Appendix I. The identification of most original samples (MOS)

Provisional identification

Provisional identification of the **original status** (i.e. bred or collected) and the **originality** (i.e. how close they are to being MOSs) of accessions is based on values in the four fields containing institute codes:

GBK	Code of institute holding the accession
DON	Code of donor's institute
COL	Code of collector's institute
BRE	Code of breeder's institute

Some examples of what can be deduced from contents of fields for holding institute, donor institute, collector institute and breeder institute are given in the following table, in which "A", "B" and "C" indicate an institute code, and a blank value under the institute indicates no data are available.

GBK	DON	COL	BRE	Comment
A				Incomplete. From where did GBK get the accession?
A		B		Incomplete. The original sample (OS) was collected, but from where did GBK get the accession?
A			B	Incomplete. The OS was bred, but from where did GBK get the accession?
A		any	any	Incomplete/invalid (could be valid if cultivar was collected from farmer's field, but not a good MOS-candidate).
A	A			Incomplete/invalid. (ignore DON because GBK cannot be same as DON).
A	B			Incomplete. Who collected or bred the OS?
A		A		Good MOS candidate. Provisionally designate the accession as MOS unless other data indicate another accession is genetically closer to the OS.
A			A	Good MOS candidate.
A	B		B	1 donation away from MOS ("one away"), but probably the best MOS-candidate if DON no longer has the material.
A	B	B		1 donation away from MOS ("more away"), but probably the best MOS-candidate if DON no longer has the material.
A	B	C		2 or more donations away from MOS ("more away"). Check DON and COL for more information.
A	B		C	2 or more donations away from MOS ("more away"). Check DON and BRE for more information.
A	B	A		Repatriated from B to A, but B's sample may be better MOS candidate. Check whether A also still retains the original.
A	B		A	Repatriated from B to A, but B's sample may be better MOS candidate. Check whether A also still retains the original.

These conditions may be summarized as follows.

A. Original status

		Breeder's institute	
		Blank	Non-blank
Collector's institute	Blank	Unknown original status	Collected accession
	Non-blank	Bred accession	Invalid data

B. Originality

Comparison of the values of GBK and DON with those of BRE and COL provisionally identifies the originality of an accession.

Special cases have to be treated first. Missing values for GBK are not allowed. It is meaningless to maintain data on an accession whose location is not known. DON should never be the same as GBK, since an institute does not donate accessions to itself. However, this error may safely be ignored by treating accessions where DON=GBK as if they had no value for DON.

If GBK is the same as either COL or BRE, then the institute is, or at least was, the holder of the Original Sample (OS). Two cases must be considered depending on the value of DON.

If DON has no value (including the invalid case DON=GBK), then the accession is provisionally classified as the MOS. This classification may be changed at a later date if it is subsequently determined that other samples with the same origin have been conserved to a higher standard, such that they are genetically more similar to the OS.

If, however, DON has a value different from GBK, the implication is that a sample of the OS must have been donated to another institute, and subsequently (directly or indirectly) donated back to the institute. Thus the accession is probably not itself an MOS, but the MOS of that accession is another accession in the same collection. The originality of the accession is provisionally classified as "with MOS".

IF GBK is not the same as COL or BRE, we must consider three cases depending on the value of DON.

- If DON has no value (including the invalid case DON=GBK), then there is no information (at least within the institute codes) on the originality of the accession. It is provisionally classified as "unknown".
- If DON has a value, and that value equals COL or BRE, then the donor's institute holds, or at least used to hold, the OS of the accession. That is, there is only one cycle of donation between the accession and its OS. It is classified as "one away".
- If DON has a value and that value is not the same as COL or BRE, then the donor's institute does not hold, and never has held, the OS of the sample. There are two or more cycles of donation between the accession and its OS. It is classified as "more away".

This algorithm may be specified more succinctly in pseudo-pascal form as follows:

```
if GBK=COL or GBK=BRE then
    if DON=null or DON=GBK then
        originality := "MOS"
    else
        originality := "with MOS"
else if DON=null or DON=GBK then
    originality := "unknown"
else if DON=COL or DON=BRE then
    originality := "one away"
else
    originality := "more away"
```

Problems

Application of the above algorithms is necessarily only provisional, for the following reasons:

- First, failure to standardize institute codes adequately will result in different codes being used for the same institute, in which case the above algorithm would generate incorrect results.
- Second, other passport data also indicate whether an accession was collected or bred. There is a need to use the other passport data to check for inconsistencies with the classification by BRE and COL, and also to assign status to accessions without values for BRE or COL.
- Third, because of variation in standards of conservation and regeneration, a donated accession may be genetically closer to the OS than the accession currently held by the institute that originally held the OS. In particular, if the OS holder failed to keep a subsample of the OS purely for conservation, and if several cycles of regeneration have been undertaken, the current sample may be genetically distinct from the OS. Therefore, when two or more accessions with the same origin are identified, their histories must be compared to confirm which sample should be the MOS.
- Fourth, special attention must be paid to the collection of advanced cultivars from the field. A value may be entered for the breeder's institute as well as the collector's institute. Such cases should not be treated as values.
- Fifth, the protocol does not take into account joint collecting expeditions, which generate two questions. How should such duplicate MOSs be identified? Which institute should be designated the primary MOS holder? The standards adopted for entering data for such collections vary between institutes. Codes for both institutes may be entered in COL, separated by a ";". The algorithm as specified above does not allow for this situation but is easily modified to do so. Alternatively, the two collaborating institutes may enter different data in their corresponding documentation systems.

The recommended procedure

To overcome the above problems, a four-step procedure is recommended as follows.

Step 1

1a   ECCDB manager provisionally classifies accessions by their original status and originality following the algorithm above.

1b   Notify curators of the results and ask them to comment on the accuracy of the classification. Particularly highlight provisional collected MOSs where the country of origin is not the country of the collecting institute, and request clarification from the curator regarding possible joint collecting expeditions. In the case of ECP/GR countries with appropriate facilities, it is preferred to assign the MOS to an institute in the country of origin.

1c   Response from curators.

Step 2

2a   Unification of institute codes. Ideally codes should follow internationally agreed (FAO) system, but this is not essential. If a universal standard is not forthcoming, each ECCDB manager may proceed by establishing a personal system in agreement with curators. ECCDB managers will tabulate all codes used in the DB, provisionally identify all codes in use for each institute, identify unique codes, and ask curators for a response on the accuracy of the classification of institute codes. Curators may then update their own contributions to the ECCDB, or the ECCDB manager may apply a translation table.

2b   Using the revised institute codes, the ECCDB managers re-classify accessions.

2c   Notify curators of the results and ask them to comment on the accuracy of the revised classification.

2d   Response from curators.

Step 3   For accessions with a value for DON (other than =GBK), the ECCDB manager traces the donation history, seeking to identify the original accession from which the accession was derived.

Step 4   Network Coordinating Group to discuss how to sort out remaining uncertainties establish
            - algorithms for identifying and resolving internal data inconsistencies,
            - alternative algorithms for identifying origins,
            - procedures for comparing the conservation history of accessions with the same origin, and therefore their probable genetic similarity to the OS.