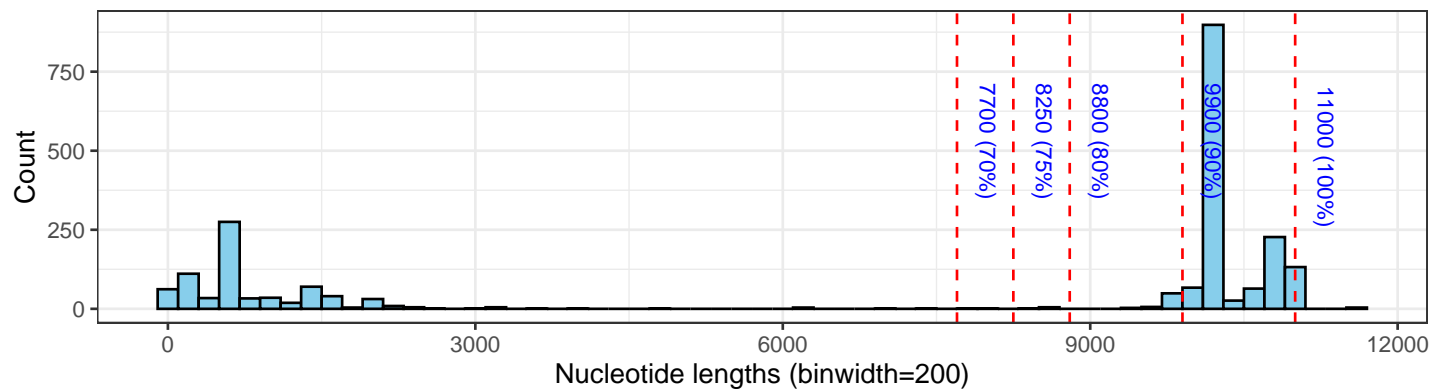


Background

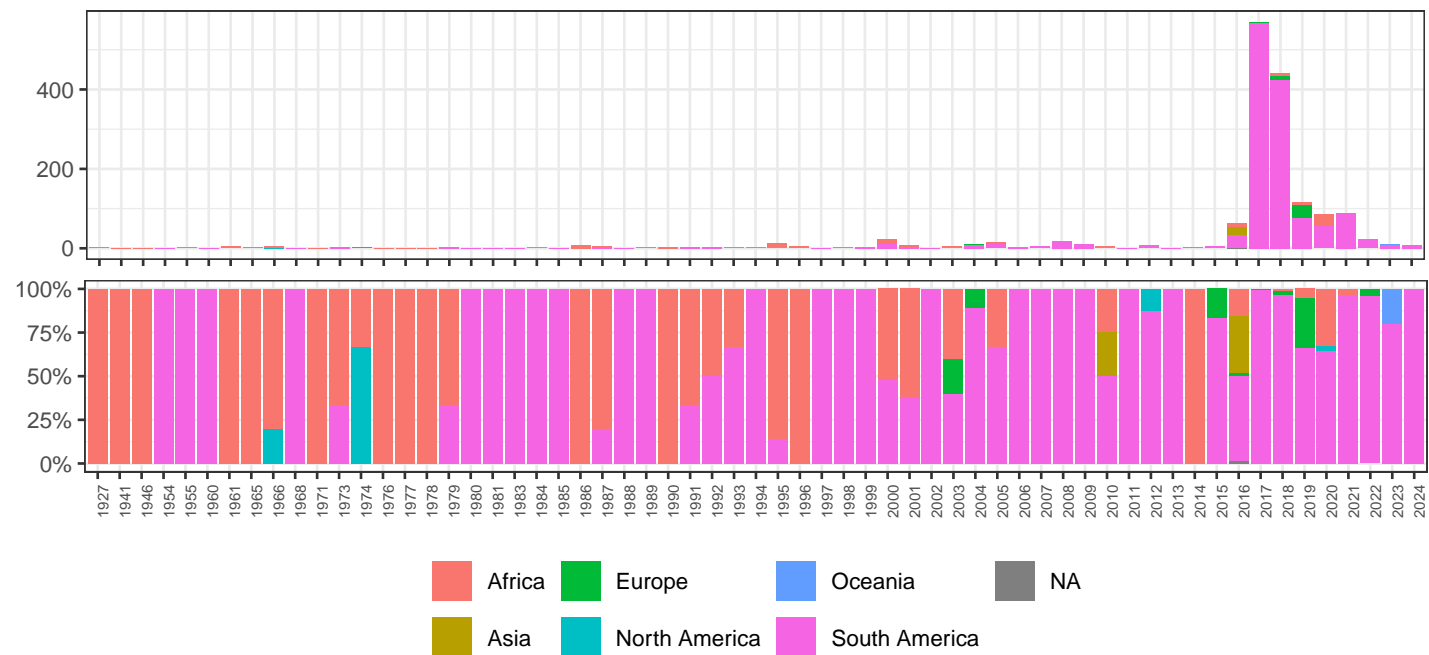
Exploratory Graphics

NCBI GenBank records

Diagnostic plot to estimate min-length filter for phylogenetic analysis (all data = 2228)



Frequency and proportion (n=1620)



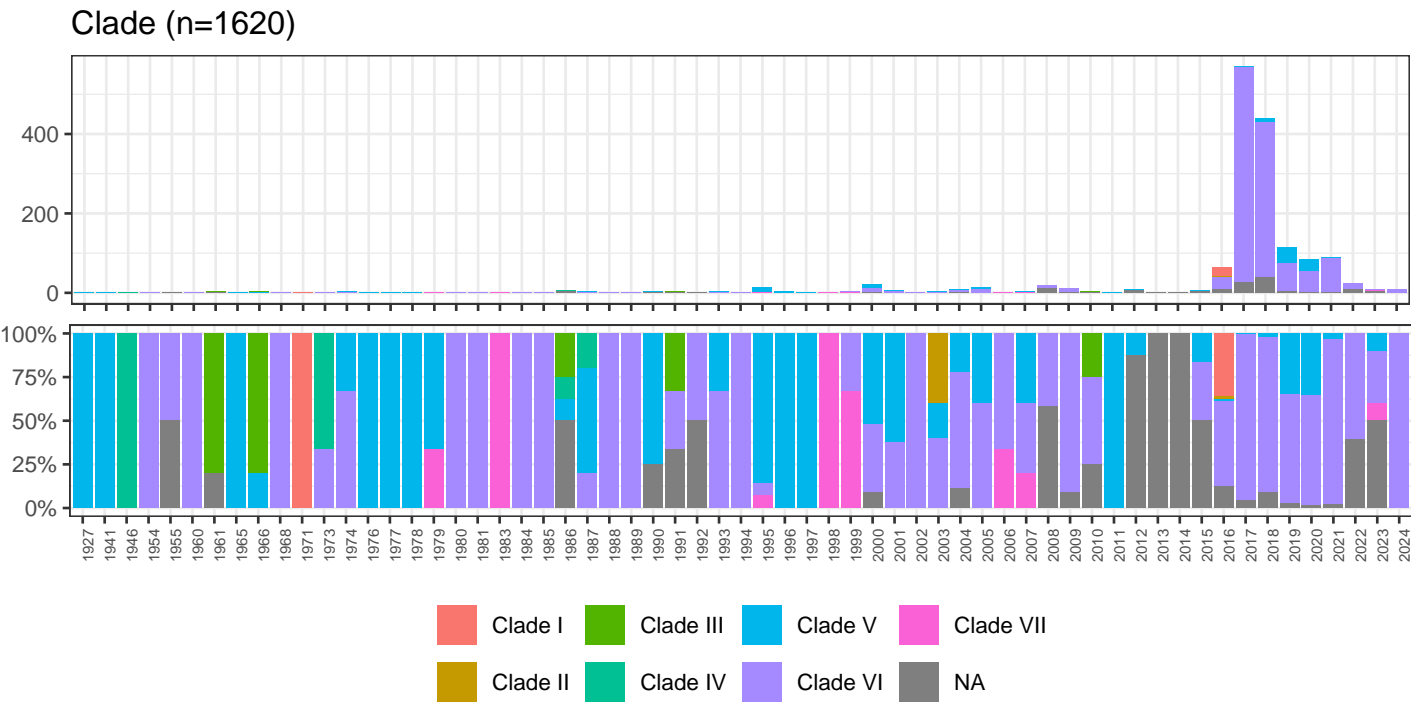


Table 1: Top 25 most frequent sequence submitters with their region and countries

authors	n	regions	countries
hill et al.	515	South America	Brazil
bryant et al.	209	South America, North America, Africa	Peru, Bolivia, Brazil, Colombia, Ecuador, Panama, Trinidad and Tobago
giovannetti et al.	169	South America	Brazil
claro et al.	88	South America	Brazil
rezende et al.	73	South America, NA	Brazil, NA
faria et al.	64	South America	Brazil
adelino et al.	54	South America	Brazil
souza et al.	46	South America	Brazil
abreu et al.	44	South America	Brazil
cunha et al.	44	South America	Brazil
salles et al.	39	South America	Brazil
wang et al.	39	NA	NA
sall et al.	37	Africa	Central African Republic, Cameroon, Côte d'Ivoire, Gambia, Mali, Nigeria, Senegal, Sierra Leone, Sudan, Togo, Zimbabwe
mutebi et al.	36	Africa, NA	Angola, Burkina Faso, Central African Republic, Ethiopia, Kenya, Nigeria, Senegal, Sierra Leone, Sudan, Togo, Zimbabwe
queiroz et al.	30	South America, North America	Brazil, Trinidad and Tobago, Venezuela, Peru, Ecuador, USA
andrade et al.	27	South America	Brazil
auguste et al.	23	South America	Trinidad and Tobago, Brazil, Peru, Bolivia, Venezuela
bakoa et al.	22	Europe	France
ciaramella et al.	20	NA	NA
cruz et al.	20	South America	Brazil
diagne et al.	20	Africa	Senegal
hilfiker et al.	20	NA	NA
de souza et al.	17	South America	Brazil
stock et al.	17	NA, Africa	NA, Senegal, Côte d'Ivoire
ajogbasile et al.	16	Africa	Nigeria