B	F1.1.1	Ontology	D. C. W.	2.11		Deprecated Label Deprecated ID	Version Tracking		
Parent Class	Field	Identifier	Definition	Guidance	Examples	IMPORTANT: Only labels and/or IDs will be	Label	ID	Description/Gui
	Colour Code Legend					deprecated, always with replacement version provided. If a term changes in its meaning, a			dance
	field name in yellow = required field name in purple = recommended					new term will be created.			
	field name in white = optional								
	Database Identifiers	GENEPIO:0001122		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					
Database Identifiers	specimen collector sample ID	GENEPIO:0001123	The user-defined name for the sample.	make it concise, unique and consistent within your lab.	prov_mpox_1234		1.0.0	1.0.0	1.0.0
				Provide the case identifer. The case ID greatly facilitates linkage between laboratory and epidemiological data. The					
			The identifier used to specify an epidemiologically detected case	case ID may be considered identifiable information. Consul	It				
Database Identifiers	case ID	GENEPIO:0100281	of disease.	the data steward before sharing.	ABCD1234		1.0.0	1.0.0	1.0.0
				Required if submission is linked to a BioProject. BioProject are an organizing tool that links together raw sequence	s				
				data, assemblies, and their associated metadata. A valid BioProject accession has prefix PRJN, PRJE or PRJD,					
				e.g., PRJNA12345 and is created once at the beginning of					
Database Identifiers	bioproject accession	GENEPIO:0001136	The INSDC accession number of the BioProject(s) to which the BioSample belongs.	a new sequencing project. Your laboratory can have one or many BioProjects.	PR.INA12345		100	100	100
				Store the accession returned from the BioSample					
Database Identifiers	biosample accession	GENEPIO:0001139	The identifier assigned to a BioSample in INSDC archives.	submission. NCBI BioSamples will have the prefix SAMN, while ENA BioSamples will have the prefix SAMEA.	SAMN14180202		1.0.0	1.0.0	1.0.0
Database Identifier	Distample decession	CENEL 10.000 F100	The labellation aboughed to a biocampio in mobile arounds.	Store the accession assigned to the submitted sequence.	G/411111100202		1.0.0	1.0.0	1.0.0
				European Nucleotide Archive (ENA) sequence accessions start with ERR. NCBI-SRA accessions start with SRR.					
			The identifier assigned to a sequence in one of the International	DNA Data Bank of Japan (DDBJ) accessions start with					
Database identifiers	INSDC sequence read accession	GENEPIO:0101203	Nucleotide Sequence Database Collaboration (INSDC) repositories.	DRR and Genome Sequence Archive (GSA) accessions start with CRR.	SRR123456, ERR123456, DRR123456, CRR123456		3.0.0	1.0.0	1.0.0
Database identifiers	INODE Sequence read accession	GENET 10:0101203	The versioned identifier assigned to an assembly or consensus	Start Will Old.	UKK123430		5.0.0	1.0.0	1.0.0
Database identifiers	INSDC assembly accession	GENEPIO:0101204	sequence in one of the International Nucleotide Sequence Database Collaboration (INSDC) repositories.	Store the versioned accession assigned to the submitted sequence e.g. the GenBank accession version.	1 7986655 1		3.0.0	1.0.0	1.0.0
Database Identifiers	GISAID virus name	GENEPIO:0101204	Identifier of the specific isolate.	Provide the GISAID EpiPox virus name, which should be w			1.0.0	1.0.0	1.0.0
			·	· ·	·				
Database Identifiers	GISAID accession	GENEPIO:0001147	The GISAID accession number assigned to the sequence.	Store the accession returned from the GISAID submission.	EPI_ISL_436489		1.0.0	1.0.0	1.0.0
	Sample collection and processing	GENEPIO:0001150							
Sample collection and processing	sample collected by	GENEPIO:0001153	The name of the agency that collected the original sample.	The name of the sample collector should be written out in f	ul BC Centre for Disease Control		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 lat e.g. johnnyblogs@lab.ca, or RespLab@lab.ca	RespLab@lab.ca		100	100	1.0.0
Sample collection and processing	sample collector contact email	GENEPIO:0001130	regarding the sample.	The mailing address should be in the format: Street number	1 0		1.0.0	1.0.0	1.0.0
Sample collection and processing	sample collector contact address	GENEPIO:0001158	The mailing address of the agency submitting the sample.	and name, City, Province/Territory, Postal Code, Country	2A2, Canada		1.0.0	1.0.0	1.0.0
				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.					
Sample collection and processing	sample collection date	GENEPIO:0001174	The date on which the sample was collected.	The date should be provided in ISO 8601 standard format "YYYY-MM-DD".	2020-03-1	6	1.0.0	1.0.0	1.0.0
		GENEPIO:0001179		ISO 8601 standard "YYYY-MM-DD"	2020-03-2	_	100	100	100
Sample collection and processing Sample collection and processing	sample received date geo loc name (country)	GENEPIO:0001179 GENEPIO:0001181	The date on which the sample was received. The country where the sample was collected.	Provide the country name from the controlled vocabulary p		20	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 where the sample was collected.	Provide the state/province/territory name from the controlle			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 co				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				1.0.0	1.0.0	1.0.0
Sample collection and processing	organism	GENEPIO:0001191 GENEPIO:0001644	Taxonomic name of the organism.	Use "Mpox virus". This value is provided in the template. N			1.0.0	1.0.0	1.0.0
Sample collection and processing Sample collection and processing	purpose of sampling	GENEPIO:0001644 GENEPIO:0001198	Identifier of the specific isolate. The reason that the sample was collected.	This identifier should be an unique, indexed, alpha-numeric As all samples are taken for diagnostic purposes, "Diagnos			1.0.0	1.0.0	1.0.0
campis concentration and processing	purpose or sampling	GENETIO.0001190	The reason that the sample was collected.	Provide an expanded description of why the sample was	Sugmosio teating [OLIVEI 10.0100002]		1.0.0	1.0.0	1.0.0
				collected using free text. The description may include the importance of the sample for a particular public health					
			The description of why the sample was collected, providing	investigation/surveillance activity/research question. If	Symptomology and history suggested				
Sample collection and processing	purpose of sampling details	GENEPIO:0001200	specific details.	details are not available, provide a null value.	Monkeypox diagnosis.		1.0.0	1.0.0	1.0.0
Sample collection and processing Sample collection and processing	anatomical material	GENEPIO:0001211 GENEPIO:0001214	A substance obtained from an anatomical part of an organism e.g. An anatomical part of an organism e.g. oropharvnx.	.1 Provide a descriptor if an anatomical material was sampled. Us Provide a descriptor if an anatomical part was sampled. Us	, ,, , , , , , ,		1.0.0	1.0.0	1.0.0
Sample collection and processing	body product	GENEPIO:0001214	A substance excreted/secreted from an organism e.g. feces, urine				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				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. swab.				1.0.0	1.0.0	1.0.0
Sample collection and processing	collection method	GENEPIO:0001241	The process used to collect the sample e.g. phlebotamy, necrops				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	s Critical for interpreting data. Select all the applicable proces	ss Specimens pooled [OBI:0600016] 5 swabs from different body sites were pooled		1.0.0	1.0.0	1.0.0
			Detailed information regarding the processing applied to a sample	Provide a free text description of any processing details	and further prepared as a single sample during				
Sample collection and processing	specimen processing details	GENEPIO:0100311	during or after receiving the sample.	applied to a sample.	library prep.		1.0.0	1.0.0	1.0.0

				Samples can play different types of roles in experiments. A sample under study in one experiment may act as a control or be a replicate of another sample in another experiment. This field is used to distinguish samples under study from controls, replicates, etc. If the sample acted as an experimental control or a replicate, select a role type from the picklist. If the sample was not a control, leave blank or	Positive experimental control				
Sample collection and processing	experimental specimen role type	GENEPIO:0100921	The type of role that the sample represents in the experiment.	select "Not Applicable".	[GENEPIO:0101018]		3.0.0	3.0.0	3.0.0
Sample collection and processing	experimental control details	GENEPIO:0100922	The details regarding the experimental control contained in the san	Provide details regarding the nature of the reference strain u	Human coronavirus 229E (HCoV-229E) spiked in sa	ample as process control	3.0.0	3.0.0	3.0.0
	Host Information	GENEPIO:0001268					1.0.0	1.0.0	1.0.0
Host Information	host (common name)	GENEPIO:0001386	The commonly used name of the host.	Common name or scientific name are required if there was a host. Both can be provided, if known. Use terms from the pick lists in the template. Common name e.g. human.	Human		1.0.0	1.0.0	1.0.0
			·	Common name or scientific name are required if there was a host. Both can be provided, if known. Use terms from the pick lists in the template. Scientific name e.g. Homo sapiens, if the sample was environmental, put "not					
Host Information	host (scientific name)	GENEPIO:0001387	The taxonomic, or scientific name of the host.	applicable	Homo sapiens [NCBITaxon:9606]		1.0.0	1.0.0	1.0.0
Host Information	host health state	GENEPIO:0001388	Health status of the host at the time of sample collection. Further details pertaining to the health or disease status of the	If known, select a value from the pick list.	Asymptomatic [NCIT:C3833]		1.0.0	1.0.0	1.0.0
Host Information	host health status details	GENEPIO:0001389	host at time of collection.	If known, select a descriptor from the pick list provided in the template.	Hospitalized [NCIT:C25179]		1.0.0	1.0.0	1.0.0
Host Information	host health outcome	GENEPIO:0001389	Disease outcome in the host.	If known, select a value from the pick list.	Recovered [NCIT:C49498]		1.0.0	1.0.0	1.0.0
Host Information	host disease	GENEPIO:0001391	The name of the disease experienced by the host.	Select "Mpox" from the pick list provided in the template. No	Mpox [MONDO:0002594]		1.0.0	1.0.0	1.0.0
				This identifier can be used to link samples from the same					
Host Information	host subject ID	GENEPIO:0001398	A unique identifier by which each host can be referred to.	individual. Caution: consult the data steward before sharing as this value may be considered identifiable information.	12345B-222		1.0.0	1.0.0	1.0.0
Host Information	host age	GENEPIO:0001392	Age of host at the time of sampling.	If known, provide age. Age-binning is also acceptable.	79		1.0.0	1.0.0	1.0.0
				If known, provide the age units used to measure the host's					
Host Information	host age unit	GENEPIO:0001393	The units used to measure the host's age.	age from the pick list.	year [UO:0000036]		1.0.0	1.0.0	1.0.0
Host Information	host age bin	GENEPIO:0001394	The age category of the host at the time of sampling.	Age bins in 10 year intervals have been provided. If a host's age cannot be specified due to provacy concerns, an age bin can be used as an alternative.	50 - 59 [GENEPIO:0100054]		1.0.0	1.0.0	1.0.0
Host Information	host gender	GENEPIO:0001395	The gender of the host at the time of sample collection.	If known, select a value from the pick list.	Male [NCIT:C46109]		1.0.0	1.0.0	1.0.0
Host Information	host residence geo_loc name (country)	GENEPIO:0001396	The country of residence of the host.	Select the country name from pick list provided in the template.	Canada [GAZ:00002560]		1.0.0	1.0.0	1.0.0
1103t IIIOIIIIaioii	nost residence geo_loc name (country)	OEINEI 10:0001330	The country of residence of the riost.	If known, provide the symptom onset date in ISO 8601	Canada (CAZ.00002300)		1.0.0	1.0.0	1.0.0
Host Information	symptom onset date	GENEPIO:0001399	The date on which the symptoms began or were first noted.	standard format "YYYY-MM-DD".	2022-05-25		1.0.0	1.0.0	1.0.0
Host Information	signs and symptoms	GENEPIO:0001400	A perceived change in function or sensation, (loss, disturbance or appearance) indicative of a disease, reported by a patient.	Select all of the symptoms experienced by the host from the pick list.	Lesion (Pustule) [NCIT:C78582], Swollen Lymph Nodes [HP:0002716], Myalgia (muscle pain) [HP:0003326]		1.0.0	1.0.0	1.0.0
Host Information	pre-existing conditions and risk factors	GENEPIO:0001401	Patient pre-existing conditions and risk factors. <i>Pre-existing condition: A medical condition that existed prior to the current infection. <inrisk a="" an="" associated="" disease="" factor:="" increased="" infection.<="" of="" or="" risk="" td="" variable="" with=""><td>Select all of the pre-existing conditions and risk factors experienced by the host from the pick list. If the desired term is missing, contact the curation team.</td><td></td><td></td><td>1.0.0</td><td>1.0.0</td><td>1.0.0</td></inrisk></i>	Select all of the pre-existing conditions and risk factors experienced by the host from the pick list. If the desired term is missing, contact the curation team.			1.0.0	1.0.0	1.0.0
Host Information	complications	GENEPIO:0001402	Patient medical complications that are believed to have occurred as a result of host disease.	Select all of the complications experienced by the host from the pick list. If the desired term is missing, contact the curation team.	Delayed wound healing (lesion healing) [MP:0002908]		1.0.0	1.0.0	1.0.0
Host Information	antiviral therapy	GENEPIO:0100580	Treatment of viral infections with agents that prevent viral replication	Provide details of all current antiviral treatment during the cu	Tecovirimat used to treat current Monkeypox infecti	on; AZT administered for concurrent HIV in	e(1.0.0	1.0.0	1.0.0
	Host vaccination information	GENEPIO:0001403					1.0.0	1.0.0	1.0.0
			The vaccination status of the host (fully vaccinated, partially						
Host vaccination information	host vaccination status	GENEPIO:0001404	vaccinated, or not vaccinated).	Select the vaccination status of the host from the pick list.	Not Vaccinated [GENEPIO:0100102]		1.0.0	1.0.0	1.0.0
Host vaccination information	number of vaccine doses received	GENEPIO:0001406	The number of doses of the vaccine recived by the host.	Record how many doses of the vaccine the host has received.	1		1.0.0	1.0.0	1.0.0
Host vaccination information	vaccination dose 1 vaccine name	GENEPIO:0100313	The name of the vaccine administered as the first dose of a vaccine regimen.	Provide the name and the corresponding manufacturer of the Smallpox vaccine administered as the first dose.	IMVAMUNE (Bayarian Nordic)				
		GENEPIO:0100313		Provide the date the first dose of Smallpox vaccine was administered. The date should be provided in ISO 8601 standard format "YYYY-MM-DD"	2022-06-01				
Host vaccination information	vaccination dose 1 vaccination date	GENEPIU:0100314	The date the first dose of a vaccine was administered. A description of the vaccines received and the administration	Free text description of the dates and vaccines administered against a particular disease/set of diseases. It is also acceptable to concatenate the individual dose	2022-06-01				
Host vaccination information	vaccination history	GENEPIO:0100321	dates of a series of vaccinations against a specific disease or a set of diseases	information (vaccine name, vaccination date) separated by semicolons	IMVAMUNE (Bavarian Nordic); 2022-06-01		100	100	100
1.000 vaccination information	vaccination natory	3EI4EI 10.0100321	Social displayed.	sumonono.	move (Davanan Nordic), 2022-00-01		1.0.0	1.0.0	1.0.0
			<u></u>						
	Host exposure information	GENEPIO:0001409	The country where the host was likely exposed to the councilies	Select the country name from the nick list provided in the			1.0.0	1.0.0	1.0.0
Host exposure information	location of exposure geo_loc name (country)	GENEPIO:0001410	agent of the illness.	Select the country name from the pick list provided in the template. Select the country name from the pick list provided in the	Canada		1.0.0	1.0.0	1.0.0
Host exposure information	location of exposure geo_loc name (country)	GENEPIO:0001410	agent of the illness.		Canada [GAZ:00002560]		1.0.0	1.0.0	1.0.0
Host exposure information	destination of most recent travel (city)	GENEPIO:0001411	The name of the city that was the destination of most recent travel.	Provide the name of the city that the host travelled to. Use this look-up service to identify the standardized term: https://www.ebi.ac.uk/ols/ontologies/gaz	New York City		1.0.0	1.0.0	1.0.0
Host exposure information	destination of most recent travel (state/province/territory)	GENEPIO:0001412	The name of the state/province/territory that was the destination of most recent travel.		,		1.0.0	1.0.0	1.0.0
riosi exposure information	desurration of most recent travel (state/province/territory)	GENEPIO:0001412	most recent travel.	Provide the name of the state/province/territory that the			1.0.0	1.0.0	1.0.0
Host exposure information	destination of most recent travel (state/province/territory)	GENEPIO:0001412	The name of the state/province/territory that was the destination of most recent travel.	host travelled to. Use this look-up service to identify the standardized term: https://www.ebi.ac.uk/ols/ontologies/gaz	California		1.0.0	1.0.0	1.0.0
Host exposure information	destination of most recent travel (country)	GENEPIO:0001413	The name of the country that was the destination of most recent travel.	Select the country name from the pick list provided in the template.	Canada		1.0.0	1.0.0	1.0.0
			The name of the country that was the destination of most recent	Select the country name from the pick list provided in the					

Part				The date of a person's most recent departure from their primary						
Part	Host exposure information	most recent travel departure date	GENEPIO:0001414	locations.	Provide the travel departure date.	2020-03-16	1	1.0.0	1.0.0	1.0.0
Marie Mari	lost exposure information	most recent travel return date	GENEPIO:0001415		Provide the travel return date.	2020-04-26	1	1.0.0	1.0.0	1.0.0
Marie September Marie Sept	Host exposure information	travel history	GENEPIO:0001416		known, separated by a comma) travelled in the last six months; can include multiple travels. Separate multiple	Canada, Vancouver; USA, Seattle; Italy, Milan	1	1.0.0	1.0.0	1.0.0
March Marc	lost exposure information		GENEPIO:0001417		Select an exposure event from the pick list provided in the template. If the desired term is missing, contact the		1	1.0.0	1.0.0	1.0.0
Section of the content of the cont										
Margane framework Marg	Host exposure information	exposure event		Event leading to exposure.	DataHarmonizer curation team.	Party [PCO:0000035]				1.0.0
Section of the Control of Contr	Host exposure information	exposure contact level	GENEPIO:0001418	The exposure transmission contact type.	Select exposure contact level from the pick-list.		1	1.0.0	1.0.0	1.0.0
Section Sect	Host exposure information	exposure contact level	GENEPIO:0001418	The exposure transmission contact type.	Select exposure contact level from the pick-list.		1	1.0.0	1.0.0	1.0.0
Marcian Care	Host exposure information	host role	GENEPIO:0001419	The role of the host in relation to the exposure setting.	in the template. If the desired term is missing, contact the	Acquaintance of case	1	1.0.0	1.0.0	1.0.0
Section Sect	Host exposure information	host role	GENEPIO:0001419	The role of the host in relation to the exposure setting	in the template. If the desired term is missing, contact the			100	100	100
Marche Desirable Marche Desi	lost exposure information	110311016	OLINEI 10.0001419	The fole of the host in relation to the exposure setting.		Acquaintance of case [OLIVEI 10.0100200]		1.0.0	1.0.0	1.0.0
Part	Host exposure information	exposure setting	GENEPIO:0001428	The setting leading to exposure.	provided in the template. If a desired term is missing, contact the DataHarmonizer curation team.	Healthcare Setting				
Material relation is a sequence of part of the content of the product of the content of part of the part of th			OFNEDIO 0004400		provided in the template. If a desired term is missing,	II III O III IOFNEDIO MANAGAN				
September information or an information or an information or an information of the control information or an information	·			,						
	Host exposure information	exposure details	GENEPIO:0001431	Additional host exposure information.	Free text description of the exposure.	Large party, many contacts	1	1.0.0	1.0.0	1.0.0
		Host reinfection information	GENEPIO:0001434					1.0.0	1.0.0	1.0.0
of melindan information prior afficial information afficial information prior afficial information prior afficial information prior afficial information affirmation affirmation affirm	Heat rainfaction information			The absence or presence of a prior Mnov infection	If known, provide information about whether the individual had a previous Mpox infection. Select a value from the pick	Dries infection		100	100	100
or inerfection information prior flages reflection information or inerfection information and professional design of the prior of the prior flages and the prior flages and the prior prior flages and the prior prior flages and the prior flag	riosi reillection illomation	phot wpox intection	GENEFIO.0100332	The absence of presence of a prior impox infection.	If known, provide information about whether the individual			1.0.0	1.0.0	1.0.0
Provide the same of the most record part infection was displayed by the most record part infection and participation was designed. Provide the part was provided in information Provide the part was preferred to provide the part was preferred to information Provide the part was preferred to the part of the part was preferred to the part of the part was preferred to the part of the p	Host reinfection information	prior Mooy infection	GENEPIO:0100532	The absence or presence of a prior Mnov infection	had a previous Mpox infection. Select a value from the pick	Prior infection [GENEPIO:0100037]		100	100	100
The allowance of presence of artifival the selected or information on the effective information of the					Provide the date that the most recent prior infection was diagnosed. Provide the prior Mpox infection date in ISO					
The absence or personal standards beatment (or a prior Moos and value from the manner of personal standard beatment for a prior Moos and value from the policy factor of standard beatment for a prior Moos and value from the policy factor of the policy factor of standard beatment (or any prior with the same of the project factor of the policy fa				The absence or presence of antiviral treatment for a prior Mpox	If known, provide information about whether the individual had a previous Mpox antiviral treatment. Select a value					
Provide a description of any attrivial treatment during pror Mpox Infection during the prior M				The absence or presence of antiviral treatment for a prior Mpox	If known, provide information about whether the individual had a previous Mpox antiviral treatment. Select a value					
Sequencing project name Sequencing by Sequencing project name Sequenced by Sequencing by Sequencing project name Sequenced by Sequencing by Sequencing project name Sequenced by Sequenced by Sequencing project name Sequenced by Sequenced by Sequenced project name	lost remiecado milomador	ргог трох аналия и сванен	GENET 10.0100334		Provide a description of any antiviral treatment administered for viral infections (not including Mpox			1.0.0	1.0.0	1.0.0
Sequencing Sequenced by intovation yname Sequenced by contact name Sequenced by contact name Sequenced by contact name Sequenced by contact demail Sequenced by contact demai	Host reinfection information	prior antiviral treatment during prior Mpox infection	GENEPIO:0100535	period.	is meant to capture concurrent treatment information.	the prior Mpox infection.				
Sequencing project name GENEPIO.0100472 The name of the project inflatitivelyorgram for which sequencing was performed. The name of the project and/or the project ID feer. If the information is unknown or cannot be provided, leave blank or provide and value. The name of the project and value. The name of the agency should be written out in full, (with submissions. If submitting specimens rather than sequencing data, please put the National Microbiotory Laboratory (NML). The name of the agency that generated the sequence. The name of the agency that generated the sequence. The name of the agency that generated the sequence. The name of the agency that generated the sequence. The name of the agency that generated the sequence. The name of the agency that generated the sequence. Provide the name of the project ID feer. If the information is unknown or cannot be provided, leave blank or provide a law write. Application of the project ID feer. The name of the agency that generated the sequence. Provide the name of the project ID feer. If the information is unknown or cannot be provided. The name of the agency that generated the sequence. Provide the name of the project ID feer. If the information is unknown or cannot be provided. If the information is unknown or cannot be provided. The name of the agency that generated the sequence. Provide the name of the project ID feer. If the information is unknown or cannot be provided. If the information is unknown or cannot be provided. If the information is unknown or cannot be provided. If the information is unknown or cannot be provided. If the name of the project ID feer. If the information is unknown or cannot be provided, leave the name of the project ID feer. If the information is unknown or cannot be provided, leave the name of the project ID feer. If the information is unknown or cannot be provided, leave the name of the project ID feer. If the information is unknown or cannot be provided, leave the name of the project ID feer. If the										
equence information sequence information sequence formation sequence formation sequence information sequence formation sequence for follow-up responsible for follow-u							1	1.0.0	1.0.0	1.0.0
equencing project name GENEPIO.0100472 The name of the agency should be written out in ful, (with more exceptions and analystus. The name of the agency should be written out in ful, (with more exceptions and analystus. The name of the agency should be written out in ful, (with more exceptions and the constant reasons multiple sequencing data, phases put the "National Microbiology Laboratory (MM."). The name of the agency should be written out in ful, (with more exceptions and the constant reasons multiple sequence in formation as equenced by GENEPIO.0100416 The name of the agency that generated the sequence. GENEPIO.0100416 The name of the agency that generated the sequence. GENEPIO.0100416 The name of the agency that generated the sequence. GENEPIO.0100416 The specific laboratory affiliation of the responsible for sequencing the information is unknown or cannot be provided, leave blank or provide a null value. Sequence information sequenced by laboratory name GENEPIO.0100470 GENEPIO.0100470 The specific laboratory affiliation of the responsible for follow-up regarding the isolater's genome. Provide the name of an individual or their job title for sequencing and individual or their job title for sequencing as equenced by contact name GENEPIO.0100472 The name of title agency submitting the sequence. The amail address of the contact responsible for follow-up regarding the sequence. The amail address of the contact responsible for follow-up regarding the sequence. The mailing address of the agency submitting the sequence to a non-institution and more and the sequence to a non-individual or their job title for sequencing and sequenced by contact and or sequenced		Sequencing	GENEPIO:0001441				1	1.0.0	1.0.0	1.0.0
Sequenced by Seque			OFNEDIO 0400470	The name of the project/initiative/program for which sequencing	If the information is unknown or cannot be provided, leave	MDOV 4050				
equencing sequenced by GENEPIO.0100416 The name of the agency that generated the sequence. Public Health Ordario (PHO) 1.00	Sequence information	sequencing project name	GENEPIO:0100472	was performed.	The name of the agency should be written out in full, (with minor exceptions) and be consistent across multiple submissions. If submitting specimens rather than	MPUX-1356	1	1.0.0	1.0.0	1.0.0
equencing sequenced by Sexual Department of the sequence of th	Sequencing	sequenced by	GENEPIO:0100416	The name of the agency that generated the sequence.	sequencing data, please put the "National Microbiology Laboratory (NML)".	Public Health Ontario (PHO)	1	1.0.0	1.0.0	1.0.0
sequenced by Sequenced by GENEPIO:0100416 The name of the agency that generated the sequence. Provide the name of the specific laboratory that that performed the sequencing in full (avoid abbreviations). If the information is unknown or cannot be provided, leave blank or provide a null value. Provide the name of the sequencing in full (avoid abbreviations). If the information is unknown or cannot be provided, leave blank or provide a null value. Provide the name of an individual or their job title. As personnel turnover may render the contact's name obsoiles, it is more prefereable to provide a job title for neuring accuracy of information and institutional memory. If the information is unknown or cannot be provided, leave blank or provide a job title for neuring accuracy of information memory. If the information is unknown or cannot be provided, leave blank or provide a job title for neuring accuracy of information memory. If the information is unknown or cannot be provided, leave blank or provide a null value. The name or title of the contact responsible for follow-up regarding the sequence. The entail address of the contact responsible for follow-up regarding the sequence and an institutional or laboration is unknown or cannot be provided, leave blank or provide a job title for neuring accuracy of information is unknown or cannot be provided. It is information is unknown or cannot be provided, leave blank or provide a job title for neuring accuracy of information is unknown or cannot be provided, leave blank or provide a job title for neuring accuracy of information memory. If the information is unknown or cannot be provided, leave blank or provide a job title for neuring accuracy of information is unknown or cannot be provided, leave blank or provided and value. Enterior Lab Manager E					minor exceptions) and be consistent across multiple					
requence information sequenced by laboratory name GENEPIO:0100470 the isolate's genome. The specific laboratory affiliation of the responsible for sequencial the information is unknown or cannot be provided, leave blank or provide a null value. Provide the name of an individual or their job title. As personnel furnover may render the contact rame chaosite, it, in more preferable to provide a job title for ensuring accuracy of information and inattitutional memory. The name or title of the contact responsible for follow-up regarding the sequence. The mail address of the contact responsible for follow-up regarding the sequence of the sequence and an among title. Provide a null value. The name or title of the contact responsible for follow-up regarding the sequence and an among title. Provide a null value. The name or title of the contact responsible for follow-up regarding the sequence and an among title of the contact responsible for follow-up regarding the sequence. The mailing address of the agency submitting the sequence. The mailing address of the agency should be written out in full, (with a sequence and blank or provided a null value. The name of the agency that submitted the sequence to a sequence and an among title, appears on the water of the agency should be written out in full, (with a sequence in survivoled. In a sequence in a seque	Sequencing	sequenced by	GENEPIO:0100416	The name of the agency that generated the sequence.	submissions.	Public Health Ontario (PHO)	1	1.0.0	1.0.0	1.0.0
Provide the name of an individual or helr job title. As personnel turnover may render the contacts mame obsolete, it is more prefereable to provide a job title for enumber of the information and institutional memory. In the name or title of the contact responsible for follow-up regarding if the information is unknown or cannot be provided, leave blank or provide a null value. The name or title of the contact responsible for follow-up regarding the sequence. The enall address of the contact responsible for follow-up regarding the sequence. The enall address of the contact responsible for follow-up regarding the sequence. The enall address of the contact responsible for follow-up regarding the sequence. The mailing address of the agency submitting the sequence. The mailing address of the agency submitting the sequence. The mailing address of the agency should be written out in full, (with minor exceptions) and be consistent across multiplie submissions. If submitting specimens rather than sequencing sequencing sequence, that submitted the sequence to a sequencing sequence will be approximated to the sequence of the agency that submitted the sequence to a sequencing sequence will be approximated the sequence of the agency that submitted the sequence to a sequencing data, please put the "National Microbiology" and sequence and applead to the "National Microbiology" and sequence and and aname, City, Province Territory, Postal Code, Country The name of the agency that submitted the sequence to a sequence and aname, City, Province Territory, Postal Code, Country The mail address of the contact responsible for follow-up the s	Saguenes information	agguenced by laboratory name	CENEDIO 0100470		performed the sequencing in full (avoid abbreviations). If the information is unknown or cannot be provided, leave	Tonn I oh		100	100	100
equence information sequenced by contact name GENEPIO:0100471 The name or title of the contact responsible for follow-up regarding in the sequence uniformation and institutional memory. The mame of the contact responsible for follow-up regarding in the sequence uniformation is unknown or cannot be provided; leave blank or provide a null value. The email address of the contact responsible for follow-up regarding the sequence. The email address of the contact responsible for follow-up regarding the sequence of the contact responsible for follow-up regarding the sequ	ocquence information	ocquenceu by laboratory name	GENEPIO:01004/0	ито томале в уетпите.		IOPP EdD		1.0.0	1.0.0	1.0.0
equencing sequenced by contact name GENEPIO:0100471 to sequence. In the sequence. In the sequence of the seque					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.					
equencing sequenced by contact email GENEPIO:0100422 regarding the sequence. e.g. johnnyblogs@lab.ca, or RespLab@lab.ca RespLab.ca	Sequence information	sequenced by contact name	GENEPIO:0100471	the sequence.	blank or provide a null value.	Enterics Lab Manager		1.0.0	1.0.0	1.0.0
equencing sequenced by contact address GENEPIO:0100423 The mailing address of the agency submitting the sequence. and name, City, Province/Territory, Postal Code, Country 16, Canada	Sequencing	sequenced by contact email	GENEPIO:0100422	The email address of the contact responsible for follow-up regarding the sequence.			1	1.0.0	1.0.0	1.0.0
The name of the agency should be written out in full, (with minor exceptions) and be consistent across multiple submissions. If submitting specimens rather than the sequence to a sequence of the agency that submitted the sequence to a sequence of the agency that submitted the sequence to a sequence of the agency that submitted the sequence to a sequence of the agency should be written out in full, (with minor submitted the sequence of the agency should be written out in full, (with minor submitted the sequence of the agency should be written out in full, (with minor submitted the sequence of the agency should be written out in full, (with minor submitted the sequence of the agency should be written out in full, (with minor submitted the sequence of the agency should be written out in full, (with minor submitted the sequence of the agency should be written out in full, (with minor submitted the sequence of the agency should be written out in full, (with minor submitted the sequence of the agency should be written out in full, (with minor submitted the sequence of the agency should be written out in full, (with minor submitted the sequence of the agency should be written out in full, (with minor submitted the sequence of the agency should be written out in full, (with minor submitted the sequence of the agency should be written out in full, (with minor submitted the sequence of the agency should be written out in full, (with minor submitted the sequence of the agency should be written out in full, (with minor submitted the sequence of the agency should be written out in full, (with minor submitted the sequence of the agency should be written out in full, (with minor submitted the sequence of the agency should be written out in full, (with minor submitted the sequence of the agency should be written out in full, (with minor submitted the sequence of the agency should be written out in full, (with minor submitted the sequence of the agency should be written out in full, (with minor submitted the sequence out in ful	Sequencing	sequenced by contact address	GENEPIO:0100423	The mailing address of the agency submitting the sequence.			1	1.0.0	1.0.0	1.0.0
equencing sequence submitted by GENEPIO.0001159 databases. GENEPIO.0001159 database. GENEPIO.0001159	· •				The name of the agency should be written out in full, (with minor exceptions) and be consistent across multiple submissions. If submitting specimens rather than					
	Sequencing	sequence submitted by	GENEPIO:0001159	database.	Laboratory (NML)".	Public Health Ontario (PHO)	1	1.0.0	1.0.0	1.0.0

				The name of the agency should be written out in full, (with minor exceptions) and be consistent across multiple submissions. If submitting specimens rather than				
Sequencing	sequence submitted by	GENEPIO:0001159	The name of the agency that submitted the sequence to a database.	sequencing data, please put the "National Microbiology Laboratory (NML)".	Public Health Ontario (PHO)	1.0.0	1.0.0	1.0.0
Sequencing	sequence submitter contact email	GENEPIO:0001165	The email address of the agency responsible for submission of the sequence.	The email address can represent a specific individual or lab e.g. johnnyblogs@lab.ca, or RespLab@lab.ca	RespLab@lab.ca	1.0.0	1.0.0	1.0.0
	·		The mailing address of the agency responsible for submission of	The mailing address should be in the format: Street number	123 Sunnybrooke St, Toronto, Ontario, M4P			
Sequencing	sequence submitter contact address	GENEPIO:0001167	the sequence.	and name, City, Province/Territory, Postal Code, Country	11.6, Canada Select "Targeted surveillance (non-random sampling)" if the specimen fits any of the following criteria: Specimens attributed to individuals with no known intimate contacts to positive cases. Specimens attributed to youth/minors <18 yrs. Specimens attributed to volunterable persons living in transient shelters or congregant settings. Specimens attributed to individuals self-identifying as "female"; For	1.0.0	1.0.0	1.0.0
Sequencing	purpose of sequencing	GENEPIO:0001445	The reason that the sample was sequenced.	The reason a sample was sequenced may provide information about potential biases in sequencing strategy.	specimens with a recent international and/or domestic travel history, please select the most appropriate lag from the following three options: Domestic travel surveillance. International travel surveillance. Travel-associated surveillance. For specimens targeted for sequencing as part of an outbreak investigation, please select: Culster/Outbreak investigation, please select: Culster/Outbreak investigation, please select causer (outbreak investigation, please select causer (outbreak investigation, please select) and other cases use: Baseline surveillance (random sampling).	1.0.0	1.0.0	1.0.0
Sequencing	purpose of sequencing		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 as 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.	Baseline surveillance (random sampling) [GENEPIO:0100005]	1.0.0	1.0.0	1.0.0
Sequencing	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: Screened due to travel history, Screened due to close contact with infected individual.	Outbreak in MSM community	1.0.0	1.0.0	1.0.0
Sequencing	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
Sequencing	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
Sequencing	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
Sequencing	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 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	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 [OBI:0002117]	1.0.0	1.0.0	1.0.0
Sequencing	sequencing instrument	GENEPIO:0001452	The model of the sequencing instrument used.	Select a sequencing instrument from the picklist provided in the template.	Oxford Nanopore MinION	1.0.0	1.0.0	1.0.0
Sequencing	sequencing instrument	GENEPIO:0001452	The model of the sequencing instrument used.	Select a sequencing instrument from the picklist provided in the template.	Oxford Nanopore MinION [GENEPIO:0100142]	1.0.0	1.0.0	1.0.0
Sequence information	sequencing flow cell version	GENEPIO:0101102	The version number of the flow cell used for generating sequence data.	Flow cells can vary in terms of design, chemistry, capacity, etc. The version of the flow cell used to generate sequence data can affect sequence quantity and quality. Record the version of the flow cell used to generate sequence data. Do not include "version" or "v" in the version number.		1.0.0	1.0.0	1.0.0
				approach. Sequencing was performed using a <fill in=""> sequencing instrument. Libraries were prepared using <fill< td=""><td>Viral sequencing was performed following a metagenomic shotgun sequencing approach. Libraries were created using Illumina DNA Prepkits, and sequence data was produced using</td><td></td><td></td><td></td></fill<></fill>	Viral sequencing was performed following a metagenomic shotgun sequencing approach. Libraries were created using Illumina DNA Prepkits, and sequence data was produced using			
Sequencing	sequencing protocol		The protocol used to generate the sequence.	in> library kit. "	Miseq Micro v2 (500 cycles) sequencing kits.	1.0.0	1.0.0	1.0.0
Sequencing	sequencing kit number		The manufacturer's kit number. The length of the DNA fragment generated by mechanical shearing or enzymatic digestion for the purposes of library	Alphanumeric value. Provide the fragment length in base pairs (do not include	AB456XYZ789	1.0.0	1.0.0	1.0.0
Sequencing	DNA fragment length		preparation. The molecular technique used to selectively capture and amplify	the units).	400	1.0.0	1.0.0	1.0.0
Sequencing	genomic target enrichment method	GENEPIO:0100966	specific regions of interest from a genome. The molecular technique used to selectively capture and amplify	Provide the name of the enrichment method	hybrid selection method	3.0.0	3.0.0	3.0.0
Sequencing	genomic target enrichment method	GENEPIO:0100966	specific regions of interest from a genome. Details that provide additional context to the molecular technique	Provide the name of the enrichment method	hybrid selection method enrichment was done using Illumina Target	3.0.0	3.0.0	3.0.0
Sequencing	genomic target enrichment method details	GENEPIO:0100967	used to selectively capture and amplify specific regions of interest from a genome.	Provide details that are applicable to the method you used.	Enrichment methodology with the Illumina DNA Prep with enrichment kit.	3.0.0	3.0.0	3.0.0
	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.	MPXV Sunrise 3.1	1.0.0	1.0.0	1.0.0
Sequencing Sequencing	amplicon size		The length of the amplicon generated by PCR amplification.	Provide the amplicon size expressed in base pairs.	300bp	1.0.0	1.0.0	1.0.0

			The name of the method used to assess whether a sequence	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				
Bioinformatics and QC metrics	quality control method name	GENEPIO:0100557	passed a predetermined quality control threshold.	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
Diominimatos ana go montos	quanty control motion voicion	SEITE 10.0100000	soquenes pussed a productorrinina quanty control uncontour.	Select a value from the pick list provided. If a desired value	1.2.0	1.0.0	1.0.0	1.0.0
				is missing, submit a new term request to the PHA4GE QC				
Bioinformatics and QC metrics	quality control determination		The determination of a quality control assessment.	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 determination	GENEPIO:0100559				1.0.0	1.0.0	1.0.0
				Select a value from the pick list provided. If a desired value				
Bioinformatics and QC metrics	quality control issues	GENEPIO:0100560	The reason contributing to, or causing, a low quality determination in a quality control assessment.	is missing, submit a new term request to the PHA4GE QC Tag GitHub issuetracker using the New Term Request form.	low average genome coverage	100	100	100
Bioinformatics and QC metrics	quality control issues	GENEPIO:0100560	in a quality control according to	ray our ab reconstruct doing the frew ferm request form.	iow avorago gonomo covorago	100	1.0.0	1.0.0
Biomornidaes and Go mories	quanty control todato	GENE: 10.010000	The details surrounding a low quality determination in a quality	Provide notes or details regarding QC results using free	CT value of 39. Low viral load. Low DNA	1.0.0	1.0.0	1.0.0
Bioinformatics and QC metrics	quality control details	GENEPIO:0100561	control assessment.	text.	concentration after amplification.	1.0.0	1.0.0	1.0.0
			The names of the software and version number used for raw data		·			
			processing such as removing barcodes, adapter trimming, filtering					
Bioinformatics and QC metrics	raw sequence data processing method	GENEPIO:0001458	etc.	Trimmomatic v. 0.38, Porechop v. 0.2.3	Porechop 0.2.3	1.0.0	1.0.0	1.0.0
			The names of the software and version number used for raw data	Describe the settlement followed by the consists of				
Bioinformatics and QC metrics	raw sequence data processing method	GENEPIO:0001458	processing such as removing barcodes, adapter trimming, filtering etc.	Trimmomatic v. 0.38, Porechop v. 0.2.3	Porechop 0.2.3	1.0.0	1.0.0	1.0.0
Biomornidaes and Go mones	Taw sequence data processing measur	GENE: 10:0001400	The method used to remove host reads from the pathogen	Provide the name and version number of the software used	1 diddiop c.E.c	1.0.0	1.0.0	1.0.0
Bioinformatics and QC metrics	dehosting method	GENEPIO:0001459	sequence.	to remove host reads.	Nanostripper	1.0.0	1.0.0	1.0.0
			The method used to remove host reads from the pathogen	Provide the name and version number of the software used				
Bioinformatics and QC metrics	dehosting method	GENEPIO:0001459	sequence.	to remove host reads.	Nanostripper	1.0.0	1.0.0	1.0.0
				Provide the deduplication software name followed by the				
Bioinformatics and QC metrics	deduplication method	GENEPIO:0100831	dataset.	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	consensus sequence name	GENEPIO:0001460	The name of the consensus sequence.	Provide the name and version number of the consensus sequence.	mpxvassembly3	1.0.0	1.0.0	1.0.0
Bioinformatics and QC metrics	consensus sequence name	GENEPIO:0001460	The name of the consensus sequence.	Provide the name and version number, with the file	mpxvassembly3	1.0.0	1.0.0	1.0.0
				extension, of the processed genome sequence file e.g. a				
				consensus sequence FASTA file or a genome assembly				
Bioinformatics and QC metrics	genome sequence file name	GENEPIO:0101715	The name of the consensus sequence file.	file.	mpxvassembly.fasta	3.0.0	3.0.0	3.0.0
					/User/Documents/ViralLab/Data/mpxvassembly.			
Bioinformatics and QC metrics	genome sequence file path	GENEPIO:0101716	The filepath of the consensus sequence file.		fasta	3.0.0	3.0.0	3.0.0
Bioinformatics and QC metrics	consensus sequence software name	GENEPIO:0001463	The name of software used to generate the consensus sequence.	Provide the name of the software used to generate the	iVar	100	100	1.0.0
Diditionnates and QC metrics	consensus sequence software frame	GENET 10.0001403		Provide the version of the software used to generate the	Ivai	1.0.0	1.0.0	1.0.0
Bioinformatics and QC metrics	consensus sequence software version	GENEPIO:0001469	sequence.	consensus sequence.	1.3	1.0.0	1.0.0	1.0.0
				Provide the name of the software used to assemble the	SPAdes Genome Assembler, Canu, wtdbg2,			
Bioinformatics and QC metrics	sequence assembly software name	GENEPIO:0100825	The name of the software used to assemble a sequence.	sequence.	velvet	1.0.0	1.0.0	1.0.0
				Provide the version of the software used to assemble the				
Bioinformatics and QC metrics	sequence assembly software version	GENEPIO:0100826	The version of the software used to assemble a sequence.		3.15.5	1.0.0	1.0.0	1.0.0
Bioinformatics and QC metrics	r1 fasto filename	GENEPIO:0001476	The user-specified filename of the r1 FASTQ file.	Provide the r1 FASTQ filename. This information aids in	ABC123 S1 L001 R1 001.fastq.gz	1.0.0	1.0.0	1.0.0
Bioinformatics and QC metrics	r i lastq illename	GENEPIO:0001476	The user-specified filename of the FT FASTQ file.	data management. Provide the r2 FASTQ filename. This information aids in	ABC123_S1_L001_R1_001.lastq.g2	1.0.0	1.0.0	1.0.0
Bioinformatics and QC metrics	r2 fasto filename	GENEPIO:0001477	The user-specified filename of the r2 FASTQ file.	data management.	ABC123 S1 L001 R2 001.fastq.gz	1.0.0	1.0.0	1.0.0
				9	/User/Documents/ViralLab/Data/ABC123_S1_L			
Bioinformatics and QC metrics	r1 fastq filepath	GENEPIO:0001478	The location of the r1 FASTQ file within a user's file system.	aids in data management.	001_R1_001.fastq.gz	1.0.0	1.0.0	1.0.0
					/User/Documents/ViralLab/Data/ABC123_S1_L			
Bioinformatics and QC metrics	r2 fastq filepath	GENEPIO:0001479	The location of the r2 FASTQ file within a user's file system.	aids in data management.	001_R2_001.fastq.gz	1.0.0	1.0.0	1.0.0
Bioinformatics and QC metrics	fast5 filename	GENEPIO:0001480	The user-specified filename of the FAST5 file.	Provide the FAST5 filename. This information aids in data management.	mpxv123seq.fast5	1.0.0	1.0.0	1.0.0
Bioinformatics and QC metrics	Tasto filename	GENEPIO:0001480	The user-specified filename of the FAST5 file.			1.0.0	1.0.0	1.0.0
Bioinformatics and QC metrics	fast5 filepath	GENEPIO:0001481	The location of the FAST5 file within a user's file system.	Provide the filepath for the FAST5 file. This information aids in data management.	/User/Documents/RespLab/Data/mpxv123seq.f ast5	1.0.0	1.0.0	1.0.0
Biomornidaes and Go mones	idoto inopatri	GENE: 10:0001401	The total number of non-unique reads generated by the	in data management.	doto	1.0.0	1.0.0	1.0.0
Bioinformatics and QC metrics	number of total reads	GENEPIO:0100827	sequencing process.	Provide a numerical value (no need to include units).	423867	1.0.0	1.0.0	1.0.0
			The number of unique reads generated by the sequencing					
Bioinformatics and QC metrics	number of unique reads	GENEPIO:0100828	process.	Provide a numerical value (no need to include units).	248236	1.0.0	1.0.0	1.0.0
			The threshold used as a cut-off for the minimum length of a read					
Bioinformatics and QC metrics	minimum post-trimming read length	GENEPIO:0100829	after trimming.	Provide a numerical value (no need to include units).	150	1.0.0	1.0.0	1.0.0
		OFNEDIO 0004474	The average number of reads representing a given nucleotide in		400x	100	100	
Bioinformatics and QC metrics	depth of coverage value	GENEPIO:0001474	the reconstructed sequence.	Provide value as a fold of coverage.	400x	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.	100x	100	100	1.0.0
District Hadisə and Go monico	departer coverage arresticia	GENE: 10:0001410	The number of total base pairs generated by the sequencing	Trovido dio directicia foia coverage.	100%	1.0.0	1.0.0	1.0.0
Bioinformatics and QC metrics	number of base pairs sequenced	GENEPIO:0001482	process.	Provide a numerical value (no need to include units).	2639019	1.0.0	1.0.0	1.0.0
			Size of the reconstructed genome described as the number of					
Bioinformatics and QC metrics	consensus genome length	GENEPIO:0001483	base pairs.	Provide a numerical value (no need to include units).	197063	1.0.0	1.0.0	1.0.0
			The length of the genome generated by assembling reads using a					
Bioinformatics and QC metrics	sequence assembly length	GENEPIO:0100846	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	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
Diomiormatics and QC Metrics	number of cortage	GENEPIO:0100937	The percentage of expected genes identified in the genome being	i Tovido d Hullichodi Value.	10	1.0.0	1.0.0	1.0.0
			sequenced. Missing genes indicate missing genomic regions	Provide the genome completeness as a percent (no need				
Bioinformatics and QC metrics	genome completeness	GENEPIO:0100844	(incompleteness) in the data.	to include units).	85	1.0.0	1.0.0	1.0.0
			The length of the shortest read that, together with other reads,					
Bioinformatics and QC metrics	N50	GENEPIO:0100938	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

Contributor acknowledgement	DataHarmonizer provenance	GENEPIO:0001518	The DataHarmonizer software and template version provenance.	The current software and template version information will be	DataHarmonizer v1.4.3, Mpox v3.3.1				
Contributor acknowledgement	authors	GENEPIO:0001517	Names of individuals contributing to the processes of sample colle-	Include the first and last names of all individuals that should	Tejinder Singh, Fei Hu, Joe Blogs				
	Contributor acknowledgement	GENEPIO:0001516							
athogen diagnostic testing	diagnostic pcr Ct value 5	GENEPIO:0100579	The Ct value result from a diagnostic SARS-CoV-2 RT-PCR test.	diagnostic RT-PCR test.	30	J	1.0.0	1.0.0	1.0.0
Pathogen diagnostic testing	gene name 5	GENEPIO:0100578	The name of the gene used in the diagnostic RT-PCR test.	from the standardized pick list. Provide the CT value of the sample from the second	RNAse P		400	400	400
athogen diagnostic testing	diagnostic pcr Ct value 4	GENEPIO:0100577	·	diagnostic RT-PCR test. Select the name of the gene used for the diagnostic PCR	21	7	1.0.0	1.0.0	1.0.0
athogen diagnostic testing	gene name 4	GENEPIO:0100576	The name of the gene used in the diagnostic RT-PCR test.	from the standardized pick list. Provide the CT value of the sample from the second	G2R_G (TNFR)		1.0.0	1.0.0	1.0.0
athogen diagnostic testing	diagnostic pcr Ct value 3	GENEPIO:0001515	·	diagnostic RT-PCR test. Select the name of the gene used for the diagnostic PCR	19		1.0.0	1.0.0	1.0.0
athogen diagnostic testing	gene name 3	GENEPIO:0001513	The name of the gene used in the diagnostic RT-PCR test.	https://bit.ly/2Sq1LbI Provide the CT value of the sample from the second	OPHA (orf B2R)		1.0.0	1.0.0	1.0.0
				Provide the full name of another gene used in an RT-PCR 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:					
Pathogen diagnostic testing	gene name 3	GENEPIO:0001513	The name of the gene used in the diagnostic RT-PCR test.	Select the name of the gene used for the diagnostic PCR from the standardized pick list.	OPHA (orf B2R)		1.0.0	1.0.0	1.0.0
Pathogen diagnostic testing	diagnostic per Ct value 2	GENEPIO:0001512	The Ct value result from a diagnostic SARS-CoV-2 RT-PCR test.	Provide the CT value of the sample from the second diagnostic RT-PCR test.	36	3	1.0.0	1.0.0	1.0.0
athogen diagnostic testing	gene name 2	GENEPIO:0001510	The name of the gene used in the diagnostic RT-PCR test.	Provide the full name of another gene used in an RT-PCR 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/2Sq1Lbl	OVP (orf 17L)		1.0.0	1.0.0	1.0.0
athogen diagnostic testing	gene name 2	GENEPIO:0001510	The name of the gene used in the diagnostic RT-PCR test.	Select the name of the gene used for the diagnostic PCR from the standardized pick list.	OVP (orf 17L)		1.0.0	1.0.0	1.0.0
athogen diagnostic testing	diagnostic per Ct value 1	GENEPIO:0001509	The Ct value result from a diagnostic SARS-CoV-2 RT-PCR test.	Provide the CT value of the sample from the diagnostic RT-PCR test.	2:	1	1.0.0	1.0.0	1.0.0
athogen diagnostic testing	gene name 1	GENEPIO:0001507	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://blt.p/2Sq1Lbl	MPX (orf B6R)		1.0.0	1.0.0	1.0.0
athogen diagnostic testing	gene name 1	GENEPIO:0001507	The name of the gene used in the diagnostic RT-PCR test.	Select the name of the gene used for the diagnostic PCR from the standardized pick list.	MPX (orf B6R)		1.0.0	1.0.0	1.0.0
athogen diagnostic testing	Pathogen diagnostic testing assay target name	GENEPIO:0001506 GENEPIO:0101206					1.0.0	1.0.0	1.0.0
							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/monkeypox-nf		3.0.0	3.0.0	3.0.0
ioinformatics 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/monkeypox-nf				
ioinformatics 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_063383.1				
ioinformatics 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	2	1.0.0	1.0.0	1.0.0
ioinformatics 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