

# MIRRI Information System (MIRRI-IS)

## Guidelines for uploading of catalogs of viruses

MIRRI ICT Task Force

(Release candidate 2, version 2022.01.26)

### Index of contents

Introduction .....	3
Identity.....	4
MIRRI Accession Number (MANDATORY) .....	4
Accession Number (MANDATORY) .....	4
Link to other sites .....	5
Other denomination .....	5
Distribution .....	6
Form of supply (MANDATORY) .....	6
Restrictions on use (MANDATORY) .....	6
Nagoya protocol restrictions and compliance conditions (MANDATORY).....	6
Strain from a Registered Collection.....	7
Risk group (MANDATORY) .....	7
Quarantine in Europe (MANDATORY) .....	7
Storage conditions .....	8
Recommended medium for growth for host cells.....	8
Propagation host.....	8
Taxonomy.....	10
Organism type (MANDATORY) .....	10
Species name (MANDATORY).....	10
Species abbreviation (MANDATORY) .....	10
Former species name.....	11
Pathotype-Serotype-Type .....	11

Reference isolate (MANDATORY).....	11
Comment on taxonomy .....	12
Identification technique .....	12
GMO (MANDATORY).....	12
GMO information .....	13
Mutant information .....	13
Gene sequences and accession numbers .....	13
Origin .....	15
Isolation host.....	15
Geographic origin.....	16
Country (MANDATORY) .....	16
History of deposit.....	18
Isolate Literature (MANDATORY) .....	18
Depositor .....	18
Date of deposit .....	18
Collected by .....	19
Date of collection (MANDATORY) .....	19
Isolated by (MANDATORY) .....	19
Date of isolation .....	20
Date of inclusion in the catalogue.....	20
Properties .....	21
Pathogenicity.....	21
Transmission by.....	21
Infectivity.....	21
Contamination.....	22
Lytic cycle .....	22
Genome integration .....	22
Remarks.....	22

## Introduction

This document specifies guidelines for uploading data of catalogs of viruses from culture collections of MIRRI partners to the Information System of MIRRI (MIRRI-IS). They were requested because the information that is needed for the description of viruses differs significantly from the information required for microbial strains (archaea, bacteria, yeasts, filamentous fungi, microalgae), which was already described in the 'Guidelines for catalog uploading' delivered by the ICT Task Force of MIRRI.

### Mandatory and recommended fields

Some data fields have been defined as mandatory. This means that in order to submit information on a given virus they must be specified: **virus descriptions missing mandatory data will not be included in the MIRRI-IS**. All non-mandatory data fields should also be provided, but if they miss, for any reason, the description of the virus will be included in the MIRRI-IS.

### Grouping of information

In this document, data fields were grouped for clarity sake on the basis of their relatedness to the following categories: **Identity, Distribution, Taxonomy, Origin, History of deposit, Properties**. Such grouping can support a better understanding of the required information, but it does not have any immediate consequence.

### Correspondences among MIRRI-IS Guidelines

Although these guidelines include data fields which are specific for viruses, there also are data fields which are the same or similar to fields described for microbial strains. In order to highlight the existing correspondences among these and previous MIRRI-IS Guidelines, they are annotated for each data field. These notes are included in the new feature "MIRRI-IS DATASET (version 2020.06.1)" where the corresponding field in the Guidelines for microbial strains is reported. If the field is not exactly the same, a short note on differences is also reported. New fields, i.e. fields which are specific for viruses, are annotated with the term "N/A".

### Field names, acronyms, short names

Short names were assigned to data fields along with their complete names. This is useful for the best management of data, e.g. in data exchange between CCs and MIRRI-IS. **Short names must be used by the CCs all times when referring to data fields**, e.g. as **headers in MS Excel files** when submitting to MIRRI-IS.

## Identity

This group includes all data fields which are needed in order to identify uniquely the isolate. This includes identifiers and names.

<b>MIRRI Accession Number (MANDATORY ON UPDATE)</b>	
Short name	mirriAccessionNumber
Description	<p>Unique identifier of the strain in the MIRRI-IS.</p> <p>It will be created on the first submission of a strain in the MIRRI-IS in a one-to-one connection with the Accession number of the strain in the CC. It is meant as a reference within the MIRRI-IS and as a unique reference for interoperability with other Life Science tools. It will include a version extension and be used as a reference for provenance issues as well.</p> <p>This information will be returned to CCs in association with the relative strain accession number. CCs must include it in their catalogue and return it to MIRRI-IS at every update.</p>
Syntax	The MIRRI Accession number will be composed by the 'MIRRI' prefix followed by 7 digits.
Values	On first submission, this field should not be compiled. At following submissions/updates, the accession number returned by MIRRI-IS should be included.
Validation	When missing, check whether the strain accession number was already included. If not, a new value will be created and returned to the collection. If the strain accession number was already submitted in the past, MIRRI-IS will retrieve the related values and assign them to the strain again.
Examples	MIRRI1234567
MIRRI-IS DATASET (version 2020.06.1)	Same as "MIRRI Accession number".

<b>Accession Number (MANDATORY)</b>	
Short name	accessionNumber
Description	Unique identifier of the strain in the CC.
Syntax	<p>CC acronym followed by a space character and a number or code. If a code is used, it cannot include spaces.</p> <p>If the current accession number is not compliant with this syntax, it must be redefined. In this case, the previous accession number must be included in the "Other denomination" field.</p>
Values	Free text, according to defined syntax.
Validation	Check that the correct syntax is used.
Examples	PLAVIT X101
MIRRI-IS DATASET (version 2020.06.1)	Same as "Accession number", but for the rule on renaming. Here, the previous accession number must be included in the "Other denomination" data field instead of the "Other culture collection numbers" field because the latter is not included in the dataset for viruses.

<b>Link to other sites</b>	
Short name	siteLinks
Description	URLs of extended descriptions of the virus isolate in other information systems. May be used in particular for connecting to the EVA-GLOBAL portal.
Syntax	Name and working URL (including method, server and full path), joined by a colon character ":". For EVA-GLOBAL, use exactly "EVA-GLOBAL" as name. For other sites, the name used must be consistent in the CC. If providing more than one site, separate them by a semicolon character ";".
Values	One or more pairs name - valid URL.
Validation	Check that the URLs are working. Report errors to the CC.
Examples	EVA-GLOBAL: <a href="https://www.european-virus-archive.com/virus/ourmia-melon-virus">https://www.european-virus-archive.com/virus/ourmia-melon-virus</a> ; Site2: <a href="https://example2.com/path/filename">https://example2.com/path/filename</a>
MIRRI-IS DATASET (version 2020.06.1)	N/A

<b>Other denomination</b>	
Short name	otherDenomination
Description	Unofficial names that have been or still are used for the strain, e.g. in publications, or a name given to the strain by the isolator before its deposit at the collection. May also include reference numbers/codes for those isolates whose accession number have been redefined according to the guidelines for the "Accession number".
Syntax	None
Values	Free text
Validation	None
Examples	TYLCV-Sar
MIRRI-IS DATASET (version 2020.06.1)	Same as "Other denomination", but for the rule on redefined accession numbers.

## Distribution

This group includes all data fields which refer to the distribution of the isolate, including both legal, organizational and biological issues.

<b>Form of supply (MANDATORY)</b>	
Short name	supplyForms
Description	The forms of supply of the strain to users.
Syntax	One or several of the allowed values, separated by a semicolon “;”.
Values	Allowed values: Agar, Cryo, Dry Ice, In vivo, Liquid Culture Medium, Lyo, Oil, Water.
Validation	Check for the validity of the format. Report errors to the CC.
Examples	In vivo
MIRRI-IS DATASET (version 2020.06.1)	Same as “Form of supply”, but for the added value ‘In vivo’.

<b>Restrictions on use (MANDATORY)</b>	
Short name	useRestrictions
Description	Report if the strain can be used for commercial development or not.
Syntax	One of the allowed values.
Values	One of the following values: 1 (no restrictions apply), 2 (for research use only), 3 (for commercial development a special agreement is requested).
Validation	Check that one and only one of the allowed values is used. Report errors to the CC.
Examples	1
MIRRI-IS DATASET (version 2020.06.1)	Same as “Restrictions on use”.

<b>Nagoya protocol restrictions and compliance conditions (MANDATORY)</b>	
Short name	nagoyaConditions
Description	Situation of the strain in relation to the Nagoya protocol.
Syntax	One of the allowed values.
Values	One of the following: 1 (“No known restrictions under the Nagoya protocol”), 2 (“Documents providing proof of legal access and terms of use available at the collection”), 3 (“Strain probably in scope, please contact the Culture Collection”).
Validation	Check that one and only one of the allowed values is used. Report errors to the CC.
Examples	1

MIRRI-IS DATASET (version 2020.06.1)	Same as “Nagoya protocol restrictions and compliance conditions”
---	--

<b>Strain from a Registered Collection</b>	
Short name	registeredCollection
Description	Strain included in the registered CC according to the <a href="#">EU Regulation 511/2014</a> . Unregistered CCc can omit this information.
Syntax	Boolean field.
Values	One of the following values: “No”, “Yes”
Validation	Check that one and only one of the allowed values is used. Report errors to the CC.
Examples	No
MIRRI-IS DATASET (version 2020.06.1)	Same as “Strain from a Registered Collection”

<b>Risk group (MANDATORY)</b>	
Short name	riskGroup
Description	Risk group according to <a href="#">EU Directive 2000/54/EC</a> and its amendments and corrections.
Syntax	One of the allowed values.
Values	Allowed values: 1, 2, 3, 4.
Validation	Check that one and only one of the allowed values is used. Report errors to the CC.
Examples	3
MIRRI-IS DATASET (version 2020.06.1)	Same as “Risk group”.

<b>Quarantine in Europe (MANDATORY)</b>	
Short name	euQuarantine
Description	Specify whether the strain is subject to quarantine according to European Directive 2000/29/CE and its amendments and corrections. The list of quarantine organisms is available in the Commission Implementing Regulation (EU) 2019/2072.
Syntax	One of the allowed values.
Values	One of the following values: “No”, “Yes”
Validation	Check that one and only one of the allowed values is used. Report errors to the CC.
Examples	Yes
MIRRI-IS DATASET (version 2020.06.1)	Same as “Quarantine in Europe”

<b>Storage conditions</b>	
Short name	storageConditions
Description	Conditions by which the isolate is stored by the CC.
Syntax	One or more of the allowed values, separated by “;”.
Values	Allowed values are “Lyo”, “Liquid nitrogen” and “-80”. More values can be added by MIRRI ICT group on demand. Requests must be submitted and accepted before the uploading of the catalog.
Validation	Check that one or more of the allowed values are included. Return errors to the CC.
Examples	Lyo
MIRRI-IS DATASET (version 2020.06.1)	N/A

<b>Recommended medium for growth for host cells</b>	
Short name	recommendedMedium
Description	The medium that is recommend for growing the isolate in the host cells.
Syntax	A textual reference, usually an acronym, to the appropriate growth medium in a table provided by the CC.
Values	CCs are invited to submit a table including a list of the growth media they use. The table should include at least an acronym and a description for each growth medium. A full description of the recipe is also welcome. All descriptions should be in English. In future versions of the MIRRI-IS dataset, a table of shared descriptions with acronyms will be provided.
Validation	Check for the presence of the textual reference in the provided table of growth media. Report errors to the CC.
Examples	LB
MIRRI-IS DATASET (version 2020.06.1)	Same as “Recommended medium for growth”, but it refers to growth in host cells.

<b>Propagation host</b>	
Short name	propagationHost
Description	The host that is used to propagate the isolate.
Syntax	The scientific name of the species, according to the corresponding reference taxonomy. It may include the strain designation, as well as other more specific designations, like cultivar, pathovar, serovar, ....
Values	All values of the reference taxonomy are allowed and cheked. Further specifications are accepted as free text, and not checked.
Validation	None.
Examples	Nicotiana tabacum Escherichia coli K21



MIRRI-IS DATASET (version 2020.06.1)	N/A
---	-----

## Taxonomy

This group includes all data fields which refer to the taxonomy of the isolate, its mutations/modifications and the methods for its identification.

<b>Organism type (MANDATORY)</b>	
Short name	organismType
Description	The type of the resource.
Syntax	Fixed value 'Virus'.
Values	'Virus'.
Validation	Check that the fixed value is specified. Report errors to the CC.
Examples	Virus
MIRRI-IS DATASET (version 2020.06.1)	Same as "Organism type", but only the value 'Virus' is allowed.

<b>Species name (MANDATORY)</b>	
Short name	speciesName
Description	Species name, as taken from the authoritative nomenclature reference of the International Committee for Virus Taxonomy ICTV.
Syntax	According to the ICTV nomenclature. In order to cope with delays in nomenclature updates, the most updated species name can be used, even when it is missing from the current version of the reference. In this case, a remark must be included in the 'Comment on taxonomy' data field.
Values	All species names included in the ICTV nomenclature, reported according to the given syntax.
Validation	Check for the correct syntax and the existence of the taxon name(s) in the reference nomenclatures. If the 'Comment on taxonomy' data field is filled in, the presence of a taxon name not present in the ICTV is allowed. CCs are reminded to revise this information as soon as the ICTV is updated. Otherwise, in case of error, it is reported to CCs.
Examples	Human polyomavirus 9 Pagavirus S05C849
MIRRI-IS DATASET (version 2020.06.1)	Similar to "Taxon name", but with reference to the ICTV nomenclature.

<b>Species abbreviation (MANDATORY)</b>	
Short name	speciesAbbreviation
Description	Virus abbreviations are short terms, often acronyms, used in substitution of the species name. They are not unique: the same abbreviation may be used for more than one virus.

	Virus abbreviations are not official ICTV designations and are not included in the ICTV reference nomenclature for viruses.
Syntax	None.
Values	Free text.
Validation	None.
Examples	HIV HPV
MIRRI-IS DATASET (version 2020.06.1)	N/A

<b>Former species name</b>	
Short name	formerSpeciesName
Description	Former species names, old names, synonymous not officially accepted. One or more separated by ";".
Syntax	None
Values	Free text
Validation	None
Examples	Watermelon mosaic virus 2
MIRRI-IS DATASET (version 2020.06.1)	N/A

<b>Pathotype-Serotype-Type</b>	
Short name	pathoSeroType
Description	Name of pathotype, serotype or strain. The ICTV is not responsible for classification and nomenclature of virus taxa below the rank of species. The classification and naming of serotypes, genotypes, strains, variants and isolates of virus species is the responsibility of acknowledged international specialist groups. There is no reference nomenclature for this information.
Syntax	This information must include a prefix ("Pathotype: ", "Serotype: " or "Strain: ") followed by the pathotype, serotype or strain designation.
Values	Free text, but for the prefixes and syntax.
Validation	Check the use of one of the allowed prefixes and the presence of the separator ":".
Examples	Pathotype: A Strain: Ohio 3A
MIRRI-IS DATASET (version 2020.06.1)	N/A

<b>Reference isolate (MANDATORY)</b>	
Short name	referenceIsolate

Description	Specifies whether it is the reference isolate.
Syntax	One of the allowed values.
Values	Allowed values: “No”, “Yes”
Validation	Check that one and only one of the allowed values is specified.
Examples	Yes
MIRRI-IS DATASET (version 2020.06.1)	N/A

<b>Comment on taxonomy</b>	
Short name	taxonomyComments
Description	Any comment and/or note on the taxonomy of the strain. It may be used, e.g., for information on new species or revised nomenclatures. It must be used when the Taxon name data field includes a name that is not present in the nomenclature reference.
Syntax	None
Values	Free text
Validation	None
Examples	This isolate species is not yet present in the ICTV
MIRRI-IS DATASET (version 2020.06.1)	Same as “Comment on taxonomy”.

<b>Identification technique</b>	
Short name	identificationTechniques
Description	All techniques used for the species identification. More than one can be specified.
Syntax	One or more of the allowed values, eventually separated by “;”.
Values	Allowed values: “ELISA”, “Western blot”, “Immuno blot”, “LAMP”, “Reverse Transcriptase PCR”, “PCR”, “Sanger sequencing”, “NGS”, “Bioassay”, “Electron microscopy”. More values can be added later by the MIRRI ICT group on demand by CCs.
Validation	Check that one or more of allowed values is specified.
Examples	PCR; Sanger sequencing
MIRRI-IS DATASET (version 2020.06.1)	N/A

<b>GMO (MANDATORY)</b>	
Short name	Gmo
Description	Specifies whether the isolate is a Genetically Modified Virus.

Syntax	One of the allowed values.
Values	The following values are allowed: “Yes”, “No”.
Validation	Check that one and only one of the allowed values is used. Report errors to the CC.
Examples	Yes
MIRRI-IS DATASET (version 2020.06.1)	Same as “GMO”.

<b>GMO information</b>	
Short name	gmoInformation
Description	Information on modifications, for genetically modified isolates only. By now, this information can be provided as free text. Future improvements of the dataset will likely foresee some syntactical rules and/or list of values.
Syntax	None
Values	Free text
Validation	None
Examples	
MIRRI-IS DATASET (version 2020.06.1)	Same as “GMO construction information”, but less demanding.

<b>Mutant information</b>	
Short name	mutantInformation
Description	Information on mutant isolates. By now, this information can be provided as free text. Future improvements of the dataset will likely foresee some syntactical rules and/or list of values.
Syntax	None
Values	Free text
Validation	None
Examples	
MIRRI-IS DATASET (version 2020.06.1)	Same as “Mutant information”.

<b>Gene sequences and accession numbers</b>	
Short name	sequences
Description	Known genomic sequences and related INSDC accession numbers of the strain. Any marker that is considered of relevance by the CC can be included. The whole genome can be specified.

	These data must be submitted in a separate table including, in distinct fields, the following information: accession number of the strain in the CC, marker name or “genome”, INSDC accession number of the marker sequence, sequence.
Syntax	Fields in the table follow different syntaxes: <u>Accession number</u> : as defined in the related field of the MIRRI-IS virus dataset. <u>Marker name</u> : the short name of the marker or the word ‘genome’. <u>INSDC accession number</u> : An INSDC accession number is an alphanumeric code made by a fixed number of letters followed by a fixed number of digits, without any separation. <u>Sequence</u> : Any valid genomic sequence.
Values	Values of fields in the table are as follows: <u>Accession number</u> : any accession number in the CC. <u>Marker name</u> : any common marker designation or the word “genome”. <u>INSDC accession number</u> : Any valid INSDC accession number. <u>Sequence</u> : Genomic sequence, any shared format, any length.
Validation	Check for the validity of the syntaxes, formats and values. Check that the sequence in INSDC actually relates to the named marker or complete genome of the virus. Report errors and discrepancies to the CC.
Examples	
MIRRI-IS DATASET (version 2020.06.1)	Similar to “Genomic sequences and accession numbers”.

## Origin

This group includes all data fields which refer to the origin of the isolate, including the geographic origin and the isolation host.

Isolation host	
Short name	isolationHost
Description	The name of the host where the virus was isolated from. The host can be animal, human, plant, cell lines, etc... Vernacular names (English or original language) can only be accepted if they were provided by depositors and cannot be unequivocally assigned to species. In this case, the species must be omitted.
Syntax	It should be expressed by using the reference nomenclature for the relative kingdom, when adequate. In this case, it should include the following information: species (latin binominal), subspecies (latin), variety (latin), cultivar or breeding line (original designation). In case of phages, it should be the bacteria species plus pathovar, according to the official nomenclature. For cell lines, the species and the cell line name should be specified.
Values	All values of the reference taxonomy are allowed. Further specifications are accepted as free text, and not checked.
Validation	None
Examples	Solanum lycopersicum cv York Pseudomonas syringae pv actinidiae
MIRRI-IS DATASET (version 2020.06.1)	N/A

Geographic origin	
Short name	geographicOrigin
Description	The locality where the sample was collected, defined with the highest possible precision. It does not include geographic coordinates.
Syntax	Reference to a separate table, which includes all localities where at least one isolate was collected. For organisms constructed in a lab, use the address of the depositor.
Values	The geographic location should be defined with the highest possible precision, but unambiguously. It should include locality, city, province, region, country. Country must be specified even if it is reported in the devoted data field. For old strains for which the geographic origin is not known, refer to the special locality 'Unknown'. Avoid specifying countries and continents only. The table should include separate fields for the geographic details. In order to improve the description of the location, you can check if it is described in GeoNames (see <a href="http://geonames.org/">http://geonames.org/</a> ) and use its 'Administrative hierarchy' to include further rows with information

	missing in the table, e.g. administrative commune and region, until you find the country.
Validation	Check for the presence of the reference in the table of localities. Check for the validity of the format and values in the table. Report errors to the CC. When some information is missing, MIRRI-IS will try to determine it through the geographical database GeoNames (see <a href="http://geonames.org/">http://geonames.org/</a> ) using the geographic origin information and will return it to the CC.
Examples	In order to insert Altafjorden, look at GeoNames. You will find it associated to the record n. 780944 whose administrative hierarchy reports Norway as country, Troms og Finnmark as adm1 and Alta as adm2. In your table you should include in separate cells Altafjorden, Alta, Troms og Finnmark, Norway.
MIRRI-IS DATASET (version 2020.06.1)	Similar to “Geographic origin”, but not mandatory, because the new field “Country” is the only mandatory information.

<b>Geographic origin coordinates</b>	
Short name	geographicCoordinates
Description	The geographic coordinates of the locality where the sample was collected.
Syntax	Geographic coordinates expressed in decimal degrees. Cardinal directions North and West are implicit and must not be reported. Conversion of latitude and longitude values from the sexagesimal format (as in 40° 26' 46") to the decimal format (40.446) can easily be achieved as follows: decimal degrees = sexagesimal degrees + (sexagesimal minutes/60) + (sexagesimal seconds/3600). For organisms constructed in a lab, use the coordinates of the depositor.
Values	For latitude and longitude, use decimal numbers from -180 to 180 for longitude and -90 to 90 latitude. Geographic coordinates can be retrieved from GeoNames (see <a href="http://geonames.org/">http://geonames.org/</a> ) on the basis of the locality.
Validation	Check for the correct syntax and value ranges.
Examples	70.05765, 23.08293
MIRRI-IS DATASET (version 2020.06.1)	

<b>Country (MANDATORY)</b>	
Short name	Country
Description	The country where the sample was collected. This information should be provided with reference to the name of the country where the locality is located now, if possible. Former country names should be specified only when the locality or its current country are unknown.



Syntax	This information must be expressed by using the <u>ISO-3166 standard</u> for country codes. The preferred set is ISO 3166-1 alpha-2 (two letters code), but ISO 3166-1 alpha-3 (three letters code) is also accepted. Former country codes must follow standard's part three ISO 3166-3 (four letters code). Only one code can be included.
Values	All codes included in ISO 3166-1 alpha-2 (two letters code), ISO 3166-1 alpha-3 (three letters code) and ISO 3166-3 (four letters code). In the examples, IT is the two letters code for Italy, GBR is the three letters code for United Kingdom, and CSHH is the four letters code for the former country Czechoslovakia.
Validation	Check for the validity of codes. Return errors to CCs.
Examples	IT GBR CSHH
MIRRI-IS DATASET (version 2020.06.1)	N/A

## History of deposit

This group includes all data fields which refer to the history of the isolate, from sampling to isolation, deposit and inclusion in the catalog.

<b>Isolate Literature (MANDATORY)</b>	
Short name	isolateLiterature
Description	<p>Information on literature on the isolate.</p> <p>In this field, only numeric references to a separate reference table must be provided. Multiple references can be included for a single strain by reporting more numeric references separated by “;”.</p> <p>The reference table must include all bibliographic fields used for citing a paper, a book, a patent, or a document available on-line, including, e.g., authors, title, journal, volume, issue, pages, editors, publishers, as separate fields.</p> <p>For publications indexed by Pubmed or having an official DOI number, collections must include in the reference table the respective identifiers, PMID and DOI, instead of all bibliographic fields: in this case, MIRRI-IS will retrieve them and complete the bibliographic data.</p>
Syntax	Zero, one or more numbers separated by a semicolon “;”.
Values	Numeric reference numbers included in the reference table.
Validation	<p>MIRRI-IS will try to extract any missing PMID and DOI on the basis of the provided information and return it to the collection.</p> <p>Any errors and inconsistencies will also be reported to the CCs.</p>
Examples	1; 5; 120
MIRRI-IS DATASET (version 2020.06.1)	Similar to “Literature”, but with revised syntax and made mandatory.

<b>Depositor</b>	
Short name	depositor
Description	Name, institute and town / country of the depositor.
Syntax	None
Values	Free text
Validation	None
Examples	<p>M. Sebold, Inst. Pasteur, Paris, France</p> <p>P. Hirsch, Inst. Allg. Mikrobiol. Univ. Kiel, Germany</p>
MIRRI-IS DATASET (version 2020.06.1)	Same as “Depositor”.

<b>Date of deposit</b>	
Short name	depositDate

Description	Date when the strain was deposited at the CC.
Syntax	May include a full date in the ISO 8601 format. YYYY-MM-DD or YYYYMMDD for full dates, YYYY-MM for year and month only, YYYY for year only. For a quick introduction see <a href="https://en.wikipedia.org/wiki/ISO_8601">https://en.wikipedia.org/wiki/ISO_8601</a> .
Values	A valid date in one of the above formats
Validation	Check for the validity of the format. Report errors to the CC.
Examples	1999-02-20
MIRRI-IS DATASET (version 2020.06.1)	Same as “Date of deposit”.

<b>Collected by</b>	
Short name	collector
Description	Name, institute and town / country of the collector.
Syntax	None
Values	Free text
Validation	None
Examples	J. Fraser, Moredun Res. Inst., Edinburgh, UK
MIRRI-IS DATASET (version 2020.06.1)	Same as “Collected by”.

<b>Date of collection (MANDATORY)</b>	
Short name	collectionDate
Description	Date when the sample was collected.
Syntax	May include a full date in the ISO 8601 format. YYYY-MM-DD or YYYYMMDD for full dates, YYYY-MM for year and month only, YYYY for year only. For a quick introduction see <a href="https://en.wikipedia.org/wiki/ISO_8601">https://en.wikipedia.org/wiki/ISO_8601</a> .
Values	A valid date in one of the above formats
Validation	Check for the validity of the format. Report errors to the CC.
Examples	1999-11-27
MIRRI-IS DATASET (version 2020.06.1)	Same as “Date of collection”, but mandatory.

<b>Isolated by (MANDATORY)</b>	
Short name	isolator
Description	Name, institute and town / country of the isolator.

Syntax	None
Values	Free text
Validation	None
Examples	I. Orskov, Ser. Inst., Copenhagen, Denmark D. Haas, Inst. Pasteur, Paris, France
MIRRI-IS DATASET (version 2020.06.1)	Same as “Isolated by”, but mandatory.

<b>Date of isolation</b>	
Short name	isolationDate
Description	Date when the strain was isolated from the sample.
Syntax	May include a full date in the ISO 8601 format. YYYY-MM-DD or YYYYMMDD for full dates, YYYY-MM for year and month only, YYYY for year only. For a quick introduction, see <a href="https://en.wikipedia.org/wiki/ISO_8601">https://en.wikipedia.org/wiki/ISO_8601</a> .
Values	A valid date in one of the above formats
Validation	Check for the validity of the format. Report errors to the CC.
Examples	2019-08-17
MIRRI-IS DATASET (version 2020.06.1)	Same as “Date of isolation”.

<b>Date of inclusion in the catalogue</b>	
Short name	accessionDate
Description	Date when the strain was included in the catalog and/or an accession number was assigned to it.
Syntax	May include a full date in the ISO 8601 format. YYYY-MM-DD or YYYYMMDD for full dates, YYYY-MM for year and month only, YYYY for year only. For a quick introduction see <a href="https://en.wikipedia.org/wiki/ISO_8601">https://en.wikipedia.org/wiki/ISO_8601</a> .
Values	A valid date in one of the above formats
Validation	Check for the validity of the format. Report errors to the CC.
Examples	1996-12-13
MIRRI-IS DATASET (version 2020.06.1)	Same as “Date of inclusion in the catalogue”.

## Properties

This group includes data fields which relates to specific properties of the isolates.

<b>Pathogenicity</b>	
Short name	pathogenicity
Description	Specifies if the isolate is pathogen or not. If it is pathogen, this field also specifies for which organisms it is pathogen among humans, animals, plants, bacteria, fungi or other organisms.
Syntax	One of the allowed values
Values	Allowed values: "Pathogen for Humans", "Pathogen for Animals", "Pathogen for Plants", "Pathogen for Bacteria", "Pathogen for Fungi", "Pathogen for other organisms", "Not pathogen".
Validation	Check that one of the allowed values is specified
Examples	Pathogen for plants
MIRRI-IS DATASET (version 2020.06.1)	N/A

<b>Transmission by</b>	
Short name	transmissionBy
Description	The way or the vector by which the virus moves to a new host in nature.
Syntax	One or more of the allowed values separated by ";". Other values can be included as free text.
Values	Allowed values: "contact", "insect", "grafting", "nematodes", "fungi", "air", "pollen", "sexual transmission", "food".
Validation	Check that one or more of allowed values is specified.
Examples	Pollen; Insect
MIRRI-IS DATASET (version 2020.06.1)	N/A

<b>Infectivity</b>	
Short name	infectivity
Description	Specifies if the infectivity of the isolate has been tested.
Syntax	One of the allowed values
Values	Allowed values: "Yes", "No"
Validation	Check that one of the allowed values is specified
Examples	Yes
MIRRI-IS DATASET (version 2020.06.1)	N/A

<b>Contamination</b>	
Short name	contamination
Description	Specifies if the material is containing other viruses (=mixed infection) or if it is in pure culture. In case of plant viruses, mitochondrial viruses or cryptic viruses are not considered as contaminants.
Syntax	One of the allowed values.
Values	Allowed values: "Yes", "No"
Validation	Check that one of the allowed values is specified.
Examples	Yes
MIRRI-IS DATASET (version 2020.06.1)	N/A

<b>Lytic cycle</b>	
Short name	lyticCycle
Description	Specifies if the virus replicate through the lytic cycle (sometimes referred to as virulent infection).
Syntax	One of the allowed values
Values	Allowed values: "Yes", "No"
Validation	Check that one of the allowed values is specified.
Examples	No
MIRRI-IS DATASET (version 2020.06.1)	N/A

<b>Genome integration</b>	
Short name	genomeIntegration
Description	Specifies if the virus/phage is integrated in the host genome.
Syntax	One of the allowed values.
Values	Allowed values: "Yes", "No"
Validation	Check that one of the allowed values is specified.
Examples	Yes
MIRRI-IS DATASET (version 2020.06.1)	N/A

<b>Remarks</b>	
Short name	Remarks
Description	Any further note that is not present in the other fields.
Syntax	None
Values	Free text

Validation	None
Examples	Assignment to strain group according to the results of the COST-88 PVY-Ringtest. Obtained from poinsettia together with poinsettia cryptic virus.
MIRRI-IS DATASET (version 2020.06.1)	Same as "Remarks"