Parent Class	Field	Ontology Identifier	Definition	Guidance	Examples	Deprecated Label	Deprecated ID	Version Tracking		
	Colour Code Legend							Label	ID	Description/Guidance
	field name in yellow = required					**IMPORTANT:** Only labels and/or IDs will be deprecated, always with replacement version provided. If a term changes in its meaning, a new term will be created.				
	field name in purple = recommended									
	field name in white = optional									
Database identifiers		**GENEPIO:0001122**								
Database identifiers	**specimen collector sample ID**	GENEPIO:0001123	The user-defined name for the sample.	Store the collector sample ID. If this number is considered identifiable information, provide an alternative ID. Be sure to store the key that maps between the original and alternative IDs for traceability and follow up if necessary. Every collector sample ID from a single submitter must be unique. It can have any format, but we suggest that you make it concise, unique and consistent within your lab.	ASDFG123			1.0.0	1.0.0	1.0.0
Database identifiers	**specimen collector subsample ID**	GENEPIO:0100752	The user-defined identifier assigned to a portion of the original sample.	Store the ID for the subsample/aliquot.	ASDFG123_12			1.0.0	1.0.0	1.0.0
Database identifiers	**pooled sample ID**	GENEPIO:0100996	The user-defined identifier assigned to a combined (pooled) set of samples.	If the sample being analyzed is the result of pooling individual samples, rename the pooled sample with a new identifier. Store the pooled sample ID.	12345AYZ			1.0.0	1.0.0	1.0.0
Database identifiers	**metagenome-assembled genome (MAG) ID**	GENEPIO:0100753	The user-defined identifier assigned to a genome reconstructed from metagenomic data.	Store the MAG ID.	XYZ1234.1			1.0.0	1.0.0	1.0.0
Database identifiers	**specimen collector project ID**	GENEPIO:0100918	The user-defined project name assigned to a sequencing project.	If the sample was collected or analyzed under the umbrella of a specific project, include the name of that project here.				1.0.0	1.0.0	1.0.0
Database identifiers	**BioProject accession**	GENEPIO:0001136	The INSDC (i.e., ENA, NCBI, or DDBJ) accession number of the BioProject(s) to which the BioSample belongs.	Store the BioProject accession number. BioProjects are an organizing tool that links together raw sequence data, assemblies, and their associated metadata. Each province will be assigned a different bioproject accession number by the National Microbiology Lab. A valid NCBI BioProject accession has prefix PRJN e.g., PRJNA12345, and is created once at the beginning of a new sequencing project.	PRJNA608651			1.0.0	1.0.0	1.0.0
Database identifiers	**BioSample accession**	GENEPIO:0001139	The identifier assigned to a BioSample in INSDC (i.e., ENA, NCBI, or DDBJ) archives.	Store the accession returned from the BioSample submission. NCBI BioSamples will have the prefix SAMN, ENA have the prefix SAMEA, DDBJ have SAMD	SAMN14180202, SAMD00000001, SAMEA			1.0.0	1.0.0	1.0.0
Database identifiers	**GenBank accession**	GENEPIO:0100754	The unique identifier assigned to an assembly or consensus sequence in GenBank archives.	Store the versioned GenBank accession assigned to the submitted sequence.	LZ986655.1			1.0.0	1.0.0	1.0.0
Database identifiers	**SRA accession**	GENEPIO:0001142	The Sequence Read Archive (SRA) identifier linking raw read data, methodological metadata and quality control metrics submitted to the INSDC.	Store the accession assigned to the submitted sequence. NCBI-SRA accessions start with SRR.	SRR11177792			1.0.0	1.0.0	1.0.0
Database identifiers	**GISAID accession**	GENEPIO:0001147	The identifier assigned to a sequence in GISAID (the Global Initiative on Sharing All Influenza Data) archives.	Store the accession assigned to the submitted sequence. GISAID accessions start with EPI.	EPI_ISL_402131			1.0.0	1.0.0	1.0.0
Database identifiers	**GISAID virus name**	GENEPIO:0100282	The user-defined GISAID virus name assigned to the sequence.	GISAID virus names should be in the format "hCoV-19/Country/Identifier/year".	hCoV-19/Canada/prov_rona_99/2020			1.0.0	1.0.0	1.0.0
Database identifiers	**ENA accession**	GENEPIO:0100755	The identifier assigned to a sequence in the European Nucleotide Archive (ENA).	Store the accession assigned to the submitted sequence. ENA sequence accessions start with ERR.	ERR123456			1.0.0	1.0.0	1.0.0
Database identifiers	**DRA accession**	GENEPIO:0100757	The identifier assigned to a sequence in DNA Data Bank of Japan (DDBJ) archives.	Store the accession assigned to the submitted sequence. DRA accessions start with DRR.	DRR123456			1.0.0	1.0.0	1.0.0
Database identifiers	**GSA accession**	GENEPIO:0100758	The identifier assigned to a BioSample in GSA (Genome Sequence Archive) archives.	Store the accession assigned to the submitted sequence. GSA accessions start with CRR.	CRR123456			1.0.0	1.0.0	1.0.0
Database identifiers	**Enterobase accession**	GENEPIO:0100759	The identifier assigned to a sequence in Enterobase archives.	Store the barcode assigned to the submitted sequence. Enterobase barcodes start with different 3 letter codes depending on the organism.	SAL_AA0019AA_ST			1.0.0	1.0.0	1.0.0
Database identifiers	**sampling site ID**	GENEPIO:0100760	The user-defined identifier assigned to a specific location from which samples are taken.	Store the ID for the site from which a sample was taken. The "site" is user defined (e.g. it may be a building and its environs, a specific entity within an environment). Please use the same site ID for all samples from a given site, regardless of when these samples were taken. Any important changes in site location, should be represented with a new site ID.	Site 12A			1.0.0	1.0.0	1.0.0
Database identifiers	**sampling event ID**	GENEPIO:0100761	The user-defined identifier assigned to a specific event during which one or more samples are taken, from one or more sites.	Store the ID for the event during which a sample or samples were taken. For example, an event could be one person taking samples from multiple sites, or multiple people taking samples from one site.	Event 120522.1			1.0.0	1.0.0	1.0.0
Sample collection and processing		GENEPIO:0001150								

Parent Class	Field Colour Code Legend	Ontology Identifier	Definition	Guidance	Examples	Deprecated Label	Deprecated ID	Version Tracking Label	ID	Description/Guidance
Sample collection and processing	**sample collection data steward name**	GENEPIO:0100762	A sample collection data field which describes the name of the individual responsible for the data governance, (meta)data usage and distribution of the sample.	Provide the name of the sample collection data steward.	Joe Bloggs			1.0.0	1.0.0	1.0.0
Sample collection and processing	**sample collected by**	GENEPIO:0001153	The name of the organization with which the sample collector is affiliated.	The name of the agency should be written out in full, (with minor exceptions) and be consistent across multiple submissions.	Public Health Agency of Canada			1.0.0	1.0.0	1.0.0
Sample collection and processing	**sample collector contact email**	GENEPIO:0001156	The email address of the contact responsible for follow-up regarding the sample.	The email address can represent a specific individual or lab e.g. johnnyblogs@lab.ca, or RespLab@lab.ca	WaterTester@facility.ca			1.0.0	1.0.0	1.0.0
Sample collection and processing	**geo_loc_name (country)**	GENEPIO:0001181	The country of origin of the sample.	If known, select a value from the pick list.	Canada			1.0.0	1.0.0	1.0.0
Sample collection and processing	**geo_loc_name (state/province/territory)**	GENEPIO:0001185	The state/province/territory of origin of the sample.	Provide the state/province/territory name from the GAZ geography ontology. Search for geography terms here: https://www.ebi.ac.uk/ols/ontologies/gaz	Western Cape			1.0.0	1.0.0	1.0.0
Sample collection and processing	**geo_loc name (county/region)**	GENEPIO:0100280	The county/region of origin of the sample.	Provide the county/region name from the GAZ geography ontology. Search for geography terms here: https://www.ebi.ac.uk/ols/ontologies/gaz	Derbyshire			1.0.0	1.0.0	1.0.0
Sample collection and processing	**geo_loc_name (city)**	GENEPIO:0001189	The city of origin of the sample.	Provide the city name from the GAZ geography ontology. Search for geography terms here: https://www.ebi.ac.uk/ols/ontologies/gaz	Vancouver			1.0.0	1.0.0	1.0.0
Sample collection and processing	**geo_loc name (site)**	GENEPIO:0100436	The name of a specific geographical location e.g. Credit River (rather than river).	Provide the name of the specific geographical site using a specific noun (a word that names a certain place, thing).	Credit River			1.0.0	1.0.0	1.0.0
Sample collection and processing	**geo_loc latitude**	GENEPIO:0100309	The latitude coordinates of the geographical location of sample collection.	Provide latitude coordinates if available. Do not use the centre of the city/region/province/state/country or the location of your agency as a proxy, as this implicates a real location and is misleading. Specify as degrees latitude in format "d[d.dddd] N|S".	38.98 N			1.0.0	1.0.0	1.0.0
Sample collection and processing	**geo_loc longitude**	GENEPIO:0100310	The longitude coordinates of the geographical location of sample collection.	Provide longitude coordinates if available. Do not use the centre of the city/region/province/state/country or the location of your agency as a proxy, as this implicates a real location and is misleading. Specify as degrees longitude in format "d[dd.dddd] W|E".	77.11 W			1.0.0	1.0.0	1.0.0
Sample collection and processing	**watershed shapefile availability**	GENEPIO:0100919	The availability status of a shapefile descriping the catchment contributing to a watershed.	Select a value from the picklist to describe whether or not a watershed shapefile would be available upon request.	watershed shapefile available			1.0.0	1.0.0	1.0.0
Sample collection and processing	**watershed shapefile filename**	GENEPIO:0100920	The name of the watershed shapefile.	Provide the shapefile filename corresponding to the watershed from which the sample was taken. If there are multiple files associated with the watershed, provide all names separated by commas.	siteAD17.shp, siteAD17.kml			1.0.0	1.0.0	1.0.0
Sample collection and processing	**organism**	GENEPIO:0001191	Taxonomic name of the organism.	Provide the official nomenclature for the organism present in the sample. Search for taxonomic names here: ncbi.nlm.nih.gov/taxonomy.	Severe acute respiratory syndrome corona			1.0.0	1.0.0	1.0.0
Sample collection and processing	**purpose of sampling**	GENEPIO:0001198	The reason that the sample was collected.	The reason a sample was collected may provide information about potential biases in sampling strategy. Provide the purpose of sampling from the picklist in the template. Most likely, the sample was collected for public health surveillance. The reason why a sample was originally collected may differ from the reason why it was selected for sequencing, which should be indicated in the "purpose of sequencing" field.	Public health surveillance			1.0.0	1.0.0	1.0.0
Sample collection and processing	**scale of sampling**	GENEPIO:0100877	The range of locations or entities sampled expressed in general terms.	Provide the scale of wastewater sampling by selecting a value from the picklist.	Community-level surveillance			1.0.0	1.0.0	1.0.0
Sample collection and processing	**sample received date**	GENEPIO:0001179	The date on which the sample was received.	Provide the sample received date in ISO 8601 format, i.e. "YYYY-MM-DD".	2020-03-28			1.0.0	1.0.0	1.0.0
Sample collection and processing	**sample collection date**	GENEPIO:0001174	The date on which the sample was collected, or sampling began for a continuous sample.	If your sample is a continuous sample please use this field to capture your start date. Sample collection date is critical for surveillance and many types of analyses. Required granularity includes year, month and day. If this date is considered identifiable information, it is acceptable to add "jitter" by adding or subtracting a calendar day (acceptable by GISAID). Alternatively, "received date" may be used as a substitute. The date should be provided in ISO 8601 standard format "YYYY-MM-DD".	2020-03-16			1.0.0	1.0.0	1.0.0
Sample collection and processing	**sample processing date**	GENEPIO:0100763	The date on which the sample was processed.	Provide the sample processed date in ISO 8601 format, i.e. "YYYY-MM-DD". The sample may be collected and processed (e.g. filtered, extraction) on the same day, or on different dates.	2020-03-16			1.0.0	1.0.0	1.0.0
Sample collection and processing	**sample collection time**	GENEPIO:0100764	The time at which the sample was collected.	If a sample was collected over a certain time period, provide the end time here and the duration of the collection in the field "sample collection duration value". Provide this time in ISO 8601 24hr format, including the time zone with respect to UTC. Format: "HH:MM:SS+[Hours+/-UTC]" e.g. 11:23pm Pacific Daylight Savings Time should be recorded as 23:23:00-07:00.	23:23:00-07:00			1.0.0	1.0.0	1.0.0

Parent Class	Field Colour Code Legend	Ontology Identifier	Definition	Guidance	Examples	Deprecated Label	Deprecated ID	Version Tracking Label	ID	Description/Guidance
Sample collection and processing	**sample collection time of day**	GENEPIO:0100765	The descriptive time of day during which the sample was collected.	If known, fill in sample collection time. Otherwise, select a value from the pick list for sample collection time of day to approximate the timing. The time of sample processing matters especially for grab samples, as fecal concentration in wastewater fluctuates over the course of the day.	Morning			1.0.0	1.0.0	1.0.0
Sample collection and processing	**sample collection time duration value**	GENEPIO:0100766	The amount of time over which the sample was collected.	Provide the numerical value of time.	4			1.0.0	1.0.0	1.0.0
Sample collection and processing	**sample collection time duration unit**	GENEPIO:0100767	The units of the time duration measurement of sample collection.	Provide the units from the pick list.	Hour			1.0.0	1.0.0	1.0.0
Sample collection and processing	**presampling activity**	GENEPIO:0100433	The activities or variables upstream of sample collection that may affect the sample.	If there was an activity that would affect the sample prior to collection (this is different than sample processing), provide the activities by selecting one or more values from the template pick list. If the information is unknown or cannot be provided, leave blank or provide a null value.	Agricultural activity			1.0.0	1.0.0	1.0.0
Sample collection and processing	**presampling activity details**	GENEPIO:0100434	The details of the activities or variables that affected the sample collected.	Briefly describe the presampling activities using free text.	Agricultural waste from large farm contributes waste to the site sampled.			1.0.0	1.0.0	1.0.0
Sample collection and processing	**sample volume measurement value**	GENEPIO:0100768	The numerical value of the volume measurement of the sample collected.	Provide the numerical value of volume.	5			1.0.0	1.0.0	1.0.0
Sample collection and processing	**sample volume measurement unit**	GENEPIO:0100769	The units of the volume measurement of the sample collected.	Provide the units from the pick list.	mL			1.0.0	1.0.0	1.0.0
Sample collection and processing	**sample storage method**	GENEPIO:0100448	The process used to store the sample.	Provide details of how the sample was stored from time of collection until time of processing. If there were issues with the cold chain storage, note those here.	The sample was placed in a tube in a cooler bag during transportation (~3 hours) to the lab site. At this point the sample was placed in storage medium and put in a -10C freezer until it was processed and extracted 5 days later.			1.0.0	1.0.0	1.0.0
Sample collection and processing	**sample storage medium**	GENEPIO:0100449	The medium in which a sample is stored.	Provide the name of the transport medium or storage medium used for this sample. If none was used, leave blank or write "None".	Cary-Blair transport medium			1.0.0	1.0.0	1.0.0
Sample collection and processing	**sample storage duration value**	GENEPIO:0101014	The numerical value of the time measurement during which a sample is in storage.	Provide the numerical value of time.	5			1.0.0	1.0.0	1.0.0
Sample collection and processing	**sample storage duration unit**	GENEPIO:0101015	The units of a measured sample storage duration.	Provide the units from the pick list.	Day			1.0.0	1.0.0	1.0.0
Sample collection and processing	**specimen processing**	GENEPIO:0001253	Any processing applied to the sample during or after receiving the sample.	Select processes from the picklist that were applied to this sample.	Centrifugation			1.0.0	1.0.0	1.0.0
Sample collection and processing	**specimen processing details**	GENEPIO:0100311	The details of the processing applied to the sample during or after receiving the sample.	Briefly describe the processes applied to the sample.	25 samples were pooled and further prepared as a single sample during library prep.			1.0.0	1.0.0	1.0.0
Sample collection and processing	**environmental site**	GENEPIO:0001232	An environmental location may describe a site in the natural or built environment e.g. contact surface, metal can, hospital, wet market, bat cave.	Provide a descriptor of the environmental site sampled. Use the picklist provided in the template. If not applicable, choose a null value.	Meat processing plant			1.0.0	1.0.0	1.0.0
Sample collection and processing	**environmental material**	GENEPIO:0001223	A substance obtained from the natural or man-made environment e.g. soil, water, sewage.	Provide a descriptor of the environmental material sampled. Use the picklist provided in the template. If not applicable, choose a null value.	Raw wastewater			1.0.0	1.0.0	1.0.0
Sample collection and processing	**environmental material properties**	GENEPIO:0100770	The properties, characteristics and qualities of a substance obtained from an environment.	Provide the environmental material properties by selecting descriptors from the pick list.	Stagnant			1.0.0	1.0.0	1.0.0
Sample collection and processing	**wastewater system type**	GENEPIO:0100771	The type or classification of a wastewater system e.g. sanitary sewer, combined sewer	Provide the classification of the wastewater system by selecting from the provided pick list.	Sanitary sewer			1.0.0	1.0.0	1.0.0
Sample collection and processing	**experimental control type**	GENEPIO:0100921	The type of control that the sample represents in the experiment.	This field is used to distinguish samples under study from controls. A sample may act as an experimental control if it is used to detect experimental errors or to ensure experimental methods are performing as expected. If the sample acted as an experimental control, select an experimental control type from the picklist. If the sample was not a control, leave blank or select "Not Applicable".	Positive experimental control			1.0.0	1.0.0	1.0.0
Sample collection and processing	**experimental control details**	GENEPIO:0100922	The details regarding the type of control that the sample represents in the experiment.	Provide details regarding the nature of the reference strain used as a control, or what is was used to monitor.	Human coronavirus 229E (HCoV-229E) spiked in sample as process control			1.0.0	1.0.0	1.0.0
Sample collection and processing	**collection device**	GENEPIO:0001234	The instrument or container used to collect the sample e.g. *grab sampler*.	Provide a descriptor of the device used for sampling. Use the picklist provided in the template. If not applicable, choose a null value.	Automatic flow-proportional sampler			1.0.0	1.0.0	1.0.0
Sample collection and processing	**collection method**	GENEPIO:0001241	The process used to collect the sample.	Provide a descriptor of the collection method used for sampling. Use the picklist provided in the template. If not applicable, choose a null value.	Automatic composite sampling			1.0.0	1.0.0	1.0.0
Sample collection and processing	**nucleic acid extraction method**	GENEPIO:0100939	The process used to extract genomic material from a sample.	Briefly describe the extraction method used.	Direct wastewater RNA capture and purification via the "Sewage, Salt, Silica and SARS-CoV-2 (4S)" method v4 found at https://www.protocols.io/view/v-4-direct-wastewater-rna-capture-and-purification-36wgq581ygk5/v4			1.0.0	1.0.0	1.0.0
Sample collection and processing	**nucleic acid extraction kit**	GENEPIO:0100772	The kit used to extract genomic material from a sample	Provide the name of the genomic extraction kit used.	QIAamp PowerFecal Pro DNA Kit			1.0.0	1.0.0	1.0.0

Parent Class	Field Colour Code Legend	Ontology Identifier	Definition	Guidance	Examples	Deprecated Label	Deprecated ID	Version Tracking Label	ID	Description/Guidance
Sample collection and processing	**endogenous control details**	GENEPIO:0100923	The description of the endogenous controls included when extracting a sample.	Provide the names of endogenous controls that were used as a reference during extraction. If relevant, include titers of these controls, as well as whether any controls were expected but not identified in the sample.	Pepper mild mottle virus (PMMoV) detected			1.0.0	1.0.0	1.0.0
Sample collection and processing	**extraction recovery efficiency measurement value**	GENEPIO:0100924	The recovery efficiency of an extraction, calculated as the amount of a synthetic or endogenous compound identified in the sample relative to the amount expected.	Provide value as a percent.	25			1.0.0	1.0.0	1.0.0
Sample collection and processing	**extraction recovery efficiency measurement method**	GENEPIO:0100925	The method by which recovery efficiency of an extraction was calculated.	Provide a brief description of how extraction recovery efficiency was measured or estimated.	Spiked in synthetic material into wastewater sample and into water control and compared recovery.			1.0.0	1.0.0	1.0.0
Environmental conditions and measurements		**GENEPIO:0100940**								
Environmental conditions and measurements	**water catchment area human population measurement value**	GENEPIO:0100773	The numerical value of the human population measurement that contributes to the composition of water in a catchment area.	Where known, provide the numerical value of population size, i.e. the number of people.	10,500			1.0.0	1.0.0	1.0.0
Environmental conditions and measurements	**water catchment area human population bin**	GENEPIO:0100774	The human population range of the water catchment that contributes effluent to a wastewater site.	Where catchment population is not well known, provide an estimation of population size by selecting a value from the picklist.	1,000 - 10,000 people			1.0.0	1.0.0	1.0.0
Environmental conditions and measurements	**water catchment area human population measurement method**	GENEPIO:0100775	The method by which a water catchment 's human population size was measured or estimated	Provide a brief description of how catchment population size was measured or estimated.	population of jurisdiction encompassing the wastewater service area			1.0.0	1.0.0	1.0.0
Environmental conditions and	**water catchment area human population density**	GENEPIO:0100776	The numerical value describing the number	Provide the numerical value of the population density in the	4			1.0.0	1.0.0	1.0.0
Environmental conditions and	**water catchment area human population density**	GENEPIO:0100777	The unit describing the number of humans	Provide the unit of the population density in the catchement	persons per Km^2			1.0.0	1.0.0	1.0.0
Environmental conditions and	**populated area type**	GENEPIO:0100778	A type of area that is populated by humans	Provide the populated area type from the pick list.	Urban area			1.0.0	1.0.0	1.0.0
Environmental conditions and	**sampling weather conditions**	GENEPIO:0100779	The state of the atmosphere at a place and	Provide the weather conditions at the time of sample	Rain			1.0.0	1.0.0	1.0.0
Environmental conditions and measurements	**presampling weather conditions**	GENEPIO:0100780	Weather conditions prior to collection that may affect the sample.	Provide the weather conditions prior to sample collection.	Drizzle			1.0.0	1.0.0	1.0.0
Environmental conditions and measurements	**precipitation measurement value**	GENEPIO:0100911	The amount of water which has fallen during a precipitation process.	Provide the quantity of precipitation in the area leading up to the time of sample collection.	12			1.0.0	1.0.0	1.0.0
Environmental conditions and measurements	**precipitation measurement unit**	GENEPIO:0100912	The units of measurement for the amount of water which has fallen during a precipitation process.	Provide the units of precipitation by selecting a value from the pick list.	inch			1.0.0	1.0.0	1.0.0
Environmental conditions and measurements	**precipitation measurement method**	GENEPIO:0100913	The process used to measure the amount of water which has fallen during a precipitation process.	Provide the name of the procedure or method used to measure precipitation.	Rain gauge over a 12 hour period prior to sample collection			1.0.0	1.0.0	1.0.0
Environmental conditions and measurements	**ambient temperature measurement value**	GENEPIO:0100935	The numerical value of a measurement of the ambient temperature.	Provide the numerical value of the measured temperature.	70			1.0.0	1.0.0	1.0.0
Environmental conditions and measurements	**ambient temperature measurement unit**	GENEPIO:0100936	The units of a measurement of the ambient temperature.	Provide the units of the measured temperature.	degree Celsius (C)			1.0.0	1.0.0	1.0.0
Environmental conditions and measurements	**pH measurement value**	GENEPIO:0001736	The measured pH value indicating the acidity or basicity(alkalinity) of an aqueous solution.	Provide the numerical value of the measured pH.	7.4			1.0.0	1.0.0	1.0.0
Environmental conditions and measurements	**pH measurement method**	GENEPIO:0100781	The process used to measure pH value.	Provide the name of the procedure or technology used to measure pH.	pH test strip (litmus test)			1.0.0	1.0.0	1.0.0
Environmental conditions and measurements	**total daily flow rate measurement value**	GENEPIO:0100905	The numerical value of a measured fluid flow rate over the course of a day.	Provide the numerical value of the measured flow rate.	10			1.0.0	1.0.0	1.0.0
Environmental conditions and measurements	**total daily flow rate measurement unit**	GENEPIO:0100906	The units of a measured fluid flow rate over the course of a day.	Provide the units of the measured flow rate by selecting a value from the pick list.	million gallons per day (MGD)			1.0.0	1.0.0	1.0.0
Environmental conditions and measurements	**total daily flow rate measurement method**	GENEPIO:0100907	The process used to measure total daily fluid flow rate.	Provide the name of the procedure or technology used to measure flow rate.	Flow meter			1.0.0	1.0.0	1.0.0
Environmental conditions and measurements	**instantaneous flow rate measurement value**	GENEPIO:0100908	The numerical value of a measured instantaneous fluid flow rate.	Provide the numerical value of the measured flow rate.	25			1.0.0	1.0.0	1.0.0
Environmental conditions and measurements	**instantaneous flow rate measurement unit**	GENEPIO:0100909	The units of a measured instantaneous fluid flow rate.	Provide the units of the measured flow rate by selecting a value from the pick list.	cubic meter per hour (m^3/h)			1.0.0	1.0.0	1.0.0
Environmental conditions and measurements	**instantaneous flow rate measurement method**	GENEPIO:0100910	The process used to measure instantaneous fluid flow rate.	Provide the name of the procedure or technology used to measure flow rate.	Flow meter			1.0.0	1.0.0	1.0.0
Environmental conditions and measurements	**turbidity measurement value**	GENEPIO:0100783	The numerical value of a measurement of turbidity.	Provide the numerical value of the measured turbidity.	0.02			1.0.0	1.0.0	1.0.0
Environmental conditions and measurements	**turbidity measurement unit**	GENEPIO:0100914	The units of a measurement of turbidity.	Provide the units of the measured turbidity by selecting a value from the pick list.	nephelometric turbidity unit (NTU)			1.0.0	1.0.0	1.0.0
Environmental conditions and measurements	**turbidity measurement method**	GENEPIO:0101013	The process used to measure turbidity.	Provide the name of the procedure or technology used to measure turbidity.	Nephelometric method			1.0.0	1.0.0	1.0.0
Environmental conditions and measurements	**dissolved oxygen measurement value**	GENEPIO:0100915	The numerical value of a measurement of dissolved oxygen.	Provide the numerical value of the measured dissolved oxygen.	5			1.0.0	1.0.0	1.0.0
Environmental conditions and measurements	**dissolved oxygen measurement unit**	GENEPIO:0100784	The units of a measurement of dissolved oxygen.	Provide the units of the measured dissolved oxygen by selecting a value from the pick list.	part per million (ppm)			1.0.0	1.0.0	1.0.0
Environmental conditions and measurements	**dissolved oxygen measurement method**	GENEPIO:0100785	The method used to measure dissolved oxygen.	Provide the name of the procedure or technology used to measure dissolved oxygen.	Dissolved oxygen meter in vertical direction			1.0.0	1.0.0	1.0.0
Environmental conditions and measurements	**oxygen reduction potential (ORP) measurement value**	GENEPIO:0100917	The numerical value of a measurement of oxygen reduction potential (ORP).	Provide the numerical value of the measured oxygen reduction potential.	-50			1.0.0	1.0.0	1.0.0
Environmental conditions and measurements	**oxygen reduction potential (ORP) measurement unit**	GENEPIO:0100786	The units of a measurement of oxygen reduction potential (ORP).	Provide the units of the measured oxygen reduction potential by selecting a value from the pick list.	milliVolt (mV)			1.0.0	1.0.0	1.0.0

Parent Class	Field Colour Code Legend	Ontology Identifier	Definition	Guidance	Examples	Deprecated Label	Deprecated ID	Version Tracking Label	ID	Description/Guidance
Environmental conditions and measurements	**oxygen reduction potential (ORP) measurement method**	GENEPIO:0100787	The method used to measure oxygen reduction potential (ORP).	Provide the name of the procedure or technology used to measure oxygen reduction potential.	ORP sensor			1.0.0	1.0.0	1.0.0
Environmental conditions and measurements	**chemical oxygen demand (COD) measurement value**	GENEPIO:0100788	The measured value from a chemical oxygen demand (COD) test.	Provide the numerical value of the COD test result.	26			1.0.0	1.0.0	1.0.0
Environmental conditions and measurements	**chemical oxygen demand (COD) measurement unit**	GENEPIO:0100789	The units associated with a value from a chemical oxygen demand (COD) test.	Provide the units of the COD test result.	milligram per liter (mg/L)			1.0.0	1.0.0	1.0.0
Environmental conditions and measurements	**chemical oxygen demand (COD) measurement method**	GENEPIO:0100790	The method used to measure chemical oxygen demand (COD).	Provide the name of the procedure or technology used to measure COD.	Hach LCK test kit			1.0.0	1.0.0	1.0.0
Environmental conditions and measurements	**carbonaceous biochemical oxygen demand (CBOD) measurement value**	GENEPIO:0100791	The numerical value of a measurement of carbonaceous biochemical oxygen demand (CBOD).	Provide the numerical value of the measured CBOD.	20			1.0.0	1.0.0	1.0.0
Environmental conditions and measurements	**carbonaceous biochemical oxygen demand (CBOD) measurement unit**	GENEPIO:0100792	The units of a measurement of carbonaceous biochemical oxygen demand (CBOD).	Provide the units of the measured CBOD by selecting a value from the pick list.	milligram per liter (mg/L)			1.0.0	1.0.0	1.0.0
Environmental conditions and measurements	**carbonaceous biochemical oxygen demand (CBOD) measurement method**	GENEPIO:0100793	The method used to measure carbonaceous biochemical oxygen demand (CBOD).	Provide the name of the procedure or technology used to measure CBOD.	CBOD measurement by optical probe			1.0.0	1.0.0	1.0.0
Environmental conditions and measurements	**total suspended solids (TSS) measurement value**	GENEPIO:0100794	The mass of suspended particulates, both organic and inorganic in a sample.	Provide the numerical value of the measured TSS.	8			1.0.0	1.0.0	1.0.0
Environmental conditions and measurements	**total suspended solids (TSS) measurement unit**	GENEPIO:0100795	The units associated with a value from a total suspended solids (TSS) test.	Provide the units of the measured TSS.	percent (%)			1.0.0	1.0.0	1.0.0
Environmental conditions and measurements	**total suspended solids (TSS) measurement method**	GENEPIO:0100796	The method used to measure total suspended solids (TSS).	Provide the name of the procedure or technology used to measure TSS.	Vacuum filter through a 2-micron filter, then oven-dried and weighed sample			1.0.0	1.0.0	1.0.0
Environmental conditions and measurements	**total dissolved solids (TDS) measurement value**	GENEPIO:0100797	The numerical value from a total dissolved solids (TDS) test.	Provide the numerical value of the measured TDS.	2			1.0.0	1.0.0	1.0.0
Environmental conditions and measurements	**total dissolved solids (TDS) measurement unit**	GENEPIO:0100798	The units associated with a value from a total dissolved solids (TDS) test.	Provide the units of the measured TDS.	percent (%)			1.0.0	1.0.0	1.0.0
Environmental conditions and measurements	**total dissolved solids (TDS) measurement method**	GENEPIO:0100799	The method used to measure total dissolved solids (TDS).	Provide the name of the procedure or technology used to measure TDS.	Subtract calculated TSS from calculated TS			1.0.0	1.0.0	1.0.0
Environmental conditions and measurements	**total solids (TS) measurement value**	GENEPIO:0100800	The numerical value from a total solids (TS) test.	Provide the numerical value of the measured TS.	10			1.0.0	1.0.0	1.0.0
Environmental conditions and measurements	**total solids (TS) measurement unit**	GENEPIO:0100801	The units associated with a value from a total solids (TS) test.	Provide the units of the measured TS.	percent (%)			1.0.0	1.0.0	1.0.0
Environmental conditions and measurements	**total solids (TS) measurement method**	GENEPIO:0100802	The method used to measure total solids (TS).	Provide the name of the procedure or technology used to measure TS.	Gravimetric method by oven drying, then weighing			1.0.0	1.0.0	1.0.0
Environmental conditions and measurements	**alkalinity measurement value**	GENEPIO:0100878	The numerical value of a measurement of alkalinity.	Provide the numerical value of the measured alkalinity.	3			1.0.0	1.0.0	1.0.0
Environmental conditions and measurements	**alkalinity measurement unit**	GENEPIO:0100879	The units of a measurement of alkalinity.	Provide the units of the measured alkalinity.	milligram per liter of calcium carbonate (mg/L CaCO3)			1.0.0	1.0.0	1.0.0
Environmental conditions and measurements	**alkalinity measurement method**	GENEPIO:0100880	The process used to measure alkalinity.	Provide the name of the procedure or technology used to measure alkalinity.	Titration method			1.0.0	1.0.0	1.0.0
Environmental conditions and measurements	**conductivity measurement value**	GENEPIO:0100916	The numerical value of a measurement of conductivity.	Provide the numerical value of the measured conductivity.	1412			1.0.0	1.0.0	1.0.0
Environmental conditions and measurements	**conductivity measurement unit**	GENEPIO:0100803	The units of a measurement of conductivity.	Provide the units of the measured conductivity.	microSiemen per centimeter (µS/cm)			1.0.0	1.0.0	1.0.0
Environmental conditions and measurements	**conductivity measurement method**	GENEPIO:0100804	The method used to measure conductivity.	Provide the name of the procedure or technology used to measure conductivity.	Conductivity electrode and meter			1.0.0	1.0.0	1.0.0
Environmental conditions and measurements	**salinity measurement value**	GENEPIO:0100805	The numerical value of a measurement of salinity.	Provide the numerical value of the measured salinity.	35			1.0.0	1.0.0	1.0.0
Environmental conditions and measurements	**salinity measurement unit**	GENEPIO:0100806	The units of a measurement of salinity.	Provide the units of the measured salinity.	practical salinity unit (PSU)			1.0.0	1.0.0	1.0.0
Environmental conditions and measurements	**salinity measurement method**	GENEPIO:0100807	The method used to measure salinity.	Provide the name of the procedure or technology used to measure salinity.	conductivity meter			1.0.0	1.0.0	1.0.0
Environmental conditions and measurements	**total nitrogen (TN) measurement value**	GENEPIO:0100808	The numerical value of a measurement of total nitrogen (TN).	Provide the numerical value of the measured TN.	120			1.0.0	1.0.0	1.0.0
Environmental conditions and measurements	**total nitrogen (TN) measurement unit**	GENEPIO:0100809	The units of a measurement of total nitrogen (TN).	Provide the units of the measured TN.	milligram per liter (mg/L)			1.0.0	1.0.0	1.0.0
Environmental conditions and measurements	**total nitrogen (TN) measurement method**	GENEPIO:0100810	The method used to measure total nitrogen (TN).	Provide the name of the procedure or technology used to measure TN.	Hach total nitrogen spectrophotometric test			1.0.0	1.0.0	1.0.0
Environmental conditions and measurements	**total phosphorus (TP) measurement value**	GENEPIO:0100811	The numerical value of a measurement of total phosphorus (TP).	Provide the numerical value of the measured TP.	2			1.0.0	1.0.0	1.0.0
Environmental conditions and measurements	**total phosphorus (TP) measurement unit**	GENEPIO:0100812	The units of a measurement of total phosphorus (TP).	Provide the units of the measured TP.	milligrams orthophosphate as phosphorus per liter (mg PO4-P/L)			1.0.0	1.0.0	1.0.0
Environmental conditions and measurements	**total phosphorus (TP) measurement method**	GENEPIO:0100813	The method used to measure total phosphorus (TP).	Provide the name of the procedure or technology used to measure TP.	Merck phosphate spectrophotometric test kit			1.0.0	1.0.0	1.0.0
Environmental conditions and measurements	**fecal contamination indicator**	GENEPIO:0100814	A gene, virus, bacteria, or substance used to measure the sanitary quality of water in regards to fecal contamination.	If a fecal contamination indicator was measured, select it from the picklist.	crAssphage			1.0.0	1.0.0	1.0.0
Environmental conditions and measurements	**fecal contamination value**	GENEPIO:0100815	The numerical value of a measurement of fecal contamination.	Provide the numerical value of the measured fecal contamination.	10			1.0.0	1.0.0	1.0.0
Environmental conditions and measurements	**fecal contamination unit**	GENEPIO:0100816	The units of a measurement of fecal contamination.	Provide the units of the measured fecal contamination.	cycle threshold (Ct) / quantification cycle (Cq)			1.0.0	1.0.0	1.0.0
Environmental conditions and measurements	**fecal contamination method**	GENEPIO:0100817	The method used to measure fecal contamination.	Provide the name of the procedure or technology used to measure fecal contamination.	quantitative PCR assay			1.0.0	1.0.0	1.0.0

Parent Class	Field Colour Code Legend	Ontology Identifier	Definition	Guidance	Examples	Deprecated Label	Deprecated ID	Version Tracking Label	ID	Description/Guidance
Environmental conditions and measurements	**fecal coliform count value**	GENEPIO:0100818	The numerical value of a measurement of fecal coliforms within a sample.	Provide the numerical value of the measured fecal coliforms.	3			1.0.0	1.0.0	1.0.0
Environmental conditions and measurements	**fecal coliform count unit**	GENEPIO:0100819	The units of a measurement of fecal coliforms.	Provide the units of the measured fecal coliforms.	most probable number per milliliter (MPN/mL)			1.0.0	1.0.0	1.0.0
Environmental conditions and measurements	**fecal coliform count method**	GENEPIO:0100820	The method used to measure fecal coliforms.	Provide the name of the procedure or technology used to measure fecal coliforms.	MPN method via serial dilutions until lack of growth			1.0.0	1.0.0	1.0.0
Environmental conditions and measurements	**urinary contamination indicator**	GENEPIO:0100837	A gene, virus, bacteria, or substance used to measure the sanitary quality of water in regards to urinary contamination.	If a urinary contamination indicator was measured, select it from the picklist.	urobilin			1.0.0	1.0.0	1.0.0
Environmental conditions and measurements	**urinary contamination value**	GENEPIO:0100838	The numerical value of a measurement of urinary contamination.	Provide the numerical value of the measured urinary contamination.	3			1.0.0	1.0.0	1.0.0
Environmental conditions and measurements	**urinary contamination unit**	GENEPIO:0100839	The units of a measurement of urinary contamination.	Provide the units of the measured urinary contamination.	nanograms per liter			1.0.0	1.0.0	1.0.0
Environmental conditions and measurements	**urinary contamination method**	GENEPIO:0100840	The method used to measure urinary contamination.	Provide the name of the procedure or technology used to measure urinary contamination.				1.0.0	1.0.0	1.0.0
Environmental conditions and measurements	**sample temperature value (at collection)**	GENEPIO:0100821	The numerical value of a measurement of temperature of a sample at collection.	Provide the numerical value of the measured temperature.	20			1.0.0	1.0.0	1.0.0
Environmental conditions and measurements	**sample temperature unit (at collection)**	GENEPIO:0100822	The units of a measurement of temperature of a sample at the time of collection.	Provide the units of the measured temperature.	degree Celsius (C)			1.0.0	1.0.0	1.0.0
Environmental conditions and measurements	**sample temperature value (when received)**	GENEPIO:0100823	The numerical value of a measurement of temperature of a sample upon receipt.	Provide the numerical value of the measured temperature.	22			1.0.0	1.0.0	1.0.0
Environmental conditions and measurements	**sample temperature unit (when received)**	GENEPIO:0100824	The units of a measurement of temperature of a sample at the time upon receipt.	Provide the units of the measured temperature.	degree Celsius (C)			1.0.0	1.0.0	1.0.0
Sequence information		GENEPIO:0001441								
Sequence information	**purpose of sequencing**	GENEPIO:0001445	The reason that the sample was sequenced.	The reason why a sample was originally collected may differ from the reason why it was selected for sequencing. The reason a sample was sequenced may provide information about potential biases in sequencing strategy. Provide the purpose of sequencing from the picklist in the template. The reason for sample collection should be indicated in the "purpose of sampling" field.	Travel-associated surveillance			1.0.0	1.0.0	1.0.0
Sequence information	**purpose of sequencing details**	GENEPIO:0001446	The description of why the sample was sequenced providing specific details.	Provide an expanded description of why the sample was sequenced using free text. The description may include the importance of the sequences for a particular public health investigation/surveillance activity/research question. Suggested standardized descriptions include: Assessing public health control measures, Determining early introductions and spread, Investigating airline-related exposures, Investigating remote regions, Investigating health care workers, Investigating schools/universities.	Investigating schools/universities			1.0.0	1.0.0	1.0.0
Sequence information	**sequenced_by**	GENEPIO:0100416	The name of the agency, organization or institution responsible for sequencing the isolate's genome.	Provide the name of the agency, organization or institution that performed the sequencing in full (avoid abbreviations). If the information is unknown or cannot be provided, leave blank or provide a null value.	Public Health Agency of Canada (PHAC) [GENEPIO:0100551]			3.0.0	3.0.0	3.0.0
Sequence information	**sequenced_by_contact_name**	GENEPIO:0100471	The name or title of the contact responsible for follow-up regarding the sequence.	Provide the name of an individual or their job title. As personnel turnover may render the contact's name obsolete, it is more prefereable to provide a job title for ensuring accuracy of information and institutional memory. If the information is unknown or cannot be provided, leave blank or provide a null value.	Enterics Lab Manager			3.0.0	3.0.0	3.0.0
Sequence information	**sequenced_by_contact_email**	GENEPIO:0100422	The email address of the contact responsible for follow-up regarding the sequence.	Provide the email associated with the listed contact. As personnel turnover may render an individual's email obsolete, it is more prefereable to provide an address for a position or lab, to ensure accuracy of information and institutional memory. If the information is unknown or cannot be provided, leave blank or provide a null value.	enterics@lab.ca			3.0.0	3.0.0	3.0.0
Sequence information	**sequence submitted by**	GENEPIO:0001159	The name of the agency that submitted the sequence to a database.	The name of the agency should be written out in full, (with minor exceptions) and be consistent across multiple submissions. For Canadian institutions submitting specimens rather than sequencing data, please put the "National Microbiology Laboratory (NML)".	Public Health Ontario (PHO)			1.0.0	1.0.0	1.0.0
Sequence information	**sequence submitter contact email**	GENEPIO:0001165	The email address of the contact responsible for follow-up regarding the sequence.	The email address can represent a specific individual or laboratory.	RespLab@lab.ca			1.0.0	1.0.0	1.0.0
Sequence information	**sequencing date**	GENEPIO:0001447	The date the sample was sequenced.	ISO 8601 standard "YYYY-MM-DD".	2020-06-22			1.0.0	1.0.0	1.0.0
Sequence information	**library ID**	GENEPIO:0001448	The user-specified identifier for the library prepared for sequencing.	The library name should be unique, and can be an autogenerated ID from your LIMS, or modification of the isolate ID.	XYZ_123345			1.0.0	1.0.0	1.0.0
Sequence information	**sequencing platform**	GENEPIO:0100473	The platform technology used to perform the sequencing.	Provide the name of the company that created the sequencing instrument by selecting a value from the template pick list. If the information is unknown or cannot be provided, leave blank or provide a null value.	Illumina [GENEPIO:0001923]			1.0.0	1.0.0	1.0.0
Sequence information	**sequencing instrument**	GENEPIO:0001452	The model of the sequencing instrument used.	Provide the model sequencing instrument by selecting a value from the template pick list. If the information is unknown or cannot be provided, leave blank or provide a null value.	Illumina HiSeq 2500 [GENEPIO:0100117]			1.0.0	1.0.0	1.0.0

Parent Class	Field Colour Code Legend	Ontology Identifier	Definition	Guidance	Examples	Deprecated Label	Deprecated ID	Version Tracking Label	ID	Description/Guidance
Sequence information	**sequencing assay type**	GENEPIO:0100997	The overarching sequencing methodology that was used to determine the sequence of a biomaterial.	Example Guidance: Provide the name of the DNA or RNA sequencing technology used in your study. If unsure refer to the protocol documentation, or provide a null value.	whole genome sequencing assay			1.0.0	1.0.0	1.0.0
Sequence information	**library preparation kit**	GENEPIO:0001450	The name of the DNA library preparation kit used to generate the library being sequenced.	Provide the name of the library preparation kit used.	Nextera XT			1.0.0	1.0.0	1.0.0
Sequence information	**sequencing protocol**	GENEPIO:0001454	The protocol or method used for sequencing.	Provide the name and version of the procedure or protocol used for sequencing. You can also provide a link to a protocol online.	https://www.protocols.io/view/ncov-2019-sequencing-protocol-bbmuik6w?version_warning=no			1.0.0	1.0.0	1.0.0
Sequence information	**amplicon pcr primer scheme**	GENEPIO:0001456	The specifications of the primers (primer sequences, binding positions, fragment size generated etc) used to generate the amplicons to be sequenced.	Provide the name and version of the primer scheme used to generate the amplicons for sequencing.	artic v3			1.0.0	1.0.0	1.0.0
Sequence information	**amplicon size**	GENEPIO:0001449	The length of the amplicon generated by PCR amplification.	Provide the amplicon size expressed in base pairs.	300			1.0.0	1.0.0	1.0.0
Sequence information	**DNA fragment length**	GENEPIO:0100843	The length of the DNA fragment generated by mechanical shearing or enzymatic digestion for the purposes of library preparation.	Provide the fragment length in base pairs (do not include the units).	400			1.0.0	1.0.0	1.0.0
Sequence information	**genomic target enrichment method**	GENEPIO:0100966	The molecular technique used to selectively capture and amplify specific regions of interest from a genome.	Provide the name of the enrichment method	hybrid selection method			1.0.0	1.0.0	1.0.0
Sequence information	**genomic target enrichment method details**	GENEPIO:0100967	Details that provide additional context to the molecular technique used to selectively capture and amplify specific regions of interest from a genome.	Provide details that are applicable to the method you used.	enrichment was done using Twist's respiratory virus research panel: https://www.twistbioscience.com/products/ngs/fixed-panels/respiratory-virus-research-panel			1.0.0	1.0.0	1.0.0
Bioinformatics and QC metrics		**GENEPIO:0001457**								
Bioinformatics and QC metrics	**quality control method name**	GENEPIO:0100557	The name of the method used to assess whether a sequence passed a predetermined quality control threshold.	Providing the name of the method used for quality control is very important for interpreting the rest of the QC information. Method names can be provided as the name of a pipeline or a link to a GitHub repository. Multiple methods should be listed and separated by a semi-colon. Do not include QC tags in other fields if no method name is provided.	ncov-tools			1.0.0	1.0.0	1.0.0
Bioinformatics and QC metrics	**quality control method version**	GENEPIO:0100558	The version number of the method used to assess whether a sequence passed a predetermined quality control threshold.	Methods updates can make big differences to their outputs. Provide the version of the method used for quality control. The version can be expressed using whatever convention the developer implements (e.g. date, semantic versioning). If multiple methods were used, record the version numbers in the same order as the method names. Separate the version numbers using a semi-colon.	1.2.3			1.0.0	1.0.0	1.0.0
Bioinformatics and QC metrics	**quality control determination**	GENEPIO:0100559	The determination of a quality control assessment.	Select a value from the pick list provided. If a desired value is missing, submit a new term request to the PHA4GE QC Tag GitHub issuetracker using the New Term Request form.	sequence failed quality control			1.0.0	1.0.0	1.0.0
Bioinformatics and QC metrics	**quality control issues**	GENEPIO:0100560	The reason contributing to, or causing, a low quality determination in a quality control assessment.	Select a value from the pick list provided. If a desired value is missing, submit a new term request to the PHA4GE QC Tag GitHub issuetracker using the New Term Request form.	low average genome coverage			1.0.0	1.0.0	1.0.0
Bioinformatics and QC metrics	**quality control details**	GENEPIO:0100561	The details surrounding a low quality determination in a quality control assessment.	Provide notes or details regarding QC results using free text.	CT value of 39. Low viral load. Low DNA concentration after amplification.			1.0.0	1.0.0	1.0.0
Bioinformatics and QC metrics	**raw sequence data processing method**	GENEPIO:0001458	The method used for raw data processing such as removing barcodes, adapter trimming, filtering etc.	Provide the software name followed by the version or a link to the github protocol e.g. Trimmomatic v. 0.38, Porechop v. 0.2.3	Porechop 0.2.3			1.0.0	1.0.0	1.0.0
Bioinformatics and QC metrics	**dehosting method**	GENEPIO:0001459	The method used to remove host reads from the pathogen sequence.	Provide the name and version number of the software used to remove host reads.	Nanostripper			1.0.0	1.0.0	1.0.0
Bioinformatics and QC metrics	**consensus sequence software name**	GENEPIO:0001463	The name of the software used to generate the consensus sequence.	Provide the name of the software used to generate the consensus sequence.	iVar			1.0.0	1.0.0	1.0.0
Bioinformatics and QC metrics	**consensus sequence software version**	GENEPIO:0001469	The version of the software used to generate the consensus sequence.	Provide the version of the software used to generate the consensus sequence.	1.3			1.0.0	1.0.0	1.0.0
Bioinformatics and QC metrics	**sequence assembly software name**	GENEPIO:0100825	The name of the software used to assemble a sequence.	Provide the name of the software used to assemble the sequence.	SPAdes Genome Assembler, Canu, wtdbg2, velvet			1.0.0	1.0.0	1.0.0
Bioinformatics and QC metrics	**sequence assembly software version**	GENEPIO:0100826	The version of the software used to assemble a sequence.	Provide the version of the software used to assemble the sequence.	3.15.5			1.0.0	1.0.0	1.0.0
Bioinformatics and QC metrics	**breadth of coverage value**	GENEPIO:0001472	The percentage of the reference genome covered by the sequenced data, to a prescribed depth.	Provide value as a percent.	95			1.0.0	1.0.0	1.0.0
Bioinformatics and QC metrics	**depth of coverage value**	GENEPIO:0001474	The average number of reads representing a given nucleotide in the reconstructed sequence.	Provide value as a fold of coverage.	400			1.0.0	1.0.0	1.0.0
Bioinformatics and QC metrics	**depth of coverage threshold**	GENEPIO:0001475	The threshold used as a cut-off for the depth of coverage.	Provide the threshold fold coverage.	100			1.0.0	1.0.0	1.0.0
Bioinformatics and QC metrics	**genome completeness**	GENEPIO:0100844	The percentage of expected genes identified in the genome being sequenced. Missing genes indicate missing genomic regions (incompleteness) in the data.	Provide the genome completeness as a percent (no need to include units).	85			1.0.0	1.0.0	1.0.0

Parent Class	Field Colour Code Legend	Ontology Identifier	Definition	Guidance	Examples	Deprecated Label	Deprecated ID	Version Tracking Label	ID	Description/Guidance
Bioinformatics and QC metrics	**number of base pairs sequenced**	GENEPIO:0001482	The number of total base pairs generated by the sequencing process.	Provide a numerical value (no need to include units).	387566			1.0.0	1.0.0	1.0.0
Bioinformatics and QC metrics	**number of total reads**	GENEPIO:0100827	The total number of non-unique reads generated by the sequencing process.	Provide a numerical value (no need to include units).	423867			1.0.0	1.0.0	1.0.0
Bioinformatics and QC metrics	**number of unique reads**	GENEPIO:0100828	The number of unique reads generated by the sequencing process.	Provide a numerical value (no need to include units).	248236			1.0.0	1.0.0	1.0.0
Bioinformatics and QC metrics	**minimum post-trimming read length**	GENEPIO:0100829	The threshold used as a cut-off for the minimum length of a read after trimming.	Provide a numerical value (no need to include units).	150			1.0.0	1.0.0	1.0.0
Bioinformatics and QC metrics	**number of contigs**	GENEPIO:0100937	The number of contigs (contiguous sequences) in a sequence assembly.	Provide a numerical value.	10			1.0.0	1.0.0	1.0.0
Bioinformatics and QC metrics	**percent Ns across total genome length**	GENEPIO:0100830	The percentage of the assembly that consists of ambiguous bases (Ns).	Provide a numerical value (no need to include units).	2			1.0.0	1.0.0	1.0.0
Bioinformatics and QC metrics	**Ns per 100 kbp**	GENEPIO:0001484	The number of ambiguous bases (Ns) normalized per 100 kilobasepairs (kbp).	Provide a numerical value (no need to include units).	342			1.0.0	1.0.0	1.0.0
Bioinformatics and QC metrics	**N50 value**	GENEPIO:0100938	The length of the shortest read that, together with other reads, represents at least 50% of the nucleotides in a set of sequences.	Provide the N50 value in Mb.	150			1.0.0	1.0.0	1.0.0
Bioinformatics and QC metrics	**percent read contamination**	GENEPIO:0100845	The percent of the total number of reads identified as contamination (not belonging to the target organism) in a sequence dataset.	Provide the percent contamination value (no need to include units).	2			1.0.0	1.0.0	1.0.0
Bioinformatics and QC metrics	**consensus genome length**	GENEPIO:0001483	The length of the genome defined by the most common nucleotides at each position.	Provide a numerical value (no need to include units).	38677			1.0.0	1.0.0	1.0.0
Bioinformatics and QC metrics	**sequence assembly length**	GENEPIO:0100846	The length of the genome generated by assembling reads using a scaffold or by reference-based mapping.	Provide a numerical value (no need to include units).	34272			1.0.0	1.0.0	1.0.0
Bioinformatics and QC metrics	**reference genome accession**	GENEPIO:0001485	A persistent, unique identifier of a genome database entry.	Provide the accession number of the reference genome.	NC_045512.2			1.0.0	1.0.0	1.0.0
Bioinformatics and QC metrics	**deduplication method**	GENEPIO:0100831	The method used to remove duplicated reads in a sequence read dataset.	Provide the deduplication software name followed by the version, or a link to a tool or method.	DeDup 0.12.8			1.0.0	1.0.0	1.0.0
Bioinformatics and QC metrics	**read mapping software name**	GENEPIO:0100832	The name of the software used to map sequence reads to a reference genome or set of reference genes.	Provide the name of the read mapping software.	Bowtie2, BWA-MEM, TopHat			1.0.0	1.0.0	1.0.0
Bioinformatics and QC metrics	**read mapping software version**	GENEPIO:0100833	The version of the software used to map sequence reads to a reference genome or set of reference genes.	Provide the version number of the read mapping software.	2.5.1			1.0.0	1.0.0	1.0.0
Bioinformatics and QC metrics	**taxonomic reference database name**	GENEPIO:0100834	The name of the taxonomic reference database used to identify the organism.	Provide the name of the taxonomic reference database.	NCBITaxon			1.0.0	1.0.0	1.0.0
Bioinformatics and QC metrics	**taxonomic reference database version**	GENEPIO:0100835	The version of the taxonomic reference database used to identify the organism.	Provide the version number of the taxonomic reference database.	1.3			1.0.0	1.0.0	1.0.0
Bioinformatics and QC metrics	**read mapping criteria**	GENEPIO:0100836	A description of the criteria used to map reads to a reference sequence.	Provide a description of the read mapping criteria.	Phred score >20			1.0.0	1.0.0	1.0.0
Bioinformatics and QC metrics	**bioinformatics protocol**	GENEPIO:0001489	A description of the overall bioinformatics strategy used.	Further details regarding the methods used to process raw data, and/or generate assemblies, and/or generate consensus sequences can. This information can be provided in an SOP or protocol or pipeline/workflow. Provide the name and version number of the protocol, or a GitHub link to a pipeline or workflow.	https://github.com/phac-nml/ncov2019-artic-nf			1.0.0	1.0.0	1.0.0
Pathogen diagnostic testing		**GENEPIO:0001506**								
Pathogen diagnostic testing	**organism**	GENEPIO:0001191	The taxonomic name of the organism.	Put the genus and species (and subspecies if applicable) if known.The standardized term can be sourced from this look-up service: https://www.ebi.ac.uk/ols4/ontologies/ncbitaxon	Severe acute respiratory syndrome coronavirus 2			1.0.0	1.0.0	1.0.0
Pathogen diagnostic testing	**gene name**	GENEPIO:0100655	The name of the gene used in the diagnostic RT-PCR test.	Provide the full name of the gene used in the test. The gene symbol (short form of gene name) can also be provided. Standardized gene names and symbols can be found in the Gene Ontology using this look-up service: https://bit.ly/2Sq1LbI	E gene (orf4)			1.0.0	1.0.0	1.0.0
Pathogen diagnostic testing	**diagnostic target presence**	GENEPIO:0100962	The binary value of the result from a diagnostic test.	Select a value from the pick list provided, to describe whether a target was determined to be present or absent within a sample.	present			1.0.0	1.0.0	1.0.0
Pathogen diagnostic testing	**diagnostic measurement value**	GENEPIO:0100963	The value of the result from a diagnostic test.	Provide the numerical result of a diagnostic test.	1000			1.0.0	1.0.0	1.0.0
Pathogen diagnostic testing	**diagnostic measurement unit**	GENEPIO:0100964	The unit of the result from a diagnostic test.	Select a value from the pick list provided, to describe the units of the given diagnostic test.	CFU/mL			1.0.0	1.0.0	1.0.0
Pathogen diagnostic testing	**diagnostic measurement method**	GENEPIO:0100965	The method by which a diagnostic result was received.	Select a value from the pick list provided to describe the method used for a given diagnostic test.	qPCR			1.0.0	1.0.0	1.0.0
Contributor acknowledgement		**GENEPIO:0001516**								
Contributor acknowledgement	**authors**	GENEPIO:0001517	Names of individuals contributing to the processes of sample collection, sequence generation, analysis, and data submission.	Include the first and last names of all individuals that should be attributed, separated by a comma.	Tejinder Singh, Fei Hu, Joe Blogs			1.0.0	1.0.0	1.0.0