FAO/BIOVERSITY MULTI-CROP PASSPORT DESCRIPTORS V.2

June 2012

This list of Multi-crop Passport Descriptors (MCPD V.2) is a revision of the original FAO/IPGRI publication released in 2001, expanded to accommodate emerging needs, such as the broader use of GPS tools, or the implementation of the International Treaty on Plant Genetic Resources for Food and Agriculture's Multilateral System for access and benefit-sharing. This MCPD V.2 list is an expansion of the first version of the MCPD, the descriptors and allowed values of the first version form a subset of those in this revision.

The 2001 list, developed jointly by Bioversity International (formerly IPGRI) and FAO, has been widely used and is considered the international standard to facilitate germplasm passport information exchange. These descriptors aim to be compatible with Bioversity's crop descriptor lists, with the descriptors used for the FAO World Information and Early Warning System (WIEWS) on plant genetic resources (PGR), and with the GENESYS global portal.

For each multi-crop passport descriptor, a brief explanation of content, coding scheme and, in parentheses, suggested fieldname are provided to assist in the computerized exchange of this type of data. Annex I provides easy access to the historical 'List of major changes' of all descriptor elements.

It is recognized that networks or groups of users may want to further expand this revised MCPD list to meet their specific needs. As long as these additions allow for an easy conversion to the format proposed in MCPD V.2, basic passport data can be exchanged worldwide in a consistent manner.

Common formatting rules

(i) If a field allows multiple values, these values should be separated by a semicolon (;) without space (e.g. Accession name: Symphony;Emma;Songino).

(ii) A field for which no value is available should be left empty (e.g. Elevation). If data are exchanged in ASCII format, a field with a missing numeric value should be left empty. If data are exchanged in a database format, missing numeric values should be represented by generic NULL values.

(iii) Dates are recorded as YYYYMMDD. If the month or day are missing, this should be indicated with hyphens or '00' [double zero]. If both (month and day) are missing, two double zeros are needed (e.g. 1975----, 19750000; 197506--, 19750600).

(iv) Country names: Three letter ISO codes are used for countries. The ISO 3166-1: Code List and the Country or the Country or area numerical codes added or changed are available online at: http://unstats.un.org/unsd/methods/m49/m49alpha.htm.

Note: The list of obsolete codes can be found at: http://en.wikipedia.org/wiki/ISO_3166-1_alpha-3#Reserved_code_elements.

(v) For institutes, the codes from FAO WIEWS should be used. The current set of Institute Codes is available from the FAO Web site (http://apps3.fao.org/wiews/wiews.jsp). From the Main Menu select: 'PGR' and 'Download'. Individual institutes can be searched for at http://apps3.fao.org/wiews/institute_query.htm?i_l=EN.

- If new Institute codes are required, they can be generated on line by national FAO WIEWS administrators (WIEWS@fao.org) or it can be requested by filling the form at: http://apps3.fao.org/wiews/newuserdialog.jsp.

- For institutes that no longer exist, please provide full details in descriptors 4.1, 4.1.1, 18.1, 22.1 and 25.1.

MULTI-CROP PASSPORT DESCRIPTORS

1. Institute code	**(INSTCODE)**

FAO WIEWS code of the institute where the accession is maintained. The codes consist of the 3-letter ISO 3166 country code of the country where the institute is located plus a number (e.g. COL001). The current set of institute codes is available from http://apps3.fao.org/wiews/wiews.jsp. For those institutes not yet having an FAO Code, or for those with 'obsolete' codes, see '*Common formatting rules (v)*'.

2. Accession number	**(ACCENUMB)**

This is the unique identifier for accessions within a genebank, and is assigned when a sample is entered into the genebank collection (e.g. 'PI 113869').

3. Collecting number	**(COLLNUMB)**

Original identifier assigned by the collector(s) of the sample, normally composed of the name or initials of the collector(s) followed by a number (e.g. 'FM9909'). This identifier is essential for identifying duplicates held in different collections.

4. Collecting institute code	**(COLLCODE)**

FAO WIEWS code of the institute collecting the sample. If the holding institute has collected the material, the collecting institute code (COLLCODE) should be the same as the holding institute code (INSTCODE). Follows INSTCODE standard. Multiple values are separated by a semicolon without space.

4.1 Collecting institute name	**(COLLNAME)**

Name of the institute collecting the sample. This descriptor should be used only if COLLCODE cannot be filled because the FAO WIEWS code for this institute is not available. Multiple values are separated by a semicolon without space

4.1.1 Collecting institute address	**(COLLINSTADDRESS)**

Address of the institute collecting the sample. This descriptor should be used only if COLLCODE cannot be filled since the FAO WIEWS code for this institute is not available. Multiple values are separated by a semicolon without space.

4.2 Collecting mission identifier	(COLLMISSID)

Identifier of the collecting mission used by the Collecting Institute (4 or 4.1) (e.g. 'CIATFOR-052', 'CN426').

5. Genus	(GENUS)

Genus name for taxon. Initial uppercase letter required.

6. Species	(SPECIES)

Specific epithet portion of the scientific name in lowercase letters. Only the following abbreviation is allowed: 'sp.'

7. Species authority	(SPAUTHOR)

Provide the authority for the species name.

8. Subtaxon	(SUBTAXA)

Subtaxon can be used to store any additional taxonomic identifier. The following abbreviations are allowed: 'subsp.' (for subspecies); 'convar.' (for convariety); 'var.' (for variety); 'f.' (for form); 'Group' (for 'cultivar group').

9. Subtaxon authority	(SUBTAUTHOR)

Provide the subtaxon authority at the most detailed taxonomic level.

10. Common crop name	(CROPNAME)

Common name of the crop. Example: 'malting barley', 'macadamia', 'maïs'.

11. Accession name	(ACCENAME)

Either a registered or other designation given to the material received, other than the donor's accession number (23) or collecting number (3). First letter uppercase. Multiple names are separated by a semicolon without space. Example: Accession name: Bogatyr;Symphony;Emma.

12. Acquisition date [YYYYMMDD]	(ACQDATE)

Date on which the accession entered the collection where YYYY is the year, MM is the month and DD is the day. Missing data (MM or DD) should be indicated with hyphens or '00' [double zero].

13. Country of origin	(ORIGCTY)

3-letter ISO 3166-1 code of the country in which the sample was originally collected (e.g. landrace, crop wild relative, farmers' variety), bred or selected (breeding lines, GMOs, segregating populations, hybrids, modern cultivars, etc.).

Note: Descriptors 14 to 16 below should be completed accordingly only if it was 'collected'.

14. Location of collecting site	(COLLSITE)

Location information below the country level that describes where the accession was collected, preferable in English. This might include the distance in kilometres and direction from the nearest town, village or map grid reference point, (e.g. 7 km south of Curitiba in the state of Parana).

15. Geographical coordinates

- For latitude and longitude descriptors, two alternative formats are proposed, but the one reported by the collecting mission should be used
- Latitude and longitude in decimal degree format with a precision of four decimal places corresponds to approximately 10 m at the Equator and describes the point-radius representation of the location, along with Geodetic datum and Coordinate uncertainty in metres.

Note: The following two mutually exclusive formats can be used for latitude:

15.1 Latitude of collecting site (Decimal degrees format)	**(DECLATITUDE)**
Latitude expressed in decimal degrees. Positive values are North of the Equator; negative values are South of the Equator (e.g. -44.6975).	
15.2 Latitude of collecting site (Degrees, Minutes, Seconds format)	**(LATITUDE)**
Degrees (2 digits) minutes (2 digits), and seconds (2 digits) followed by N (North) or S (South) (e.g. 103020S). Every missing digit (minutes or seconds) should be indicated with a hyphen. Leading zeros are required (e.g. 10----S; 011530N; 4531--S).	

Note: The following two mutually exclusive formats can be used for longitude:

15.3 Longitude of collecting site (Decimal degrees format)	**(DECLONGITUDE)**
Longitude expressed in decimal degrees. Positive values are East of the Greenwich Meridian; negative values are West of the Greenwich Meridian (e.g. +120.9123).	
15.4 Longitude of collecting site (Degrees, Minutes, Seconds format)	**(LONGITUDE)**
Degrees (3 digits), minutes (2 digits), and seconds (2 digits) followed by E (East) or W (West) (e.g. 0762510W). Every missing digit (minutes or seconds) should be indicated with a hyphen. Leading zeros are required (e.g. 076----W).	
15.5 Coordinate uncertainty [m]	**(COORDUNCERT)**
Uncertainty associated with the coordinates in metres. Leave the value empty if the uncertainty is unknown.	
15.6 Coordinate datum	**(COORDDATUM)**
The geodetic datum or spatial reference system upon which the coordinates given in decimal latitude and decimal longitude are based (e.g. WGS84, ETRS89, NAD83). The GPS uses the WGS84 datum.	
15.7 Georeferencing method	**(GEOREFMETH)**
The georeferencing method used (GPS, determined from map, gazetteer, or estimated using software). Leave the value empty if georeferencing method is not known.	
16. Elevation of collecting site [masl]	**(ELEVATION)**
Elevation of collecting site expressed in metres above sea level. Negative values are allowed.	
17. Collecting date of sample [YYYYMMDD]	**(COLLDATE)**
Collecting date of the sample, where YYYY is the year, MM is the month and DD is the day. Missing data (MM or DD) should be indicated with hyphens or '00' [double zero].	
18. Breeding institute code	**(BREDCODE)**
FAO WIEWS code of the institute that has bred the material. If the holding institute has bred the material, the breeding institute code (BREDCODE) should be the same as the holding institute code (INSTCODE). Follows INSTCODE standard. Multiple values are separated by a semicolon without space.	
18.1 Breeding institute name	**(BREDNAME)**
Name of the institute (or person) that bred the material. This descriptor should be used only if BREDCODE cannot be filled because the FAO WIEWS code for this institute is not available. Multiple names are separated by a semicolon without space.	

19. Biological status of accession	(SAMPSTAT)

The coding scheme proposed can be used at 3 different levels of detail: either by using the general codes (in boldface) such as 100, 200, 300, 400, or by using the more specific codes such as 110, 120, etc.

100) Wild

    110)   Natural

    120)   Semi-natural/wild

    130)   Semi-natural/sown

200) Weedy

300) Traditional cultivar/landrace

400) Breeding/research material

    410)   Breeder's line

    411)   Synthetic population

    412)   Hybrid

    413)   Founder stock/base population

    414)   Inbred line (parent of hybrid cultivar)

    415)   Segregating population

    416)   Clonal selection

    420)   Genetic stock

    421)   Mutant (e.g. induced/insertion mutants, tilling populations)

    422)   Cytogenetic stocks (e.g. chromosome addition/substitution, aneuploids, amphiploids)

    423)   Other genetic stocks (e.g. mapping populations)

500) Advanced or improved cultivar (conventional breeding methods)

600) GMO (by genetic engineering)

999) Other (Elaborate in REMARKS field)

20. Ancestral data	(ANCEST)

Information about either pedigree or other description of ancestral information (e.g. parent variety in case of mutant or selection). For example a pedigree 'Hanna/7*Atlas//Turk/8*Atlas' or a description 'mutation found in Hanna', 'selection from Irene' or 'cross involving amongst others Hanna and Irene'.

21.	Collecting/acquisition source	(COLLSRC)

The coding scheme proposed can be used at 2 different levels of detail: either by using the general codes (in boldface) such as 10, 20, 30, 40, etc., or by using the more specific codes, such as 11, 12, etc.

- **10)** **Wild habitat**
  - 11) Forest or woodland
  - 12) Shrubland
  - 13) Grassland
  - 14) Desert or tundra
  - 15) Aquatic habitat
- **20)** **Farm or cultivated habitat**
  - 21) Field
  - 22) Orchard
  - 23) Backyard, kitchen or home garden (urban, peri-urban or rural)
  - 24) Fallow land
  - 25) Pasture
  - 26) Farm store
  - 27) Threshing floor
  - 28) Park
- **30)** **Market or shop**
- **40)** **Institute, Experimental station, Research organization, Genebank**
- **50)** **Seed company**
- **60)** **Weedy, disturbed or ruderal habitat**
  - 61) Roadside
  - 62) Field margin
- **99)** **Other** (Elaborate in REMARKS field)

22. Donor institute code	(DONORCODE)

FAO WIEWS code of the donor institute. Follows INSTCODE standard

22.1 Donor institute name	(DONORNAME)

Name of the donor institute (or person). This descriptor should be used only if DONORCODE cannot be filled because the FAO WIEWS code for this institute is not available.

23. Donor accession number	(DONORNUMB)

Identifier assigned to an accession by the donor. Follows ACCENUMB standard.

24. Other identifiers associated with the accession	(OTHERNUMB)

Any other identifiers known to exist in other collections for this accession. Use the following format: INSTCODE:ACCENUMB;INSTCODE:identifier;… INSTCODE and identifier are separated by a colon without space. Pairs of INSTCODE and identifier are separated by a semicolon without space. When the institute is not known, the identifier should be preceded by a colon.

25. Location of safety duplicates	(DUPLSITE)

FAO WIEWS code of the institute(s) where a safety duplicate of the accession is maintained. Multiple values are separated by a semicolon without space. Follows INSTCODE standard.

25.1 Institute maintaining safety duplicates	**(DUPLINSTNAME)**

Name of the institute where a safety duplicate of the accession is maintained. Multiple values are separated by a semicolon without space.

26. Type of germplasm storage	**(STORAGE)**

If germplasm is maintained under different types of storage, multiple choices are allowed, separated by a semicolon (e.g. 20;30). (Refer to FAO/IPGRI Genebank Standards 1994 for details on storage type.)

- **10) Seed collection**
  - 11) Short term
  - 12) Medium term
  - 13) Long term
- **20) Field collection**
- **30) *In vitro* collection**
- **40) Cryopreserved collection**
- **50) DNA collection**
- **99) Other** (elaborate in REMARKS field)

27. MLS status of the accession	**(MLSSTAT)**

The status of an accession with regards to the Multilateral System (MLS) of the International Treaty on Plant Genetic Resources for Food and Agriculture. Leave the value empty if the status is not known

- 0   No (not included)
- 1   Yes (included)
- 99   Other (elaborate in REMARKS field, e.g. 'under development')

28. Remarks	**(REMARKS)**

The remarks field is used to add notes or to elaborate on descriptors with value 99 or 999 (= Other). Prefix remarks with the field name they refer to and a colon (:) without space (e.g. COLLSRC:riverside). Distinct remarks referring to different fields are separated by semicolons without space.

List of major changes in MCPD V.2 (2012)

The revision process to update the FAO/IPGRI List of Multi-Crop Passport Descriptors (MCPD, 2001), has involved consultation with more than 300 people from 187 institutions in 87 countries.

The review process indicated that: (a) there were no major concerns with the 2001 version, thereby supporting its validity; (b) some additional flexibility with existing descriptors was considered necessary; and (c) a few new descriptors could enhance the MCPD. All these amendments are listed below.

Previous versions are available upon request from Adriana Alercia at Bioversity International <mailto:a.alercia@cgiar.org>.

List of major changes in MCPD V.2 (2012) compared with the 2001 version:

Title

It now reads FAO/Bioversity (formerly FAO/IPGRI) to reflect the name change to Bioversity International.

Introduction

Text related to this updated version was amended reflecting main changes, including addition of this historical list of changes, as an Annex.

Common formatting rules
- Explanation of coding schemes related to new descriptors is provided to assist in the computerized exchange of data.
- **Country names**: New ISO Country codes Web link is provided, along with a Web link for obsolete codes (e.g. SUN, DDR).
- **FAO WIEWS Institute codes:** The Web link for the current set of institute codes has been updated. A link is provided to an electronic form for generating new institute codes online.
- **Missing data:** For those descriptors related to 'Dates', missing data can be indicated with hyphens or '00'[double zero] to ensure consistent data transfer whatever system is used.
- **Latitude and longitude:** Two formats (Latitude: DDMMSSH or +/-DD.DDDD; Longitude: DDDMMSSE/W or +/-DDD.DDDD) are allowed, but only one of the two coordinate formats should be used (the one that was reported by the collecting mission).

Modifications

NOTE: New descriptors added are underlined.

1.   Institute code
Relevant FAO Web link updated
New link to request FAO institute code form provided

2.   Accession number
'Number' replaced by 'identifier'
Deleted 'collection'

3.   Collecting number
'Number' replaced by 'identifier' within the text, since it is not necessarily a number

4.   Collecting institute code
Multiple values allowed

4.1   Collecting institute name
New descriptor added

4.1.1 Collecting institute address
New descriptor added

4.2   Collecting mission identifier
New descriptor added

8.   Subtaxon
Added 'Group' for cultivar groups

10.   Common crop name
Deleted 'preferable in English' since not all crops and wild relatives have one in English

11.   Accession name
Description text has been refined

12.   Acquisition date [YYYYMMDD]
Missing data can also be indicated with '00' (double zero as alternative to hyphens)

13.   Country of origin
Text has been modified with better definition

15. Geographical coordinates

Conversion formula for latitude and longitude has been deleted

The following new fields have been added:

    15.1 Latitude of collecting site (Decimal degrees format)

    15.3 Longitude of collecting site (Decimal degrees format)

    15.5 Coordinate uncertainty [m]

    15.6 Coordinate datum

    15.7 Georeferencing method

(Old 18)   **17.   Collecting date of sample** [YYYYMMDD]

Missing data can also be indicated with '00' [double zero]

(Old 19)   **18.   Breeding institute code**

Multiple values allowed

18.1 Breeding institute name

New field added

(Old 20)   **19.   Biological status of accession**

New state added: '130 = Semi-natural or sown'

Refined original state: '420 = Mutant or genetic stock', with better definition

New descriptor state added: '600 = GMO (by genetic engineering)'

22.1 Donor institute name

New descriptor added

(Old 24)   **23.   Donor accession number**

'Number' replaced by 'identifier'

(Old 25)   **24.   Other identifiers associated with the accession**

Title and text modified, replacing 'numbers' by 'identifiers'

(Old 26)   **25.   Location of safety duplicates**

Multiple values allowed

25.1 Institute maintaining safety duplicates

New descriptor added

(Old 27)   **26.   Type of germplasm storage**

New descriptor state added: '50 = DNA collection'

27.   MLS status of the accession

New descriptor added

28.   Remarks

Changed example

OPEN QUESTION

Global unique identifier

A global unique identifier or persistent identifier (GUID) such as the Digital Object Identifier (DOI), or the Life Science Identifier (LSID), would be useful to build automatic services that integrate germplasm data and would enable the linkage required to identify accessions and other genotype entities across different information systems.

The topic was raised by a few respondents in the survey for this revision, but further international consultation would be needed to provide a consensus.