

Generic data collection template for ECPGR EVA networks

General information

This data collection template was designed to be applicable to ALL active EVA networks, which are evaluating different crops (e.g. cereals, vegetables) under different conditions (e.g. field, greenhouse, lab). Therefore, some fields may not be relevant for all networks (based on their experimental setup) and thus remain empty. Using a common template ensures that the EURISCO-EVA intranet can handle upload of data templates for all experiments.

The generic template has been split into two files which together collect all information pertinent to an experiment (one set of accessions of a crop type) in one place, including relevant metadata (on plant material, experiment setup and treatments) and phenotypic data (guidance on the scoring methods and scales of all traits under investigation and the actual data collection worksheet). In the framework of EVA information on plant material and traits is provided centrally in part A, while partners are required to complete information on their experiment metadata (setup and treatments, where applicable) and the actual phenotypic data in part B.

The EVA-ID (identifying plant material), Trial-ID (identifying single experiments), TreatmentName (identifying types of treatments) and TraitAcronyms (identifying individual traits) are unique identifiers and centrally assigned by the EVA coordinator.

Mandatory parameters are in bold font.

Part A: Experiment metadata

In this file are collected the centrally provided metadata on plant material, experiment setup and trait descriptions applicable to one set of experiments on one set of accessions for a specific network. Some networks evaluate multiple sets in parallel, however, for ease of data management each replicate of one set is considered a separate experiment and linked through the metadata.

Do not modify any information in Part A!!! This metadata is centrally provided by the EVA coordinator for information only and should be consulted for guidance on scoring of different traits and identity of plant material in the experiment, with data recorded in Part B. Any suggestions for modifications/additions should be communicated to the EVA coordinator.

Worksheet 1_Plant material

"This worksheet collects relevant passport information to identify material under evaluation in the EVA network and links it to the original accessions conserved in European genebanks. Depending on the crops worked at in the different EVA networks, the material evaluated can be the original accession (e.g. landrace, population) or a derived accession (E.g. single-seed-descent SSD lines, test cross populations). The fields in this worksheet are the minimum

requirements for identifying material, which is not being maintained in a genebank. If the material is maintained in a genebank, the EVA-ID + the fields given by the MCPD2 format should be used. The combination of Genus, AccessionNumber and InstitutionCode is used to identify plant material in a genebank and allows identification within EURISCO. In case the material was received from a genebank, these fields are mandatory.

Additional passport information provided by holding institutes is available to EVA partners within the network SharePoint and on EURISCO. "

MaterialType distinguishes different types of material in the EVA project, identifying derived materials in addition to the accessions provided from genebanks. Multiple materials can be derived from one original parent accession. Possible types are: original accession (material used in the same way as received from genebanks, e.g. landraces), SSD (single seed descent-line, derived from original accession), cross (testcross population, derived from original accession), check (controls used in experiments), tester (anonymized parent for crosses).

Depending on the experimental design of each network, different types of material are being used in the trials.

Partners are requested to use the EVA-ID in their experiments for coding plant material.

Table: Worksheet 1 Plant material

EVA-ID	InstitutionCode	Genus	Species	Crop	AccessionNumber	DOI	AccessionNumberParent	EVA_ID_PARENT	DOI-Parent
<i>Unique identifier of germplasm accession within the project, centrally provided</i>	<i>FAO-WIEWS code of the institute maintaining the original material, following MCPD, or EVA institution code for proprietary or derived material, where applicable. Use TBD for material where institute is unclear.</i>	<i>Genus name for taxon. Initial uppercase letter required.</i>	<i>Specific epithet portion of the scientific name in lowercase letters. Only the following abbreviation is allowed: 'sp.'</i>	<i>common English name of crop under evaluation, e.g. barley, wheat, carrot</i>	<i>Identifier for material in EVA. For original accessions, this is the genebank accession number. For derived material (SSD, crosses) these are linked to parent accessions provided by genebanks (original accessions)</i>	<i>A permanent unique identifier for the plant material, following MCPD. Include where available.</i>	<i>Identifier for parent accessions of derived material (SSD); as provided by holding genebank, following MCPD.</i>	<i>EVA identifier of parent accessions. Mostly useful for those that are not in EURISCO. (included where available)</i>	<i>A permanent unique identifier for parent material, e.g. for the original genebank accession the SSD line was selected from. Include where available.</i>
EVA_Hv_00630	ITA382	Hordeum	vulgare	barley	5678_SSD		5678	EVA_Hv_00632	
EVA_Hv_00632	ITA382	Hordeum	vulgare	barley	5678				
EVA_Hv_00546	DEU146	Hordeum	vulgare	barley	HOR 19653 BRG	10.25642/IPK/GBIS/7825430	HOR 19653		10.25642/IPK/GBIS/245636
EVA_Zm_tester01	EVA_CHE002	Zea	mays	maize	tester01				
EVA_Zm_00631	EVA_CHE002	Zea	mays	maize	EVA_Zm_00001XEVA_Zm_tester01				
EVA_xx_check1	TBD	Hordeum	vulgare	barley	check01				

MaterialType	Female Parent	MaleParent	in EURISCO	Accession name	ProvenanceCountry	SampleStatus	RegistrationYear	CollectionYear	Remarks	SearchVisible
<i>Identifier for the type of material in the EVA project, multiple materials can be derived from one original parent accession. possible types are: <u>original accession</u> (material used in the same way as received from genebanks, e.g. landraces), <u>SSD</u> (single seed descent-line, derived from original accession), <u>cross</u> (testcross population, derived from original accession), <u>check</u> (controls used in experiments), <u>tester</u> (anonymized parent for crosses).</i>	<i><u>only relevant for MaterialType "cross", EVA-ID of female parent</u></i>	<i><u>only relevant for MaterialType "cross", EVA-ID of male parent</u></i>	<i>information on whether material is already included in EURISCO (yes/no)</i>	<i>Either a registered or other designation given to the material received, other than the donor's accession number (column E).</i>	<i>ISO3 code of the country in which the material was collected or bred. For derived material this can be the country of the parent material or remain empty.</i>	<i>Biological status of accession following MCPD format. Derived material is classified as 420 (genetic stock, SSD lines) or 400 (breeding/research materials, crosses).</i>	<i>(for varieties)</i>	<i>(for landraces or wild populations)</i>	<i>free text remarks on material.</i>	<i>This indicates which plant material will be searchable in the database. Excluded are e.g. private checks, parent materials and testers that are not evaluated by themselves.</i>
SSD			N	selection from cultivar		420 (genetic stock)			selection SSD generated from	y
original accession			Y	Aquilall	ITA	300 (or whatever was provided by genebank)			original genebank accession	y
SSD			Y		ROU	420 (genetic stock)				y
tester			N			410 (Breeder's line)				n
cross	EVA_Zm_00001	EVA_Zm_tester01	N	testcross	ITA	400 (Breeding/Research material)			testcross	y
check			N	name of check		400 or 500 (registered variety)				y

Worksheets 2a_experiment metadata and 2b_experiment metadata lab

Metadata is collected for the different experiments to enable comparison between trials. Depending on the crop type and typical experimental set-up in the field or under controlled conditions (greenhouse/lab), parameters outlined in this sheet may be adjusted and not all are required.

Where appropriate, we are using some additional parameters to more easily group related experiments, such as:

- Crop (common name, e.g. barley, wheat, carrot)
- ExperimentType (e.g. field, lab)
- ExperimentSubtype (where necessary, e.g. winter/spring, Eval A etc)
- ExperimentGroup (allows grouping according to evaluation sets/regions, e.g. Set1)

In addition, information is provided for trial location, including GPS coordinates and physical information, experiment design and planting information.

Most metadata information on trials has been centrally collected by the EVA Coordinator, and is provided for each planned trial for information only. In Part B, some specific information on experiments should be provided, this will complement the existing information. These fields are highlighted in yellow in Part A.

If the provided metadata for an experiment is incorrect, please contact the EVA Coordinator. You will also be able to modify metadata for experiments of your organization directly in the EURISCO-EVA intranet.

Table: Worksheet 2a Experiment metadata field

TrialID	Name	Crop	ExperimentType	Experiment Subtype	ExperimentGroup	Organisation	ContactPerson
unique identifier for each experiment, provided by the EVA coordinator	descriptive name of the experiment set-up	common name, e.g. barley, wheat, carrot	e.g. field, greenhouse	where necessary	allows grouping according to evaluation sets/regions as necessary	company/institute/group organising the trial	person in charge of trial/providing data, include email if relevant

EVA_crop_trial# barley field trial barley field southern set 1 Institute of Plant Breeding and Genetic Resources, ELGO-DIMITRA,
2020/2021

Country	Location	Site	Latitude	Longitude	HeightAboveSeaLevel	LongTermMeanOfPrecipitation	LongTermMeanOfTemperature	SoilType
country in which the field trial is located	city or region of the trial	name of the field	GPS coordinates [preferably in decimal degrees, e.g. 40.741895]	GPS coordinates [preferably in decimal degrees, e.g. -73.989308]	[m]	[mm] used for describing the location, not for analysis	[°C] used for describing the location, not for analysis	e.g. following FAO soil classification

Greece Thermi-Thessaloniki Thermi 40°53'52.1" N 23°00'55.0" W 19 Sandy Clay

Experimental design	PlotLength	PlotWidth	Number of plots	Number of Rows Per Plot	Distance Between Rows Within Plots	Distance Between Rows Between Plots	Sowing Depth	Sowing Density Count	Sowing Density	Sowing Density Count Area	Remarks
[1 or 2 block]	[m]	[m]	[replicates]		[m]	[m]	[mm]	[number of seeds/m]	[kg of seeds/ha]	[number of seeds/m ²]	free text to provide relevant info on trial, e.g. deviations from protocol, problems with scoring

2 2 0.25 320 (1st set+ check/diffuser) 1 0.25 0.25 50 50 seeds/row; Hand sowing

Worksheet 2b Experiment metadata lab

TrialID	Name	Crop	ExperimentType	ExperimentSubtype	ExperimentGroup	Organisation	ContactPerson	Country	Location	Site
unique identifier for each experiment, provided by the EVA coordinator	descriptive name of the experiment set-up	common name, e.g. barley, wheat, carrot	e.g. lab, greenhouse	where necessary	allows grouping according to evaluation sets/regions as necessary	company/institute/group organising the trial	person in charge of trial/providing data, include email if relevant	country in which the field trial is located	city or region of the trial	name of the greenhouse/growth facility, as applicable
EVA_crop_trial #	biotic stress [pathogen name]	carrot	lab			JKI		Germany	Quedlinburg	JKI

Experimental design	RelativeHumidity	PotDiameter	PotDepth	SoilType	SoilVolume	LightDuration [h/d]	LightIntensity	AverageTemperature Day	AverageTemperatureNight	Irrigation	Remarks
[1 or 2 block]	[%]	[cm]	[cm]	e.g. following FAO soil classification	[litres]	[m]	[$\mu\text{mol}/\text{m}^2\text{s}$]	[°C]	[°C]	[mm]	
	20	20	15			6		22	18		...

Worksheets 3_Traits

The traits worksheets provide guidance on the protocols and allowed scoring values of the traits under evaluation. To allow effective comparison between trials, methods should be followed and only allowed values used to record phenotypic data

Worksheet 3a_Definition of Traits contains a brief summary of scoring methods for all traits in the experiment, and is provided centrally. For additional detail partners are referred to the scoring protocols developed within the different networks. Worksheets 3b_Range of values (rating) and 3c_Range of values (metric) list the allowed values for scoring scales and measurements of the specific traits.

Table: Worksheet 3a Definition of traits

TraitAcronym	TraitName	Description	Unit	Type	CropOntologyTerm	Remarks
(unique identifier, centrally provided)		brief summary of method, for additional detail refer to scoring protocol	(e.g. n for scoring scales; ASCII unit for measurements)	(use only: score, date, measurement, text)	where available	
EWB_1000	1000 kernel weight (g);	g; determined as the weight of 1000 grains sampled from 100% clean harvest	g	measurement	CO 321:0000025	optional
EWB_WBG	powdery mildew (<i>Blumeria graminis f. sp. tritici</i>)	IPGRI descriptor 8.2.4. Average of percentage of infected leaves per plot, symptom expression as 1-9 scores; 1= least symptoms 9= most symptoms	n	score	CO 321:0000939	see standard protocols in shared folder
EWB_Y	yield	t ha ⁻¹ ; Grain yield is measured by harvesting each plot and converting the weight to tons per hectare, based on the plot area that was harvested.	t/ha	measurement	CO 321:0000013	optional

Table: Worksheet 3b Range of values (rating)

TraitAcronym	TraitName	RatingScore	Value	Remarks
(unique identifier, centrally provided)		only the below values are allowed for specific traits	(characteristic associated with that rating score)	
EWB_WBG	powdery mildew (<i>Blumeria graminis f. sp. tritici</i>)	1	no symptoms (resistant)	
EWB_WBG	powdery mildew (<i>Blumeria graminis f. sp. tritici</i>)	2	less than 1% leaf surface affected	
EWB_WBG	powdery mildew (<i>Blumeria graminis f. sp. tritici</i>)	3	less than 3% leaf surface affected	
EWB_WBG	powdery mildew (<i>Blumeria graminis f. sp. tritici</i>)	4	less than 5% leaf surface affected	
EWB_WBG	powdery mildew (<i>Blumeria graminis f. sp. tritici</i>)	5	~10% leaf surface affected	
EWB_WBG	powdery mildew (<i>Blumeria graminis f. sp. tritici</i>)	6	~20-30% leaf surface affected	
EWB_WBG	powdery mildew (<i>Blumeria graminis f. sp. tritici</i>)	7	~40-50% leaf surface affected	
EWB_WBG	powdery mildew (<i>Blumeria graminis f. sp. tritici</i>)	8	~60% leaf surface affected	
EWB_WBG	powdery mildew (<i>Blumeria graminis f. sp. tritici</i>)	9	~70% leaf surface affected	

Table: Worksheet 3c Range of values (metric)

TraitAcronym	TraitName	MinimumValue	MaximumValue	Unit	Remarks
(unique identifier, centrally provided)		only values between the minimum and maximum values specified below are allowed for specific traits			
EWB_1000	1000 kernel weight (g);	0	100	g	g; determined as the weight of 1000 grains sampled from 100% clean harvest
EWB_Y	yield	0	10	t/ha	t ha ⁻¹ ; Grain yield is measured by harvesting each plot and converting the weight to tons per hectare, based on the plot area that was harvested.

Part B: Data collection template

This file contains the worksheets partners should use to record specific experiment metadata, information on treatments applied in the experiment and the actual phenotypic data. Partners are reminded to use Part A as a reference document for recording data in their experiments.

Each experiment (as identified by unique Trial-ID) should be recorded in a separate file. Files should be saved using the as filename the unique Trial-ID assigned to each experiment.

Cells for which no information is available should either remain blank or include NA (not available) - these values will be ignored during upload. Note that required fields (with titles in bold font) cannot remain blank.

The completed worksheets can be directly uploaded to the EURISCO-EVA intranet database using the upload tool. However, partners are required to follow the below guidelines in completing the data files to ensure that the uploader recognizes the format of the provided data.

Worksheet 2c_specific metadata

In this worksheet, partners are requested to provide specific information on their experiment, which may not be available in advance. Please refer to sheet 2a or 2b in part A for information provided by partners before the trials. Note that all metadata for experiments can be modified by the responsible organization directly in the EURISCO-EVA intranet.

The following should be provided:

- TrialID (refer to sheet 2a/2b in part A to identify the unique Trial-ID for your experiment)
- ContactPerson (person in charge of trial/providing data, include email if relevant)
- ExperimentStart [yyyy-mm-dd]
- HarvestDate [yyyy-mm-dd] (where applicable)
- ExperimentEnd [yyyy-mm-dd]

- Remarks (free text to provide relevant info on trial, e.g. deviations from protocol, problems with scoring)

Table: Worksheet 2c specific metadata

TrialID	ContactPerson	ExperimentStart	HarvestDate	ExperimentEnd	Remarks
(refer to sheet 2a/2b in part A to identify the unique Trial-ID for your experiment)	person in charge of trial/providing data, include email if relevant	sowing date [yyyy-mm-dd]	[yyyy-mm-dd]	[yyyy-mm-dd]	free text to provide relevant info on trial, e.g. deviations from protocol, problems with scoring
EVA_crop_trial#	John Doe <j.doe@EVApartner.com>				hail storm on 2023-08-10 reduced crop yield

Worksheet 2d_experiment treatment

This worksheet collects information on treatments applied during the experiments. In order to facilitate comparison between experiments in different locations, information on these should be as complete as possible. Several different treatments could be recorded during experiments in field, greenhouse or lab environments, using consistent naming for the TreatmentNames, with the Remark field used to provide specific detail on treatments. Please only use the following TreatmentNames:

- HerbicideTreatment
- FungicideTreatment
- InsecticideTreatment
- FertilizerTreatment
- GrowthRegulatorTreatment
- IrrigationTreatment
- RainoutShelterTreatment
- DiseaseInoculationTreatment

Additional TreatmentNames can be added by the EVA coordinator as per network needs.

For each of these treatments, a set of parameters should be stored, as applicable, with rows added for treatments as necessary.

- TreatmentName (e.g. HerbicideTreatment; ensure consistent naming with allowed names as listed above)
- TreatmentNo (running number of treatment)
- StartDate [yyyy-mm-dd]
- EndDate [yyyy-mm-dd]

- BBCHScale (where relevant)
- Amount
- Unit
- Product
- Remarks (Free text, use this field to specify treatments, e.g. Macronutrient, KPN etc. for FertilizerTreatment)

Table: Worksheet 2d experiment treatment

TreatmentName	TreatmentNo	StartDate	EndDate	BBCHScale	Amount	Unit	Product	Remarks
ensure consistent naming with allowed names as listed	(running number of treatment)	[yyyy-mm-dd]	[yyyy-mm-dd]	(where relevant)				(Free text, use this field to specify treatments, e.g. Macronutrient, KPN etc. for FertilizerTreatment)
HerbicideTreatment	1	2018-05-30	2018-05-30	30	0.3	l/m ²	Biscaya	...
FungicideTreatment	2	2018-06-01	2018-06-01	20	40	kg/ha	calcium ammonium nitrate	...
InsecticideTreatment	3	2018-07-24	2018-07-24	20	70	kg/ha	ammonium sulphate nitrate	...
FertilizerTreatment	4	2018-07-24	2018-07-24	30	0.5	l/m ²	Medax Top	...

Worksheet 4_obs values

All observation data should be collected in this sheet, which may be adapted to the experiments relevant for different crop types and experimental set-ups. For practical purposes in the field, simplified versions could be developed, but collected data will need to be transformed to the below format for data transfer and storage in the EURISCO-EVA intranet.

The sheet starts with a part to record the field/experiment layout, which should be provided by partners, where relevant. For each replicate plot of an accession (EVA-ID), a separate row should be recorded, additional rows can be added as needed.

- Plot (continuous numbering of experimental plots within the trial)
- Row (row number of a given plot in field layout)
- Column (column number of a given plot in field layout)
- Replicate (for block design)

- EVA-ID (see sheet 1_Plant material in part A)
- Control (accessions used as controls are indicated by "C")
- Remark (free text to record any relevant observations of the plot, for example if the material is segregating for a specific trait, details can be recorded here. 4000 char max.)

This layout information is followed by any number of pairs of additional columns for each trait scored in the experiment, maintaining the order of columns as below:

- TraitAcronym (See sheets 3_Traits in part A)
- ObservationDate [yyyy-mm-dd] of the trait scoring

Additional timepoints for some traits (e.g. in timecourse experiments) can be recorded by inserting columns for TraitAcronym and ObservationDate as needed, using the same TraitAcronym.

The observed values recorded under each TraitAcronym must only use allowed values as specified in sheets 3b and 3c in part A.

The observation date must use the format yyyy-mm-dd and can only be omitted in case of traits that only consist of a date value, e.g. flowering date.

If a trait was scored in the experiment, but for a certain accession/plant/plot no value could be obtained (e.g. technical reasons or dead plot), the respective table cell must be marked with ND (not detected).

If a trait was not scored in the experiment, the respective table cells must be marked with NA (not available) or left empty for the whole column. Cells with NA and empty cells will be ignored by the uploader.

Table: Worksheet 4_obs values

Plot	Row	Column	Replicate	EVA-ID	Control	Remark	EWB_1000	ObservationDate	EWB_WBG	ObservationDate	EWB_Y	ObservationDate
							Trait 1		Trait 2		Trait 3	
continuous numbering of experimental plots	row number of a given plot in field layout	column number of a given plot in field layout	number the replicate plots for each accession in the trial	(unique identifier of germplasm accession within the project), centrally provided	indicate accessions used as controls as "C"	free text to record any relevant observations, for example if the material is segregating for a specific trait, details can be recorded here. 4000 char max.	1000-kernel weight [g]	[yyyy-mm-dd]	powdery mildew (Blumeria graminis f. sp. tritici)	[yyyy-mm-dd]	yield [t/ha]	[yyyy-mm-dd]
1	1	1	1	EVA_Ta_00554		e.g. accession segregating for trait Exx_T1: 2 (60%), 3 (40%); Exx_T2: 3 (75%), 2 (25%) etc...	89	2022-07-18	1	2022-06-22	7.5	2022-07-18
2	1	2	1	EVA_Ta_00555			75	2022-07-18	3	2022-06-22	6.48	2022-07-18
3	1	3	1	EVA_Ta_00556			82	2022-07-18	9	2022-06-22	9.14	2022-07-18
4	1	4	1	EVA_Ta_00557			91	2022-07-18	5	2022-06-22	8.14	2022-07-18
5	2	4	1	EVA_Ta_Toras	C		95	2022-07-18	7	2022-06-22	8.54	2022-07-18
6	2	3	2	EVA_Ta_00554			87	2022-07-18	2	2022-06-22	6.8	2022-07-18
7	2	2	2	EVA_Ta_00555			81	2022-07-18	2	2022-06-22	8.23	2022-07-18
8	2	1	2	EVA_Ta_00556			79	2022-07-18	8	2022-06-22	8.56	2022-07-18
9	3	1	2	EVA_Ta_00557			89	2022-07-18	4	2022-06-22	9.87	2022-07-18
10	3	2	2	EVA_Ta_Toras	C		90	2022-07-18	6	2022-06-22	8.63	2022-07-18