Table 4*:* Recommended metadata standards to be collected for SARS-CoV-2 sequencing data

Metadata	Label	Details	Potential Analyses
Tier 1: Core metadata	Sample Identification number		
	Sample type	Examples: "sputum", "blood", "serum", "saliva", "stool", "nasopharyngeal swab", "wastewater"	
	Sample collection date		Introduction and evolutionary rates
	Country of collection		Introduction and transmission routes, using BEAST (Bayesian Evolutionary Analysis Sampling Tree)
	State/province of collection		
	Originating diagnostic lab	Where the clinical specimen or virus isolate was first obtained	
	Sequence submitting lab	Where sequence data have been generated	Sequencing capacity assessment
	Sampling method	Part of routine surveillance or focused sampling, representative or targeted sampling	
	Host	e.g. human, animal (specifics), environment, unknown	Transmission routes
Tier 2: Descriptive metadata	Age		Risk factors
	Sex	e.g. male, female, other, unknown	Risk factors
	Race and/or ethnicity*		Risk factors
	Heath worker status	e.g. yes, no, unknown. See HW definition in surveillance protocol for Health Workers	Transmission routes, risk factors
	Travel History	Location(s) and timing	Introductions and transmission routes
Tier 3: Metadata for characterization	RT-PCR assay used (if any)		
	RT-PCR Ct value (if any)		
	Symptomatic	e.g. yes, no, unknown	Severity analysis
	Vaccination status (for humans)	Date of vaccination (dose 1 and/or dose 2, as needed), vaccine type, source of information (documented evidence such as vaccine register or vaccine card versus recall)	Vaccine failure
	Date of symptom onset		Delay between onset and sequence submission
	Hospitalization status	e.g. ever hospitalized, never hospitalized, unknown	Severity analysis
	Admission to intensive care unit (ICU)	e.g. yes, no, unknown	Severity analysis
	Mechanical ventilation	e.g. yes, no, unknown	Severity analysis
	Outcome	Deceased/recovered	Severity analysis
	Past history of SARS-CoV-2 infection and date		Reinfection risk
	Therapeutics received	COVID-19-specific	Therapeutic failure
	Location of exposure, link to known cluster/ outbreak		Cluster/outbreak analysis, transmission routes
	Contact with known animal reservoir	e.g. yes, no, unknown; and type(s) of animal(s)	Transmission routes
	Comorbidities	List comorbidities known to increase COVID-19 severity	Risk factors

*This item should be used in regard to local context and individual data collection laws