Evaluating the performance of qPCR primers for the detection of *Xanthomonas fragariae* with ROC curve analysis

William Turechek[1] & John Hartung[2]

[1]USDA-ARS Subtropical Plant Pathology Unit, Fort Pierce, FL; [2]USDA-ARS Fruit Laboratory, Beltsville, MD

Introduction

Angular leaf spot is an important disease of cultivated strawberry. The disease is transmitted primarily through systemically infected nursery stock. This creates problems for nurseries wishing to export plants to Europe because of quarantine restrictions. Currently, field inspections for symptoms are used to certify plants free of *X. fragariae*, but visual inspection is not useful for detecting plants infected systemically. To detect systemic infections, PCR is the desired tool because of its sensitivity, specificity, and ease of use. In previous work, we developed three sets of real-time PCR primers and probes (q241, q245, q295) and determined optimal reaction conditions for use of these primers for the detection of *X. fragariae* in strawberry crown tissue. The objectives of this study were to: 1) Evaluate the performance of the three primer pairs with Receiver Operating Characteristic (ROC) curve analysis; and 2) Provide information for the selection of optimal cutpoints for each primer set.

Materials and Methods

Primer sensitivity and specificity: The sensitivities of the qPCR primer pairs were determined individually from 10-fold serial dilutions of: 1) genomic DNA extracted from pure cultures of *X. fragariae*, 2) whole cell bacteria, and 3) DNA extracts from mixtures of strawberry crown tissue extract with whole cell bacteria. qPCR was performed on all dilution series to determine the sensitivity and specificity of the qPCR primer pairs. *X. fragariae* strains Xf3, Xf6, and Xf128 were used in all experiments. All reactions were performed in triplicate for each primer pair, and were run with a *X. fragariae* Xf3 positive control and a master-mix-only negative control.

ROC curve analysis: The data were partitioned into 2 groups based on whether they exceeded a predefined pathogen density or threshold, D_{thresh}. A '*case*', denoted as *D*+, is defined as a sample known to exceed the threshold value. A '*control*', denoted as *D*–, is a sample that is known not to exceed the threshold value. In this study, cases were prepared through serial dilutions, the controls were samples where no target DNA or bacteria was added or, where higher tolerances were permissible, the controls were samples where $D < D_{thresh}$. Cases and controls were partitioned according to one of three values of *D*: 0 (no tolerance), 10, and 100 cells. In performing qPCR, samples are classified as positive (*T*+) when the Ct value was less than an arbitrary cycle value called the cutpoint, T_{cp}. By convention, the cutpoint is often set to 35. I.e., if the fluorescence from the reaction does not exceed the background fluorescence by the 35th cycle, the sample is considered to be void of the target DNA and the sample is declared negative (*T*–), irrespective of its true status as *D*+ or *D*–. However, Ct values can often exceed 35 when very small quantities of the target are present, resulting in false negative classifications. Conversely, if a tolerance other than 0 is selected, Ct values less than 35 are to be expected. These shortcomings may be resolved by selecting a cutpoint other than 35 and this can be done through ROC curve analysis.

ROC curves: ROC curves were created for each predictor as a plot of the sensitivity versus '1-specificity' by allowing the cutpoint (T_{cp}), to vary over the continuous range of probable values. The '*sensitivity*' is the conditional probability $P(T+|D+)$ and was calculated by dividing the number of true positive reactions (*T*+) by the total number of cases. The '*specificity*' is the conditional probability $P(T-|D-)$ and was calculated by dividing the number of true negatives by the total number of controls. The area under the ROC curve (AUROC) and its standard error were calculated for each curve as implemented in Sigma Plot's (v. 10) ROC curve analysis routine. The calculations are based on the generalization that the AUROC curve is derived by dividing the Mann-Whitney *U*-statistic by the product of the sample sizes of the cases and controls. Pairwise comparisons of the AUROC curves were also done using Sigma Plot (v. 10). If the sensitivity and specificity are considered equally important, such that the goal is to minimize the overall error rate, then the cutpoint that comes closest to the point (0,1) is considered the optimal cutpoint. This cutpoint, T_J, was identified for each of ROC curve produced by calculating Youden's index: *J* = sensitivity + specificity – 1.

Results & Conclusions

The number of data sets classified as cases and controls, the Mann-Whitney *U*-statistic, the AUROC, its standard error and associated *z*-test statistic, the 95% confidence limits and Youden's index for real-time reactions are shown in the table below. Reactions with the largest AUROC are considered superior. No one primer pair was significantly superior according to analysis of their AUROCs. The optimal cutpoint for that reaction, as determined by *J*, was dependent upon the primer pair and D_{thresh}. However, it is clear the Ct value should be adjusted from 35 depending upon PCR target and the tolerance. These results are likely machine specific and should be repeated under different settings.

qPCR Test	D_{thresh}	Case	Control	*U*	AUROC	se_{AUC}	*z*	CI_{Lower}	CI_{Upper}	*J*
Purified DNA q241	0	45	9	381	0.9407	0.031	4.15	0.7713	0.9852	35.64
	10	36	18	628	0.9691	0.021	5.58	0.8625	0.9931	34.27
	100	27	27	716	0.9822	0.013	6.08	0.8931	0.9970	31.89
Purified DNA q245	0	45	9	357	0.8815	0.048	3.59	0.6920	0.9572	34.80
	10	36	18	560	0.8642	0.060	4.33	0.7191	0.9370	34.11
	100	27	27	692	0.9492	0.035	5.66	0.8419	0.9839	33.32
Purified DNA q295	0	45	9	400.5	0.9889	0.011	4.60	0.8470	0.9992	∞
	10	36	18	613.5	0.9468	0.029	5.31	0.8286	0.9837	34.26
	100	27	27	700.5	0.9609	0.032	5.81	0.8592	0.9892	32.88
Whole Cells q241	0	45	9	400.5	0.9889	0.011	4.60	0.8470	0.9992	∞
	10	36	18	531.5	0.8202	0.068	3.81	0.6667	0.9080	37.14
	100	27	27	641	0.8793	0.052	4.78	0.7480	0.9440	34.13
Whole Cells q245	0	45	9	378	0.9333	0.026	4.07	0.7607	0.9822	∞
	10	36	18	534	0.8241	0.058	3.85	0.6711	0.9106	37.41
	100	27	27	671	0.9204	0.035	5.31	0.8015	0.9687	34.80
Whole Cells q295	0	45	9	351	0.8667	0.049	3.45	0.6737	0.9492	35.31
	10	36	18	545.5	0.8418	0.052	4.06	0.6920	0.9225	34.38
	100	27	27	689	0.9451	0.028	5.61	0.8359	0.9818	34.03
Plant/Bacteria q241	0	45	9	355.5	0.8778	0.032	3.55	0.6874	0.9553	∞
	10	36	18	619.5	0.9560	0.025	5.42	0.8423	0.9878	37.44
	100	27	27	715	0.9808	0.013	6.07	0.8908	0.9966	34.64
Plant/Bacteria q245	0	45	9	364.5	0.9000	0.030	3.76	0.7157	0.9668	∞
	10	36	18	637.5	0.9838	0.015	5.75	0.8865	0.9977	37.17
	100	27	27	720	0.9877	0.010	6.15	0.9025	0.9984	34.88
Plant/Bacteria q295	0	45	9	373.5	0.9222	0.027	3.97	0.7453	0.9773	∞
	10	36	18	631.5	0.9745	0.018	5.64	0.8711	0.9950	35.56
	100	27	27	719	0.9863	0.010	6.13	0.9001	0.9981	34.65

Figure 1: ROC curves for qPCR reactions performed on serial dilutions of **A, B, C,** genomic DNA extracted from pure cultures of *Xf*, **D, E, F,** whole cells of *Xf*, and **G, H, I,** genomic DNA extracted from mixtures of crown tissue and bacteria. Reactions were conducted at the tolerance thresholds (*Dthresh*) **A, D, G,** *D* = 0, **B, E, H,** *D* = 10 and **C, F, I,** *D* = 100. The larger symbols shown in each figure represent the cutpoints identified by Youden's index as the optimal cutpoint and are listed in the Table.