Supplemental Material

Phylogeography of SARS-CoV-2 pandemic in Spain: a story of multiple introductions, micro-geographic stratification, founder effects, and super-spreaders

Alberto Gómez-Carballa, Xabier Bello, Jacobo Pardo-Seco, Pérez del Molino-Bernal ML, Federico Martinón-Torres, Antonio Salas

Figure S1. Accumulative counts of Spanish genomes in GISAID.

Figure S2. Features of HQ vs LQ datasets related to quality of sequence data.

Figure S3. Frequency of main haplogroups by Spanish regions.

Figure S4 to S10. Lineages observed in Spanish regions. Only those with >30 genomes analyzed are shown. Large regional maps were downloaded from https://d-maps.com/ and can be used freely according to https://d-maps.com/conditions.php?lang=es. Map of a) Andalusia: https://d-maps.com/carte.php?num_car=13251&lang=es; b) Basque Country: https://d-maps.com/carte.php?num_car=13481&lang=pt; c) Catalonia: https://d-maps.com/carte.php?num_car=13853&lang=es; d) Galicia: https://d-maps.com/carte.php?num_car=14017&lang=es; e) Madrid: https://d-maps.com/carte.php?num_car=14101&lang=es; f) Navarra: https://d-maps.com/carte.php?num_car=14301&lang=es; and g) Valencian Community: https://d-maps.com/carte.php?num_car=14399&lang=es. Small Spanish maps were downloaded from Wikipedia.

Figure S11. Sampling points of interpolated maps shown in Figure 2 (main text).

Figure S12. Diversity indices for different Spanish regions. (A) Haplotype diversity, (B) nucleotide diversity, and (C) accumulated sequences. Note, discontinuity in sampling occurred from 15 April, however this did not impact our study as most analyses focused on the beginning of the pandemic and the time right after lockdown on 15 March.

Figure S13. Network of all Spanish genomes.

Figure S14. Boxplot and density plot of genome lifespan of identical haplotypes in GISAID database.

Figure S15. Interpolation map of Spanish haplogroup A2 frequencies, which roughly matches frequency of D614G. Map on top indicates sampling points.

Figure S16 to S21. ML phylogenies and TMRCA for most common Spanish clades.

Figure S1.

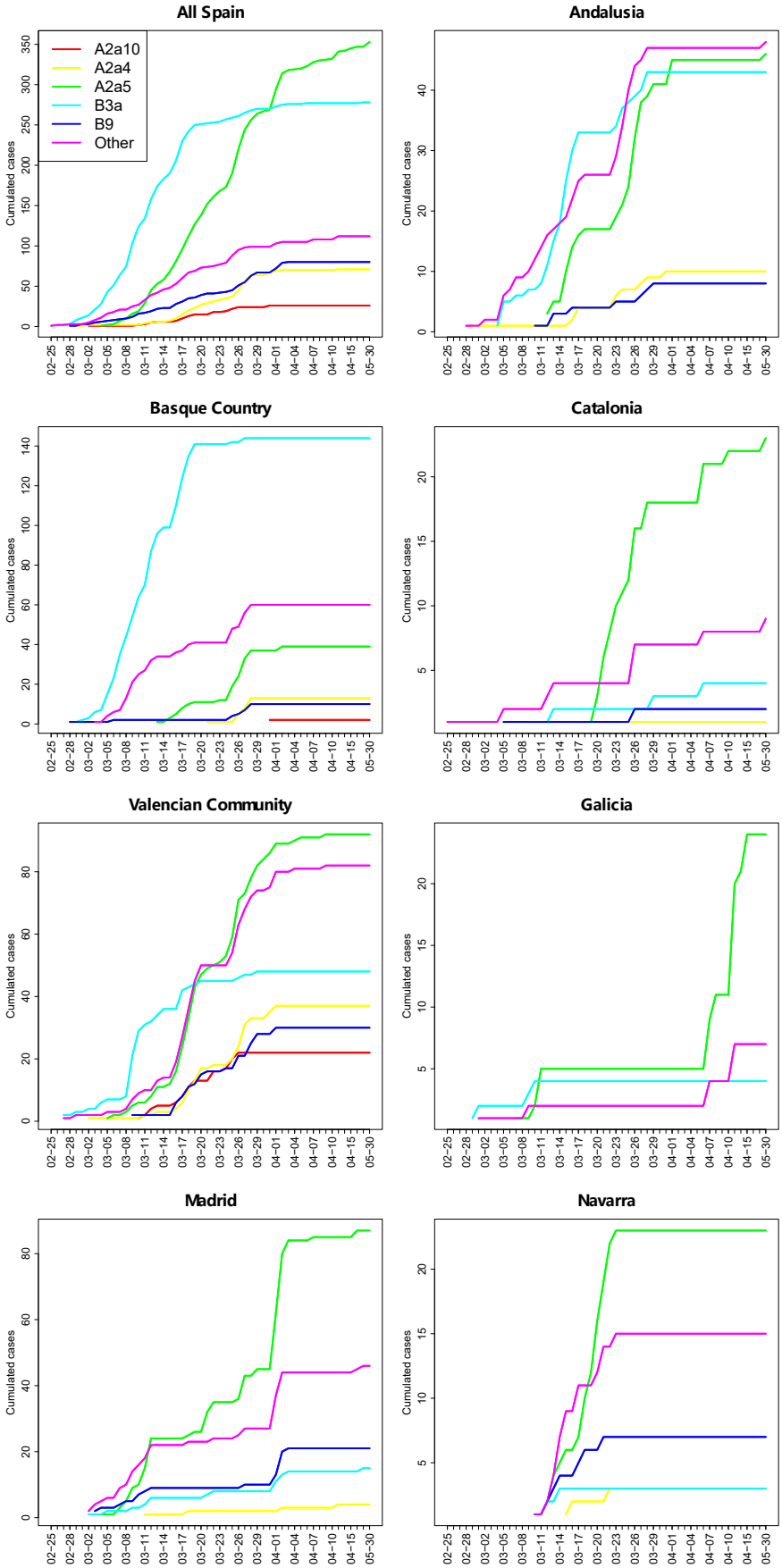


Figure S2.

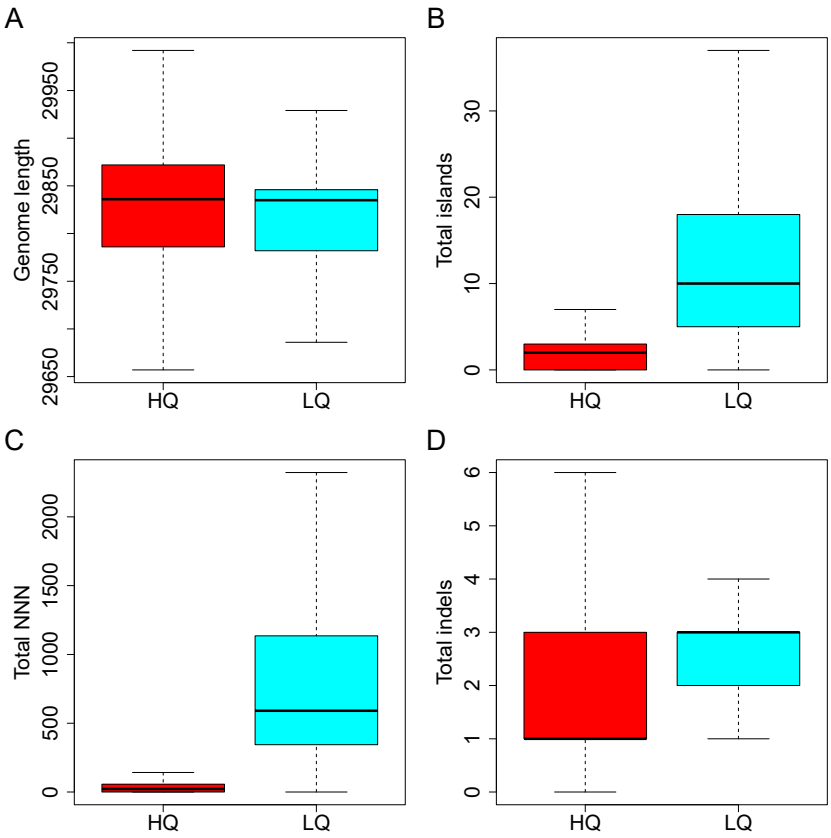

A
B
C
D
Genome length
29650
29750
29850
29950
Total islands
0
10
20
30
Total NNN
0
500
1000
1500
2000
Total indels
0
1
2
3
4
5
6
HQ
LQ

Figure S3.

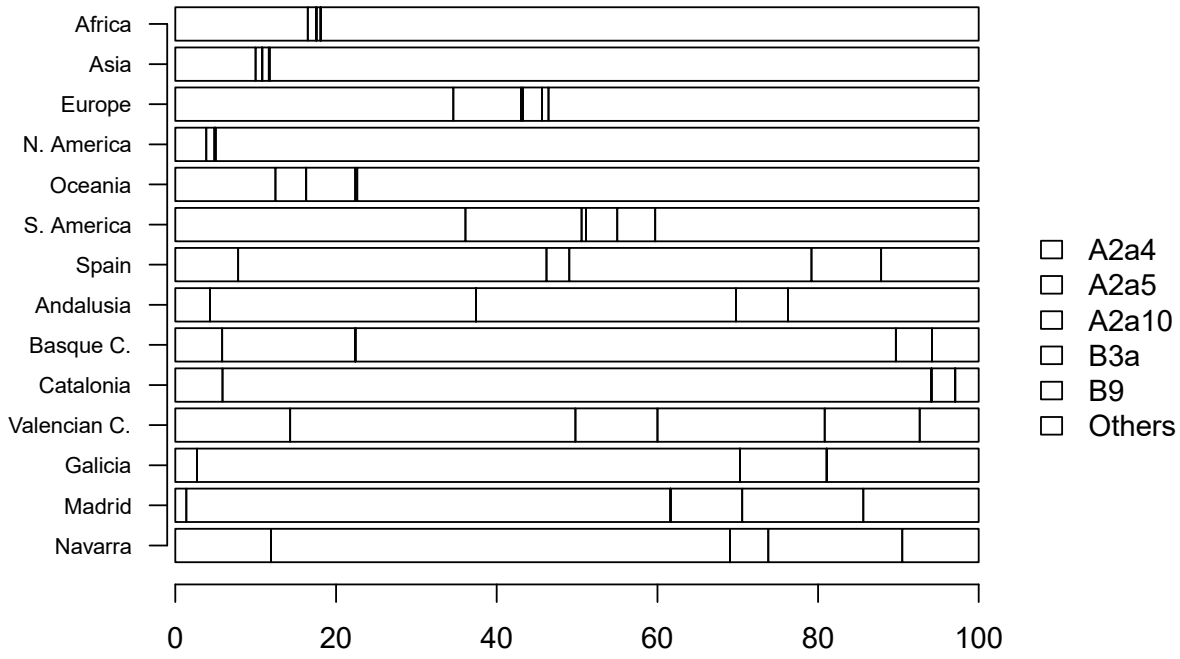

Africa
Asia
Europe
N. America
Oceania
S. America
Spain
Andalusia
Basque C.
Catalonia
Valencian C.
Galicia
Madrid
Navarra
0
20
40
60
80
100
A2a4
A2a5
A2a10
B3a
B9
Others

Figure S4

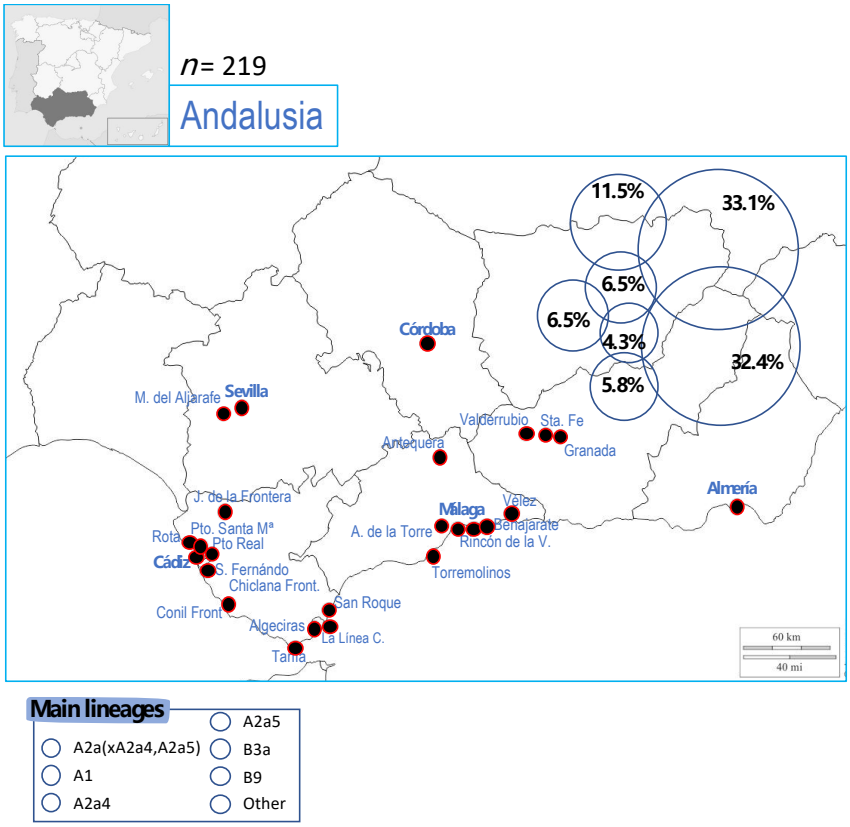


Figure S5

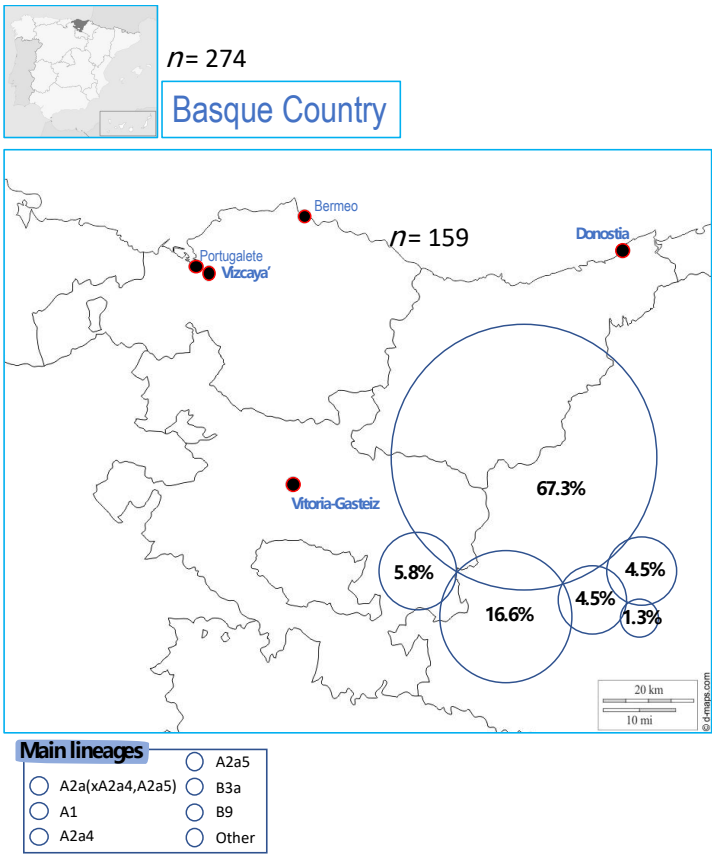


Figure S6

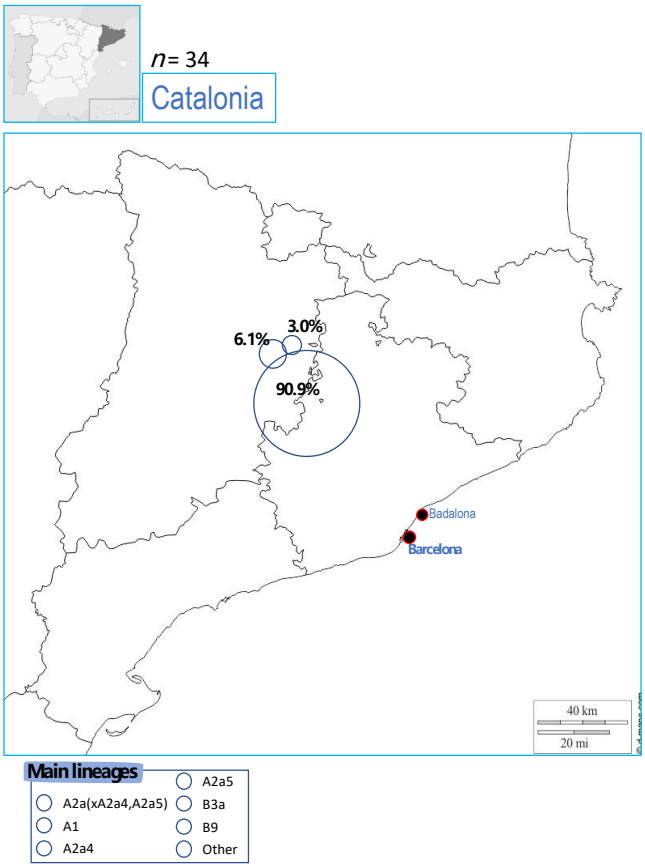

n = 34
Catalonia
6.1%
3.0%
90.9%
Badalona
Barcelona
40 km
20 mi
Main lineages
A2a(xA2a4,A2a5)
A1
A2a4
A2a5
B3a
B9
Other

Figure S7

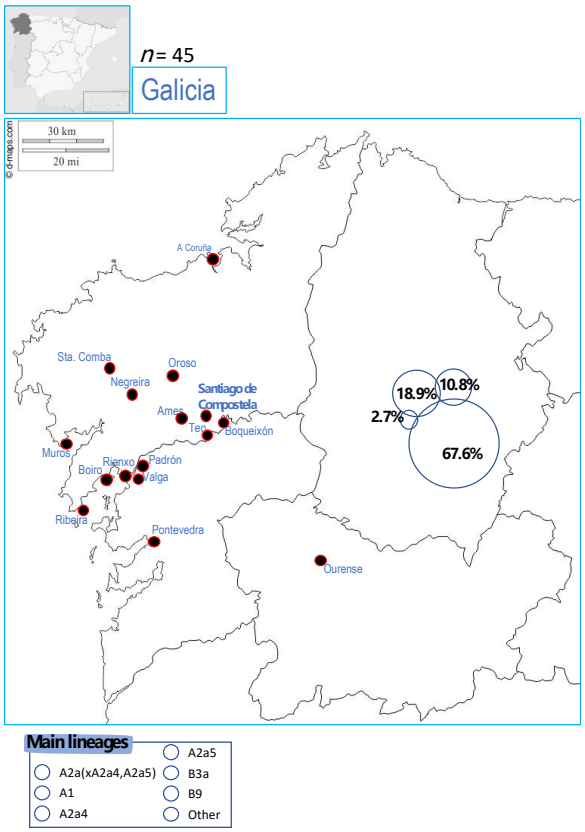

n = 45
Galicia
30 km
20 mi
A Coruña
Sta. Comba
Oroso
Negreira
Santiago de Compostela
Ames
Teo
Boqueixón
Muros
Rianxo
Padrón
Boiro
Valga
Ribeira
Pontevedra
Ourense
18.9%
10.8%
2.7%
67.6%
Main lineages
A2a(xA2a4,A2a5)
A1
A2a4
A2a5
B3a
B9
Other

Figure S8

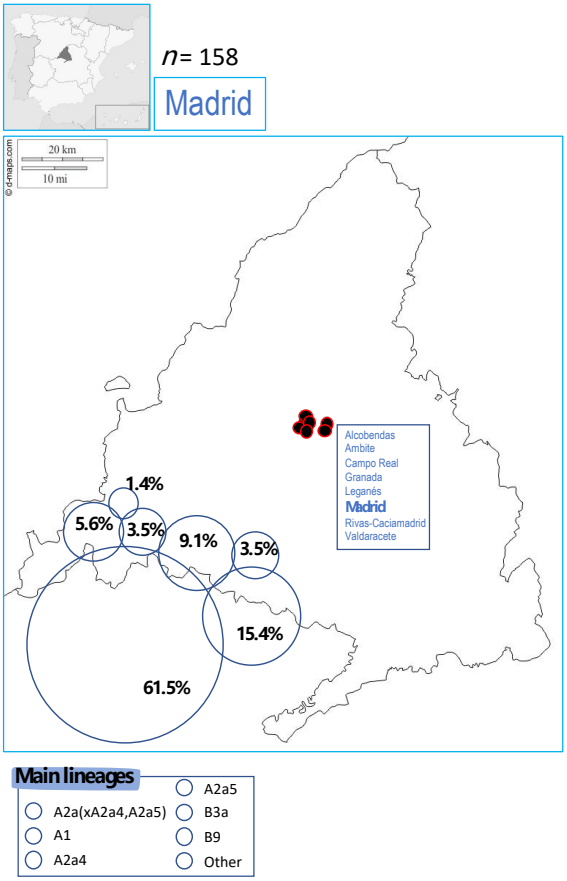


Figure S9

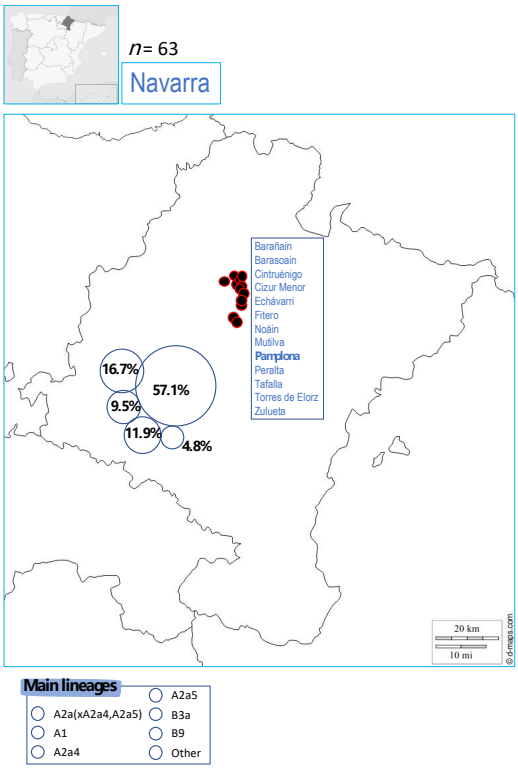


Figure S10

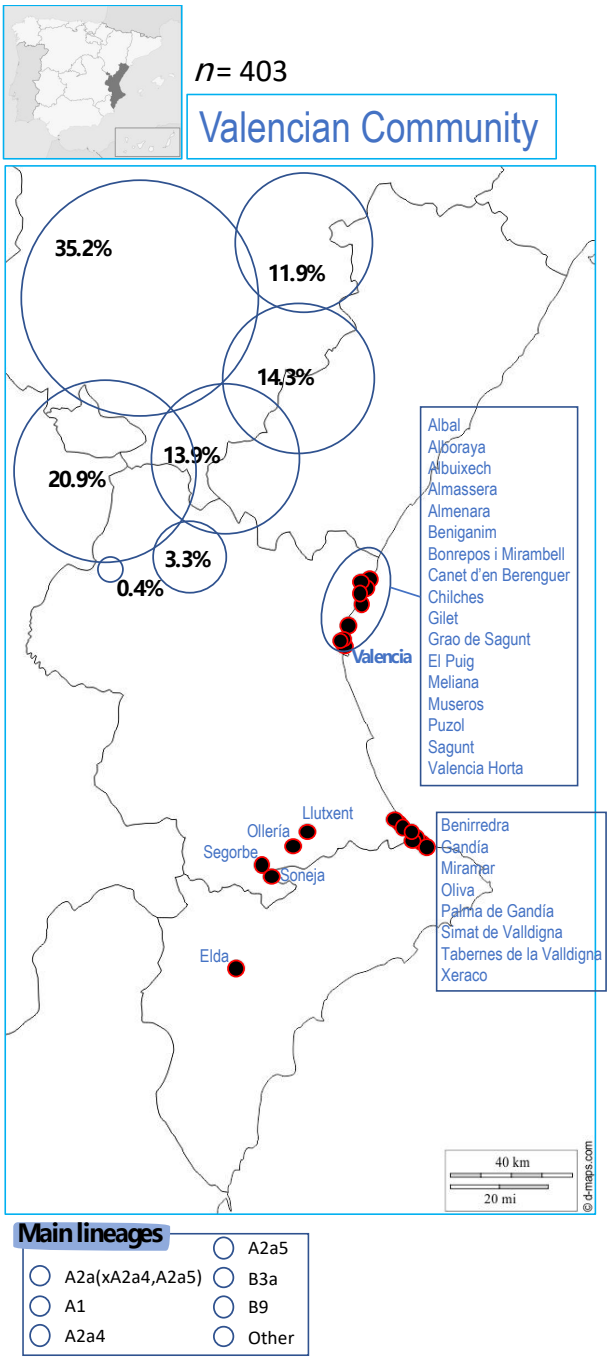

n= 403
Valencian Community
35.2%
11.9%
14.3%
13.9%
20.9%
3.3%
0.4%
Albal
Alboraya
Albuixech
Almassera
Almenara
Beniganim
Bonrepos i Mirambell
Canet d'en Berenguer
Chilches
Gilet
Grao de Sagunt
El Puig
Meliana
Museros
Puzol
Sagunt
Valencia Horta
Valencia
Llutxent
Ollería
Segorbe
Soneja
Benirredra
Gandía
Miramar
Oliva
Palma de Gandía
Simat de Valldigna
Tabernes de la Valldigna
Xeraco
Elda
40 km
20 mi
© d-maps.com
Main lineages
A2a(xA2a4,A2a5)
A1
A2a4
A2a5
B3a
B9
Other

Figure S11.

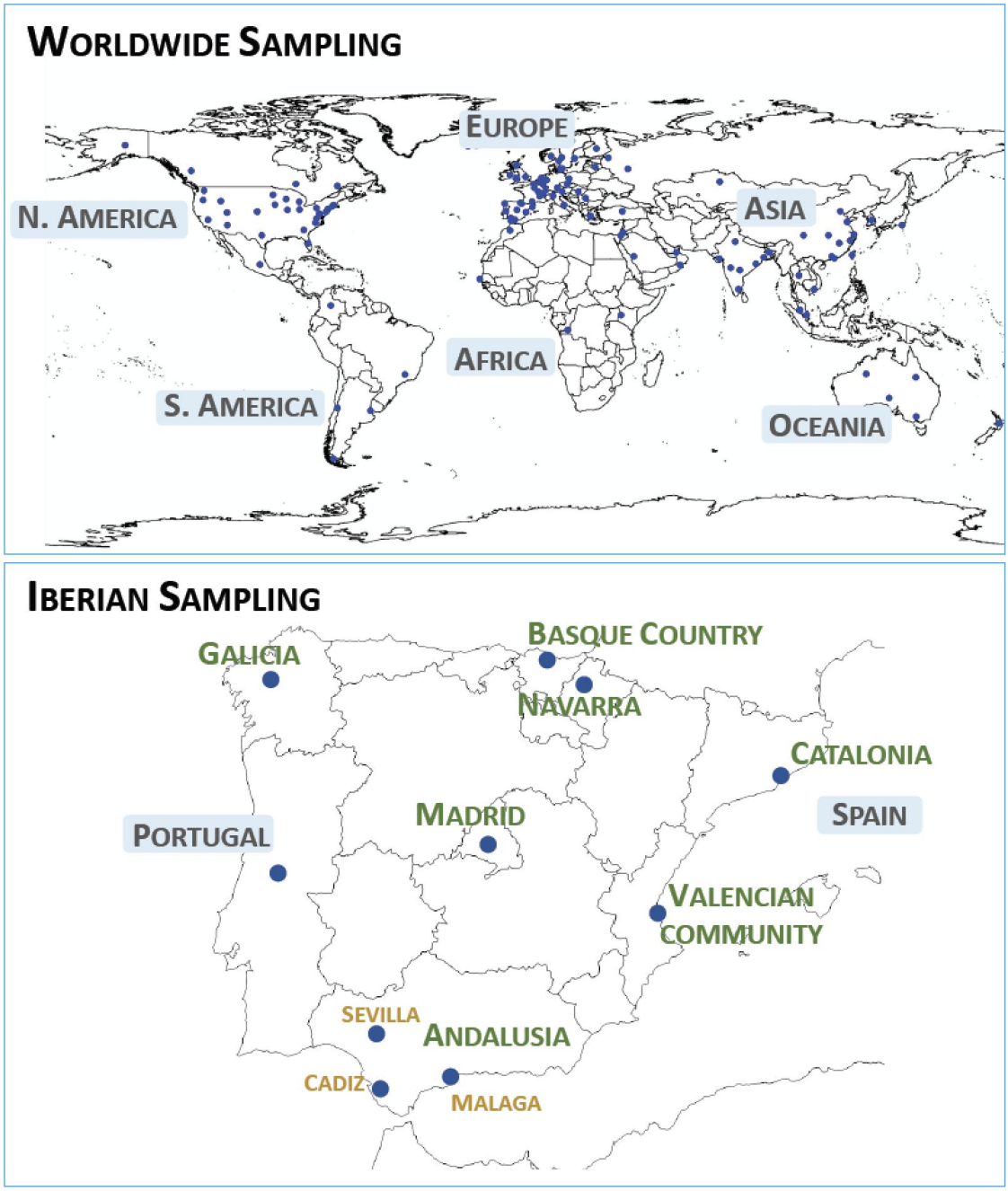

WORLDWIDE SAMPLING
EUROPE
N. AMERICA
ASIA
AFRICA
S. AMERICA
OCEANIA
IBERIAN SAMPLING
BASQUE COUNTRY
GALICIA
NAVARRA
CATALONIA
MADRID
SPAIN
PORTUGAL
VALENCIAN COMMUNITY
SEVILLA
ANDALUSIA
CADIZ
MALAGA

Figure S12.

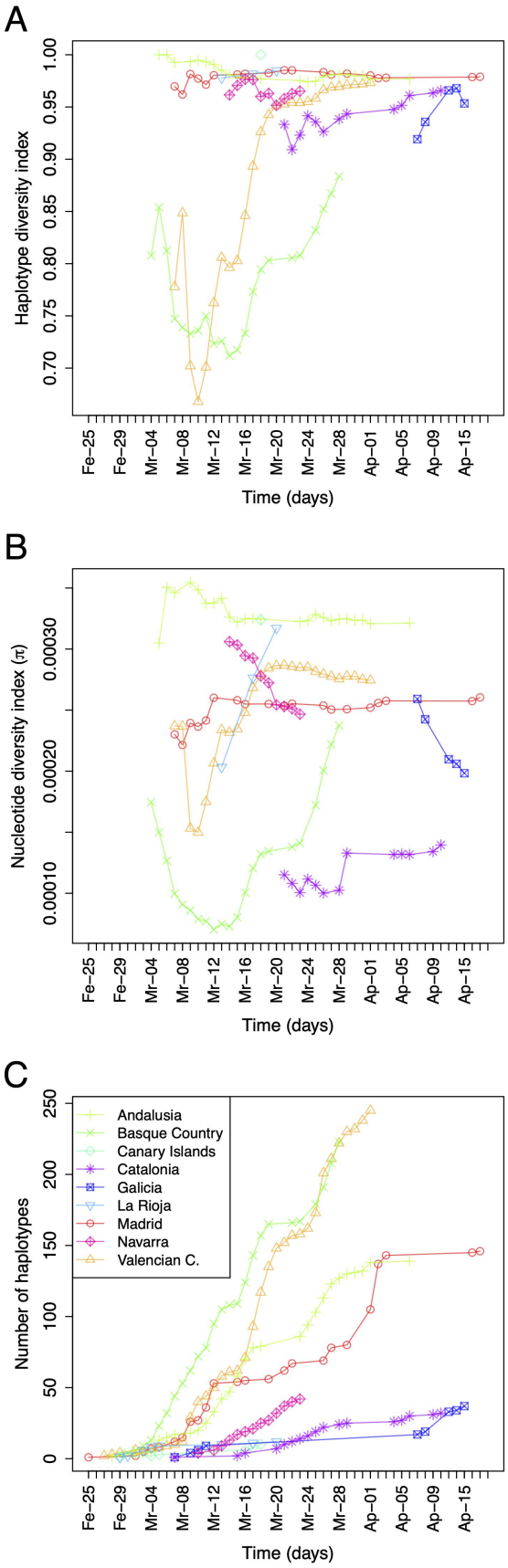

A
Haplotype diversity index
0.70
0.75
0.80
0.85
0.90
0.95
1.00
Fe–25
Fe–29
Mr–04
Mr–08
Mr–12
Mr–16
Mr–20
Mr–24
Mr–28
Ap–01
Ap–05
Ap–09
Ap–15
Time (days)
B
Nucleotide diversity index (π)
0.00010
0.00020
0.00030
Time (days)
C
Number of haplotypes
0
50
100
150
200
250
Andalusia
Basque Country
Canary Islands
Catalonia
Galicia
La Rioja
Madrid
Navarra
Valencian C.
Time (days)

Figure S13.

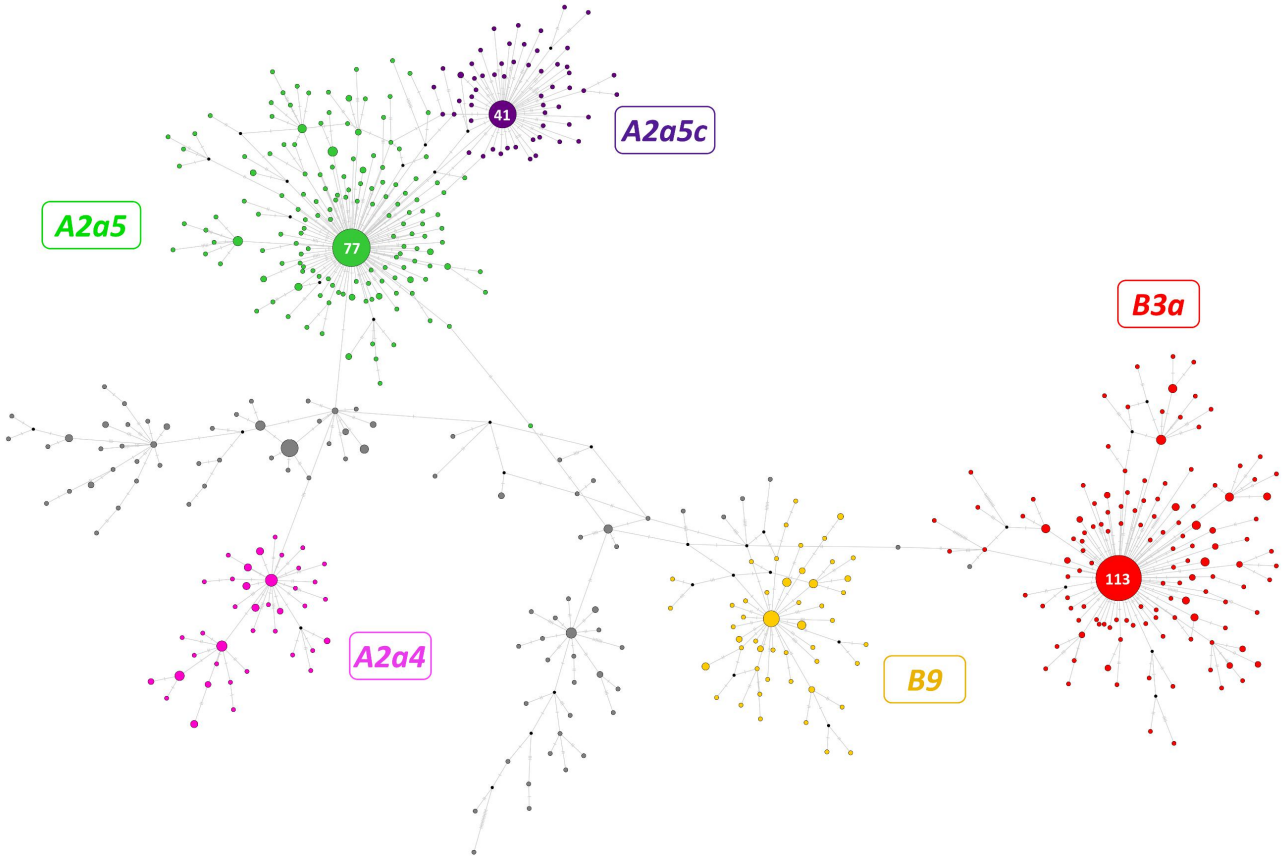


Figure S14.

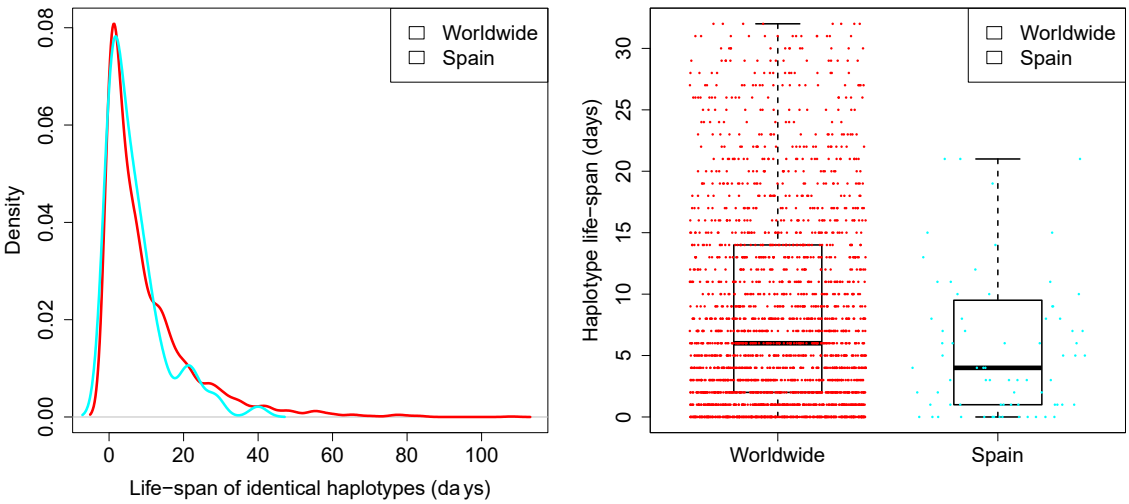
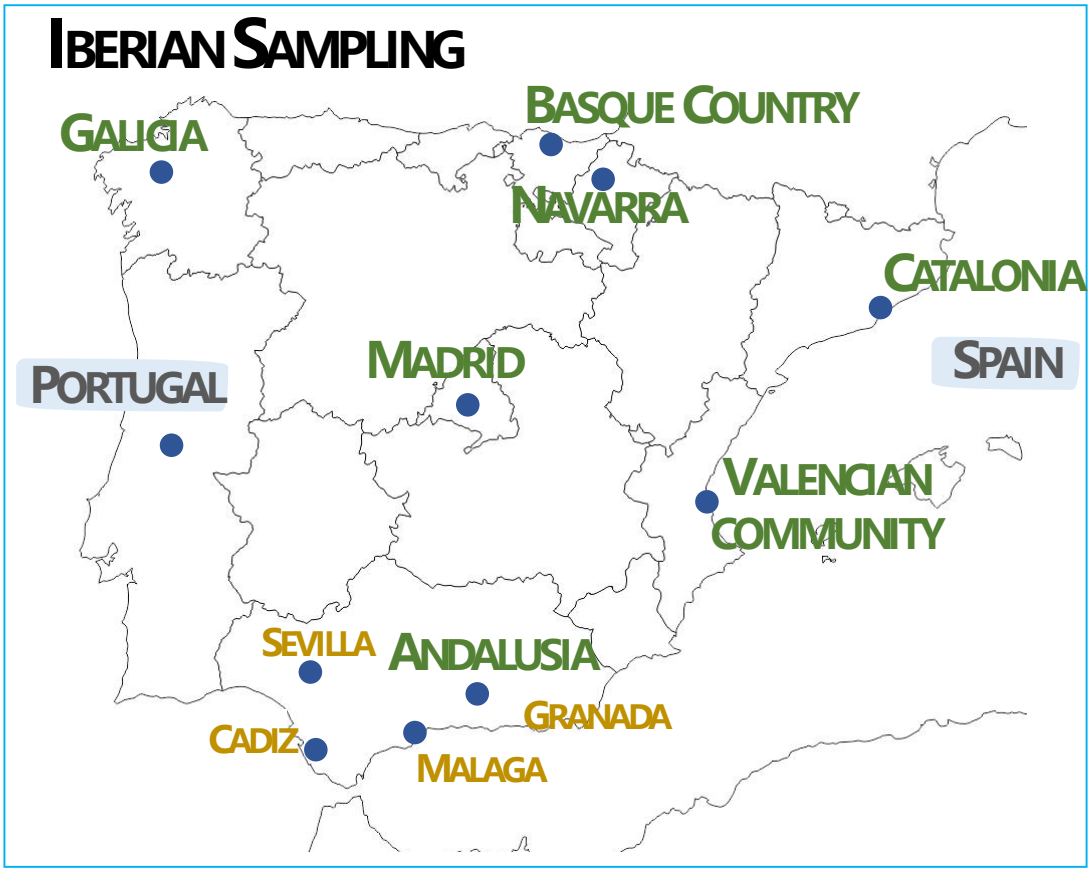


Figure S15.


IBERIAN SAMPLING
BASQUE COUNTRY
GALICIA
NAVARRA
CATALONIA
MADRID
SPAIN
PORTUGAL
VALENCIAN COMMUNITY
SEVILLA
ANDALUSIA
GRANADA
CADIZ
MALAGA

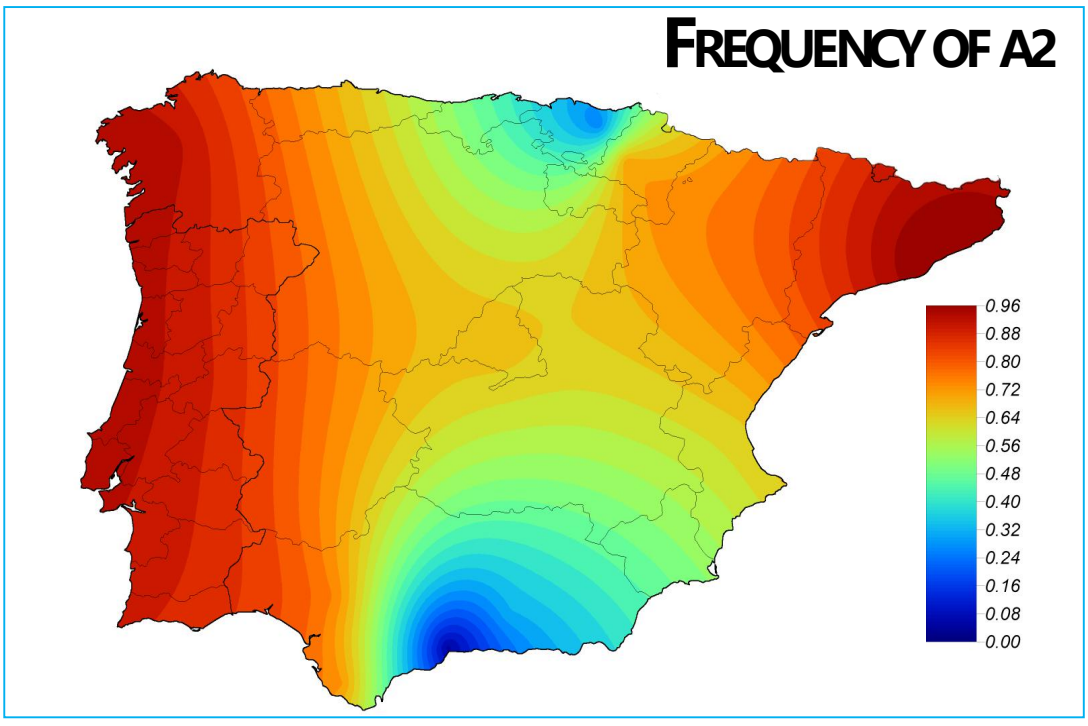

FREQUENCY OF A2
0.96
0.88
0.80
0.72
0.64
0.56
0.48
0.40
0.32
0.24
0.16
0.08
0.00

Figure S16.

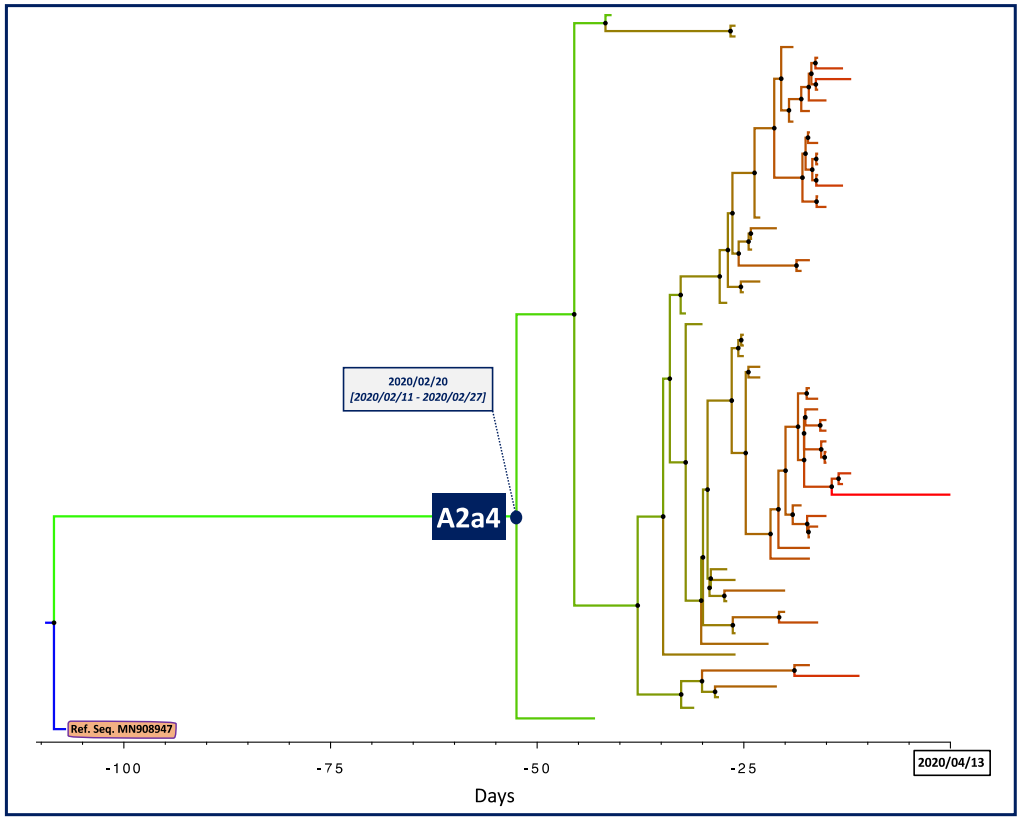


Figure S17.

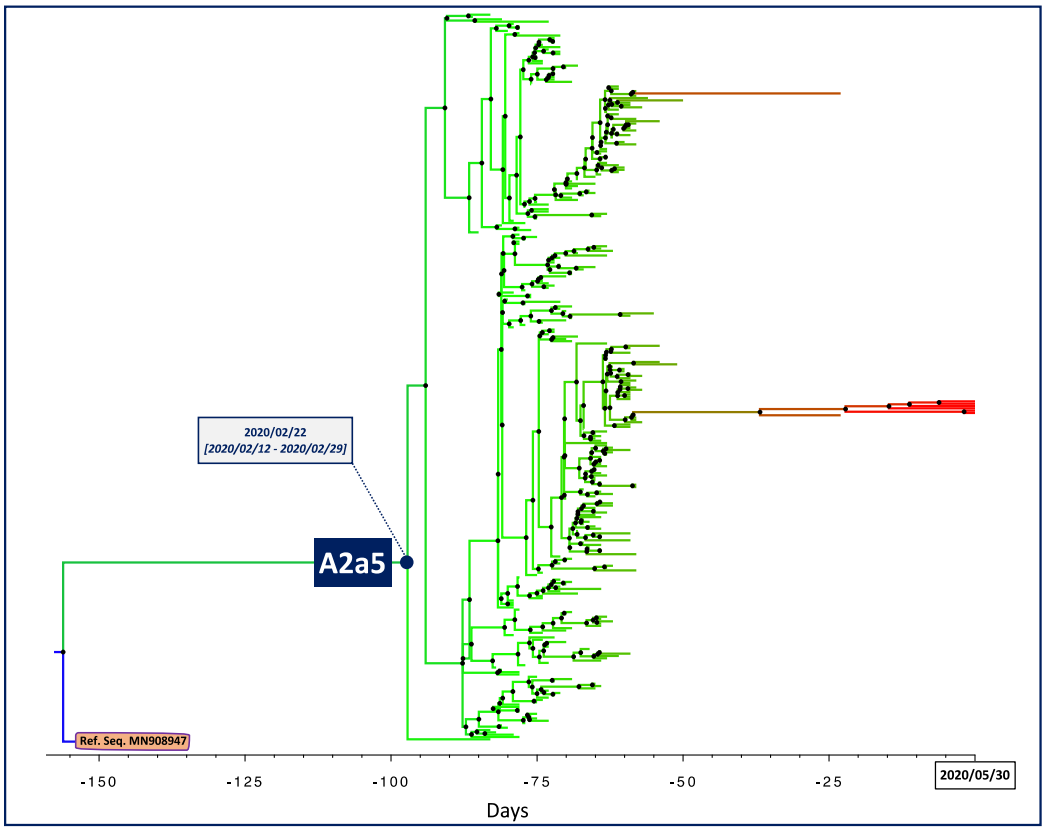


Figure S18.

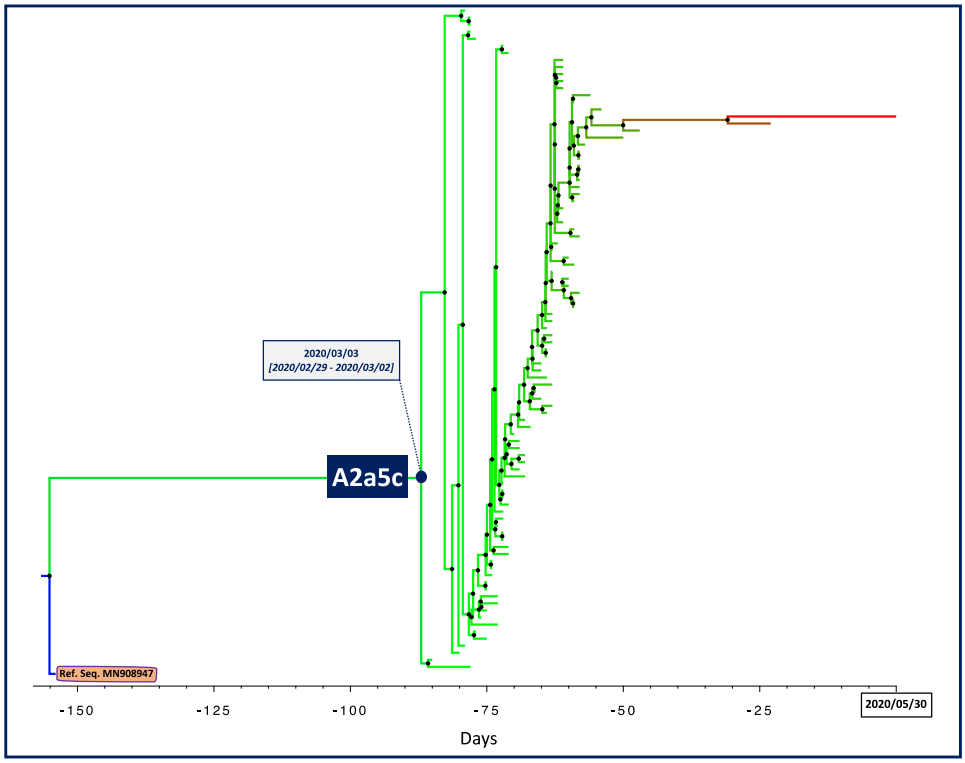


Figure S19.

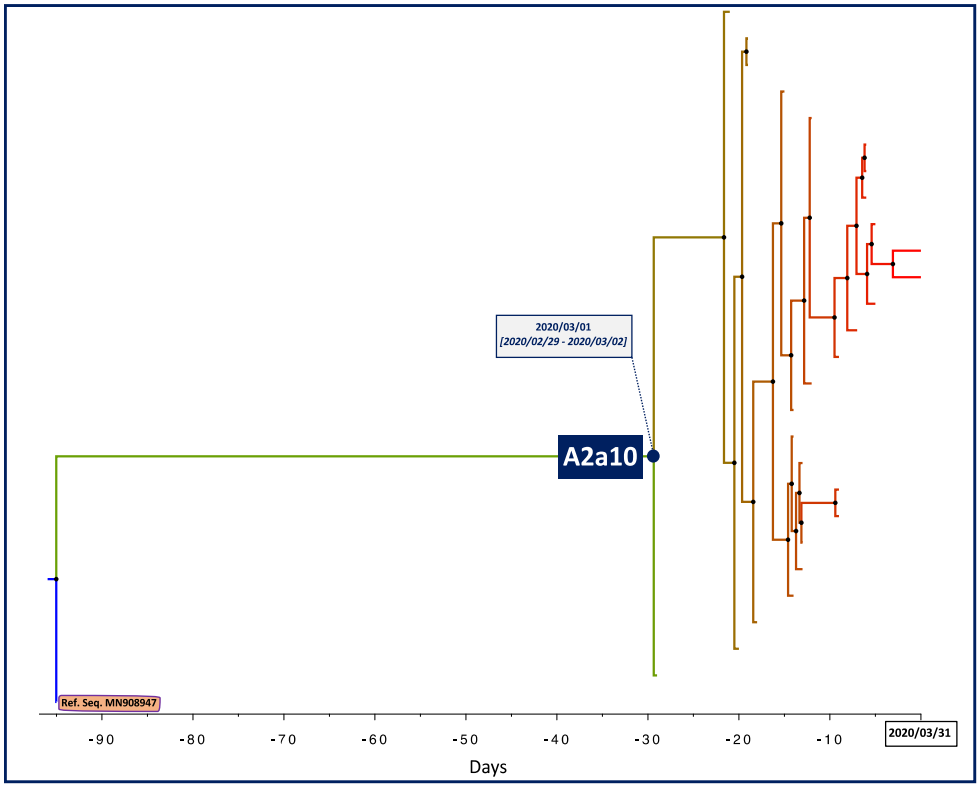


Figure S20.

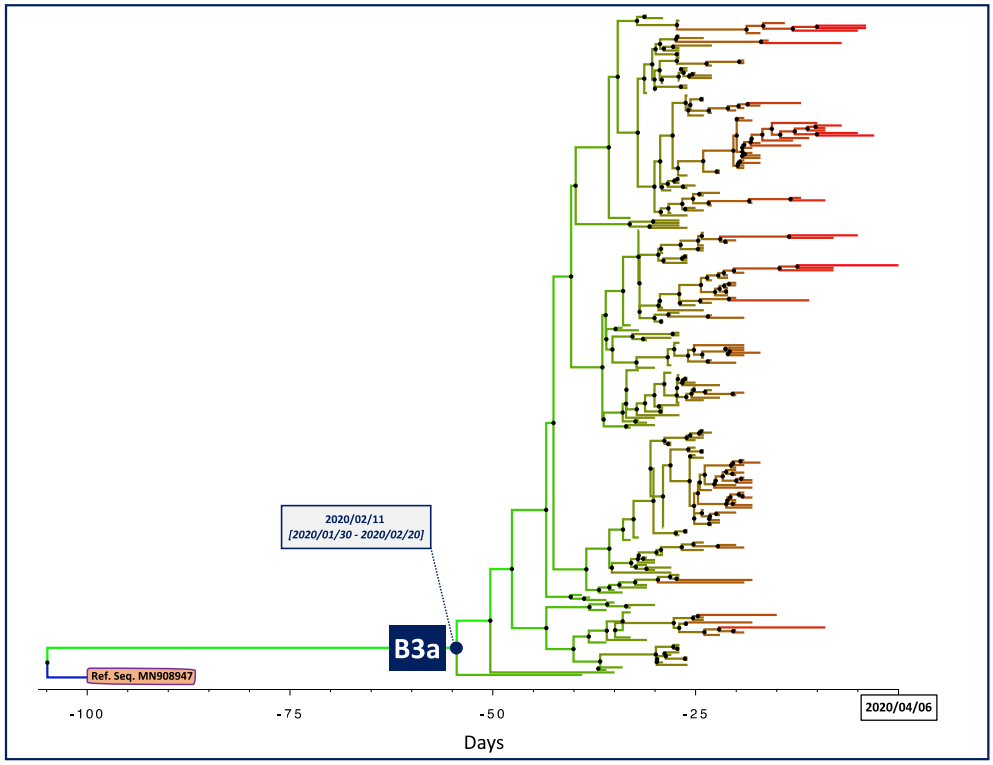
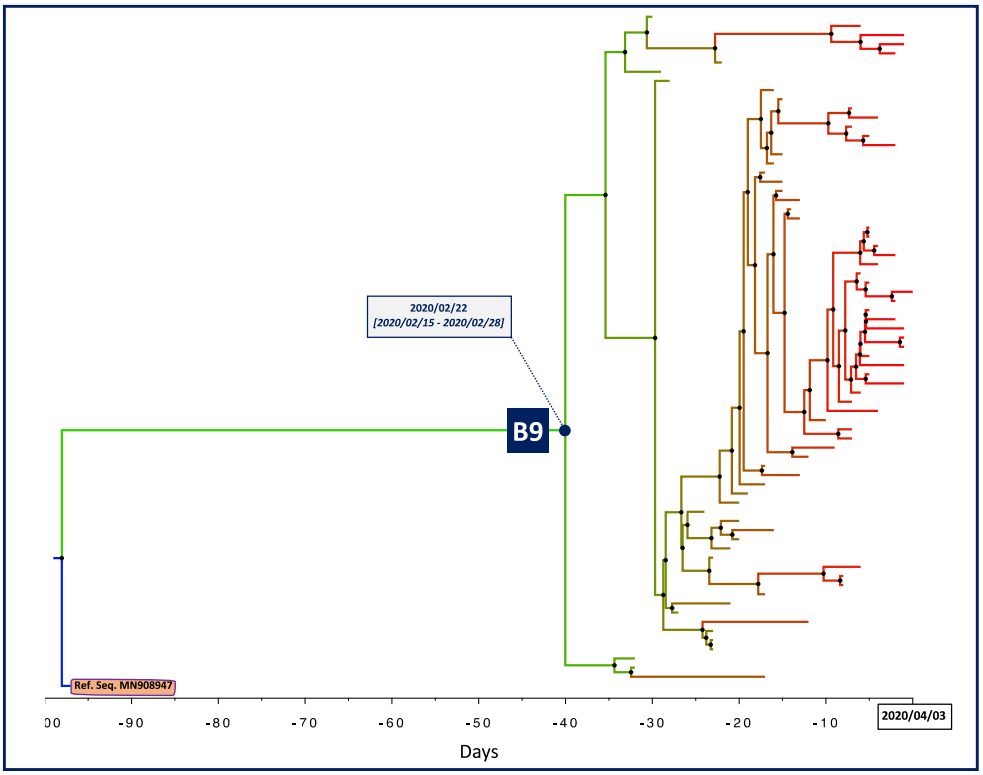



Figure S21.