Eur J Plant Pathol (2008) 122:539–547
DOI 10.1007/s10658-008-9322-1

Inter-laboratory evaluation of a duplex RT-PCR method using crude extracts for the simultaneous detection of Prune dwarf virus and *Prunus* necrotic ringspot virus

S. Massart · Y. Brostaux · L. Barbarossa · V. César · M. Cieslinska · O. Dutrecq · F. Fonseca · R. Guillem · A. Laviña · A. Olmos · S. Steyer · T. Wetzel · J. Kummert · M. H. Jijakli

Received: 20 September 2007 / Accepted: 10 April 2008 / Published online: 13 May 2008
© KNPV 2008

Abstract The operational capacity of a duplex RT-PCR method for simultaneous detection of Prune dwarf virus (PDV) and *Prunus* necrotic ringspot virus (PNRSV) has been established by nine European laboratories. A total of 576 samples from *Prunus* trees with known sanitary status, corresponding to 32 samples in two repetitions for each laboratory, were analysed. The level of sensitivity achieved by the

S. Massart · J. Kummert · M. H. Jijakli (✉)
Plant Pathology Unit,
Gembloux Agricultural University (FUSAGx),
Passage des Déportés, 2,
5030 Gembloux, Belgium
e-mail: jijakli.h@fsagx.ac.be

Y. Brostaux
Unité de Statistique, Informatique et Mathématique appliquées (SIMa), FUSAGx,
Gembloux, Belgium

L. Barbarossa
CNR-Istituto di Virologia Vegetale-Sezione di Bari,
Via Amendola 165/A,
70126 Bari, Italy

V. César
Corder ASBL, Unité de Phytopathologie,
Croix du Sud, 2/3,
1348 Louvain-La-Neuve, Belgium

M. Cieslinska
Virology Laboratory,
Research Institute of Pomology and Floriculture,
Pomologiczna 18,
96-100 Skierniewice, Poland

O. Dutrecq
DNAlis sprl,
rue de la Croix-Rouge, 20,
5032 Gembloux, Belgium

F. Fonseca
Molecular Plant Virology Lab,
University of Algarve, CDCTPV,
Campus de Gambelas,
Gambelas 8005-139 Faro, Portugal

R. Guillem
Service Régional de la protection des végétaux de la DRAF Aquitaine,
51 rue Kiéser,
33077 Bordeaux, France

A. Laviña
Deparmento de Protecció Vegetal,
Institut de Recerca i Tecnologia Agroalimentaries,
Ctra. de Cabrils s/n,
08348 Cabrils, Barcelona, Spain

A. Olmos
Instituto Valenciano de Investigaciones Agrarias (IVIA),
Carretera de Moncada a Náquera km 5,
46113 Moncada, Valencia, Spain

S. Steyer
Département Lutte biologique et Ressources phytogénétiques, Centre Wallon de Recherches Agronomiques,
Rue de Liroux 4,
5030 Gembloux, Belgium

T. Wetzel
Institute for Plant Research, RLP Agroscience, AlPlanta,
Breitenweg 71,
67435 Neustadt an der Weinstrasse, Germany

method was 98.3% for PDV and 90.4% for PNRSV. The specificity was 87.4% for PDV and 94.3% for PNRSV. The unilateral 95% confidence intervals were calculated for all these values. Cohen's Kappa coefficient of repeatability and reproducibility of the technique indicated a strong agreement between data. Likelihood ratios were 7.50 (positive) and 0.02 (negative) for PDV. For PNRSV, the positive likelihood ratio was 15.00 while the negative likelihood ratio was 0.11. In addition, post-test probabilities of infection were calculated to manage the risk associated with the routine use of this method. This allows an accurate test result interpretation to facilitate the integration of this new technique into a certification scheme.

Keywords Diagnostic · Inter-laboratory · Plant virus · Ring-test

Introduction

Prune dwarf virus (PDV) and *Prunus* necrotic ringspot virus (PNRSV) are members of the genus *Ilarvirus*, family *Bromoviridae*, naturally infecting the majority of *Prunus* species. These viruses are transmitted mechanically, through grafting, pollen and seeds. PDV causes necrotic and chlorotic spots on cherry leaves and stunting on peach and plum trees (Desvignes 1999). It has been reported to cause about 43% losses in fruit yield (Scott et al. 2001). Although PNRSV infections may be latent, growth reductions ranging from 12% to 33% have been estimated (Albertini et al. 1993; Pine 1964). A virus combination between PDV and PNRSV known to induce peach stunt infection (PSD) may occur showing premature defoliation, bark splitting and increased watersprout production, with fruit yields reduced up to 60% (Scott et al. 2001). Consequently the use of healthy plant material is a requirement to prevent virus spread in woody crops. In this context certification schemes worldwide are being established with the objective of identifying healthy sources for propagation. Numerous molecular methods based on hybridization or PCR have been developed to detect these viruses at low concentration limits (Crosslin et al. 1992; Hammond et al. 1998; Rosner et al. 1997; Sánchez-Navarro et al. 2005; Spiegel et al. 1996).

The reverse transcription/polymerase chain reaction (RT/PCR) is a powerful tool for the detection and/or characterisation of plant viruses with RNA genomes (Henson and French 1993). Despite its potential and the numerous protocols developed so far, RT-PCR is still less used than immunological or biological methods for the primary diagnosis of plant pathogens in certification schemes. The main advantage of RT-PCR in the detection of fruit tree viruses is its high sensitivity. Another great advantage is the possibility to simultaneously detect several targets in a single reaction (James et al. 2006). This property is particularly interesting for PDV and PNRSV as they are often present in mixed infections and their absence has to be controlled in phytosanitary certification programmes. A duplex protocol, based on primer pairs published by Kummert et al. (2000), was previously optimized for the simultaneous detection of both viruses. The polyvalence of the primers was evaluated at the FUSAGx laboratory on 63 isolates of PDV and 101 isolates of PNRSV, respectively (Kummert et al. 2000; Massart et al., unpublished data).

The reluctance to adopt the new molecular techniques in plant-virus routine diagnostics may be partly explained by the lack of confidence in the obtained results and the poor knowledge on the repeatability, reproducibility and limits of the RT-PCR methods. Such knowledge can be obtained through the inter-laboratory evaluation of the methodology. Although many RT-PCR protocols have so far been developed to detect plant viruses, very few of them have been thoroughly validated through an inter-laboratory evaluation involving reference laboratories.

This study reports the inter-laboratory evaluation involving nine European laboratories of a duplex RT-PCR method using plant crude extracts as template for the molecular diagnosis of PDV and PNRSV. The objective of this validation was to estimate the likelihood ratios of this method and the post-test probabilities of infection, to provide useful information to manage the risk associated with the use of this method in routine analysis. Additionally, this validation aimed to write and select a comprehensive and detailed protocol accompanying this duplex RT-PCR test.

Materials and methods

Sample preparation

Samples were prepared in the Plant Pathology Unit of the Gembloux Agricultural University (FUSAGx).

One year-old shoots from plum and cherry trees were sampled in the reference orchard of the FUSAGx and the CRA-W in Gembloux. A small portion of bark tissue was removed from the shoot. Vascular tissues (0.2 g) were sampled using a razor blade and placed in nylon mesh bags (Bioreba, Reinach, Switzerland) with 2 ml of cold KAJI buffer (DNAlis, Gembloux, Belgium). The samples were ground for 10 s using a Homex (Bioreba). Plant crude extracts were immediately put on ice and further conserved at −20°C. Pooled samples were prepared by mixing one volume of raw extract from an infected tree with either 7 or 79 volumes of raw extract from a healthy tree. Crude extracts were further diluted 10 times in fresh Ultrapure water (Invitrogen, LaJolla, USA) and divided into identical aliquots of 50 µl.

Sample distribution

Thirty-two samples were prepared in two repetitions for each laboratory and coded to ensure a double blind test, e.g. neither the sample status nor the correspondence between repetitions was known by the laboratories. The frozen samples were sent in solid CO_2 to the nine participant laboratories (randomly numbered A to I) of the inter-laboratory validation. All the reagents for PCR amplification were provided along with the samples.

RT-PCR

Hundred times diluted crude extracts were prepared by each partner by adding 450 µl of fresh Ultrapure water to the received samples (50 µl).The duplex RT-PCR was carried out using the One-step RT-PCR kit (Qiagen, Hiden, Germany) and specific primers (Eurogentec, Liège, Belgium) The primers PNRSV-10F (5′-TTCTTGAAGGACCAACCGAGAGG-3′) and PNRSV-10R (5′-GCTAACGCAGGTAA GATTTCCAAGC-3′) were used to amplify a 348 bp fragment of PNRSV coat protein. The primers PDV-17F (5′-CGAAGTCTATTTCCGAGTGGATGC-3′) and PDV 12R (5′-CACTGGCTTGTTTCGCTGT GAAC-3′) were used to amplify a 303 bp fragment of the PDV coat protein. The reaction mix consisted on: RT-PCR buffer 1×, 400 nM of each primer, 400 nM dNTPs, 1 µl of enzyme mix, 2 µl of RNAse-free water and 5 µl of 100× diluted crude extract in a total volume of 25 µl. Thermal cycle protocol consisted of a first step of 30 min at 50°C for reverse-transcription, followed by 15 min at 95°C for polymerase activation and 40 cycles at 95°C for 45 s, 55°C for 45 s and 72°C for 1 min with an final step at 72°C for 10 min. PCR products were subjected to electrophoresis in agarose gel (2% *v*/*v*) and stained with ethidium bromide for visualisation of amplicons. The following thermocyclers were used in the laboratories: PTC 200 (MJ Research), GeneAmp® PCR System 2700 and 2720 (Applied Biosystem), Mastercycler® 5341 and Gradient (Eppendorf), I-cycler (Biorad).

Statistical analysis

Sensitivity and specificity of the method were estimated according to Altman and Bland (1994). Sensitivity was the proportion of true positives that were correctly identified by the method (number true positive/number of infected samples). Specificity was the proportion of true negatives that were correctly identified by the method (number true negative/number of healthy samples). Both parameters were estimated independently for each repetition in each laboratory for both viruses separately. Unilateral 95% confidence intervals (CI) were calculated for the global estimation of both parameters for each virus, using the Agresti-Coull method (Agresti and Coull 1998). These confidence intervals included the real value of the corresponding parameter in 95% of the trials. In other words, for a determinate assay, there was 95% probability that the real value of the parameter for the corresponding virus was higher than the lower limit of the CI.

To study the pooling effect on the sensitivity of the tests, a generalized linear model (GLM), corrected for the laboratory and sample effects, was fitted to obtain a linear relationship between the dilution factor and the sensitivity. The linearity relationship was tested using the GLM through a χ^2 test detecting significant deviation of the linearity with significance level at $P<0.05$. The influence of the sample pooling on the sensitivity was further evaluated through another χ^2 test with significance level at $P<0.05$.

Repeatability and reproducibility of the RT-PCR test were estimated through the concordance analysis and the calculation of Cohen's Kappa coefficients (Cohen 1960), which measure the agreement of a classification between raters, taking into account the

agreement occurring by chance. Kappa coefficient represents how much better is an agreement than would result from chance only. The following criteria, based on the interpretation of Landis and Koch (1977) was used: 0.00 to 0.20: no agreement; 0.21 to 0.40: weak agreement; 0.41 to 0.60: moderate agreement; 0.61 to 0.80: strong agreement; and 0.81 to 1.00, almost perfect agreement. Repeatability was evaluated by averaging the Kappa agreement between repetitions in each laboratory, and the reproducibility by averaging those values between all pairs of repetitions whatever their origin (Light 1971).

The relationship between pre- and post-test probability of infection is accurately described by Baye's theorem (Fagan 1975). As the prevalence of the infection in the samples of this ring-test could be different from the prevalence of the infection in routine analysis, the likelihood ratios were used to estimate the post-test probability of the method. The likelihood ratio incorporates both the sensitivity and specificity of the test and provides a direct estimate of how much a test result will change the odds [odd = probability/(1−probability)] of being infected for the individual within its population. The positive likelihood ratio is calculated as follows: Sensitivity/(1− Specificity), while the negative likelihood ratio corresponds to (1−Sensitivity)/Specificity. The post-test probability was further calculated using Fagan's monogram (Fagan 1975) or the following formulae: post-test odd = pre-test odd x likelihood ratio.

Results

Discarded values

All the samples arrived still frozen in all the laboratories. The results were sent to FUSAGx for interpretation. Figure 1 shows an example of results obtained after gel electrophoresis with 10 samples: both viruses were detected in samples no 1, 5, 6 and 7, PNRSV was detected in sample no 4° and PDV in sample no °10, no virus was detected in samples no 2, 3, 8 and 9. For one laboratory (D), the second repetition was discarded because all the samples were contaminated. Obviously, this result came from contamination of the PCR reaction mix, most probably through PCR products from previous PCR reactions. Additionally, during the first evaluation of the results, it appeared that one laboratory (E) did not process the sample on ice before RT-PCR. As this was a major deviation of the protocol, the results of this laboratory were not included in the global analysis.

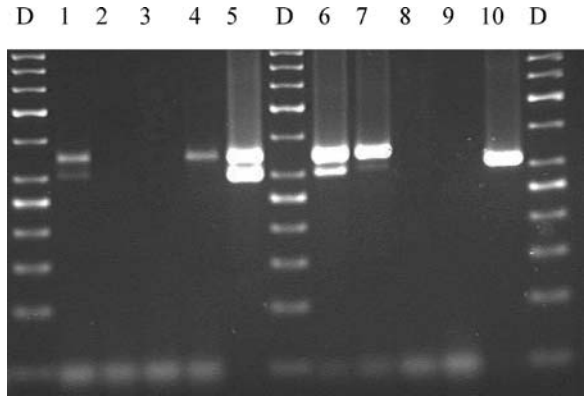

Fig. 1 Detection of PDV and/or PNRSV after gel electrophoresis of PCR products. *D* O'GeneRuler™ 50 bp DNA Ladder (Fermentas); *Lanes no 1 to 10* trees analysed during the ring-test assay

Study of the pooling effect

The effect of sample pooling on the sensitivity of detection was evaluated for each virus separately (Table 1) through the establishment of a generalized linear model. The analysis was done on 18 samples, corresponding to six samples with a double infection at 1×, 8× and 80× pooling. The data were transformed and the best linear relationship was obtained using a decimal logarithmic transformation of the dilution factor and a logit link function for the probability of detection. No deviation from the linear hypothesis was detected (PDV, $\chi^2_{1dl} = 0.807$, P-value=0.369; PNRSV, $\chi^2_{1dl} = 0.224$, P-value=0.636). There was no significant influence of the decimal logarithm of

Table 1 Evaluation of the effect of sample pooling on the sensitivity of detection (%) by statistical analysis through a generalized linear model

Virus	Sensitivity (%)			*P*-value
	1×	8×	80×	
PDV	98	98	100	0.360
PNRSV	81	97	100	<0.001

1× doubly-infected Non-pooled sample; *8*× 1 doubly-infected sample pooled with 7 healthy samples; *80*× 1 doubly-infected sample pooled with 79 healthy samples

pooling on the detection sensitivity for PDV ($\chi^2_{1dl} = 0.839$, P-value=0.360). The average sensitivity for these samples was 98.7%. For PNRSV, a very highly significant influence of the decimal logarithm of pooling on the detection sensitivity was observed ($\chi^2_{1dl} = 24.456$, P-value<0.001). Surprisingly, the detection probability rose with the pooling factor ($\beta_{\log(dilution)}$=2.995).

Specificity and sensitivity

Table 2 shows the estimated specificity and sensitivity together with the lower limit of the corresponding unilateral confidence intervals (C.I.$_{95\%}$). Whatever the parameter, the estimations were based on more than 200 values coming from the eight laboratories. For PDV, the estimated specificity and sensitivity were 87.4% (C.I.$_{95\%}$=83.7%) and 98.3% (C.I.$_{95\%}$=96.2%), respectively. For PNRSV, the specificity and sensitivity were estimated at 94.3% (C.I.$_{95\%}$=91.0%) and 90.4% (C.I.$_{95\%}$=87.1%), respectively.

Repeatability analysis

Table 3 shows the Cohen's Kappa coefficients of repeatability calculated for each laboratory and the average. The concordance analysis revealed that the average repeatability was 90% for both viruses. The average Kappa's coefficient of Cohen was 79% for both viruses. According to Landis and Koch (1977), these results indicate a strong agreement, or concordance, between the values obtained from two independent repetitions made in the same laboratory. An almost perfect agreement was obtained by Laboratory A, C and I for both viruses and by Laboratory F for PNRSV only. A strong agreement was obtained by laboratories B and G for both viruses and by laboratory F for PDV. A moderate agreement was obtained by Laboratory H.

Table 2 Estimation (%) of the specificity and sensitivity of the RT-PCR method and calculation of the confidence interval for each value

Virus	Specificity			Sensitivity		
	n	Estimation	*C.I.*	*n*	Estimation	*C.I.*
PDV	240	87.4	83.7	270	98.3	96.2
PNRSV	300	94.3	91.0	210	90.4	87.1

n Number of observations, *C.I.* lower limit of the unilateral 95% confidence interval

Table 3 Cohen's Kappa coefficients of repeatability and reproducibility (%) of the duplex RT-PCR detection method

Virus	Repeatability								Reproducibility
	A	B	C	F	G	H	I	Average	
PDV	94	71	100	71	77	54	88	79	75
PNRSV	100	65	88	82	70	50	94	79	77

A to *I* laboratories

Reproducibility analysis

The reproducibility of the method calculated according to the concordance method was estimated at 93% for PDV and PNRSV. The Cohen's Kappa coefficients of reproducibility calculated by Lights method (Table 3) were estimated at 75% and 77% for PDV and PNRSV, respectively. These Kappa values indicated a strong agreement between all the repetitions, whatever the laboratory.

Likelihood ratios and post-test predictive value

The positive likelihood ratios for PDV and PNRSV were 7.5 and 15, respectively. The negative likelihood ratios for PDV and PNRSV were 0.02 and 0.11, respectively. The likelihood ratio can be combined with the incidence of the infection to determine the post-test probability of the infection. Therefore, the post-test probability of infection versus the incidence of the infection was simulated (Fig. 2a,b).

Discussion

An inter-laboratory evaluation should be performed under the range of conditions in which they are likely to be used in practice (Banoo et al. 2006). So, this study was designed to mimic practical analysis conditions in a routine certification laboratory. Samples were selected among trees with various degrees of infection: healthy or infected with high or low titre, long-time or recently infected, plum or cherry. Moreover, they were processed by a simple, quick and reliable preparation method based on crude extract preparation in an optimized buffer, avoiding

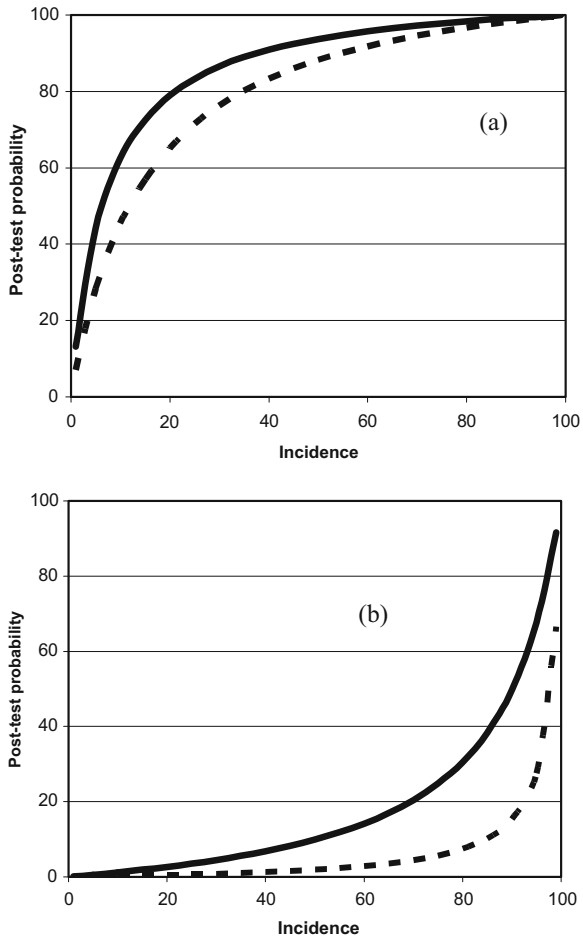

Fig. 2 Post-test probability of an individual being infected after a positive (**a**) or negative (**b**) result with the duplex RT-PCR protocol, depending on the virus incidence in the population. *Continuous curves* PNRSV and *discontinuous curve* PDV

the need of cost- and labour-intensive RNA extraction and purification protocols.

Compared to most of the previous inter-laboratory assays of (RT-)PCR detection methods (Jeffries and James 2005; Lopez et al. 2006; Paton et al. 2000), the experiment presented three additional difficulties: (1) the processing of vascular tissues (more complicated to process than leaves), (2) the use of crude extract preparations, and (3) the simultaneous identification of two different RNA targets.

The sample pooling is an interesting alternative to reduce the cost analysis or to increase the number of analysed samples within a certification programme. This pooling may correspond to various leaves from the same tree, various shoots from a single batch from the nursery, and heterogeneous samples in a low infection prevalence population. Nevertheless, this pooling can not be done without evaluating its effect on the sensitivity of detection. Only doubly-infected trees were used to study the pooling effect (worst case scenario). The statistical analysis revealed no effect of pooling for PDV. The observed effect on PNRSV was surprising as the sensitivity rose with pooling level. This effect was due to two trees with a very intense signal for PDV and no signal for PNRSV in the non-pooled sample. It is suggested that a competition phenomenon occurred between both targets, PDV being very abundant and limiting amplification of PNRSV amplicon. This effect disappeared when samples were pooled and both virus concentrations decreased. When the statistical analysis was done without these two samples, there was no significant effect of pooling on the sensitivity (P=0.191).

The specificity of the duplex RT-PCR protocol was estimated at 87% (PDV) and 94% (PNRSV). Other inter-laboratory evaluations of (RT)-PCR protocols for single plant or human pathogen detection obtained specificity values ranging from 82% to 100% (Jeffries and James 2005; Josefsen et al. 2004; Lopez et al. 2006; Malorny et al. 2003; Taha et al. 2005). The estimated sensitivities of the duplex RT-PCR protocol were 98% and 90% for PDV and PNRSV, respectively. During previous inter-laboratory evaluations of (RT)-PCR protocols for single plant, animal or human pathogen detection, the sensitivity values ranged between 72% to 96% (Jeffries and James 2005; Josefsen et al. 2004; Paton et al. 2000; Taha et al. 2005; Truyen et al. 2006) or 38% to 85% (Lopez et al. 2006). So, while using crude extracts and detecting two viruses simultaneously, the specificity and the sensitivity values were very high and equal or higher than the average values published in the literature for detection of single pathogens using purified RNA or DNA.

Repeatability and reproducibility of a laboratory test between different laboratories are crucial characteristics but, unlike specificity and sensitivity, they are rarely taken into account when evaluating the usefulness of a test. The Kappa coefficients of repeatability and reproducibility indicated a strong concordance between repetitions within or between the laboratories. Interestingly, the Kappa coefficients of reproducibility and average repeatability were very similar whatever the virus, indicating a strong inter-laboratory

agreement. This means that the variability of the duplex protocol observed between two repetitions within a laboratory is similar to those observed between repetitions from various laboratories. In other words, the laboratory had a negligible influence on the result obtained. Nevertheless, the problem encountered with laboratory E and the full contamination observed in laboratory D exemplified the crucial importance of (1) having a clear and detailed protocol underlining the important tips for the method, and (2) following carefully all the preventive measure to avoid carry-over contamination. The protocol was therefore adapted to ensure comprehensive and non-ambiguous understanding of the instructions. The current version of the detailed protocol is available on request.

An assay is considered valid only to the extent that its results accurately predict, by inference, the true status of the samples being examined (Ferris et al. 2006). In fact, any test may not prove presence or absence of infection, but its result can give more accurate probability of its presence or absence (post-test probability). Sensitivity and specificity have limited use in day-to-day diagnostic practice. A more useful approach is to combine the sensitivity and specificity results into single measures representing how much more likely a sample testing positive or negative corresponds to an infected sample than to a healthy one. These are known as the likelihood ratios for a positive test (LR+) and for a negative test (LR−). A likelihood ratio of 1 indicated a non-informative test, while the higher the LR+ and the lower the LR−, the better the test. A diagnostic method with a LR+ (LR−) higher (lower) than 10 (0.1) is considered as very good and led to a large change in pre- to post-test probability (Akobeng 2007). The LR+ for PNRSV (15.8) is therefore very good as well as the LR− for PNRSV (0.11) and PDV (0.02). The positive likelihood ratio for PDV (7.5) indicates that the test led to moderate change in pre- to post-test probability. The LR+ and LR− can be further combined with pre-test probability (prevalence) to calculate the post-test probability for a positive and a negative result, respectively. The calculation of post-test probabilities of infection provides useful information to manage the risks associated with a method and to integrate it adequately in certification schemes. The concept of pre- and post-test probability may be explained using three hypothetical populations with virus incidences of 25%, 50% and 75%, respectively. Using the LR+ and LR−, the post-test probability of being infected for a sample with a positive or a negative result can be calculated (Table 4). As an example, in a population with a PDV prevalence of 50%, the probability of a tested individual being really infected after a positive result rises from 50% (pre-test probability) to 88% (post-test probability). After a negative result, there is only a 2% probability that the tree is infected by PDV.

In epidemiological studies, the PDV incidence ranged from 0.4% to 17% (Jarrar et al. 2001; Herrera and Madariaga 2002; Dominguez et al. 1998). With these values, the maximum probability that a sample testing negative is infected is 0.41%. This very low probability makes this RT-PCR protocol a method of choice as a first screening technique. This technique will drastically reduce the release of infected plant material. Nevertheless, for the same prevalence, the maximum probability that a sample testing positive is really infected is only 61%. This result underlined the absolute necessity to further confirm a positive result by an independent diagnostic technique.

According to literature data, the PNRSV incidence is more variable. Usually, the incidence ranged between 5.6% to 10% (Myrta et al. 2002; Herrera and Madariaga 2002; Dominguez et al. 1998) while a peak of 46% has been observed (Jarrar et al. 2001). A sample testing negative has therefore an infection probability of approximately 1% (low incidence) and 8% (high incidence). Clearly, the risk of releasing plant material infected by PNRSV is higher than for PDV. Nevertheless, it is acceptable (lower than 1%) for samples coming from a population with low infection prevalence, and should be managed on a case-by-case basis. On the other hand, the probability that a positive result corresponds to an infected sample ranged from 94% (high incidence) to 47%

Table 4 Post-test probabilities of PDV or PNRSV infection for a sample after a positive or negative result of the test, depending on three levels of virus incidence within the studied population

Incidence (%)	Post-test probabilities (%)			
	Positive result		Negative result	
	PDV	PNRSV	PDV	PNRSV
25	71.4	83.3	0.7	3.5
50	88.2	93.8	1.9	9.9
75	95.7	97.8	5.6	24.8

(low incidence). As for PDV, it should be mandatory to further confirm a positive result by an independent diagnostic technique.

In conclusion, the inter-laboratory evaluation of a molecular diagnostic protocol provides very useful information on the performance and limits of the method. This information is crucial to adequately (1) integrate the protocol within a certification scheme, (2) use the protocol in routine analysis and (3) assess and manage the risk linked to the results of the test. This inter-laboratory assay confirmed the high specificity and sensitivity of the duplex RT-PCR protocol to detect PDV and PNRSV. It also showed the reproducibility of the protocol in various laboratories. Finally, the calculation of post-test probability values underlined the great potential of this molecular protocol as a first screening technique whose positive results must be confirmed by independent methodology.

Acknowledgements We thank Frédéric Vansteen and Sylvie Vanmellaert for the excellent sample management and technical assistance. This work has been funded by the General Management of Agriculture (DGA) of the Walloon Region Ministry (Belgium).

References

Agresti, A., & Coull, B. A. (1998). Approximate is better than "Exact" for interval estimation of binomial proportions. *The American Statistician*, *52*(2), 119–126.

Akobeng, A. (2007). Understanding diagnostic tests 2, likelihood ratios, pre-and post-test probabilities and their use of clinical practice. *Acta Paediatrica*, *96*, 644–647.

Albertini, A., Giunchedi, L., Dradi, D., & Benini, A. (1993). Effetto di infezioni virali su alcune variet!a di ciliegio. *Rivista di Frutticultura*, *2*, 61–64.

Altman, D., & Bland, J. M. (1994). Statistics note, diagnostic tests 1: sensitivity and specificity. *BMJ*, *308*, 1552.

Banoo, S., Bell, D., Bossuyt, P., Herring, A., Mabey, D., Poole, F., et al. (2006). Evaluation of diagnostic tests for infectious diseases, general principles. *Nature Reviews Microbiology*, *5*(11), S17–S29.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, *20*, 37–46.

Crosslin, J. M., Hammond, R. W., & Hammerschlag, F. A. (1992). Detection of *Prunus* necrotic ringspot virus serotypes in herbaceous and *Prunus* hosts with a complementary RNA probe. *Plant Disease*, *76*, 1132–1136.

Desvignes, J. (Ed.) (1999). Virus infections of fruit trees. Paris: Editions CTIFL.

Dominguez, S., Aparicio, F., Sanchez-Navarro, J., Cano, A., Pallas, V., & Garcia-Bruntun, J. (1998). Studies on the incidence of Ilarviruses and *Apple chlorotic leaf spot virus* (ACLSV) in apricot trees in the Murcia region (Spain) using serological and molecular hybridization methods. *Acta Horticulturae*, *472*, 203–210.

Fagan, T. (1975). Nomogram for Bayes theorem. *New England Journal of Medicine*, *293*, 257.

Ferris, N., King, D., Reid, S., Hutchings, G., Shaw, A., Paton, D., et al. (2006). Foot-and-mouth disease virus: A first inter-laboratory comparison trial to evaluate virus isolation and RT-PCR detection methods. *Veterinary Microbiology*, *117*, 130–140.

Hammond, R. W., Howell, W. E., Mink, G. I., & Crosslin, J. M. (1998). Strain-specific polymerase chain reaction assays for discrimination of *Prunus* necrotic ringspot virus isolates. *Acta Horticulturae*, *472*, 235–242.

Henson, J. M., & French, R. (1993). The polymerase chain reaction and plant disease diagnosis. *Annual Review in Phytopathoogy*, *31*, 81–109.

Herrera, M., & Madariaga, V. (2002). Incidencia de *Prunus necrotic ringspot virus* (PNRSV), *Prune dwarf virus* (PDV), *Tomato ringspot virus* (ToRSV)-*Y plum pox virus* (PPV) en viveros de frutales de carozo de la zona central de Chile. *Agricultura Tecnica*, *62*, 38–45.

James, D., Varga, A., Pallas, V., & Candresse, T. (2006). Strategies for simultaneous detection of multiple plant viruses. *Canadian Journal of Plant Pathology*, *28*, 16–29.

Jarrar, S., Myrta, A., di Terlizzi, B., & Savino, V. (2001). Viruses of stone fruits in Palestine. *Acta Horticulturae*, *550*, 245–248.

Jeffries, C., & James, C. (2005). Development of an EU protocol for the detection and diagnosis of Potato spindle tuber pospiviroid. *Bulletin OEPP/EPPO*, *35*, 125–132.

Josefsen, M., Cook, N., D'Agostino, M., Hansen, F., Wagner, M., Demnerova, K., et al. (2004). Validation of a PCR-based method for detectionc of food-borne thermotolerant campylobacters in a multicenter collaborative trial. *Applied and Environmental Microbiology*, *70*, 4379–4383.

Kummert, J., Vendrame, M., Steyer, S., & Lepoivre, P. (2000). Development of routine RT-PCR tests for certification of fruit tree multiplication material. *Bulletin OEPP/EPPO*, *30*, 441–448.

Landis, R., & Koch, G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*, 159–174.

Light, R. J. (1971). Measures of response agreement for qualitative data: Some generalizations and alternatives. *Psychological Bulletin*, *76*, 365–377.

Lopez, M., Llop, P., Gorris, M., Penalver, J., Donat, V., Cambra, M., et al. (2006). European protocol for diagnosis of *Erwinia amylovora*. *Acta Horticulturae*, *704*, 99–103.

Malorny, B., Tassios, P., Radström, P., Cook, N., Wagner, M., & Hoorfar, J. (2003). Standardization of diagnostic PCR for the detection of foodborne pathogens. *International Journal of Food Microbiology*, *83*, 39–48.

Myrta, A., di Terlizzi, B., Pallas, V., & Savino, V. (2002). Stato sanitario dell'albicocco nei paesi del bacino mediterraneo. *Rivista di Frutticoltura e di ortofloricoltura*, *64*, 91–95.

Paton, D., McGoldrick, A., Bensaude, E., Belak, S., Mittelholzer, C., Koenen, F., et al. (2000). Classical *Swine fever virus*, a second ring test to evaluate RT-PCR detection methods. *Veterinary Microbiology*, *77*, 71–81.

Pine, T. S. (1964). Influence of *Necrotic ringspot virus* on growth and yield of peach trees. *Phytopathology*, *54*, 504–505.

Rosner, A., Maslenin, L., & Spiegel, S. (1997). The use of short and long PCR products for improved detection of *Prunus necrotic ringspot virus* in woody plants. *Journal of Virological Methods*, *67*, 135–141.

Sanchez-Navarro, J. A., Aparicio, F., Herranz, M. C., Minafra, A., Myrta, A., & Pallas, V. (2005). Simultaneous detection and identification of eight stone fruit viruses by one-step RT-PCR. *European Journal of Plant Pathology*, *111*, 77–84.

Scott, S. W., Zimmerman, M. T., Yilmaz, S., Zehr, E. I., & Bachman, E. (2001). The interaction between *Prunus necrotic ringspot virus* and *Prune dwarf virus* in peach stunt disease. *Acta Horticulturae*, *550*, 229–236.

Spiegel, S., Scott, S. W., Bowman-Vance, V., Tam, Y., Galiakparov, N. N., & Rosner, A. (1996). Improved detection of *Prunus necrotic ringspot virus* by the polymerase chain reaction. *European Journal of Plant Pathology*, *102*, 681–685.

Taha, M., Alonso, J., Cafferkey, M., Caugant, D., Clarke, S., Diggle, M., et al. (2005). Interlaboratory comparison of PCR-based identification and genogrouping of neisseria meningitidis. *Journal of Clinical Microbiology*, *43*, 144–149.

Truyen, U., Wilhem, S., Genzow, M., & Schagemann, G. (2006). *Porcine reproductive* and *Respiratory syndrome virus* (PRRSV), a ring test performed in germany to assess RT-PCR detection methods. *Journal of Veterinary Medicine*, *53*, 68–74.