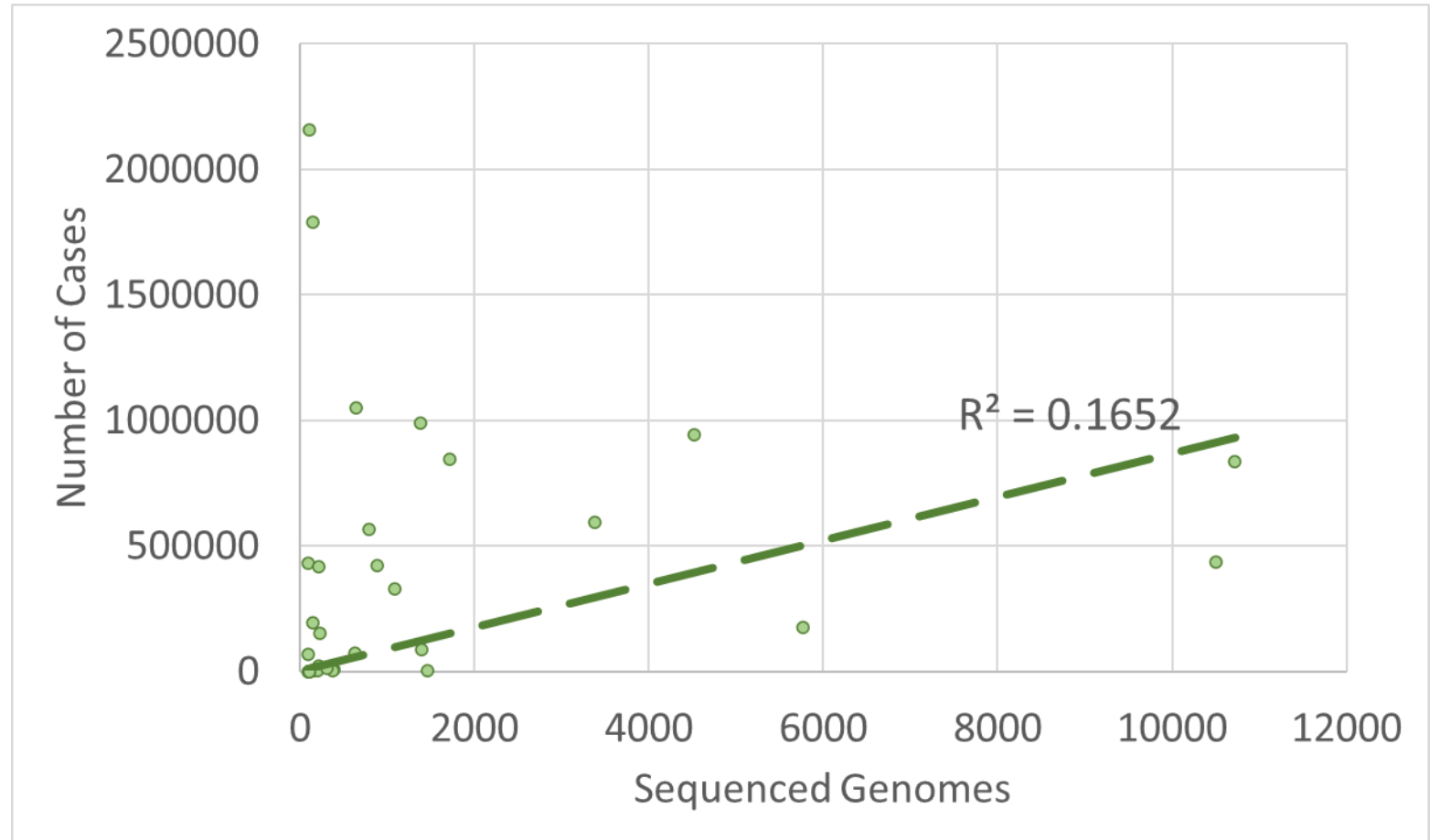


Global Geographic and Temporal Analysis of SARS-CoV-2 Haplotypes Normalized by COVID-19 Cases

23-09-2020

There is not a correlation between the number of cases and sequenced genomes

Each point represent a month-continent combination



Normalization of frequencies by number of COVID-19 cases

$$NRF_p = \frac{\sum_{m,c} (RFP_{m,c} \times NC_{m,c})}{WNC}$$

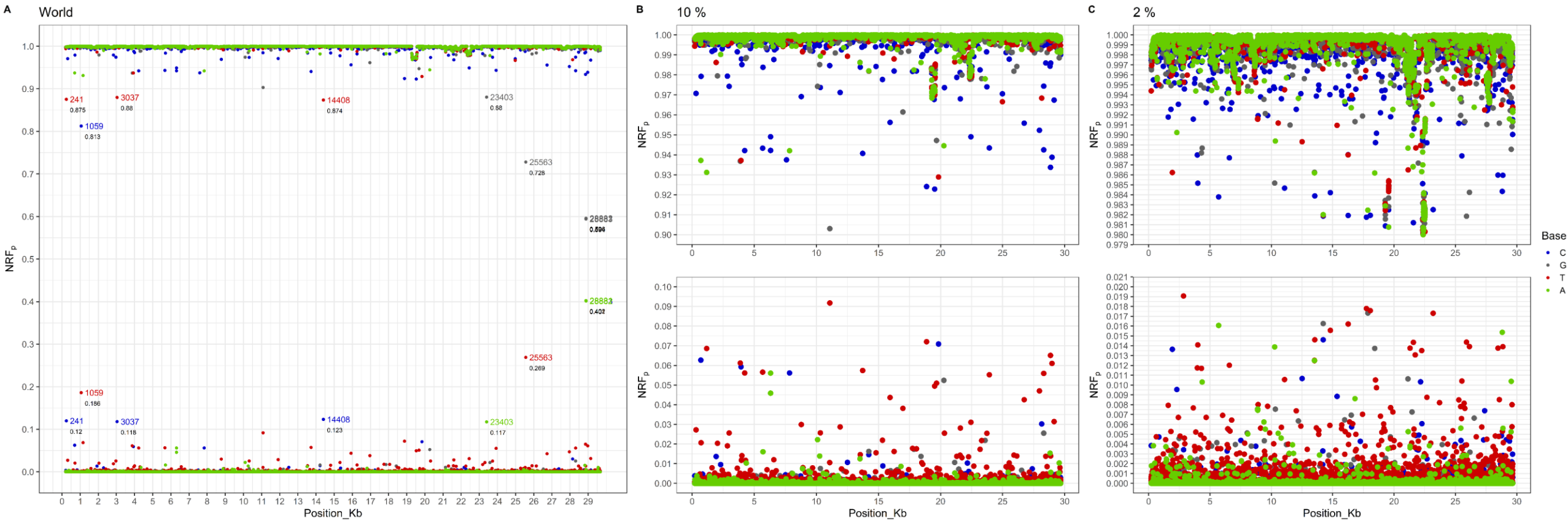
$RFP_{m,c}$ = Relative Frequency on genomic Position (p) in determined month (m) and continent (c).

$NC_{m,c}$ = Number of Cases in determined month (m) and continent (c).

WNC = World total Number of Cases.

NRF_p = Normalized Relative Frequency on genomic position (p).

NRFp analysis in each SARS-CoV-2 genomic position

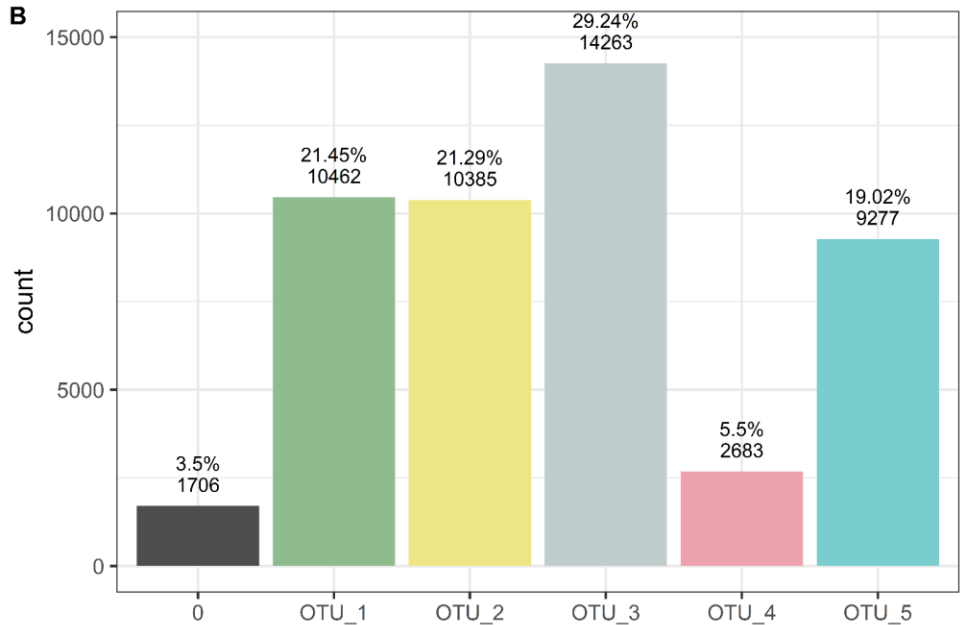


Haplotypes identification and phylogenomic analysis of haplotypes

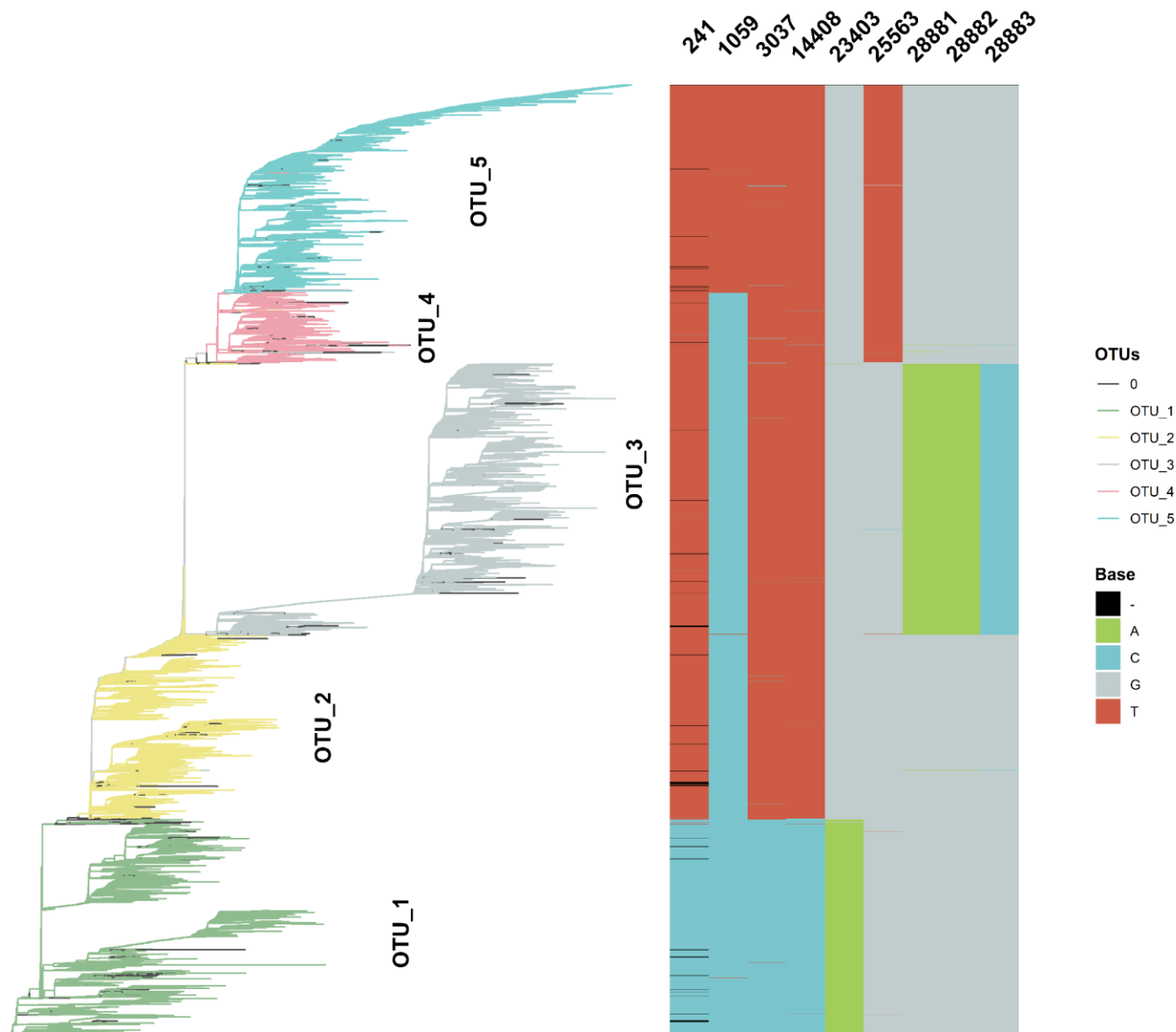
A

Position	OTU_1	OTU_2	OTU_3	OTU_4	OTU_5	Region	AA_change
241	C	T	T	T	T	5'UTR	''
1059	C	C	C	C	T	Nsp2	T85I
3037	C	T	T	T	T	Nsp3	Syn
14408	C	T	T	T	T	Nsp12	P323L
23403	A	G	G	G	G	S	D614G
25563	G	G	G	T	T	Orf8	S84L
28881	G	G	A	G	G	N	R203K
28882	G	G	A	G	G	N	R203K
28883	G	G	C	G	G	N	G204R

B

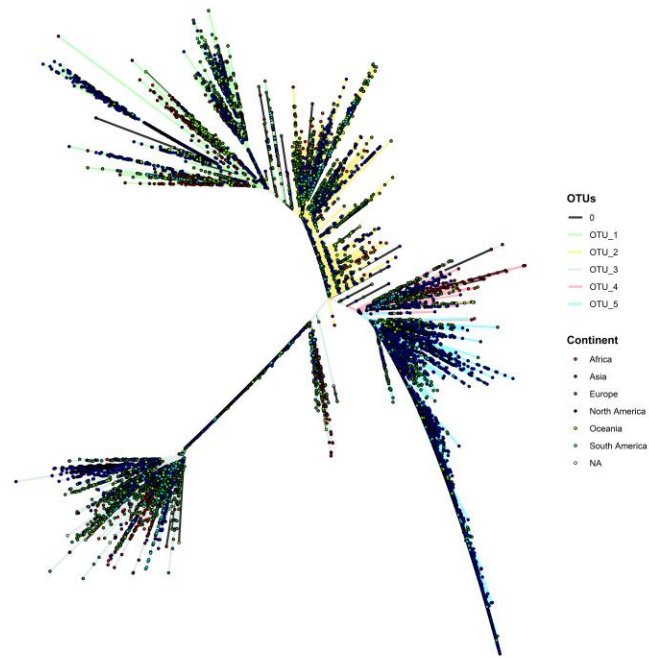


C

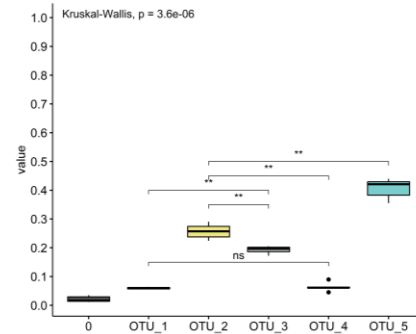


Geographical Analysis of Haplotypes

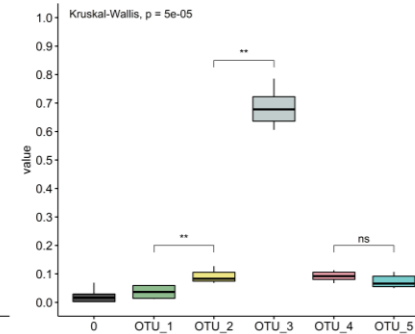
A



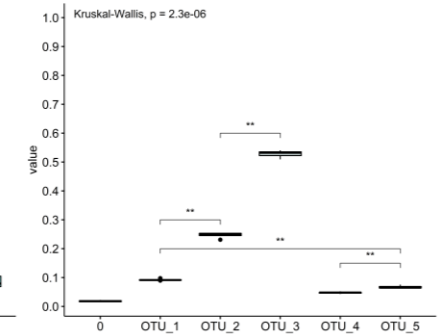
B North America



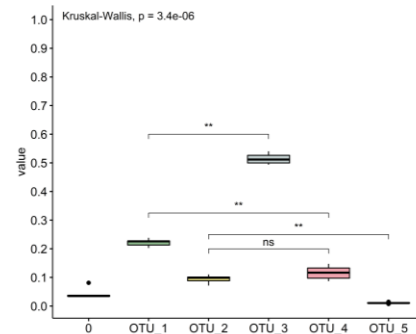
C South America



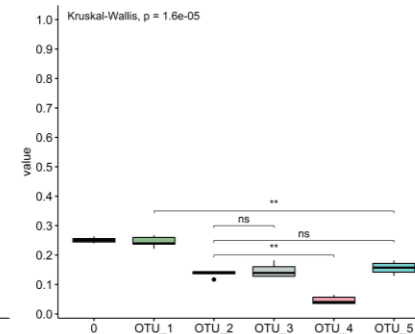
D Europe



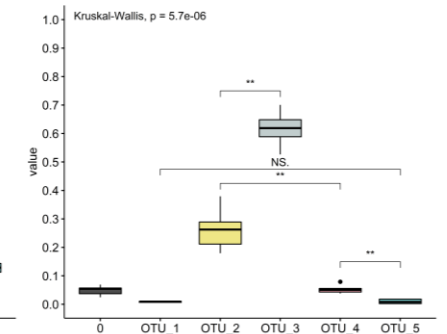
E Asia



F Oceania

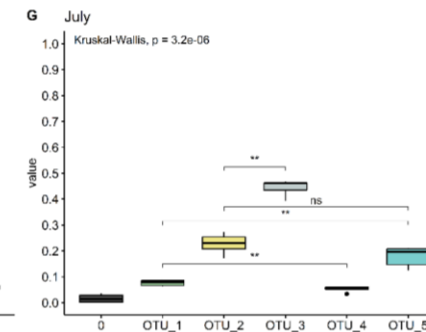
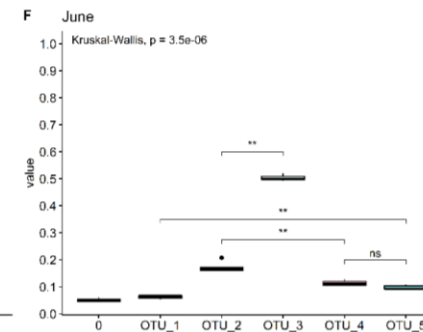
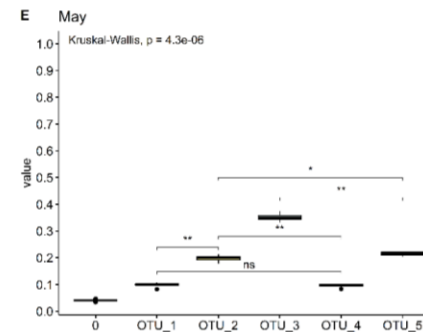
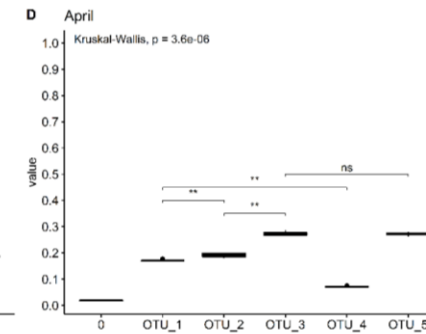
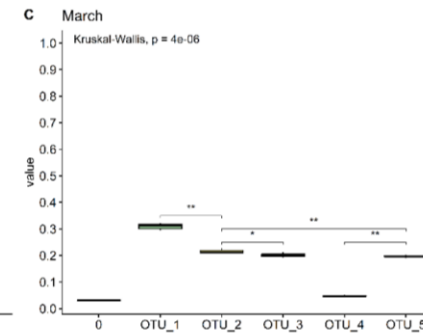
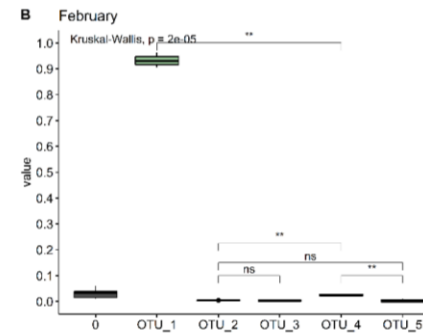
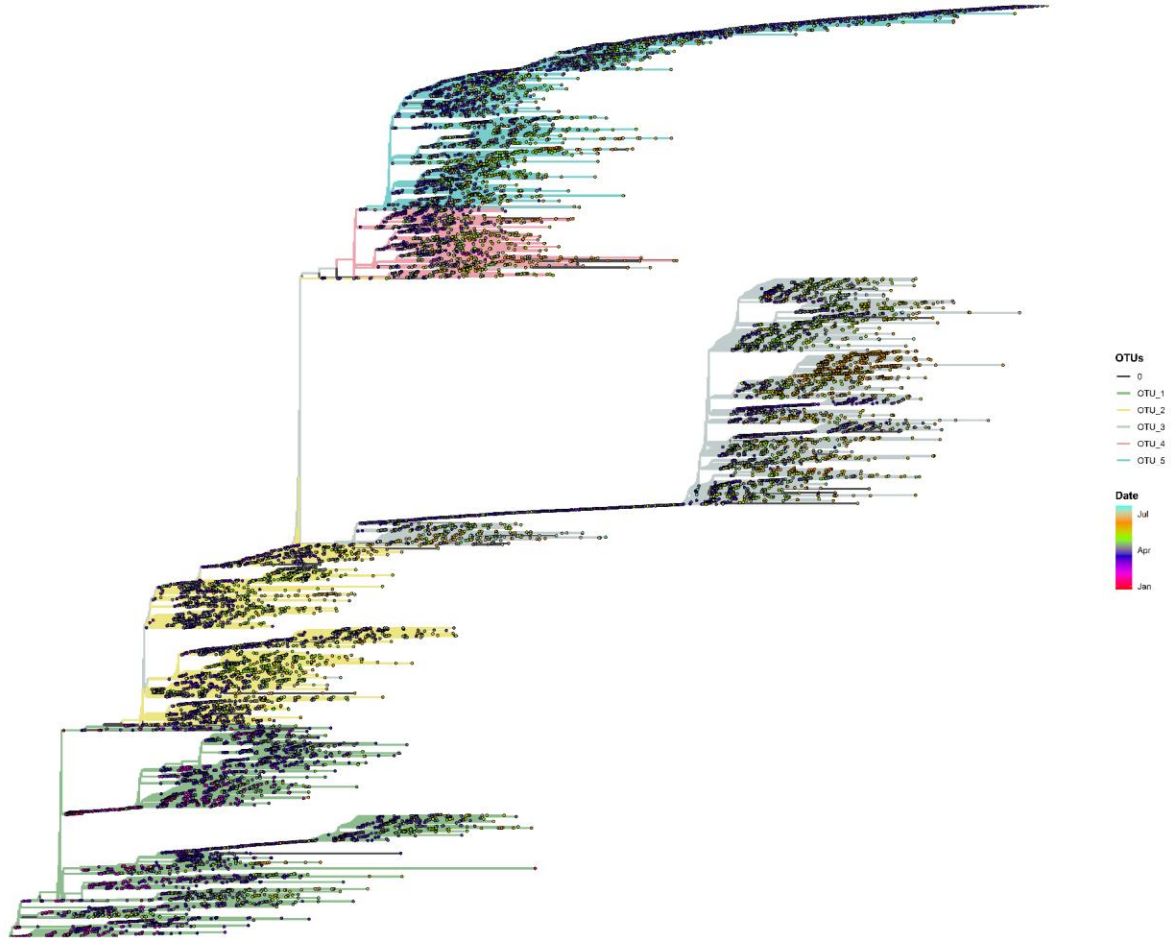


G Africa



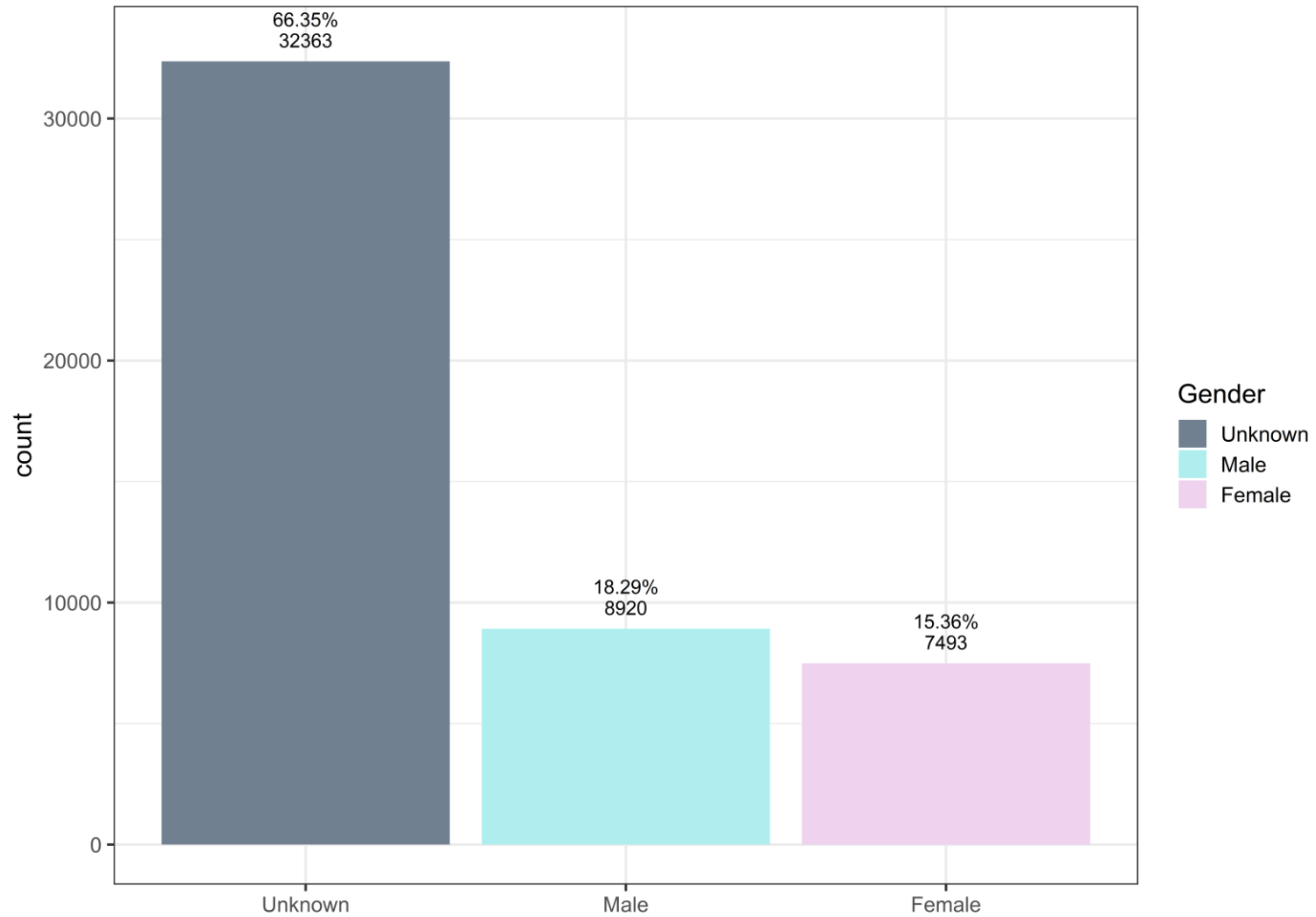
Global temporal analysis of haplotypes

A



Some problems...

Scarce patient data publicly available

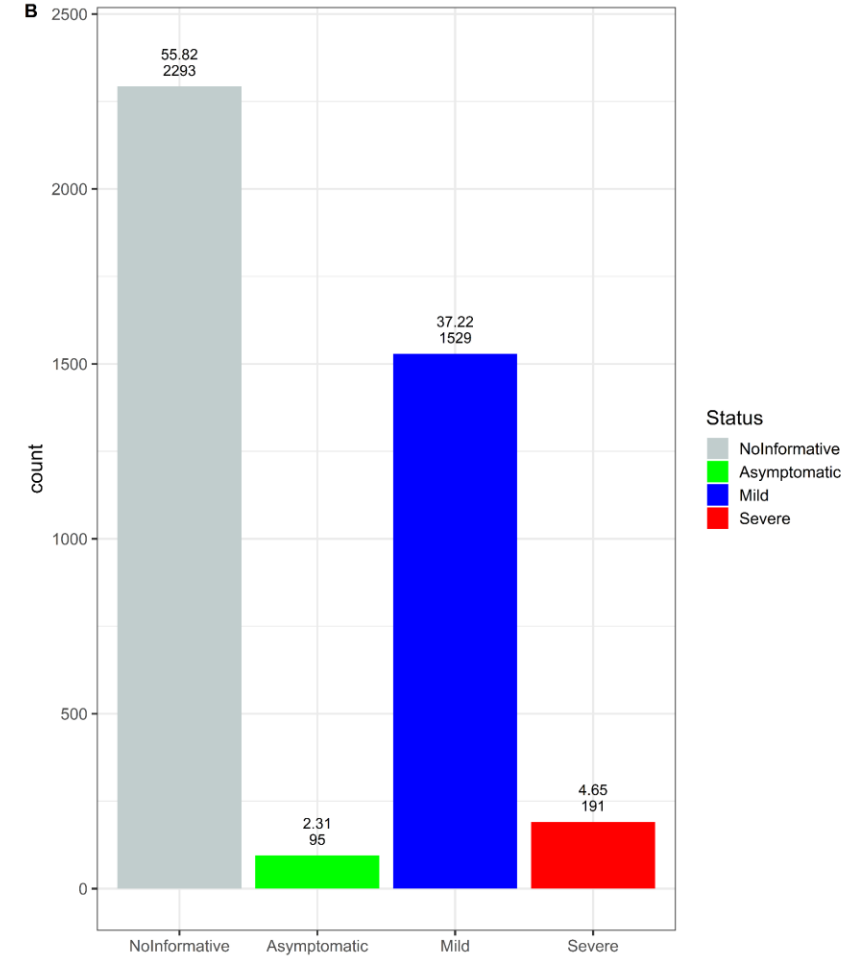


Incorrectly organized patient status data

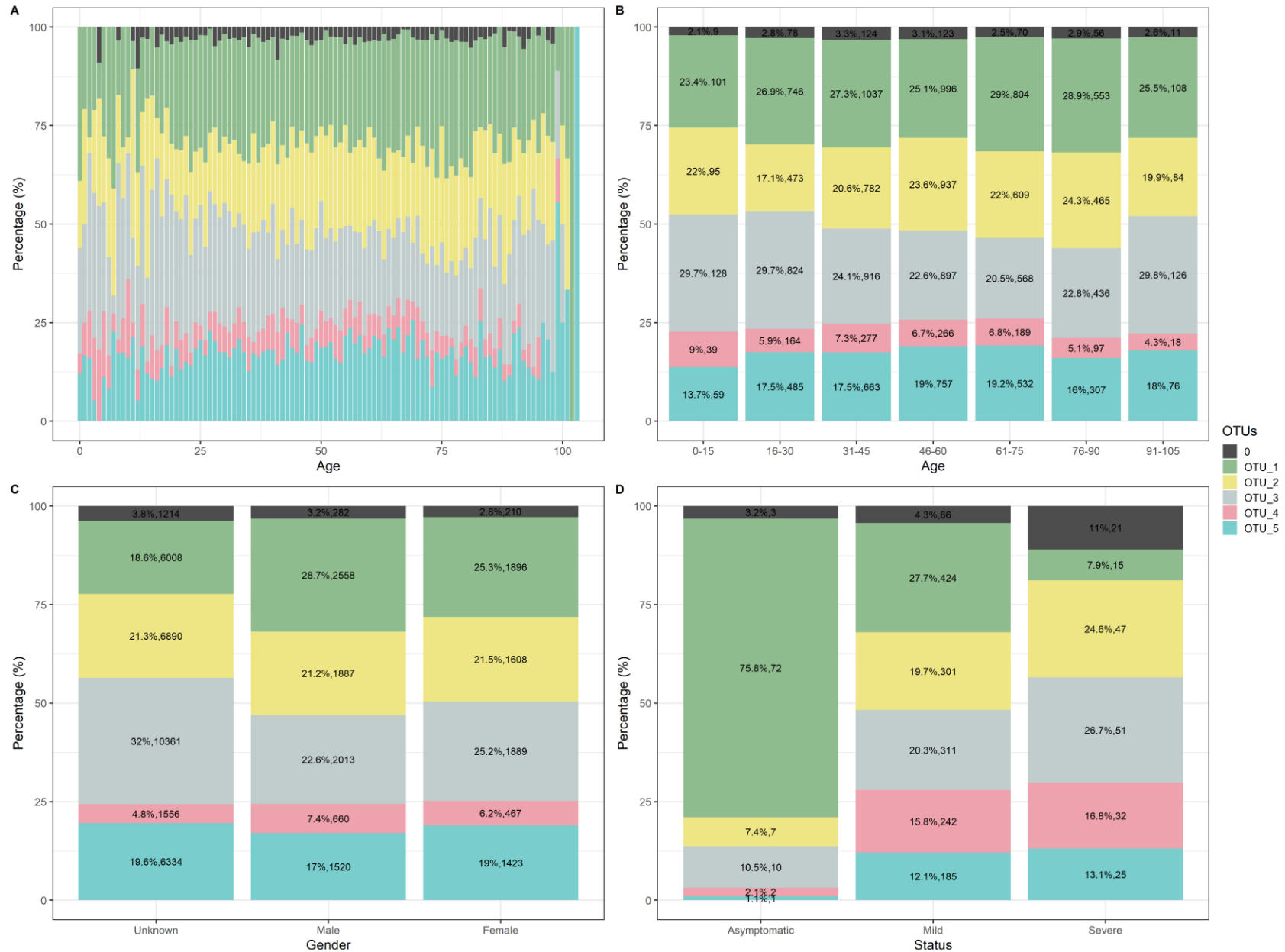
A

Group	Gisaid Categories
Not Informative	Released,Alive,Discharged,Discharged after recovery,Live,Mild/Contact exposure/Asymptomatic, Oro-pharyngeal swab,Recovered,Recovered and Released,Symptomatic,Recovering,Unknown,Z039, Pneumonia (chest X-ray),Screening,Physician network,Not hospitalized,Epidemiology Study,Cured,DAMA, Mild / Contact exposure / Asymptomatic
Asymptomatic	Asymptomatic,Asymptomatic/Released,Asymptomatic,identified as positive during preoperation investigation
Mild	EHPAD,Mild symptoms inpatient for observation,Mild,at home,Moderate / Outpatient,Mild case, Mild clinical signs without hospitalization,Quarantined,Home,Hospitalized,Hospitalized or to be hospitalized, In-hospital.Initially hospitalized, but now improved and discharged,inpatient,Isolation, Mild symptoms (fever, cardiovascular disorders),Hospitalized, oxygenotherapy, diarrhea,Hospitalized, released,ICD-10 Disease: J06.9 Acute upper respiratory infection, unspecified, ICD-10 Disease: Z20.8 Contact with and exposure to other communicable diseases, ICD-10 Disease:J18.9 Pneumonia, unspecified organism, ICD-10 Disease: Z03.8 Observation for other suspected diseases and conditions,Live, acute respiratory infection,Live, mild symptoms, at home,Live, physical examination, Still hospitalized,symptomatic,Symptoms indicative of upper respiratory infection
Severe	ICU,Hospitalized in ICU,Intensive Care Unit,Deceased,Hospitalized/Deceased,Death

B



Relative frequencies of number of genomes by patient data

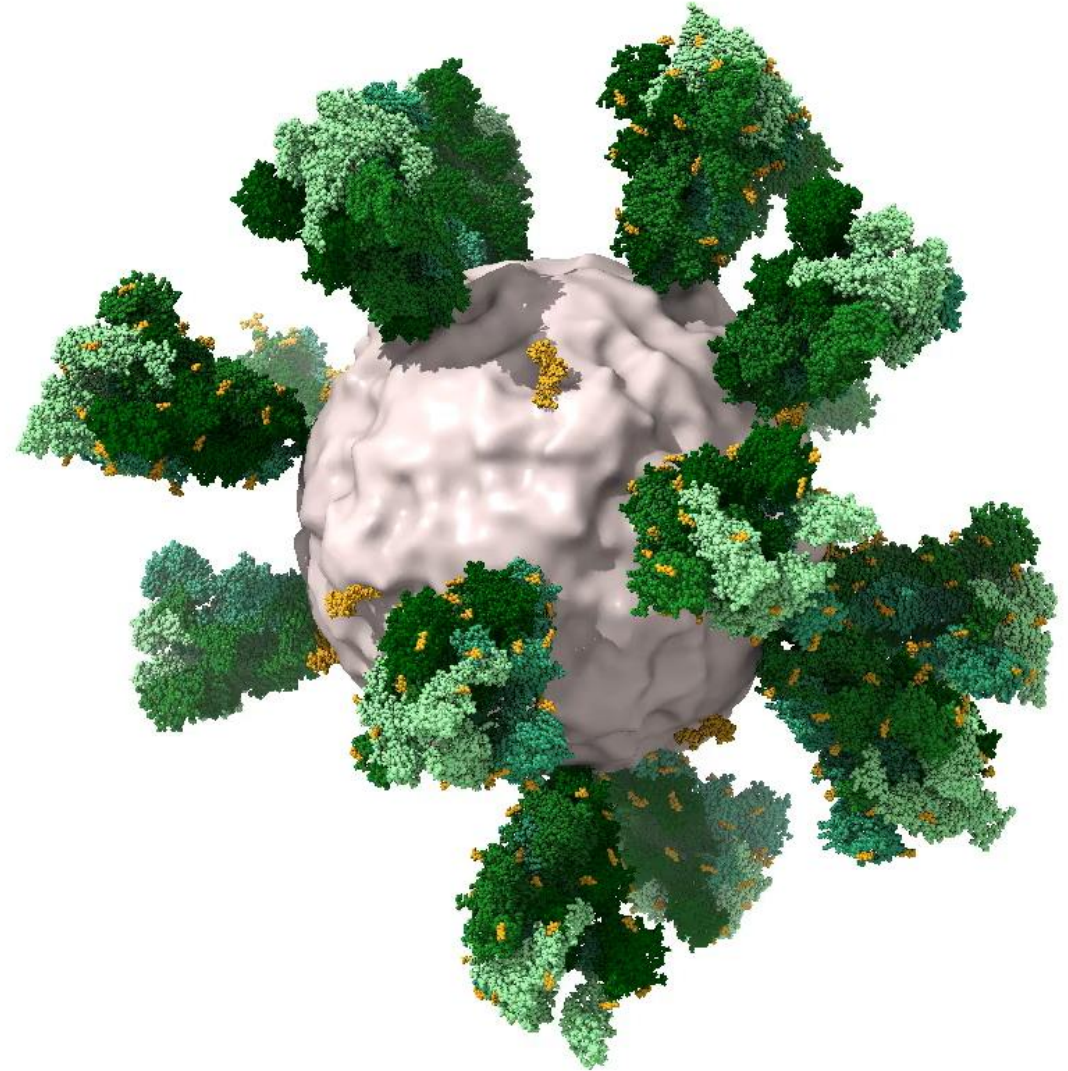


Trying to improve the collection of data related to Patient Status

We created the web-service:

SARS-CoV-2 HaploFinder_URP

<http://sarscov2haplofinder.urp.edu.pe/>



Gracias!