Sampling Protocols and Risk of Error Significance in Molecular Detection Tests for Fruit Trees Certification

J. Kummert, M. Malice, S. Marbot
and P. Lepoivre
Unité de Phytopathologie
Faculté Universitaire des Sciences
Agronomiques
B-5030 Gembloux
Belgium

S. Steyer and R. Oger
Département de Lutte Biologique et
Ressources Phytogénétiques et
Département de Biométrie
Centre de Recherches Agronomique
B-5030 Gembloux
Belgium

Keywords: ACLSV, ASPV, apple-trees, RT-PCR, real-time RT-PCR

Abstract

In the frame of certification, the sampling and detection protocols satisfy the quality standards by considering the probability to erroneously declare that an infected plant is virus-free (risk β). In this context, the control of the risk β essentially lies on two factors: the adequacy of the detection technique towards the variability of the targeted virus and the representativity of the sample. The significance of this second factor has been addressed by the study of the detection of *Apple stem pitting virus* (ASPV) and *Apple chlorotic leaf spot virus* (ACLSV) from crude extracts of infected apple trees tissues by real time RT-PCR. Eight leaves harvested at different dates during the growing period were separately tested. For each date and each virus, the frequency distribution of the number of positive leaves was established and the probability distribution deduced.

The risk of non-detection (risk β) is represented by the theoretical probability to find no positive leaf on the eight leaves tested. This risk was generally below 0.1%, with a maximum value of 0.8% at one date for ACLSV. For the protocol used, the detection of ACLSV and ASPV seems thus reliable during the period concerned. However, the distribution of the virus in the crown of a tree is not a static phenomenon and the risk of non-detection depends on internal and external factors; it notably varies according to the number of leaves tested per tree and the multiplication rate of the virus in the tree. For the same number of leaves tested, the risk of non-detection increases if the fraction of infected leaves decreases.

INTRODUCTION

Whatever may be the sensitivity and the specificity of the techniques used to detect the presence of a pathogen, they are only able to reveal what is present in the sample. The reliability of the results significantly depends on the size of the sample, particularly when using techniques like PCR performed from very little aliquots of sample preparations. It is thus necessary to accurately define the protocols of application of the detection techniques and, particularly, the modalities of sampling and treatment of the samples taking into account the technical potentialities of the tests and the characteristics of the biological material.

A diagnostic test is a test of hypothesis for which two risks of error exist at the end of analysis: a healthy tree is declared infected (false positive = risk α) and a virus-infected tree is declared healthy (false negative = risk β). In the frame of certification, diagnostic and sampling protocols have essentially to consider the risk β (false negative) or the probability to erroneously conclude that an infected plant is virus-free. In this context, the control of the risk β rests on two main factors: the appropriateness of the detection technique towards the pathogenic agent concerned and its variability, and the representativity of the sample. The first factor is controlled by application of the selected detection technique to a population of infected trees or samples representing the variability of the isolates of the pathogen. This is made during the validation phase of the different selected RT-PCR protocols. Here we will focus on the second factor concerning

the representativity of the sample.

In order to respond to the requirements of the certification schemes, it is necessary to define with a maximum of precision the mode of sampling during the prospections in the nuclear stocks, propagation stocks and nurseries, taking into account the technical characteristics of the applied tests as well as the distribution of the virus in the infected tree, the type of tissue or organ taken and the period at which it has been taken. To secure the risk β, it is necessary to determine the minimum number of samples to be taken in function of the virus/plant pair considered [homogeneous or heterogeneous distribution of the infection, type (tree, twig) and nature (leaves, bark tissue, …) of the sample material and period of sampling].

As the number of observed cases with 0 positive leaves among the 8 leaves tested is practically null, the evaluation of the risk of error linked to the analysis of the sample is based on the adjustment of the frequency distribution of the number of items tested positive among the total number of samples taken on the same tree (x_i). If the number of trees (N) is sufficient, to each value of x_i corresponds a probability (Pi) and we can define a probability distribution for each plant/virus pair. To evaluate the risk of false negative, as an approximation, the observed probability distribution of the number of positive items is adjusted to a theoretical probability distribution by means of a X^2 test. This theoretical distribution is then used to calculate the risk corresponding to the first class of the distribution (no positive sample among all the units taken and tested). This risk, expressed in %, is represented by the probability to have a positive sample in the N samples tested, and must be lower than the acceptable risk.

The study of sampling in relation to the calculation of the risk of error linked to the detection of the virus has been tackled by taking *Apple stem pitting virus* (ASPV) and *Apple chlorotic leaf spot virus* (ACLSV) with rather homogeneous and heterogeneous distributions, respectively, in apple trees.

MATERIAL AND METHODS

Plant Material

Twenty-seven trees responding positively for ACLSV and ASPV (and also mostly for ASGV) to our real time RT-PCR protocols were selected in the original orchard with the collection of apple varieties of the Department of Phytogenetical Ressources of the Agricultural Research Centre of the Walloon Ministry of Agriculture, in Gembloux (Belgium). Eight leaves were collected on these trees at predetermined sites (the four cardinal points, at two levels in the crown) at five dates during the growing period of the trees (May and June 2002).

Sample Preparation

Each leaf was separately ground in 10 volumes of extraction buffer (Tris 2.4g/l, NaCl 8g/l, PVP K25 20g/l, KCl 0.2 g/l – pH 7.4) with the Pollähne press. After a clarification by centrifugation at 10000 rpm for 5 min, the extracts were diluted 10 times in the same buffer and 2 µl aliquots were added directly in the PCR mix. Healthy control samples consisting in leaves taken from M 106 rootstocks were included in each series of 8 samples to control the absence of contamination between samples during grinding with the Pollähne press.

RT-PCR Amplification

The RT-PCR amplification was with the One-Step RT-PCR kit (Qiagen, Hiden, Germany) with the primers pairs ACLSV5F-8R and ASPV4F-4R (Kummert et al., 2000) and the 3' MGB probes 26 (ACLSV) and 148 (ASPV) (Salmon et al., 2002a and b). Reactions (25 µl) were incubated for transcription (30 min at 50°C), activation of the *Taq*-polymerase (15 min at 95°C) and amplification for 50 cycles (15 s at 95°C, 45 s at 55°C and 1 min at 60°C).

Detection of the Amplification Products

The Gene Amp 5700 Sequence Detection System (Applied Biosystems, Foster City, USA) was used for thermal cycling and recording of the fluorescence emitted during the extension phase of each amplification cycle. Results were analyzed with the 5700 SDS software. After amplification, aliquots were analyzed by electrophoresis in precast ethidium bromide stained 2% agarose gels in TAE buffer using the E-gel 96 system (Invitrogen, Life Technologies).

RESULTS AND DISCUSSION

The experiments conducted for the evaluation of the risk of false negative results (non detection of the virus present in the tree) consisted in the RT-PCR analysis of crude extracts from eight leaves taken on virus-infected trees. Analysis of the amplification products was made either by real time measurement of the fluorescence emission generated by the hydrolysis of a specific MGB probe (Fig. 1A), or by electrophoresis of the end products obtained in agarose gels stained with ethidium bromide (Fig. 1B). The result of each experiment is expressed as the number of leaves giving a positive response in RT-PCR among the eight leaves tested (x_i). For each virus and date, this experiment was repeated for each of the 27 trees tested. We can then record the frequency associated to each value of x_i and present the results as frequency distributions of the different numbers of positive leaves for each date and virus. Figure 2 shows examples of these frequency distributions. In order to evaluate the risk of false negative, we tried to adjust a theoretical probability distribution to the observed frequency distribution by using a X^2 test. A problem for this adjustment resulted from the over representation of the last class (8 positive leaves/tree). If this class is not considered, a binomial distribution could be adjusted to the observed distributions. The probability outline of a binomial distribution is realized by a set of random experiments, identical and stochastically independent, for each of which two totally exclusive events are associated; our test for which each leaf is healthy or infected can effectively respond to this outline. The theoretical distribution, after having being adjusted to the observations, was used to evaluate the risk of false negative results for the RT-PCR tests. For the sampling protocol used, this risk is observed if none of the 8 leaves appears positive and corresponds to the first class of the probability distribution (P_0). Table 1 presents these risks of non-detection of ACLSV and ASPV expressed in % (risk $\beta = P_0.100$).

According to the results reported in Table 1, in our work conditions (samples taken in the four cardinal directions and at two levels in the crown of the tree), the risk of having a tree for which the 8 leaves are negative is generally very low (<0.1%), at least during the growing leaf period. Only for the tests made on the material collected at June 24, during a period of very hot weather, the risk was higher (0.8% for ACLSV and 0.3% for ASPV).

The statistical analysis of the results for each virus and each date of sampling allowed the evaluation of the risk of non-detection linked to the analysis by RT-PCR and the evolution of the risk during the period of sampling. This showed that the distribution of the virus in the tree crown evolve quantitatively and qualitatively during the season. The risk of error may thus also vary continuously in relation to internal and external factors. It would be possible to develop a dynamic model for the evolution of the risk of non-detecting the infection of a tree by a specific virus if a mathematical relation could be determined between the level of infection of the tree and the main factors influencing this level. This, however, would need a high number of observations and analysis on the distribution of the virus in the tree, allowing the identification of the factors acting on this evolution and the relations linking them to the level of infection. A more exhaustive study of the different factors influencing the repartition of the viruses in the tree and of their impact on the run risk of non detection should thus be carried out.

According our results, the RT-PCR on crude extracts of apple leaves seems reliable during the period considered. However, these results are valid only for the sampling protocol used, because the risk of non-detection depends on the number of

leaves taken per tree (n) and the level of infection of the tree characterized by the fraction of infected leaves (p). The formula $P_0 = (1-p)^n$ links these variables and the evolution of the risk of non detection in function of the fraction of infected leaves, when 8 leaves are taken on each tree, is shown Figure 3. For a same number of leaves sampled, the risk of false negative results increases if the fraction of infected leaves decreases.

The same approach has to be used if separated trees, instead of leaves, are considered. In this context, we have to take into account that the infection level is supposed to be very low in nurseries where certified material is multiplied. The formula presented allows determination of the number of samples to be taken to reach a defined risk of error. According to this formula, this number of samples depends on the level of infection for which a hypothesis has to be drawn. For a defined risk, the number of samples to be taken increases if the level of infection decreases.

The application of statistical analysis to the problem of sampling ought to allow a better understanding of the definition of the regulations related to the phyto-sanitary control of plant products. For example, if we consider an orchard of 1000 trees, 5% of which are infected, and a sampling of only 0.5% of the trees as required by our actually proposed regulation for fruit tree certification, the probability that no virus-infected tree exists in the five trees from where samples were collected would be 77%. This demonstrates that the fraction of trees to be indexed according to the legislation is to low. The situation is similar for many other certification regulations. The use of pooled samples, coupled to more sensitive techniques like PCR, would enable to better reconcile the requirements of the certification regulations and the statistical validity of the sampling techniques used.

ACKNOWLEDGEMENTS

This research was funded by the Belgian Ministry of Small Enterprises, Traders and Agriculture – DG6 (Research and Development) and is actually funded by the Ministry of Agriculture of the Walloon Region.

Literature Cited

Kummert, J., Vendramme, M., Steyer, S. and Lepoivre, P. 2000. Development of routine RT-PCR tests for certification of fruit tree multiplication material. Bulletin OEPP/EPPO Bulletin 30:441-448.

Salmon, M.A., Vendramme, M., Kummert, J. and Lepoivre, P. 2002a. Detection of apple chlorotic leaf spot virus using a 5' nuclease assay with a fluorescent 3' minor groove binder-DNA probe. J. Virol. Methods 104:99-106.

Salmon, M.A., Vendramme, M., Kummert, J. and Lepoivre, P. 2002b. Rapid and homogenous detection of apple stem pitting virus by RT-PCR and a fluorogenic 3' minor groove binder-DNA probe. Eur. J. Plant Pathol. 108:755-762.

Tables

Table 1. Values of the risk of error β (false negative results = probability to have 0 infected leaf, expressed in % = P_0;100)

Date	ACLSV		ASPV	
	Real time	Gel electrophoresis	Real time	Gel electrophoresis
13/05	<0.1 (25)[1]	<0.1 (27)	<0.1 (27)	<0.1 (27)
27/05	<0.1 (27)	<0.1 (27)	<0.1 (18)	<0.1 (18)
10/06	<0.1 (25)	<0.1 (19)	<0.1 (7)	<0.1 (17)
24/06 [2]	-[3]	0.8 (17)	-	0.3 (5)

[1] Number of trees considered, [2] Samples harvested after a period of very hot days

[3] Unworkable results; fluorescence levels very low and variable from cycle to cycle (device or software problem)

Figures

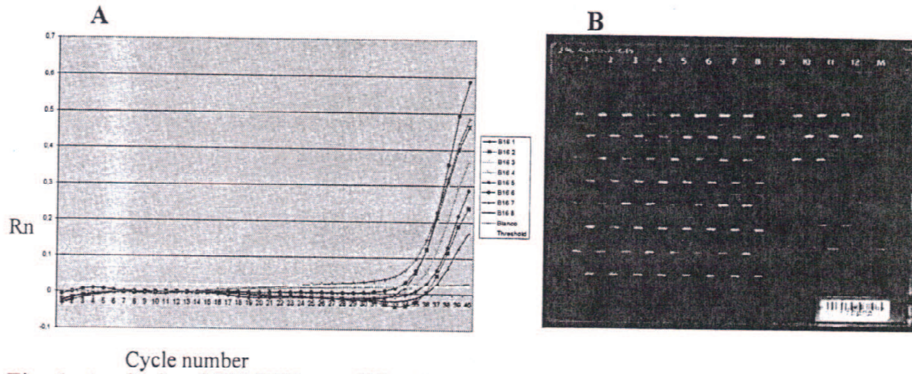

Fig. 1. Analysis of RT-PCR amplification products obtained from crude extracts of the 8 leaves harvested on virus infected apple trees. A. Real time detection of amplification products obtained with ASPV primers for the 8 leaves of tree B16 (harvesting data 13/05/2002). B. Agarose gel electrophoretic analysis of amplification products obtained with ASPV primers from leaves of trees A17, A56, A64, A65, A66, A68, B6, B15 (wells 1 to 8 of rows A to H from top to bottom) and from leaves of tree B16 (wells 10, 11, 12 of row A, wells 10, 11, 12 of row B, and wells 10, 11 of row C). Wells of column 9 contained negative controls. (Harvesting data 24/06/2002).

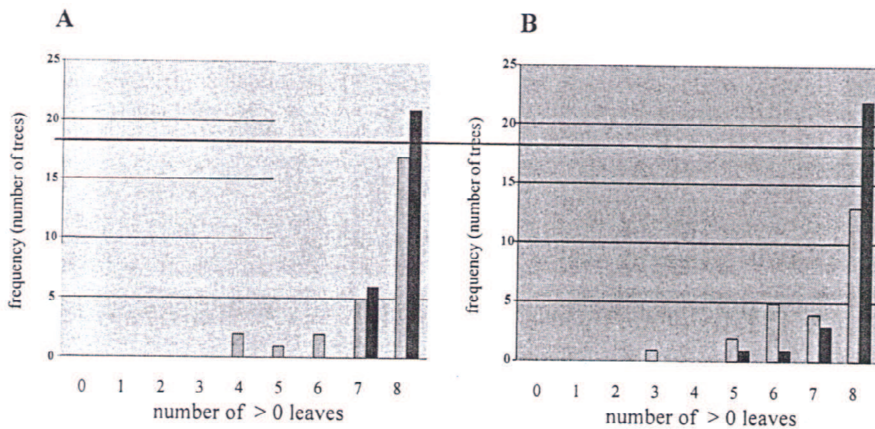

Fig. 2. Frequency distribution of the number of leaves tested positive among the 8 leaves harvested per tree when analyzed by RT-PCR with real-time fluorescence detection (□) or agarose gel electrophoresis (■). Leaves harvested on May 13 and tested for the presence of ASPV (A) or ACLSV (B).

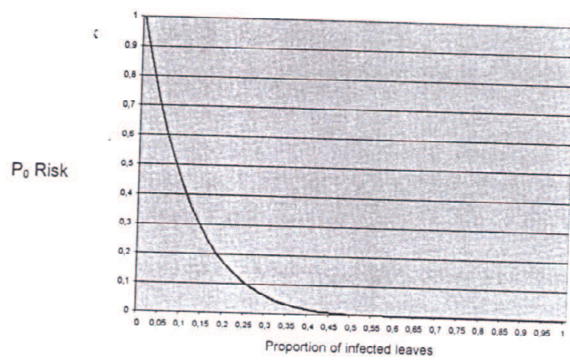

Fig. 3. Evolution of the risk of non detection (P_0) as a function of the fraction of infected leaves when 8 leaves are sampled on a tree.