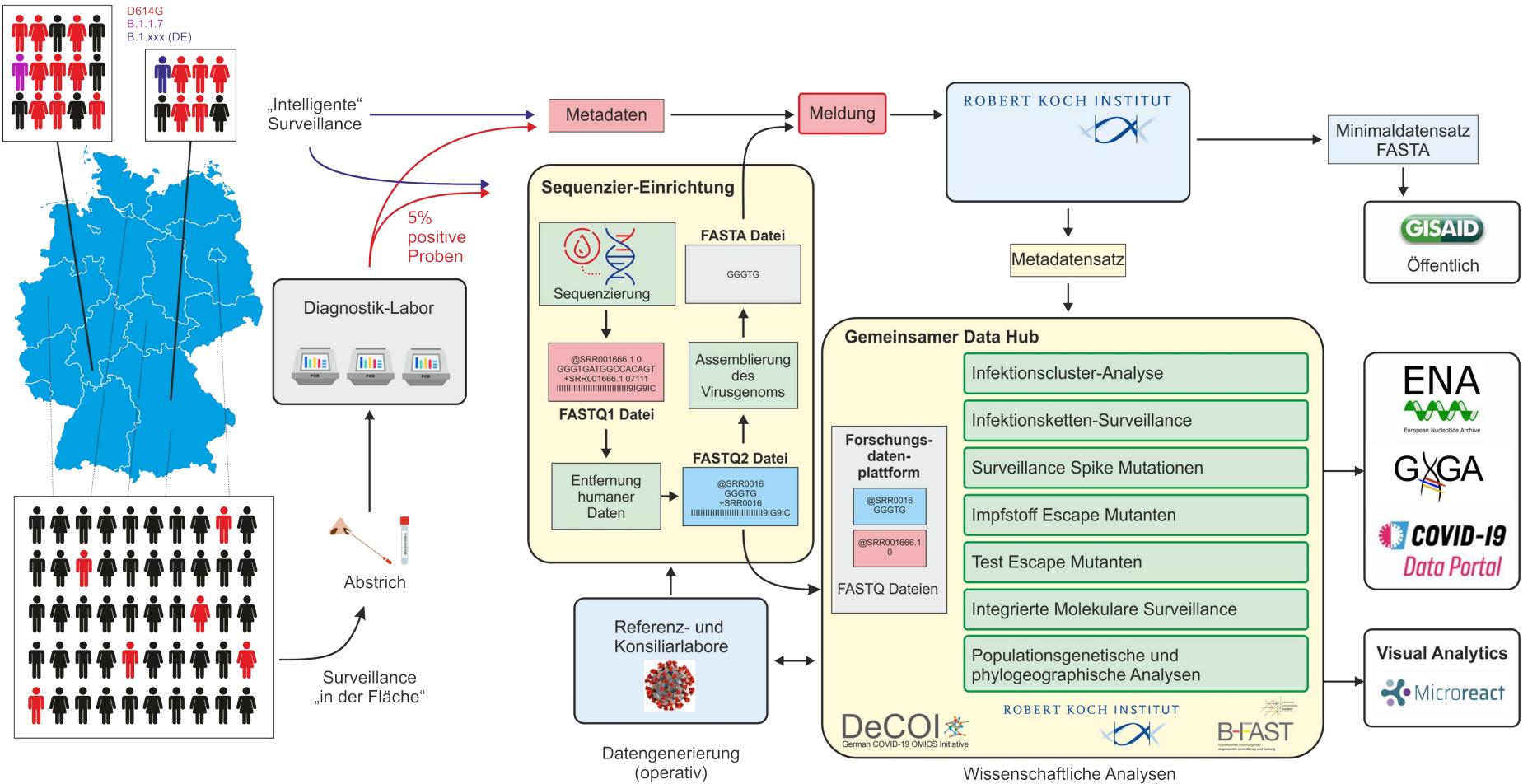


Distribution of low-AF variants across 32 Samples





Outlook: Federated Molecular Surveillance in Germany



Thanks!



Tübingen (NCCT, Microbiology, Human Genetics, QBiC):

Michaela Pogoda
Angel Angelov
Michael Sonnabend
Marie Gauder
Silke Peter
Daniela Bezdan
Nicolas Casadei
Olaf Riess

Düsseldorf:

Alexander Dilthey
Maximilian Damagnez
Ashley-Jane Duplessis
Torsten Feldt
Torsten Houwaart
Lisanna Hülse
Björn Jensen
Malte Kohns Vasconcelos
Susanne Kolbe-Busch
Jessica Nicolai
Klaus Pfeffer
Daniel Strelow
Jörg Timm
Andreas Walker
Tobias Wienemann

Twist:

Tina Han
Ulrike Harms
Oliver Stephan

Tübingen Bioinfo:

Caspar Gross
Leon Schütz
Jakob Admard
Ersoy Kocak

Tübingen Virology:

Tina Ganzenmüller
Angelika Iftner
Thomas Iftner
Tübingen ICU:
Siri Göpel
Michael Bitzer

TU Munich

(Virology, Computer Science):
Elisabeth Esser
Julien Gagneur
Dieter Hoffmann
Alexander Karollus
Christian Mertes
Thomas Michler
Ulrike Protzer
Eva-Christina Schulte

Charité:

Michael Hummel

CeBiTec Bielefeld:

David Brandt
Tobias Busche
Markus Haak
Jörn Kalinowski
Levin-Joe Klages
Alexander Sczyrba
Marina Simunovic
Svenja Vinke

Ev. Klinikum Bethel, Bielefeld:

Dr. Christiane Scherer

MVZ Diamedis, Sennestadt:

Dr. Thomas Diedrich

MVZ Eberhard und Partner, Dortmund:

Dr. Frank-Jörg Vorhölter

Saarland University:

Kathrin Kattler
Jörn Walter

DeCOI & NGS-CN:*

RKI:

Oliver Drechsel
Stephan Fuchs
Max von Kleist
Djin Ye Oh
Aleksandar Radonic
Maureen Smith
Andrea Theumer
Wiep van der Toorn
Maria Trofimova

Max Planck Institute SHH:

Denise Kühnert
Ariane Weber

DRESDEN-concept

Genome Center:

Andreas Dahl
Katrin Sameith-Lauber
Sylke Winkler

HZI:

Adrian Fritz
Alice McHardy

Cologne:

Konrad Förstner



Deutsche
Forschungsgemeinschaft

“Detecting known and novel mutations in a host's SARS-CoV-2 population by ultra-deep sequencing with unique molecular barcodes”

Questions?

