	MIAPPE				
line #	MIAPPE Check list	Definition	Example	Format	Cardinality
DM-1	**Investigation**	Investigations are research programmes with defined aims. They can exist at various scales (for example, they could encompass a grant-funded programme of work, the various components comprising a peer-reviewed publication, or a single experiment).			1 per MIAPPE submission
DM-2	**Investigation unique ID**	Identifier comprising the unique name of the institution/database hosting the submission of the investigation data, and the accession number of the investigation in that institution.	EBI:12345678	Unique identifier	0-1
DM-3	**Investigation title**	Human-readable string summarising the investigation.	Adaptation of Maize to Temperate Climates: Mid-Density Genome-Wide Association Genetics and Diversity Patterns Reveal Key Genomic Regions, with a Major Contribution of the Vgt2 (ZCN8) Locus.	Free text (short)	1
DM-4	**Investigation description**	Human-readable text describing the investigation in more detail.	The migration of maize from tropical to temperate climates was accompanied by a dramatic evolution in flowering time. To gain insight into the genetic architecture of this adaptive trait, we conducted a 50K SNP-based genome-wide association and diversity investigation on a panel of tropical and temperate American and European representatives.	Free text	0-1
DM-5	**Submission date**	Date of submission of the dataset presently being described to a host repository.	2012-12-17	Date/Time (ISO 8601, optional time zone)	0-1
DM-6	**Public release date**	Date of first public release of the dataset presently being described.	2013-02-25	Date/Time (ISO 8601, optional time zone)	0-1
DM-7	**License**	License for the reuse of the data associated with this investigation. The Creative Commons licenses cover most use cases and are recommended.	CC BY-SA 4.0, Unreported	Unique identifier	0-1
DM-8	**MIAPPE version**	The version of MIAPPE used.	1.1	Version number	1
DM-9	**Associated publication**	An identifier for a literature publication where the investigation is described. Use of DOIs is recommended.	doi:10.1371/journal.pone.0071377	DOI	0+
DM-10	**Study**	A study (or experiment) comprises a series of assays (or measurements) of one or more types, undertaken to answer a particular biological question.			1+ per investigation
DM-11	**Study unique ID**	Unique identifier comprising the name or identifier for the institution/database hosting the submission of the study data, and the identifier of the study in that institution.	EBI:12345678 http://phenome-fppn.fr/maugio/2013/t2351	Unique identifier	0-1
DM-12	**Study title**	Human-readable text summarising the study	2002 evaluation of flowering time for a panel of 375 maize lines at the experimental station of Maugio (France).	Free text (short)	1
DM-13	**Study description**	Human-readable text describing the study	2002 evaluation of male and female flowering time for a panel of 375 maize lines representing the worldwide genetic diversity at the experimental station of Maugio, France.	Free text	0-1
DM-14	**Start date of study**	Date and, if relevant, time when the experiment started	2002-04-04 2006-09-27T10:23:21+00:00	Date/Time (ISO 8601, optional time zone)	1
DM-15	**End date of study**	Date and, if relevant, time when the experiment ended	2002-11-27	Date/Time (ISO 8601, optional time zone)	0-1
DM-16	**Contact institution**	Name and address of the institution responsible for the study.	UMR de Génétique Végétale, INRA – Université Paris-Sud – CNRS, Gif-sur-Yvette, France	Free text (short)	1
DM-17	**Geographic location (country)**	The country where the experiment took place, either as a full name or preferably as a 2-letter code.	FR	Country name or 2-letter code (ISO 3166)	1
DM-18	**Experimental site name**	The name of the natural site, experimental field, greenhouse, phenotyping facility, etc. where the experiment took place.	INRA, UE Diascope - Chemin de Mezouls - Domaine expérimental de Melgueil - 34130 Mauguio - France	Free text (short)	1
DM-19	**Geographic location (latitude)**	Latitude of the experimental site in degrees, in decimal format.	+43.619264	Degrees in the decimal format (ISO 6709)	0-1 (1 if longitude is provided)
DM-20	**Geographic location (longitude)**	Longitude of the experimental site in degrees, in decimal format.	+3.967454	Degrees in the decimal format (ISO 6709)	0-1 (1 if latitude is provided)
DM-21	**Geographic location (altitude)**	Altitude of the experimental site, provided in metres (m).	100 m	Numeric + unit abbreviation	0-1
DM-22	**Description of the experimental design**	Short description of the experimental design, possibly including statistical design. In specific cases, e.g. legacy datasets or data computed from several studies, the experimental design can be "unknown"/"NA", "aggregated/reduced data", or simply 'none'.	Lines were repeated twice at each location using a complete block design. In order to limit competition effects, each block was organized into four sub-blocks corresponding to earliness groups based on a priori information.	Free text	1
DM-23	**Type of experimental design**	Type of experimental design of the study, in the form of an accession number from the Crop Ontology.	CO_715:0000145	Crop Ontology term (subclass of "CO_715:0000003")	0-1
DM-24	**Observation unit level hierarchy**	Hierarchy of the different levels of repetitions between each others	block>rep>plot	Formatted text (level>level)	0-1
DM-25	**Observation unit description**	General description of the observation units in the study.	Observation units consisted in individual plots themselves consisting of a row of 15 plants at a density of approximately six plants per square meter. NA	Free text	1
DM-26	**Description of growth facility**	Short description of the facility in which the study was carried out.	field environment condition NA	Free text (short)	1
DM-27	**Type of growth facility**	Type of growth facility in which the study was carried out, in the form of an accession number from the Crop Ontology.	CO_715:0000162	Crop Ontology term (subclass of "CO_715:0000005")	0-1
DM-28	**Cultural practices**	General description of the cultural practices of the study.	Irrigation was applied according needs during summer to prevent water stress.	Free text	0-1
DM-29	**Map of experimental design**	Representation of the experimental design.	https://urgi.versailles.inra.fr/files/ephesis/181000503/181000503_plan.xls	URL or File name (of gis or tabular file like csv or tsv)	0+
DM-30	**Person**	A human involved in the investigation or specifically any of its studies.			1+ per investigation / 0+ per study
DM-31	**Person name**	The name of the person (either full name or as used in scientific publications)	Ines Chaves	Name	1
DM-32	**Person email**	The electronic mail address of the person.	ichaves@itqb.unl.pt	email address	0-1
DM-33	**Person ID**	An identifier for the data submitter. If that submitter is an individual, ORCID identifiers are recommended.	orcid.org/0000-0001-6494-0008; orcid.org/0000-0002-7054-800X	Unique identifier	0-1
DM-34	**Person role**	Type of contribution of the person to the investigation	data submitter; author; corresponding author	Free text (short)	1+
DM-35	**Person affiliation**	The institution the person belongs to	ITQB, Portugal; grid.10772.33	Free text (short)	1+
DM-36	**Data File**	A file or digital object holding observation data recorded during one or more assays of the study, typically in tabular form. Multiple data files may be provided per study, and each file can include observations for several observation units and several observed variables.			0+ per study
DM-37	**Data file link**	Link to the data file (or digital object) in a public database or in a persistent institutional repository; or identifier of the data file when submitted together with the MIAPPE submission.	http://www.ebi.ac.uk/arrayexpress/experiments/E-GEOD-32551/	URL or File name	1
DM-38	**Data file description**	Description of the format of the data file. May be a standard file format name, or a description of organization of the data in a tabular file.	FASTA tab-delimited column headers headers: 1. A 2. B 3. C	Free text (short)	1
DM-39	**Data file version**	The version of the dataset (the actual data).	1.0	Software version number	1
DM-40	**Biological Material**	The biological material being studied (e.g. plants grown from a certain bag or seed, or plants grown in a particular field). The original source of that material (e.g., the seeds or the original plant cloned) is called the material source, which, when held by a material repository, should have its stock identified.			1+ per study; 0+ per observation unit
DM-41	**Biological material ID**	Code used to identify the biological material in the data file. Should be unique within the Investigation. Can correspond to experimental plant ID, seed lot ID, etc… This material identification is different from a BiosampleID which corresponds to Observation Unit or Samples sections below.	INRA:W95115_inra_2001; INRA:inra_kernel_2351; Rothamsted:rres_GK090847	Unique identifier	1
DM-42	**Organism**	An identifier for the organism at the species level. Use of the NCBI taxon ID is recommended.	NCBI:4577	Unique identifier	1
DM-43	**Genus**	Genus name for the organism under study, according to standard scientific nomenclature.	Zea Solanum	Genus name	0-1
DM-44	**Species**	Species name (formally: specific epithet) for the organism under study, according to standard scientific nomenclature.	mays lycosperium x pennellii	Species name	0-1
DM-44	**Infraspecific name**	Name of any subtaxa level, including variety, crossing name, etc. It can be used to store any additional taxonomic identifier. Either free text description or key-value pair list format (the key is the name of the rank and the value is the value of the rank). Ranks can be among the following terms: subspecies, cultivar, variety, subvariety, convariety, group, subgroup, hybrid, line, form, subform. For MCPD compliance, the following abbreviations are allowed: 'subsp.' (subspecies); 'convar.' (convariety); 'var.' (variety); 'f.' (form); 'Group' (cultivar group).	vinifera Pinot noir B73 subspecies:vinifera ; cultivar:Pinot noir var:B73 subsp. vinifera var. Pinot Noir var. B73	Free text, or key-value pair list, or MCPD-compliant format	0-1
DM-45	**Biological material latitude**	Latitude of the studied biological material. [Alternative identifier for in situ material]	+39.067	Degrees in the decimal format (ISO 6709)	0-1 (1 if longitude is provided)
DM-46	**Biological material longitude**	Longitude of the studied biological material. [Alternative identifier for in situ material]	-8.73	Degrees in the decimal format (ISO 6709)	0-1 (1 if latitude is provided)
DM-47	**Biological material altitude**	Altitude of the studied biological material, provided in meters (m). [Alternative identifier for in situ material]	10 m	Numeric + unit abbreviation	0-1
DM-48	**Biological material coordinates uncertainty**	Circular uncertainty of the coordinates, preferably provided in meters (m). [Alternative identifier for in situ material]	200 m	Numeric	0-1
DM-49	**Biological material preprocessing**	Description of any process or treatment applied uniformly to the biological material, prior to the study itself. Can be provided as free text or as an accession number from a suitable controlled vocabulary.	EO:0007210 - PVY(NTN); transplanted from study http://phenome-fppn.fr/maugio/2013/t2351 observation unit ID: pot:894	Plant Environment Ontology and/or free text	0+

DM-50	**Material source ID (Holding institute/stock centre, accession)**	An identifier for the source of the biological material, in the form of a key-value pair comprising the name/identifier of the repository from which the material was sourced plus the accession number of the repository for that material. Where an accession number has not been assigned, but the material has been derived from the crossing of known accessions, the material can be defined as follows: "mother_accession X father_accession", or, if father is unknown, as "mother_accession X UNKNOWN". For in situ material, the region of provenance may be used when an accession is not available.	INRA:W95115_inra ICNF:PNB-RPI	Unique identifier	0-1
DM-51	**Material source DOI**	Digital Object Identifier (DOI) of the material source	doi:10.15454/1.4658436467893904E12	DOI	0-1
DM-52	**Material source latitude**	Latitude of the material source. [Alternative identifier for in situ material]	+39.067	Degrees in the decimal format (ISO 6709)	0-1 (1 if longitude is provided)
DM-53	**Material source longitude**	Longitude of the material source. [Alternative identifier for in situ material]	-8.73	Degrees in the decimal format (ISO 6709)	0-1 (1 if latitude is provided)
DM-54	**Material source altitude**	Altitude of the material source, provided in metres (m). [Alternative identifier for in situ material]	10 m	Numeric + unit abbreviation	0-1
DM-55	**Material source coordinates uncertainty**	Circular uncertainty of the coordinates, provided in meters (m). [Alternative identifier for in situ material]	200 m	Numeric + unit abbreviation	0-1
DM-56	**Material source description**	Description of the material source	Branches were collected from a 10-year-old tree growing in a progeny trial established in a loamy brown earth soil.	Free text	0-1
DM-57	**Environment**	Environmental parameters that were kept constant throughout the study and did not change between observation units or assays. Environment characteristics that vary over time, i.e. environmental variables, should be recorded as Observed Variables (see below).			0-1 per study
DM-58	**Environment parameter**	Name of the environment parameter constant within the experiment.	sowing density rooting medium composition; pH	Free text (see Appendix I)	1+
DM-59	**Environment parameter value**	Value of the environment parameter (defined above) constant within the experiment.	300 seeds per m2 Clay 50% plus sand; 6.5	Free text	1 per parameter
DM-60	**Experimental Factor**	The object of a study is to ascertain the impact of one or more factors on the biological material. Thus, a factor is, by definition a condition that varies between observation units, which may be biotic (pest, disease interaction) or abiotic (treatment and cultural practice) in nature. Depending on the level of the data, an experimental factor can be either "what is the factor applied to the plant" (i.e. Unwatered), or the "environmental characterisation" (i.e. if no rain on unwatered plant : Drought ; if rain on unwatered plant: Irrigated)			0+ per study; 0+ per observation unit
DM-61	**Experimental Factor type**	Name/Acronym of the experimental factor.	Watering	Free text (see Appendix II)	1
DM-62	**Experimental Factor description**	Free text description of the experimental factor. This includes all relevant treatments planification and protocol planned for all the plants targeted by a given experimental factor.	Daily watering 1 L per plant.	Free text	0-1
DM-63	**Experimental Factor values**	List of possible values for the factor.	Watered; Unwatered	Free text	2+ per factor
DM-64	**Event**	An event is discrete occurrence at a particular time in the experiment (which can be natural, such as rain, or unnatural, such as planting, watering, etc). Events may be the realization of Factors or parts of Factors, or may be confounding to Factors. Can be applied at the whole study level or to only a subset of observation			0+ per study/observation unit
DM-65	**Event type**	Short name of the event.	Planting Fertilizing	Free text (short)	1
DM-66	**Event accession number**	Accession number of the event type in a suitable controlled vocabulary (Crop Ontology).	CO_715:0000007 CO_715:0000011	Crop Ontology term (subclass of CO_715:0000006)	0-1
DM-67	**Event description**	Description of the event, including details such as amount applied and possibly duration of the event.	Sowing using seed drill Fertilizer application: Ammonium nitrate at 3 kg/m2	Free text	0-1
DM-68	**Event date**	Date and time of the event.	2006-09-27T10:23:21+00:00 2006-10-27; 2006-11-13; 2016-11-21	Date/Time (ISO 8601, optional time zone)	1+
DM-69	**Observation Unit**	Observation units are objects that are subject to instances of observation and measurement. An observation unit comprises one or more plants, and/or their environment. There can be pure environment observation units with no plants. Synonym: Experimental unit.			1+ per study
DM-70	**Observation unit ID**	Identifier used to identify the observation unit in data files containing the values observed or measured on that unit. Must be locally unique.	plot:894	Unique identifier	1
DM-71	**Observation unit type**	Type of observation unit in textual form, usually one of the following: block, sub-block, plot, plant, study, pot, replication or replicate, individual. Use of other observation unit types is possible but not recommended. The observation unit type cannot be used to indicate sub-plant levels. However, observations can still be made on the sub-plant level, as long as the details are indicated in the associated observed variable (see observed variables). Alternatively, it is possible to use samples for more detailed tracing of sub-plant units, attaching the observations to them instead.	plot	Free text	1
DM-72	**External ID**	Identifier for the observation unit in a persistent repository, comprises the name of the repository and the identifier of the observation unit therein. The EBI Biosamples repository can be used. URI are recommended when possible.	Biosamples:SAMEA4202911	Unique identifier	0+
DM-73	**Spatial distribution**	Type and value of a spatial coordinate (georeference or relative) or level of observation (plot 45, subblock 7, block 2) provided as a key-value pair of the form type:value. Levels of observation must be consistent with those listed in the Study section.	Latitude:+2.341; row:4 ; X:3; Y:6; Xm:35; Ym:65; Block:1; Plot:894	Formatted text (Key:value)	0+
DM-74	**Observation Unit factor value**	List of values for each factor applied to the observation unit.	Watered	Free text	0+
DM-75	**Sample**	A sample is a portion of plant tissue harvested, non-harvested or extracted from an observation unit for the purpose of sub-plant observations and/or molecular studies. A sample must be used when there is a physical sample that needs to be stored and traced. Otherwise, observations made at the sub-plant level should be recorded as plant level observations using the observed variables to characterize the object of the observation (e.g. Berry sugar content, Fruit weight, Grain Protein content, Leaf 1 width, Leaf 2 width, Leaf 2 length).			0+ per observation unit
DM-76	**Sample ID**	Unique identifier for the sample.	CEA:BE00034067	Unique identifier	1
DM-77	**Plant structure development stage**	The stage in the life of a plant structure during which the sample was taken, in the form of an accession number to a suitable controlled vocabulary (Plant Ontology, BBCH scale)	PO:0025094 BBCH-17	Plant Ontology term (subclass or PO:0009012) or BBCH scale term	0-1
DM-78	**Plant anatomical entity**	A description of the plant part (e.g. leaf) or the plant product (e.g. resin) from which the sample was taken, in the form of an accession number to a suitable controlled vocabulary (Plant Ontology).	PO:0000003 PO:0025161	Plant Ontology term (subclass of PO:0025131)	1
DM-79	**Sample description**	Any information not captured by the other sample fields, including quantification, sample treatments and processing.	Distal part of the leaf ; 100 mg of roots taken from 10 roots at 20°C, conserved in vacuum at 20 mM NaCl salinity, stored at -60 °C to -85 °C.	Free text	0-1
DM-80	**Collection date**	The date and time when the sample was collected / harvested	2005-08-15T15:52:01+00:00	Date/Time	1
DM-81	**External ID**	An identifier for the sample in a persistent repository, comprising the name of the repository and the accession number of the observation unit therein. Submission to the EBI Biosamples repository is recommended. URI are recommended when possible.	Biosamples:SAMEA4202911	Unique identifier	0+
DM-82	**Observed Variable**	An observed variable describes how a measurement has been made. It typically takes the form of a measured characteristic of the observation unit (plant or environmental trait), associated to the method and unit of measurement. Multiple variables with the same combination of trait, method and scale can be used in association with different plant parts (leaf 1, leaf 2), when this distinction is necessary for observations referring to different parts of the same observation unit.			1+ per study
DM-83	**Variable ID**	Code used to identify the variable in the data file. We recommend using a variable definition from the Crop Ontology where possible. Otherwise, the Crop Ontology naming convention is recommended: <trait abbreviation>_<method abbreviation>_<scale abbreviation>). A variable ID must be unique within a given investigation.	Ant_Cmp_Cday	Unique identifier	1
DM-84	**Variable name**	Name of the variable.	Anthesis computed in growing degree days	Free text	0-1
DM-85	**Variable accession number**	Accession number of the variable in the Crop Ontology	CO_322:0000794	Crop Ontology term	0-1
DM-86	**Trait**	Name of the (plant or environmental) trait under observation	Anthesis time Reproductive growth time	Free text	1
DM-87	**Trait accession number**	Accession number of the trait in a suitable controlled vocabulary (Crop Ontology, Trait Ontology).	CO_322:0000030 TO:0000366	Term from Plant Trait Ontology, Crop Ontology, or XML Environment Ontology	0-1
DM-88	**Method**	Name of the method of observation	Growing degree days to anthesis	Free text	1
DM-89	**Method accession number**	Accession number of the method in a suitable controlled vocabulary (Crop Ontology, Trait Ontology).	CO_322:0000189	Term from Plant Trait Ontology, Crop Ontology, or XML Environment Ontology	0-1
DM-90	**Method description**	Textual description of the method, which may extend a method defined in an external reference with specific parameters, e.g. growth stage, inoculation precise organ (leaf number)	Days to anthesis for male flowering was measured in thermal time (GDD: growing degree-days) according to Ritchie J, NeSmith D (1991;Temperature and crop development. Modeling plant and soil systems American Society of Agronomy Madison, Wisconsin USA) with TBASE=8°C and T0=30°C. Plant height was measured at 5 years with a ruler, one year after Botritis inoculation.	Free text	0-1
DM-91	**Reference associated to the method**	URI/DOI of reference describing the method.	http://doi.org/10.2134/agronmonogr31.c2	URI or DOI	0-1
DM-92	**Scale**	Name of the scale associated with the variable	°C day	Unique identifier	1
DM-93	**Scale accession number**	Accession number of the scale in a suitable controlled vocabulary (Crop Ontology).	CO_322:0000510	Crop Ontology term	0-1
DM-94	**Time scale**	Name of the scale or unit of time with which observations of this type were recorded in the data file (for time series studies).	Growing degree day (GDD) Date/Time	Free text	0+