This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement N° 773139

Grant agreement N. 773139

DELIVERABLE N° 3.1 – V2

Title: List of criteria the reference materials have to meet for use in validation studies

Validation of diagnostic tests to support plant health

Due date:	Month 8
Actual submission date	01-03-2019 (Month 11)
Start date of the project	01-05-2018
Deliverable lead contractor (organization name)	NIB, WR
Participants (Partners short names)	WR, EPPO, WBF, UNITO, ANSES, NVWA, IPADLAB, SEDIAG, NIB (Task Leader)
Author(s) in alphabetical order	Chappé, A.-M. (ANSES), Chabirand, A. (ANSES), Dahlin, P. (Agroscope), de Krom, C. (NVWA), Dreo, T. (NIB, Task leader), Gentit, P. (ANSES), Laurenson, L. (FERA), Petter, F. (EPPO), Spadaro, D. (Unito), van der Vlugt, R. (WR, WP leader), van Veen, E. (NVWA), Westenberg, M. (NVWA).
Contact for queries	Tanja.dreo@nib.si; rene.vandervlugt@wur.nl
Level of dissemination	Public
Type of deliverable	Report

Abstract:
A list of general minimum criteria was developed to be used in preparation of reference materials to be used in interlaboratory testing, including validations through test performance studies. The list was based on previously identified criteria in standards including ISO standards, EPPO guidelines, deliverables of relevant projects, related fields, and own experience of project partners. Several additional criteria were proposed: 'availability', 'purity' and 'commutability'. These criteria result from a series of discussions, taking into account previous work and relevant international standards. Where relevant, each criterion was first defined as a series of levels from the highest to lowest ranking with the lowest ranking considered to be the minimum. Depending on the intended use, the reference material may need to fulfil higher levels of selected criteria or some criteria may not be relevant at all. The criteria were tested by the organizers of the TPS in round 1 of validation within VALITEST project which include bacteria, viruses, nematodes and fungi. The systematic and structured approach of describing RMs was found to be useful in promoting transparent descriptions of the RMs used and comparability of TPSs, as well as a step toward implementing FAIR data principles. The criteria defined here are applicable, potentially with minor modification, beyond the scope of this deliverable and project to other pests and other uses of RMs.

Partners involved: WR, UNITO, WBF, EPPO, NIB, NVWA, ANSES, IPADLAB, SEDIAG

HISTORY OF CHANGES		
Version	Publication date	Change
1.0	01 March 2019	Initial version
2.0	11 February 2020	Following the review of the project, correction of the spelling mistakes.

Contents

List of Figures

List of Tables

1 Purpose

The purpose of this deliverable is to report on the development of a minimum list of criteria for reference materials as relevant to the WP1 of the project i.e. organisation of a test performance study (TPS) addressing the following task: "for a maximum of 6 of the selected pest/test/matrix combinations included in the first test performance study (TPS) round of WP1 and covering as much as possible the different pest/test/matrix combinations, the criteria for the reference materials to be used in the validation studies will be identified for different types/classes of reference materials e.g. their homogeneity, stability, purity and commutability, the extent to which they need to resemble actual samples. The associated acceptance values for the criteria will be recommended reflecting the intended purpose of the tests".

2 Terms, abbreviations and definitions

Certified reference material (CRM): Reference material derived from a source that certifies the authenticity of the material. Preferably material should come from an internationally recognised source such as a national reference collection. The material should have a unique identification code allowing traceability and the name of the person who certifies its authenticity. Details of how the material was authenticated should also be supplied. If appropriate, information about its activity (e.g. pathogenicity, antigenic properties) under specified conditions should also be supplied along with any related uncertainty at a stated level of confidence (EPPO PM 7/76 (5)).

Commutability: a characteristic describing the extent to which they resemble actual samples. A reference material would be considered commutable when a measurement produces the same result as it does for an authentic sample that contained the same analyte concentration. Because the term commutability is used in different fields various definitions are available (see e.g. review by Vesper *et al*., 2007).

Interlaboratory tests or **interlaboratory comparisons** include both test performance studies and proficiency tests (EPPO PM 7/122 (1)).

NACs: nucleic acids which, both DNA and RNA.

Optical density (OD): optical density, a measurement which can be used to indirectly determine concentrations of bacteria in a suspension.

Proficiency test (PT): Evaluation of participant performance against pre-established criteria by means of interlaboratory comparisons.

Reference material (RM): Material appropriate to the test and diagnosis being performed such as live cultures, infected plant material, DNA/RNA preparations, images of a diagnostic quality or mounted specimens. The reference material used should be documented and appropriate for the test and diagnosis being performed. It should be ensured that it has the features for which it was selected, for example expressing a desired antigen for use in serological diagnosis or display specific physical features (e.g. sporulation) if used for morphological diagnosis (EPPO PM 7/76 (5)). Further definitions are provided by international standards and are included in the sections of this document where relevant.

Test performance study (TPS): Evaluation of the performance of one or more tests by two or more laboratories using defined samples (evaluation of a test). A TPS is also referred to as ring tests or collaborative trials (EPPO PM 7/76 (5)). Test performance study is part of validation studies and usually follows in-house validation of tests.

3 Methodology

The task related to this deliverable was to prepare a list of criteria for the production of reference materials to be used in interlaboratory studies. A series of guided discussions were held in the form of teleconferences to further define the task steps and identify the gaps in the current resources/documents. The methodology approach included sourcing data and information from several key references including international standards (ISO) and guidelines (EPPO), deliverables of most relevant previous projects (e.g. Q-Collect, VirusCollect – Roenhorst *et al*., 2017), other sources (e.g. databases listing minimum metadata to accompany samples like 'minimum metadata in metagenomics') and own laboratory experience with preparation of positive controls and test items (samples) for validations and test performance studies (Figure 1). In the first steps a list of potential descriptors of reference materials was created with the idea of structuring and systematizing the way we describe it. The individual descriptors were then selected and grouped into broader descriptors and different levels identified for each (Descriptors for reference material).

Among the levels the minimum level identified was considered as a required criterion. However, as described further, the specific levels of criteria depend on the scope of use and the purpose of the validations/TPS for which the material required was prepared.

As a final step the criteria were discussed with the organizers of the test performance studies, round 1, undertaken within the WP1 of the VALITEST project. Through this the feedback on the criteria and applicability of the list was collected. The deliverable text summarizes the discussions and their conclusions.

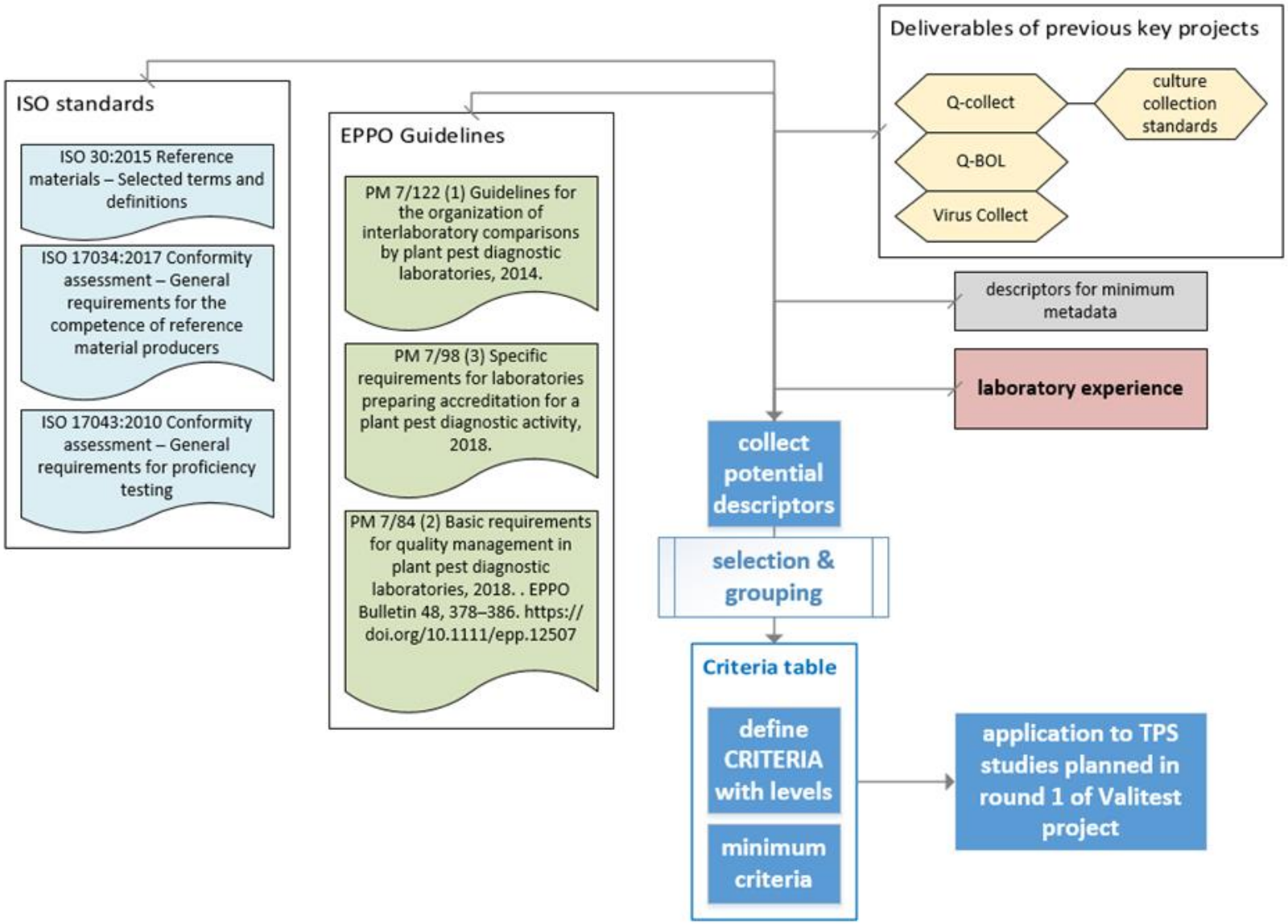

Figure 1: Methodology of task with key data and information sources. TPS = test performance study.

4 Reference materials – state of the art in plant health

Reference material (RM), as defined by the ISO is any material, sufficiently homogeneous and stable with respect to one or more specified properties, which has been established to be fit for its intended use in a measurement process (ISO/Guide 30:2015). The term 'reference material' is a generic term, of which properties can be quantitative or qualitative. Uses may include the calibration of a measurement system, assessment of a measurement procedure, assigning values to other materials, and quality control. Correspondingly, a certified reference material (CRM) is defined as being characterized by a metrological valid procedure for one or more specified properties, accompanied by a certificate that provides the value of the specified property, its associated uncertainty, and a statement of metrological traceability. The international vocabulary of metrology (VIM) provides the following definition of the reference material: "material, sufficiently homogeneous and stable with reference to specified properties, which has been established to be fit for its intended use in measurement or in examination of nominal properties" (Anonymous, 2008).

The EPPO guidelines which for many years serve as a practical interpretation of international standards specifically for the plant health field, lay the foundations and provide the following **definition of reference material** in the **EPPO PM 7/76 (5)**: **material appropriate to the test and diagnosis** being performed such as live cultures, infested plant material, DNA/RNA preparations, images of a diagnostic quality or mounted specimens. The reference material used should be documented and appropriate to the test and diagnosis being performed. It should be ensured that the material used is producing the features for which it was selected, for example expressing a desired antigen for use in serological diagnosis, or display specific physical features (e.g. sporulation) if used for morphological diagnosis.

*Reference materials **provide essential traceability in testing** and are used, for example (i) for detection and identification, (ii) to demonstrate the accuracy of results, (iii) to calibrate or verify equipment, (iv) to monitor laboratory performance, (v) to validate or verify tests and (vi) to enable comparison of tests (EPPO PM 7/84 (2)). The production of materials used for these purposes and similar activities in the field of plant health remains for the largest part limited to in-house production for own purposes by the testing laboratories themselves or by the companies providing positive/negative controls as part of their kits.*

Generally, **it is perceived that a reference material is more reliable when it is sourced from an entity which certifies the authenticity of the material**. Examples include international culture collections of harmful organisms which work according to commonly agreed standards to ensure the authenticity and quality of material. Examples of agreed upon standards include guidelines developed by the World Federation of Culture Collections and recommendations developed within the Q-Collect project. Such collections are an invaluable source of material which can be used further to prepare defined reference material.

It is worth noting that collections are well established for culturable and non-fastidious bacteria. Some fastidious bacteria already present a challenge as demonstrated e.g. with *Xylella fastidiosa*, a slow growing and fastidious bacteria of which strains are repeatedly lost from collections. There are no established collections of unculturable bacteria presenting a serious bottleneck for diagnostics and research. Collections of viruses are fewer, none of them are complete and not all viruses can be maintained or stored stably for prolonged periods of time. There are a few large international collections of fungi which can provide material. On the other hand, phytoplasmas need to be maintained in living plants and therefore such collections face more challenges in providing material on demand. There are several collections of nematodes.

The EU-funded project **Q-collect** (KBBE-CALL 7-612712, http://www.q-collect.eu/) identified significant differences in the number of specimens available in collections of other pests. To that end Q-collect developed **minimum quality guidelines for EU reference collections** of quarantine plant pests and invasive plants. However, at the moment it is not clear to what extent the collections in various fields have adopted these guidelines. Due to the sheer amount and

pioneering work of projects like Q-Collect and Q-BOL (EU Grant Agreement 226482), dealing with reliable sources of positive material and focusing on culture collection **the term 'reference material' was sometimes used interchangeably with pure cultures or specimens obtained from reference collections**.

Overall, there is **no agreement on what kind of material can indeed be called reference material in the plant health field**. It is perhaps a credit to the laboratories to avoid using the term altogether to avoid confusion even when the material is thoroughly characterized and prepared with the purposes to be used in activities requiring reference material. During discussions and **for the purpose of the task of defining a list of criteria for reference materials, the term is considered to encompass also the subsequent steps of preparing material** for e.g. test performance studies (TPS) or validations during which often **a reference culture collection material is mixed with a matrix to produced test items**. The relationship of the reference culture collection materials and the test items (samples) is schematically shown in Figure 2. The process itself can involve many sequential steps requiring expertise, traceability and quality control with the initial source of the target pest (bacterium) being only the first step. Similar test items are prepared for validation studies.

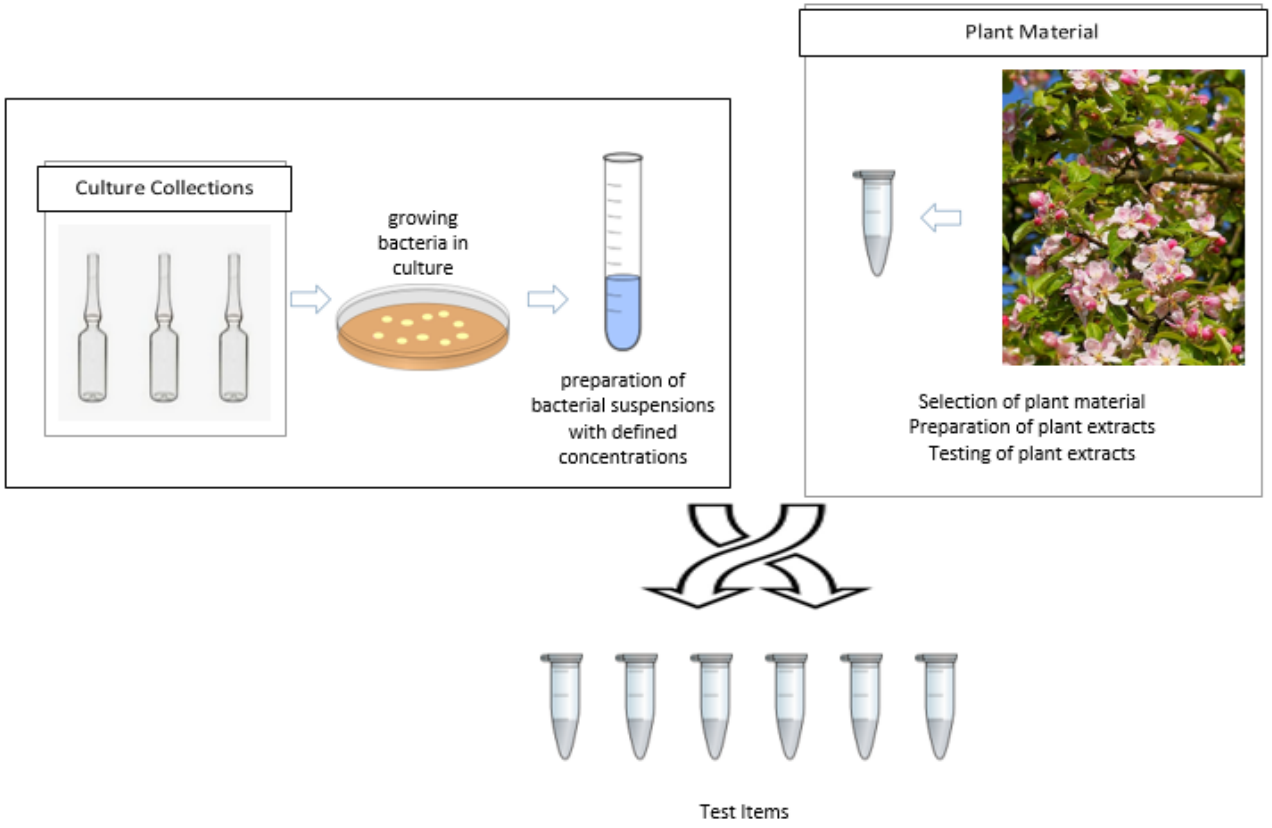

Figure 2: A schematic representation of a typical process of preparing test items (samples) for an interlaboratory study in bacteriology. Such material could be considered 'reference material'. In this example the test items are a mixture of defined bacterial suspensions and plant extracts prepared from plant material. The source of bacterial culture can be an international/reference culture collection.

5 Descriptors for reference material (RM) for validations and TPSs

A defined set of descriptors is a basis for efficient manipulation of data coming from e.g. interlaboratory studies or in-house validations. They allow for a more systematic collecting of data, improve data comparability and can be used in constructing databases. Systematic use of descriptors and their implementation also contributes to data (results) being

findable, accessible, interoperable and re-usable (characteristics commonly known as FAIR data principles; Wilkinson *et al*., 2016).

The descriptors and their levels, ranked from the highest quality to the minimum required quality as identified are described below. Together the list provides a way on how to describe the reference material improving the comparability of information on its intended use, preparation and quality control.

Previously identified descriptors and work was taken into account building on the results of Q-Collect, Q-BOL and VirusCollect projects which mainly touched on the first stage of preparation of reference materials. Also, EPPO guidelines and international standards were considered.

5.1 Scope/intended use

Scope or the intended use is a vital part of defining criteria for the reference materials and should be defined prior to its preparation. An example of the scope suitable for an activity in WP1 of the project would be: 'Preparation of reference materials for the validation/TPS of detection methods for *Erwinia amylovora* in symptomatic plant material'.

Scope may differ considerably among studies. The requirements would consequently be quite different for a study aiming to determine analytical sensitivity of a test versus a study aiming to solely address analytical specificity with a number of bacterial isolates. The intended use thus predefines the type of material suitable (see also further under Commutability).

5.2 Availability

The use of biological material is governed by conditions upon which it was obtained defined by e.g. Nagoya protocol requirements, material transfer agreements and other. While challenging in cases where data for the specimens is incomplete (particularly for older specimens) and hinders obtaining permissions, the reference material producers are advised to comply with the necessary permissions.

The source material e.g. culture collection specimens used to prepare reference material should be as widely available as possible to permit its use by the community and over time. This would suggest that material from more established collections, usually maintained in several of such collections, is preferable to material originating from a working collection.

5.3 Identity

The reference material should be clearly identified and characterized at least to the extent ensuring its correct identification. If the specimen used originates from and is available in several reference collections working according to commonly agreed quality standards this provides additional guarantee into its quality.

At minimum, the material should be thoroughly identified following accepted diagnostic protocols (when available) to ensure it is properly identified. The list of tests used for its identification should be clear from its description.

5.4 Traceability

Traceability can be considered as an aspect of both identity and availability in the sense that it may provide some additional guarantees to the correct identity of the materials used to prepare the reference material and its future availability. Also, it may be more widely available that e.g. material only available in working culture collections or material sourced at one time from natural environment.

However, it is to be noted that many materials may only be available in working collections or sourced from research environment when the target pests is emerging, is not widely present or its maintenance in laboratory environment is

not feasible. Additionally, working culture collections can sometimes maintain larger collections of specific isolates with the desired qualities.

In many cases, to prepare reference material the target (corresponding to a collection specimen) is mixed with presumably healthy/non-contaminated matrix. This matrix can be plant material, water, soil. To ensure traceability it is important to properly identify and describe the matrix as well and ensure its proper identification. Collecting meta-data otherwise used for samples is suitable.

5.5 Commutability level

Commutability in its narrower sense describes the extent to which the reference material prepared is similar to the actual samples i.e. its exchangeability by virtue of being replaceable. The type of test item chosen reflects the intended use of the reference material. For molecular tests it may range from a simple synthetic DNA through various purified nucleic acids, their mixture to naturally contaminated samples (Figure 3). In general, the more complex material is, the more similar it is to the actual samples.

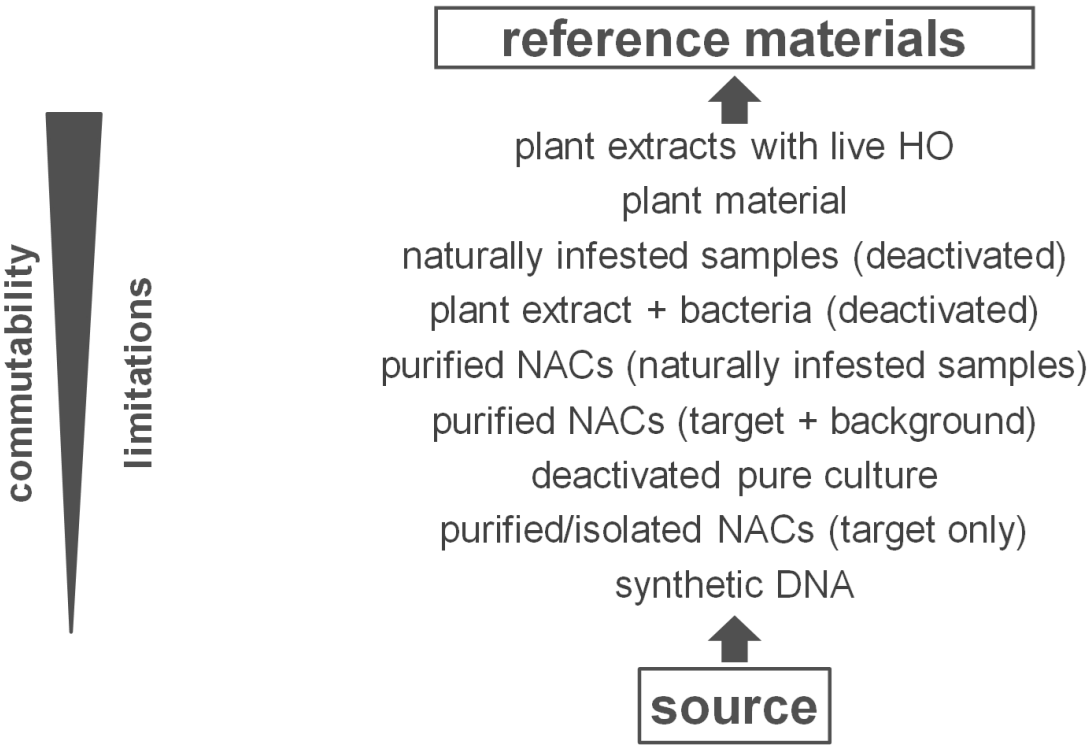

Figure 3: A range of different types of reference materials used for molecular testing. In general, increased commutability increases also the complexity of the material. NACs = nucleic acids, DNA or RNA.

At this stage, the commutability of the test items is proposed to be described through describing the type of the samples e.g. synthetic DNA or naturally infected and symptomatic plant material. The following types were identified during discussion to be suitable for inclusion in the table of criteria:

- naturally infested plant material
- artificially infested plant material
- spiked plant material
- purified organisms
- purified nucleic acids
- synthetic nucleic acids

5.6 Homogeneity

Homogeneity, as it pertains to the plant health field, has been previously defined by the EPPO guidelines (EPPO PM 7/122 (1)) and are summarized here.

The guidelines require the producers of the reference material to ensure their homogeneity i.e. the materials should be as homogenous as possible because this may affect the outcome of validations. An inhomogeneous material introduces additional measurement uncertainty which, if unrecognized, may be inappropriately transferred to the test (in validations and TPS) or the proficiency of the participant (in proficiency tests).

The assessment of homogeneity and stability should be performed by the same laboratory (generally the organizer) using the same analytical method or methods (where relevant) and measuring the same characteristic of the samples. The test used for homogeneity testing should be a standardized (or validated) test that can be implemented in the laboratory. The procedures for the assessment of homogeneity and stability should be documented. An example of reporting on homogeneity is shown in Table 1.

Table 1: An example of reporting on homogeneity of test items used in a proficiency test for detection of *Erwinia amylovora*. For homogeneity testing ten randomly selected aliquots of samples prepared for proficiency test were selected for each concentration level. From the selected samples, DNA was extracted, and tested in three technical repeats (wells) each using real-time PCR assay (Pirc *et al*. (2009) targeting amsC gene), analysing 2 µL of DNA in each reaction. The results were in concordance with the true values for all samples and concentration levels with coefficients of variations below 2% (Dreo & Pirc, 2019).

Concentration level [cells/mL]	Spiked plant extract	Sample ID	Min(Cq) - Max(Cq)	Average(Cq) ± CV
Ea amsC probe-primer set (Pirc et al. 2009)				
0	D729/18	Ea 1	neg (45)	NA
0	D729/18	Ea 2	neg (45)	NA
0	D778/18	Ea 3	neg (45)	NA
0	D778/18	Ea 4	neg (45)	NA
3xE8	D729/18	Ea 5	17,5 - 18,1	17,84 ± 0,010
3xE7	D729/18	Ea 6	21,6 - 22,1	21,78 ± 0,007
3xE8	D778/18	Ea 7	17,2 - 18,6	18,12 ± 0,018

In some cases, it is not feasible for samples to be subjected to homogeneity and stability testing. Such cases would include, for example, when limited material is available to prepare samples.

The EPPO Guidelines PM 7/122 provide further information on the assessment of homogeneity for different types of material also stating: *the current available guidelines recommend to test a minimum of 10 randomly chosen samples (for each pest/matrix/infestation level, including negative samples) in duplicate (e.g. ISO 13528). Some laboratories use the square root (rounded up) of the total number of samples. Based on current experience and depending on the method used it is recognized that this is not always feasible because of the multiple pest/matrix/infestation level combinations. Therefore, number of samples included in the homogeneity testing may be reduced if suitable data are available from previous homogeneity testing on similar samples prepared by the same procedures or according to the expertise of the organizer. The choice of the number of samples should be documented.*

5.7 Stability

Stability, as it pertains to the plant health field, has been previously defined by the EPPO guidelines and are summarized here (text in italics):

Samples should be demonstrated to be sufficiently stable to ensure that they will not undergo any significant change throughout the conduct of the interlaboratory comparison, including storage and transport conditions. When required,

stability testing should be conducted in conditions that mimic transport and storage conditions. An example of data on the short term and long term stability reported for test items is shown in Table 2 and Table 3, respectively. As an alternative, samples can be sent to the participant with the most challenging environmental or transport conditions and returned unopened for testing.

For test performance studies reagents are usually also provided to participants in addition to samples. Stability of those reagents which have an influence on the outcome of the test should be verified following the same procedures. A stability check may be performed on samples held by the organizer. This should be done after the deadline for performing analyses by the participants, in order to verify that the stability of samples has been maintained throughout the interlaboratory comparison. Some pest stages are known to be stable over long periods (e.g. Globodera spp. cysts, or fungal spores) in such case stability testing is not needed.

… the number of aliquots to be tested for stability also depends on the quantity of the reference material produced. Provided that the producer performs risk analysis and demonstrates that there is no adverse effect expected, the number of aliquots to be tested may be decreased.

Table 2: An example of results of short-term stability testing of aliquots of test items of *Erwinia amylovora* for the purpose of proficiency testing. The test items were stored at different temperatures for 1 week. Three aliquots per concentration level and spiked plant extract, were tested in three technical repeats (wells) with real-time PCR assays developed by Pirc *et al*. (2009) targeting amsC gene, after one week of incubation at temperature below -15 °C, 2-8 °C and 25 °C. Cq = cycle of threshold, CV = coefficient of variation, NA = not applicable. The different temperatures represent different transport conditions with room temperature also simulating extended degradation over longer time when kept frozen. Source: Dreo & Pirc, 2019.

Concentration level [cells/mL]	Spiked plant extract	Sample ID	T < -15 °C		2 - 8 °C		25 °C	
			Min(Cq) - Max(Cq)	Average(Cq) ± CV	Min(Cq) - Max(Cq)	Average(Cq) ± CV	Min(Cq) - Max(Cq)	Average(Cq) ± CV
Ea amsC probe-primer set (Pirc et al. 2009)								
0	D729/18	Ea 1	neg (45)	NA	neg (45)	NA	neg (45)	NA
0	D729/18	Ea 2	neg (45)	NA	neg (45)	NA	neg (45)	NA
0	D778/18	Ea 3	neg (45)	NA	neg (45)	NA	neg (45)	NA
0	D778/18	Ea 4	neg (45)	NA	neg (45)	NA	neg (45)	NA
3xE8	D729/18	Ea 5	17,1 - 17,6	17,33 ± 0,010	17,1 - 17,2	17,17 ± 0,002	16,9 - 17,6	17,22 ± 0,016
3xE7	D729/18	Ea 6	20,4 - 20,6	20,51 ± 0,004	20,3 - 20,6	20,44 ± 0,005	20,4 - 20,6	20,45 ± 0,003
3xE8	D778/18	Ea 7	17,1 - 17,4	17,22 ± 0,007	17,2 - 17,5	17,34 ± 0,006	17,0 - 17,6	17,24 ± 0,013

Table 3: An example of results of long-term stability testing of aliquots of test items (spiked plant extracts) of *Erwinia amylovora* for the purpose of proficiency testing. Sample aliquots were stored at temperature < -15 °C. Three aliquots per concentration level and spiked plant extract were tested in three technical repeats (wells) with real-time PCR assays developed by Pirc *et al*. (2009) targeting amsC gene, after 1, 5 and 10 weeks. Week 10 corresponds to the conclusions of the proficiency test. Cq = cycle of threshold, CV = coefficient of variation, NA = not applicable. Source: Dreo & Pirc, 2019.

Concentration level [cells/mL]	Spiked plant extract	Sample ID	Week 1		Week 5		Week 10	
			Min(Cq) - Max(Cq)	Average(Cq) ± CV	Min(Cq) - Max(Cq)	Average(Cq) ± CV	Min(Cq) - Max(Cq)	Average(Cq) ± CV
Ea amsC probe-primer set (Pirc et al. 2009)								
0	D729/18	Ea 1	neg (45)	NA	neg (45)	NA	neg (45)	NA
0	D729/18	Ea 2	neg (45)	NA	neg (45)	NA	neg (45)	NA
0	D778/18	Ea 3	neg (45)	NA	neg (45)	NA	neg (45)	NA
0	D778/18	Ea 4	neg (45)	NA	neg (45)	NA	neg (45)	NA
3xE8	D729/18	Ea 5	17,1 - 17,6	17,33 ± 0,010	17,3 - 17,6	17,47 ± 0,008	16,8 - 17,5	17,17 ± 0,011
3xE7	D729/18	Ea 6	20,4 - 20,6	20,51 ± 0,004	20,5 - 20,7	20,66 ± 0,003	20,6 - 21,0	20,78 ± 0,007
3xE8	D778/18	Ea 7	17,1 - 17,4	17,22 ± 0,007	17,5 - 17,7	17,56 ± 0,005	17,2 - 17,4	17,28 ± 0,005

5.8 Assigned values

EPPO Guidelines PM 7/122 state that "*The organizer has to define/establish assigned values for samples, i.e. value attributed to a particular property of an interlaboratory test sample. In the plant health field, assigned values correspond to the expected result of the test (pest present or absent, concentration of the pest, morphological characteristics of the specimen, etc.). In some cases, the assigned value may be declared as 'undetermined' (e.g. samples yielding an OD between positive and negative threshold in ELISA, specimens presenting overlapping morphological characters)*".

While it may not be always feasible or indeed necessary, it is possible to provide information on the quantity of the target pest in the reference material.

In preparing the reference material the biologically relevant concentrations should be taken into account if known.

Based on the different types of assigned values defined by EPPO, the different levels of the quantity descriptions were defined:

- Known amount i.e. absolute quantification of the target pest and/or its components (e.g. DNA copy numbers).
- Level of concentration (high/medium/low) known (as determined through use of at least one semi-quantitative or quantitative test)
- Qualitative status known (positive/negative above the determined limit of detection using at least one test)
- Consensus values from participants in proficiency test. Rules for definition of these values from the participants' results should be defined: statistical methods, outliers' effect (e.g. a virology interlaboratory comparison may assign the values this way). Uncertainty of assigned values should be defined.
- Originating from plants with known health status with a recent test result (a given period of time depends on the plant-pest combination and previous experience)

The quantity can be reported in different ways however, it should reflect the way it was determined. E.g. if turbidity measurements were used to determine bacterial cells concentrations this should be reported as 'the turbidity of XY which corresponds to xy cells/unit'. Similarly, if the concentration was determined through colony counts, the concentration should be reported as colony forming units/unit also stating the media and growth conditions used.

For molecular methods digital PCR (dPCR) is a suitable method which has been used to determine copy numbers of the target pest's nucleic acids even in cases of unculturable pests. Protocols were designed to characterize the reference material using digital PCR and to assign a reference value to the concentration of target sequences. As a method enabling absolute quantification without the need for standards, digital PCR is the method currently used as a higher order method in several metrological projects based on nucleic acid detection in the clinical field. An important advantage of dPCR over quantitative, or real-time, PCR (qPCR) is that it can be used to absolutely quantify target's concentrations without the need for calibration, which simplifies both experimentation and data comparability. An example of assigning concentration to plant samples containing phytoplasmas using digital PCR (droplet format) is shown in Figure 4 and Table 4. The approach is easily transferrable to other types of pests.

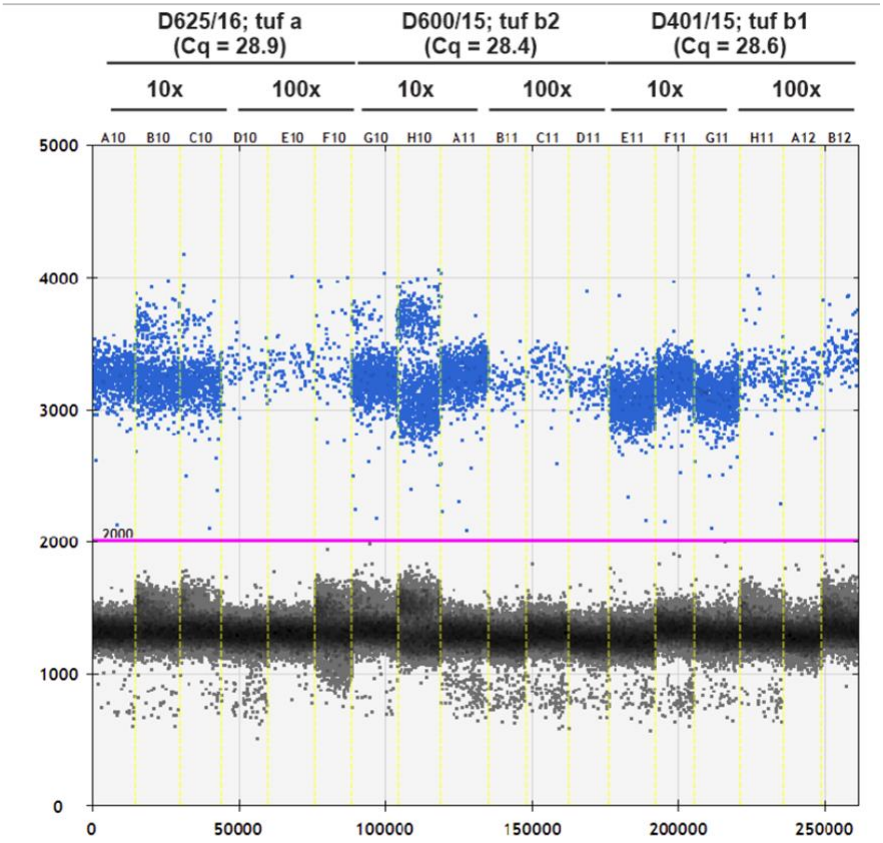

Figure 4: An example of using digital PCR analysis to determine the target copy numbers of Bois Noir phytoplasma in three samples of grapevine with the purpose to characterize the reference material. Digital PCR results are shown. Each sample was analysed in decimal dilutions (10x and 100x diluted in molecular grade water), each in three replicates. For each sample droplets are depicted according to the event (droplet number as read, x-axis) and their fluorescence in FAM channel (y-axis). The threshold discriminating between negative and positive droplets (pink line) was set manually at 2,000 relative fluorescence unit. Type of phytoplasma (tuf-type) and the average Cq value as determined in qPCR using Hren *et al*. (2007) system on 10-fold diluted DNA sample.

Table 4: An example of using digital PCR analysis to determine the target copy numbers of Bois Noir phytoplasma in three samples of grapevine with the purpose to characterize the reference material. Values assigned to the DNA samples expressed as copies (cps) or logarithm of copies (log cps) per stated volume of the original samples (undiluted DNA extracts, dilution factor taken into account). Min = minimum value, Max = maximum value, CV = coefficient of variation. Average Cq values as determined in qPCR using Hren *et al*. (2007) on 10-fold diluted DNA sample are shown.

Plant extract ID	Cq (BN)	cps/µL DNA		log(cps/mL DNA)	
		Min - Max	Average ± CV	Min - Max	Average ± CV
D625/16	28.9	1250 - 1525	1351.7 ± 0.071	6.10 - 6.18	6.13 ± 0.0046
D600/15	28.4	1675 - 2300	1972.2 ± 0.126	6.22 - 6.36	6.29 ± 0.0091
D401/15	28.6	2075 - 2600	2245.8 ± 0.091	6.32 - 6.41	6.35 ± 0.0055

5.9 Purity

Purity was defined as a ratio of target pest versus non-targets, particularly non-targets interfering with a test. The following levels were identified:

- absence of non-targets
- known ratio of target VS non-target interfering with the test - high
- known ratio of target VS non-target interfering with the test - medium
- known ratio of target VS non-target interfering with the test – low

The purity of material can be determined in different ways. At minimum, when reference material is prepared via spiking of targets into presumably healthy plant material, the absence of the target from the said plant material should be confirmed through testing.

It is worth noting, that often, assessment of purity cannot be absolute. It is determined via testing of a number of aliquots. Depending on the risk analysis the number of aliquots to be tested should be confirmed.

In some cases, the use of (high throughput) sequencing may be warranted to thoroughly characterize the reference material or its components to determine its level of purity.

5.10 Other descriptors/criteria considered

Several other descriptors/criteria were considered but were excluded from the list of criteria:

- Criteria for rejection: these were considered to be covered by the use of other criteria. Following good practice to define the necessary characteristics of the reference material in advance, the material should be rejected if it does not meet the pre-defined criteria (also recognized by EPPO PM 7/122 with regards to homogeneity and stability). It can however, be re-purposed. E.g. a material found to not be homogenous enough with respect to the pest concentration may still be used as positive isolation controls.
- quality control and quality assurance for delivery: this was considered to include e.g. calibration, recounting, distributions etc. It is considered that this is currently covered by standard quality systems in place according to e.g. EPPO guidelines (PM 7/84 and PM 7/98). It is to be expected that the laboratories working within an accreditation system (e.g. ISO 17025) will require the producers of reference materials to have quality systems at a similar level. In the future it is to be expected that the producers of reference materials and organizers of proficiency test in plant health will be required to adhere to corresponding ISO standards. However, at this stage requiring accreditation would significantly hinder preparation of reference materials.

6 Criteria for reference material

Based on the descriptors, the minimum criteria for the reference materials were defined and are provided here in the form of a checklist (Table 5). It is important to note, that since criteria are inherently linked to the intended use of the reference materials and may be test-specific, the criteria may be different for different uses and should be defined by the producer.

Table 5: List of criteria for reference materials. Minimum criteria to be fulfilled are shown in bold. The list was devised via a series of discussions, taking into account previous work and relevant international standards. Where relevant, each criterion was first defined as a series of levels from the highest to lowest ranking with the lowest ranking considered to be the minimum. The criteria are specific for the intended use, pest and test specific. Depending on the intended use, the reference material may need to fulfil higher levels of selected criteria or some criteria may not be relevant at all. RM = reference material.

Descriptor	Value	Minimum criterium
Intended use	**should be defined (in this case it equales preparation of RM for the scope of the individual TPS)**	**yes**
Identity	**identified to the level of internationally recognized diagnostic protocols**	**yes**
Traceability	traceability to a specimen from a reference culture collection	no
	traceability to a specimen from a working culture collection	no
	traceability provided for the target pest and matrix used (the latter if relevant)	**yes**
Commutability level	naturally infested plant material	no
	artificially infested plant material	no
	spiked plant material	no
	purified organisms	no
	purified nucleic acids	no
	synthetic nucleic acids	**yes**
Homogeneity	**homogenous**	**yes**
Stability	**stable**	**yes**
	stability - short term	no
	stability - long term	no
Assigned value	absolute concentration known	no
	level of concentration known (high/medium/low)	no
	qualitative status known (above LOD level)	no
	originating from plants with known health statues with a recent test result (a given period of time depends on the plant-pest combination and previous experience)	**yes**
Purity	absence of non-targets	no
	absence of interferring non-targets	no
	known ratio of target VS non-target interferring with the test - high	no
	known ratio of target VS non-target interferring with the test - medium	no
	known ratio of target VS non-target interferring with the test - low	**yes**

Also, as noted before, not all levels of all criteria are possible or feasible with all types of target organisms in matrices. While naturally contaminated samples are often considered as the ultimate goal for interlaboratory studies, a lot of

useful data has been gathered using other types of samples including synthetic DNA in particular when the material containing target pests is difficult to obtain. Indeed, different types of samples may identify different issues in test performance or proficiency of participants and can thus provide additional data.

It was considered that the basic quality assurance necessary to conduct the preparation of reference materials is currently covered by standard quality systems in place according to e.g. EPPO guidelines (PM 7/84 and PM 7/98). It is to be expected that the laboratories working within an accreditation system (e.g. ISO 17025) will require the producers of reference materials to have quality systems at a similar level. In the future it is to be expected that the producers of reference materials and organizers of proficiency test in plant health will be required to adhere to corresponding ISO standards. However, at this stage requiring accreditation would significantly hinder preparation of reference materials. This criterion is a good candidate for inclusions in the checklist of minimum criteria.

7 Application of list of criteria to different types of organisms

The list of descriptors which includes the minimum criteria i.e. the lowest level of each descriptor (where levels are relevant) was applied to the pests included in the round 1 of test performance studies. The organisers of the TPSs were asked to describe the reference material as it would be used. The summary is presented in Table 6. Note, that since this activity was done at a planning stage and is primarily a theoretical exercise to assess the suitability of the list of criteria, the actual reference materials (test items) used in the TPS may differ.

The availability criteria is excluded from the table because it has been identified and added after receiving some of the answers from the TPS organisers. The organisers aware of this criterion have complied with it. Overall, purity, as a new criterion would need additional background data to allow for it to be filled in by all organisers.

Based on the table we can conclude that all planned reference materials fit and in many cases surpass the minimum criteria for RM as identified in this deliverable.

Table 6: List of descriptors and criteria as applied to the organisms included in the round 1 of the test performance studies. The organisms included are: Ea = *Erwinia amylovora*, Pstew = *Pantoea stewartii* subsp. *stewartii*, Bxyl = *Bursaphelenchus xylophilus*, PPV = Plum pox virus, CTV = citrus tristeza virus, and Fcirc = *Fusarium circinatum*. Where several levels are available for a descriptor, the lowest corresponds to the minimum criterion for the specific descriptor. Legend: x = yes; (x) = relative determination and test results used as indicators for assigned values.

Descriptor	Value	Ea	Pstew	Bxyl	PPV	CTV	Fcirc
Intended use	should be defined (in this case it equales preparation of RM for the scope of the individual TPS)	x	x	x	x	x	x
Identity	identified to the level of internationally recognized diagnostic protocols (when available)	x	x	x	x	x	x
Traceability	traceability to a specimen from a reference culture collection	x	x	x			x
	traceability to a specimen from a working culture collection	x	x	x	x	x	
	traceability provided for the target pest and matrix used (the latter if relevant)	x	x	x		x	
Commutability level	naturally infested plant material	x			x	x	
	artificially infested plant material				x	x	
	spiked plant material	x	x	x			
	purified organisms						x
	purified nucleic acids	x	x	x			
	synthetic nucleic acids						
Homogeneity	homogenous	x	x	x	x	x	x
Stability	stable	x	x	x	x	x	x
	stability - short term						
	stability - long term						
Assigned value	absolute concentration known	x	x	x		(x)	
	level of concentration known (high/medium/low)	x	x	x	x	(x)	
	qualitative status known (above LOD level)			x	x	x	
	originating from plants with known health statues with a recent test result (a given period of time depends on the plant-pest combination and previous experience)	x	x			x	x
Purity	absence of non-targets			x		x	x
	absence of interferring non-targets	x	x		x		
	known ratio of target VS non-target interferring with the test - high				x		
	known ratio of target VS non-target interferring with the test - medium						
	known ratio of target VS non-target interferring with the test - low						

8 Limitations and future perspectives

The descriptors of reference material were selected and, where applicable, different levels were defined for them (Table 5). Through discussions the minimum criteria to be fulfilled for the reference material to be used in test performance studies. Collaboration with the TPS organisers from WP1 of the project was sought to apply the criteria list to the planned reference material.

The following text summarizes key comments and conclusions identified during selection of criteria and their application to the TPSs:

- Two criteria were added to those previously identified: 'availability' and 'purity'. As expected, some of the criteria which were newly introduced e.g. 'purity' required more background information and may benefit from further discussions during round 2 of TPSs.
- The checklist table provided to the organisers of the TPS of round 1 presumed one column per organism and target. However, different types of reference materials are planned within each of the TPS. It may be more suitable to describe each different RM separately. Also, when the reference material is prepared by mixing components e.g. target organisms with matrix, information should be provided for both.
- For the criterion 'identity' it was considered necessary that the identity of the pest in the RM is determined following internationally recognized protocols. However, when these are not available, the producer should select tests which ensure correct identification and are mutually recognizable as suitable among laboratories.
- For the criteria of homogeneity and stability the existing guidelines (EPPO 7/122 (1)) provide detailed instructions. In determining stability, there is often distinction made between determining long-term stability of material stored under recommended conditions and short-term stability determined under recommended and non-optimal conditions. However, as a final conclusion the material should be stable. However, the two levels are considered complementary to each other and are kept for future consideration.

As the current deliverable encompassed one round of practical application to actual RMs, it is to be expected that further improvements may be identified during round 2. The current descriptors and criteria were focused on the reference materials as most relevant to the TPS organised. However, the list may be applicable with minor modification well beyond the current scope of the VALITEST project e.g. for morphological tests in entomology or some very specific tests and for preparation of reference materials for other purposes.

References

Anonymous, 2008. International Vocabulary of Metrology – Basic and General Concepts and Associated Terms (VIM 3rd edition)

Camloh, M., Dreo, T., Gruden, K., Mehle, N., Milavec, M., Žel, J., 2015. Nucleic-acid analysis in new fields of metrology, in: 17th International Congress of Metrology. Presented at the 17th International Congress of Metrology, EDP Sciences, p. 06005. https://doi.org/10.1051/metrology/20150006005.

DREO, Tanja, ALIČ, Špela, MEHLE, Nataša, DERMASTIA, Marina. Preparation of defined reference material for molecular testing of ˝bois noir˝ phytoplasma. V: DERMASTIA, Marina (ed.). Proceedings of the 5th European Bois Noir Workshop, City Hotel, Ljubljana, Slovenia, 18-19 September 2018, 5th European Bois Noir Workshop, City Hotel, Ljubljana, Slovenia, 18-19 September 2018. Ljubljana: National Institute of Biology. 2018, str. 1-4. https://www2.cd-cc.si/Skripte/boisn/BOISNOIR2018/papers/a29.pdf.

DREO, Tanja, PIRC, Manca. Final report on the 'NIB Proficiency Test Round 2018-01': proficiency test for the molecular and/or serological detection of *Erwinia amylovora*. Ljubljana: National Institute of Biology, 2019. 17 pp.

EPPO (2018) PM 7/76 (5) Use of EPPO Diagnostic protocols. Bulletin OEPP/EPPO Bulletin 48 (3): 373–377.

EPPO (2018) PM 7/84 (2) Basic requirements for quality management in plant pest diagnosis laboratories. Bulletin OEPP/EPPO Bulletin 48 (3): 378–386.

EPPO (2018) PM 7/98 (3) Specific requirements for laboratories preparing accreditation for a plant pest diagnostic activity. Bulletin OEPP/EPPO Bulletin 48 (3): 387–404.

Hren, M., Boben, J., Rotter, A., Kralj, P., Gruden, K., Ravnikar, M. (2007). Real-time PCR detection systems for "flavescence dorée" and "bois noir" phytoplasma in grapevine: a comparison with the conventional PCR detection system and their application in diagnostics. Plant Pathology, 56, 785-796.

ISO 17034:2016 - General requirements for the competence of reference material producers [WWW Document], n.d. URL https://www.iso.org/standard/29357.html (accessed 10.8.18).

ISO/IEC 17043:2010 - Conformity assessment -- General requirements for proficiency testing [WWW Document], n.d. URL https://www.iso.org/standard/29366.html (accessed 10.8.18).

JCGM 200:2012 (JCGM 200:2008 with minor corrections).

Pirc, M., Ravnikar, M., Tomlinson, J., Dreo, T., 2009. Improved fireblight diagnostics using quantitative real-time PCR detection of *Erwinia amylovora* chromosomal DNA. Plant Pathology 58, 872–881. https://doi.org/10.1111/j.1365-3059.2009.02083.x

PM 7/122 (1) Guidelines for the organization of interlaboratory comparisons by plant pest diagnostic laboratories, 2014. EPPO Bulletin 44, 390–399. https://doi.org/10.1111/epp.12162

Roenhorst, J.W., Lacomme, C., Nisbet, C., Leichtfried, T., Menzel, W., Winter, S., Vlugt, R.A.A. van der, 2017. Euphresco project VirusCollect – fulfilling the need for a common collection of plant viruses and viroids for reference. EPPO Bulletin 47, 41–47. https://doi.org/10.1111/epp.12353.

Vesper, H.W., Miller, W.G., Myers, G.L., 2007. Reference Materials and Commutability. Clin Biochem Rev 28, 139–147.

Wilkinson, M.D., Dumontier, M., Aalbersberg, Ij.J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L.B., Bourne, P.E., Bouwman, J., Brookes, A.J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C.T., Finkers, R., Gonzalez-Beltran, A., Gray, A.J.G., Groth, P., Goble, C., Grethe, J.S., Heringa, J., 't Hoen, P.A.C., Hooft, R., Kuhn, T., Kok, R., Kok, J., Lusher, S.J., Martone, M.E., Mons, A., Packer, A.L., Persson, B.,

Rocca-Serra, P., Roos, M., van Schaik, R., Sansone, S.-A., Schultes, E., Sengstag, T., Slater, T., Strawn, G., Swertz, M.A., Thompson, M., van der Lei, J., van Mulligen, E., Velterop, J., Waagmeester, A., Wittenburg, P., Wolstencroft, K., Zhao, J., Mons, B., 2016. The FAIR Guiding Principles for scientific data management and stewardship. Scientific Data 3, 160018. https://doi.org/10.1038/sdata.2016.18.

World Federation for Culture Collections [WWW Document], n.d. URL http://www.wfcc.info/ (accessed 2.28.19).