
This is to certify that the

thesis entitled

XANTHOMONAS PATHOVARS IDENTIFICATION THROUGH A NEURAL NETWORK-BASED GENOMIC FINGERPRINT CLASSIFICATION SYSTEM

presented by

Fei Ni Tuang

has been accepted towards fulfillment
of the requirements for

M.S degree in Biosystems Engineering

EC Flores Jr
Major professor

Date June 26, 1998

| DATE DUE | DATE DUE | DATE DUE |
|---|---|---|
| MAR 2 3 2000 | | |
| | | |
| | | |
| | | |
| | | |

# *XANTHOMONAS* PATHOVARS IDENTIFICATION THROUGH A NEURAL NETWORK-BASED GENOMIC FINGERPRINT CLASSIFICATION SYSTEM

By

Fei Ni Tuang

A THESIS

Submitted to
Michigan State University
In partial fulfillment of the requirements
for the degree of

MASTER OF SCIENCE

Department of Agricultural Engineering

1998

ABSTRACT

*XANTHOMONAS* PATHOVARS IDENTIFICATION THROUGH A NEURAL NETWORK-BASED GENOMIC FINGERPRINT CLASSIFICATION SYSTEM

By

Fei Ni Tuang

A genomic fingerprint classification system was developed to identify 63 *Xanthomonas* pathovars. Three sets of genomic fingerprints generated from repetitive DNA sequence-based polymerase chain reaction (rep-PCR), using BOX, ERIC and REP primers, were used in this research. In addition, a fourth set of BER fingerprints was formed by linearly combining the BOX, ERIC and REP fingerprints. Mean and wavelet filter techniques were used to reduce noise on the fingerprints. Several backpropagation neural network (BPN) classifiers were trained using the BOX, ERIC, REP and BER original fingerprints and filtered fingerprints. Both mean and wavelet filtering helped improve the recognition rates. Wavelet filtering was better at reducing misclassification error rates, and mean filtering was better at reducing false rejection error rates. The average top-2 recognition rates of BOX, ERIC, REP and BER BPN classifiers were 95%, 93%, 92% and 98%, respectively. By combining the results of the BOX, ERIC and REP BPN classifiers with the lowest misclassification error rates, a top-1 recognition rate of 95% was achieved together with a misclassification error rate of 0.57% and a false rejection rate of 4.3%.

## ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my committee members: Dr. Evangelyn Alocilja (Biosystems Engineering), Dr. John Gerrish (Biosystems Engineering), and Dr. John Weng (Computer Science) for their insight, guidance and support in the course of this study.

My sincere thanks to Jan Rademaker and Dr. Frans de Bruijn of MSU DOE-Plant Research Laboratory for providing the genomic fingerprint data and giving me the opportunity to learn the rep-PCR genomic fingerprinting technique in the laboratory.

I am also appreciative for the funding provided by Dr. Robert von Bernuth, through the Manure Management and Nutrient Balance for Dairy project, and the fellowship provided by the College of Agriculture and Natural Resources.

I would like to thank the faculty, staff and students of the Department of Agricultural Engineering for providing a challenging learning environment. Special thanks to my friends at MSU: Thomas Moen, Tse-chia Yu, Erica Yang, Ismail Kavdir, Pankaj Jagtap, Jim Schäper and Takako Inagaki for their support and encouragement.

Finally, a very special thank you to my parents and my entire family for their never-ending guidance, support, encouragement and love.

## TABLE OF CONTENTS

## LIST OF TABLES

## LIST OF FIGURES

## LIST OF ABBREVIATIONS

DNA ..... Deoxyribonucleic acid

PCR ..... Polymerase chain reaction

rep-PCR ..... Repetitive DNA sequence-based PCR

BPN ..... Backpropagation neural network

GUI ..... Graphical user interface

Pathovar ..... Pathogenic variant

bp ..... Base pairs

IE ..... Information Engineering

ERD ..... Entity relationship diagram

DFD ..... Data flow diagram

SSE ..... Sum square error

SQL ..... Structured Query Language

# INTRODUCTION

Bacteria are one of the smallest life forms on Earth, but they have significant impacts on large organisms and on the ecosystem as a whole. Some bacteria, such as the *Xanthomonas* species, are plant pathogens. They can cause diseases in a wide variety of agriculturally and economically important plants. *Xanthomonas* infections of fruits and vegetables cause considerable loss in yield (Hayward, 1993). Identification of the bacterial species and pathogenic variants (pathovars) will lead to disease control, and prevention of losses in crop yield. In addition to its crucial role in pathological studies, bacterial identification also plays an important role in ecological studies, biodiversity assessment and microbial screening for industrial application.

Among the latest developments in bacterial taxonomy are: (1) molecular identification techniques based on genotypic features and (2) computer-assisted classification with numerical analysis (Vauterin et al., 1993). Repetitive deoxyribonucleic acid (DNA) sequence-based polymerase chain reaction (rep-PCR) genomic fingerprinting technique is capable of rapidly generating large quantities of highly discriminated DNA fingerprints. The fingerprinting method, coupled with computer-assisted cluster analysis, was found to successfully classify and identify *Xanthomonas* pathovars (Rademaker et al., 1997). In cluster analysis, a similarity matrix of the different fingerprints is used to generate a dendrogram. In the dendrogram, strains of bacteria belonging to the same pathovar are clustered together. A new bacterial strain is identified by comparing its

fingerprint to fingerprints of other pathovars in a reference library. The bacterial strain is assigned to the pathovar with the closest match based on the dendrogram (Rademaker and de Bruijn, 1997).

Backpropagation neural network (BPN) is a type of artificial neural network that is suitable for use as a pattern classifier. The BPN is a feedforward network made up of a collection of interconnected nodes or processing elements (Castleman, 1996). It can be trained to recognize and classify complex patterns. Using a supervised training approach, training data consisting of input and desired output pairs are presented to the network repetitively, and the connection weights are adjusted until the error between the network output and desired output converges (Haykin, 1994). A network can be used to recognize patterns for which it is trained.

The use of rep-PCR genomic fingerprints with BPN can be an efficient approach to identify newly isolated bacterial strains. A trained network can identify bacterial pathovars based on its internal representation, without comparing them to other fingerprints in the reference library. The massive computation needed to learn the classifications is done only once during the training phase. The advantage of using the BPN approach over the cluster analysis approach will be appreciated when the size of the reference library and the number of strains that needs to be identified are large.

Noise and variability in the genomic fingerprints are factors affecting the accuracy of BPN classification. Noise is introduced in the process of fingerprinting, and is caused by some physical, chemical and biological factors in PCR reactions, gel electrophoresis and photography. Filtering techniques such as local averaging mean filter

and wavelet analysis offer possible noise reduction. Classification accuracy may be improved with the use of these filtering techniques.

Since the amount of data to be organized and processed for BPN training, testing, and classification, is large, an information system is essential. Graphical visualization capability is also necessary for visualizing fingerprint patterns and examining within class variations. In addition to a classifier component, a good classification system should have an integrated database, data processing and data visualization components, and a graphical user interface (GUI) that links every component together and serves as a means for the user to communicate with the system. This will lead to a more efficient and standardized classification procedure since every component is built into the system.

## Objectives

The objectives of this research are:

1. To build a rep-PCR genomic fingerprint classification system that is rapid, accurate and easy to use.
2. To investigate the effect of noise reduction on classification accuracy.
3. To evaluate the performance of the several BPN classifiers in identifying *Xanthomonas* pathovars.

# Chapter I

# LITERATURE REVIEW

## 1.1 Bacterial Taxonomy and Classification Systems

Taxonomy is the study of the classification of living things (King and Stanfield, 1997). It involves the study of relationships among living things and their arrangement into categories. Bacterial classification is important for several reasons. Firstly, it serves as a means of summarizing and cataloguing information about an organism. Secondly, it forms the basis for the recognition of new isolates. Thirdly, it provides an insight into the origin and evolutionary pathways of bacteria and higher organisms (Priest and Austin, 1993). Fourthly, it is fundamental in understanding ecological processes in a biosphere with diverse species and complex biological interactions. Fifthly, it plays an important role in disease control and microbial screening for industrial applications (Towner and Cockayne, 1993).

To effectively meet the classification objectives, a classification method must satisfy the following three criteria: high stability, good predictivity and high objectivity (Priest and Austin, 1993). While biological classifications are subject to change in time (this is unavoidable due to the discovery of species or the collection of new data), there are still some measures that can be taken to enhance classification stability. Firstly, all relevant objects should be included before trying to form groups. Secondly, the

significant characteristics of the objects and all their variations should be included in the study (Pankhurst, 1991). To enhance predictivity and to include more generalizations about the taxa involved, classifications should be based on high information content. To increase objectivity, the development of a classification should be empirical, reproducible and scientifically based (Priest and Austin, 1993). While the development of a classification system is often data dependent, these criteria and measures provide some general guidelines.

The development of bacterial taxonomy can be traced back to the seventeenth century in the work by Morison in 1672 and Ray in 1688 (Pankhurst, 1991). The early classifications were based on morphological information as obtained from light microscopy. Following the interest in morphology, pure culture techniques were developed to study bacterial metabolism and physiological information (Priest and Austin, 1993). The traditional bacterial classifications were mainly based on phenotypic properties such as biochemical, nutritional and physiological features. In the 1960s and early 1970s, numerical taxonomy was developed to provide more objective analysis and to automate the process (Goodfellow et al., 1997). Some computer-assisted numerical classification methods include multi-access and hypertext key, expert systems (Edwards and Morse, 1995), matching and clustering using similarity coefficient, maximum likelihood classification, discriminant analysis (Pankhurst, 1991) and neural networks (Weeks and Gaston, 1995). With computer-assisted classification, a comparison of large numbers of features and strains is allowed (Vauterin et al., 1993). The numerical methods greatly enhance classification accuracy and efficiency.

In a survey paper by Wilkins et al. (1996), two statistical methods, including parametric Gaussian multivariate statistical method (GAUSS) and non-parametric K-nearest neighbour method (KNN), and four artificial neural network paradigms, including multilayer perceptron networks (MLP), learning vector quantization networks (LVQ), radial basis function network (RBF) and asymmetric radial basis function network (ARBF), were compared for their ability to identify phytoplankton species from flow cytometry data. From the results of the studies, all classifiers had similar identification success, with recognition rate within 3.4% of each other. The relative merits of each classifier were compared. The MLP was found to be reliable, robust, simple-to-use and rapid in operation, but had a longer training time in comparison with other methods. Training is not required for the KNN classifier. However, it is slow in operation and requires storage of a large representative training set (Wilkins et al., 1996). The MLP is the most popular class of multilayer feed-forward network, which is also called backpropagation neural network (BPN) after its back-propagation learning algorithm (Jain et al., 1996).

Recent advances in molecular biology make possible the classification of bacteria using detailed genotypic properties. Some nucleic acid methods, such as DNA-DNA hybridization, DNA-RNA hybridization and nucleic acid sequencing, are widely used in microbial taxonomy (Macdonell and Colwell, 1985). In addition, rep-PCR technique was also found to be reliable, reproducible and rapid in generating highly discriminated fingerprints for bacterial identification and classification (Rademaker and de Bruijn, 1997). In contrast to the traditional classifications that are based mainly on a few

phenotypic properties, the molecular identification and typing methods lead to more general and natural classifications of bacterial species (Towner and Cockayne, 1993).

## 1.2 The Genus *Xanthomonas*

*Xanthomonas* species are plant-associated bacteria. *Xanthomonas* infections occur on at least 124 monocotyledonous and 268 dicotyledonous plant species (Leyns et al., 1984). Some examples of the host plants are cotton (Zomorodian and Rudolph, 1993), rice (Mew, 1993), mung bean (Vidaver, 1993), soybean (Hokawat and Rudolph, 1993), citrus fruit (Stall and Civero, 1993), cabbage, broccoli, cauliflower (Schaad and alvarez, 1993), forage grasses (Leyns, 1993), tomato, pepper (Stall, 1993), apricot, cherry, almond, peach, nectarine, plum (Civerolo and Hattingh, 1993), poplar (Ride, 1993), strawberry (Rat, 1993), sugarcane (Rott, 1993), barley, rye, wheat (Duveiller and Maraite, 1993) and mango (Pruvost and Manicom, 1993). Most of these are important agricultural crops and fruits grown in different parts of the world. The disease symptoms include necrosis, gummosis and vascular or parenchymatous diseases on leaves, stems or fruits of the plants. The disease may cause significant loss in yield (Hayward, 1993).

*Xanthomonas* cells are Gram-negative rods, measuring 0.2-0.6μm by 0.8-2.9μm (Swings et al., 1993). When cultivated on common growth media, most *Xanthomonas* strains form yellow, water-insoluble pigments called Xanthomonadins (Swings et al., 1993). Most *Xanthomonas* species have a polar flagellum and are strictly aerobic chemoorganotrophs (Leyns et al., 1984). The genus *Xanthomonas* was proposed by Dowson in 1939 and is classified under the superkingdom *Prokaryota*, kingdom *Monera*, subkingdom *Eubacteria*, division *Gracilicutes*, phylum *Proteobacteria* and subclass *Gamma* (Margulis and Schwartz, 1998; Stackebrandt et al., 1988). Seven species are

recognized within the genus *Xanthomonas*, namely: *X. albilineans*, *X. axonopodis*, *X. campestris*, *X. fragariae*, *X. populi*, *X. maltophilia* (not a plant pathogenic species) and *X. oryzae*. Nearly every new species and subspecies or pathogenic variant (pathovar) of *Xanthomonas* was named after the host plant (Vauterin et al., 1993). As a result, some of classifications may not reflect genomic relationships. *Xanthomonas* was reclassified by Vauterin et al. (1995) based on a comprehensive study of 183 strains using DNA-DNA hybridization technique. The genus was shown to comprise 20 genomic species. Both the genomic relationships and the needs of plant pathologists for a rational nomenclature are taken into account in the new classification (Vauterin et al., 1995).

### 1.3 rep-PCR Genomic Fingerprinting

rep-PCR genomic fingerprinting is a DNA amplification-based technique for generating bacterial fingerprints that can be used for species identification. It consists of the following steps: (1) bacterial cells and DNA isolation, (2) rep-PCR amplification, (3) gel electrophoresis, (4) gel image capture and (5) image processing and analysis (Rademaker and de Bruijn, 1997).

Polymerase chain reaction (PCR) was devised by Mullis in 1984 for amplifying specific DNA sequences (Watson et al., 1992). A PCR cycle consists of three stages, namely, strand separation, hybridization of primers, and DNA synthesis (Stryer, 1995). To prepare the mixture for PCR, two oligonucleotide primers, a DNA polymerase, and all four deoxyribonucleoside triphosphates (dNTPs) are added to the target DNA. When heated at a temperature of about 94°C for one minute, the double-stranded DNA molecules separate completely, forming single strands that become the templates for the primers and DNA polymerase. The solution is then cooled to about 50°C, a temperature

that allows the primers to anneal to the complementary sequences in the DNA molecules. With the primed templates generated, the temperature is then raised to about 65°C, the optimal temperature for the DNA polymerase. 65°C is maintained for eight minutes for the DNA synthesis to proceed. These three steps are repeated for 30 cycles. All newly synthesized strands serve as templates in each successive cycle; DNA sequences are amplified as a result. The amplification is a billionfold after 30 cycles (Watson et al., 1992; Stryer, 1995; Rademaker and de Bruijn, 1997).

In rep-PCR, DNA primers complementary to naturally occurring, highly conserved repetitive DNA sequences present in the genomes of most Gram-negative and several Gram-positive bacteria are used. They include three families of repetitive DNA sequences, namely the 35-40 base pairs (bp) repetitive extragenic palindromic (REP) sequence, the 124-127 bp enterobacterial repetitive intergenic consensus (ERIC) sequence, and the 154 bp BOX element. The use of these primers in PCR leads to the selective amplification of distinct genomic regions between REP, ERIC or BOX elements. The collective name for this protocol is called rep-PCR, and individually can be referred to as REP-PCR, ERIC-PCR and BOX-PCR with respect to the different primers (Rademaker and de Bruijn, 1997).

Since the amplified fragments of DNA are of different sizes, they can be resolved in a gel matrix using gel electrophoresis to generate fingerprints (Rademaker and de Bruijn, 1997). In gel electrophoresis, negatively charged DNA molecules move from the negative pole toward the positive pole of the gel. The separation depends on the molecular weight of the molecules. Smaller molecules move faster than larger molecules. At the end of the gel electrophoresis, which usually runs for 18 hours, distinct bands of

DNA can be found (Hartl, 1994; Rademaker and de Bruijn, 1997). The rep-PCR genomic fingerprints generated through PCR and gel electrophoresis permit differentiation to the species, subspecies and strain level (Rademaker and de Bruijn, 1997).

The last step of rep-PCR genomic fingerprinting involves making a hard copy of the fingerprint. The gel is stained and a photograph of the gel is taken. The gel image is processed and analyzed using numerical imaging tools (Rademaker and de Bruijn, 1997). GelCompare (Applied Maths, Kortrijk, Belgium) is a software package designed for gel image processing and analysis.

## 1.4 Digital Image Processing

Image digitization is the first step towards machine pattern recognition. Physical images are converted into numerical representations by digitization. In the process, an image is divided into small regions called pixels and at each pixel location, the brightness of the image is sampled and quantized into a numerical gray level (Castleman, 1996). A digital image can be seen as an array of different gray levels.

Digital images are used as inputs for pattern recognition systems. The design of a pattern recognition system includes the following five steps, namely: image segmentation, feature selection, classifier design, classifier training and performance evaluation. Image segmentation involves the isolation of individual objects to be recognized from the rest of the scene, using thresholding and edge detection techniques. Feature selection involves choosing object properties that can best distinguish the various classes of objects among themselves (Castleman, 1996).

Several techniques can be used to enhance image quality and reduce undesired noise. Mean filter is one of the simplest linear filters. It is implemented by a local

averaging operation where the value of each pixel is replaced by the average of all the values in the local neighborhood of size N (Jain et al., 1995):

$$h(i) = \frac{1}{N} \sum_{k=i-M}^{i+M} f(k) \tag{1.1}$$

where h is the filtered image, f is the original image, i is the pixel location and $M = \lfloor \frac{N}{2} \rfloor$. The size of the neighborhood N is directly proportional to the amount of filtering. However, there is a trade-off between noise reduction and image detail. Sharp detail in the image is often lost and steep changes will be blurred into gradual changes as a result of linear filtering (Jain et al., 1995).

Wavelet analysis offers another technique for noise reduction. Wavelet multiresolution analysis divides the frequencies of an image into octave bands, and the image can be analyzed at different scales (Strang and Nguyen, 1996). The wavelet de-noising procedure consists of the following three steps: (1) multilevel wavelet decomposition of the image, (2) detail coefficients thresholding at each level, and (3) wavelet reconstruction based on the original approximation coefficients and modified detail coefficients (Donoho, 1995).

Discrete wavelet transform involves the decomposition of an image space into orthogonal subspaces (Strang and Nguyen, 1996). Let $V_0 \supset V_1 \supset V_2 \supset \ldots \supset V_J$ be a sequence of J nested subspaces, from fine to coarse scales. At each level j, the low frequency approximation of an image is captured in the scaling subspace $V_j$, and the high frequency detail of an image is captured in the complement wavelet subspace $W_j$. Each subspace $V_{j-1}$ is the direct sum of the coarser subspace $V_j$ and its orthogonal complement $W_j$. The multiresolution decomposition of $V_0$, the image space, can be written as

$$V_0 = W_1 \oplus W_2 \oplus \ldots \oplus W_J \oplus V_J \tag{1.2}$$

(Liang, 1997). A wavelet decomposition tree is shown in Figure 1.1, where S is the image, and $A_j$ and $D_j$ represent the approximations and details, respectively (Misiti et al., 1996).

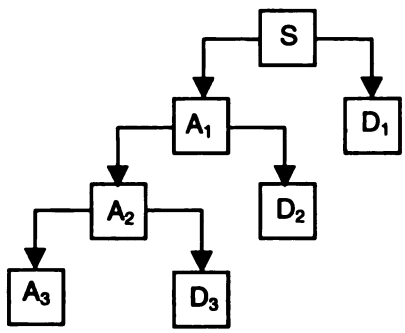



**Figure 1.1** Wavelet Multilevel Decomposition Tree

$\{\phi_{j,k}\}$ and $\{\psi_{j,k}\}$ form the orthonormal bases for the subspaces $V_j$ and $W_j$, respectively. A family of scaling basis functions can be obtained through dilations and translations of a scaling function as shown in Equation 1.3:

$$\phi_{j,k} = 2^{-j/2}\phi(2^{-j}x - k) \tag{1.3}$$

where $\phi$ is the scaling function, $x$ is the pixel location, and $j$ and $k$ are integers representing the scales and positions (Misiti et al., 1996). Similarly, a family of wavelet basis functions can be obtained as shown in Equation 1.4:

$$\psi_{j,k} = 2^{-j/2}\psi(2^{-j}x - k) \tag{1.4}$$

where $\psi$ is the mother wavelet and $x$, $j$ and $k$ are as previously defined. Some examples of wavelet families are Haar, Daubechies, Biorthogonal, Coiflets, Symlets, Morlet, Mexican Hat and Meyer (Misiti et al., 1996). The wavelet basis functions are dually localized in both time and frequency domains (Barclay and Bonner, 1997).

The approximation wavelet transform coefficients $a_{j,k}$ are given by the inner product $<s, \phi_{j,k}>$ and the detail wavelet transform coefficients $c_{j,k}$ are given by the inner product $<s, \psi_{j,k}>$, where $s$ is the image (Antonini et al., 1992). The discrete wavelet transform can be implemented as a filtering algorithm using the following equations:

$$c_{j,k} = \sum_n g_{n-2k} a_{j-1,n} \tag{1.5}$$

$$a_{j,k} = \sum_n h_{n-2k} a_{j-1,n} \tag{1.6}$$

which demonstrate that the coefficients $c_{j,k}$ can be obtained by convolving the coefficients $a_{j-1,k}$ with $g_{-n}$ followed by decimation, in which only the even indexed elements are kept. Similarly, the coefficients $a_{j,k}$ can be obtained by convolving the coefficients $a_{j-1,k}$ with $h_{-n}$ followed by decimation (Liang, 1997). $g_{-n}$ and $h_{-n}$ are high-pass and low pass filters, respectively.

After the decomposition at each level, thresholding can be applied to high frequency detail coefficients to suppress high frequency noise. With a selected threshold value, any detail coefficient whose value is below the threshold will be set to zero (Donoho and Johnstone, 1995). The image can then be reconstructed with the noise-reduced wavelet transform coefficients using inverse wavelet transform. The reconstruction is shown as follows:

$$a_{j-1,n} = \sum_k g_{2k-n} c_{j,k} + \sum_k h_{2k-n} a_{j,k} \tag{1.7}$$

which demonstrates that $a_{j-1,n}$ can be obtained by upsampling (zero-padding) $c_{j,k}$ and $a_{j,k}$, followed by convolving $c_{j,k}$ and $a_{j,k}$ with $g_n$ and $h_n$ respectively, and taking the sum of the convolutions (Strang and Nguyen, 1996). $g_n$ and $h_n$ are the reconstruction high-pass and low-pass filters, respectively. The image can be reconstructed by passing the wavelet

transformed coefficients through a series of the same reconstruction filters (Misiti et al., 1996).

As compared with other noise-filtering techniques, wavelet analysis does not reduce the sharpness of edges, as long as the wavelet transform of the noise is smaller than the significant features in the image (Weaver et al., 1991).

## 1.5 Backpropagation Neural Network

The neural network approach has been applied successfully for pattern classification in several areas of biological research. Examples of such applications are the protein classification artificial neural system (Wu et al., 1992), backpropagation and counter propagation neural networks for phylogenetic classification of ribosomal RNA sequences (Wu and Shivakumar, 1994), classification of mass spectra (Curry and Rumelhart, 1990), identification of marine phytoplankton species using flow cytometric data (Body et al., 1994), and identification of gram-negative rods using phenotypic characters (Schindler et al., 1994).

Inspired by the structure of a human brain, a BPN is made up of interconnected basic processing units called neurons (Rumelhart et al., 1994). The model of a neuron is shown in Figure 1.2.

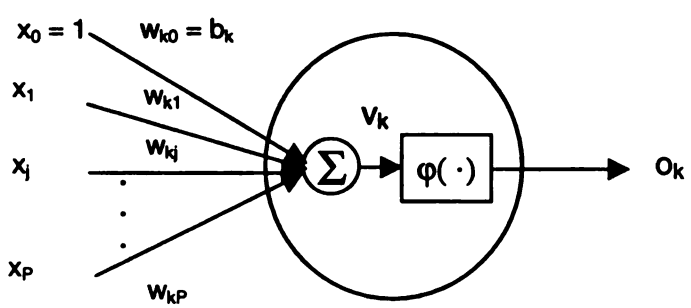



**Figure 1.2** Model of a Neuron

For the $k^{th}$ neuron,

$$v_k = \sum_{j=0}^{P} w_{kj} x_j \tag{1.8}$$

$$o_k = \varphi(v_k) = \frac{1}{1+e^{-v_k}} \tag{1.9}$$

where $b_k$ is the bias, $x_j$ is the $j^{th}$ input, $w_{kj}$ is the connection weight between the $j^{th}$ input and the neuron, $v_k$ is the activation potential, $\varphi(.)$ is the activation function, and $o_k$ is the output of the neuron (Haykin, 1994). The bias represents the internal resting value of a neuron (Rumelhart et al., 1994). The activation function defines the output of a neuron with respect to the activation potential. The nonlinear sigmoid function shown in Equation 1.9 is the most commonly used activation function in BPN (Haykin, 1994).

A BPN is formed by arranging the neurons into a multilayer feedforward network. The structure of a BPN is shown in Figure 1.3. The BPN is made up of an input layer, one or several hidden layers and an output layer (Haykin, 1994). Neurons in input, hidden and output layers are called input units, hidden units and output units, respectively (Rumelhart et al., 1994)

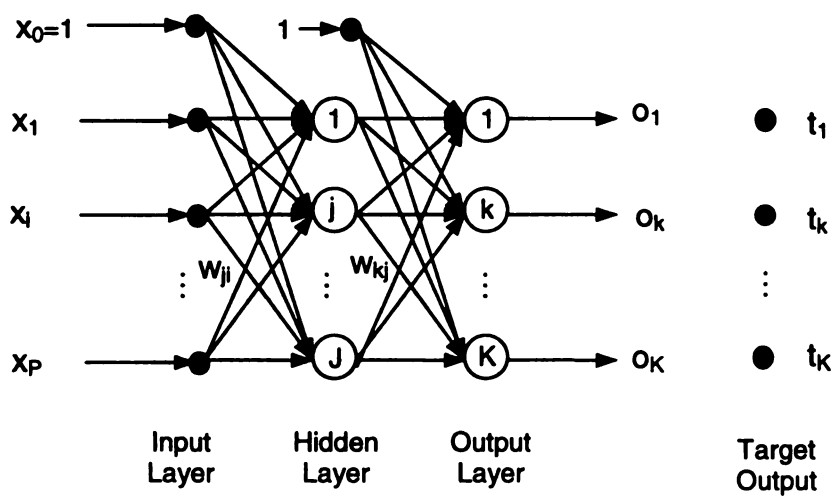



**Figure 1.3** Structure of a BPN

The purpose of the BPN training is to find a set of connection weights that enables the network to perform the desired nonlinear mapping from input space to the output space (Rumelhart et al., 1994; Skapura, 1996). The supervised training of a BPN consists of two passes, a forward pass and a backward pass. In the forward pass, the input signals from the input units propagate through the layers of the network and generate the actual network outputs at each of the output units. The error signal at the $k^{th}$ output unit for the $n^{th}$ training pattern is defined by

$$e_k(n) = t_k(n) - o_k(n) \tag{1.10}$$

where t is the target output and o is the network output (Haykin, 1994).

The energy of the network is a function of the error signals, and it can be expressed on pattern-by-pattern basis as:

$$E(n) = \tfrac{1}{2}\sum_{k=1}^{K} e_k^{\,2}(n) \tag{1.11}$$

where K is the number of output units (Haykin, 1994). In the backward pass, the output layer weights and hidden layer weights are adjusted based on the error signals so as to minimize the energy of the network. The weight update is defined by the delta rule:

$$\Delta w_{ji}(n) = -\eta \frac{\partial E(n)}{\partial w_{ji}(n)} + \alpha \Delta w_{ji}(n-1) \tag{1.12}$$

where $\eta$ is the learning rate, $\alpha$ is the momentum, and the use of the minus sign before the gradient term accounts for gradient descent in weight space (Haykin, 1994). For an output unit with sigmoidal activation function, the weight update equation is

$$w_{kj}(n+1) = w_{kj}(n) + \alpha[w_{kj}(n) - w_{kj}(n-1)] + \eta \delta_k(n) o_j(n) \tag{1.13}$$

where $\delta_k(n) = e_k(n)o_k(n)[1-o_k(n)]$. Similarly, for a hidden unit with sigmoidal activation function, the weight update equation is

$$w_{ji}(n+1) = w_{ji}(n) + \alpha[w_{ji}(n) - w_{ji}(n-1)] + \eta\delta_j(n)o_i(n) \qquad (1.14)$$

where $\delta_j(n) = o_j(n)[1-o_j(n)]\sum_k \delta_k(n)w_{kj}(n)$, and $\delta_k$ and $\delta_j$ are the local gradients of neurons in the output and hidden layers, respectively (Haykin, 1994). Randomly selected training patterns are presented to the BPN at each training iteration. The training is run repetitively until the mean square error over the entire training set converges to an acceptable minimal value.

An alternative to the pattern mode training described above is the batch mode training. One complete presentation of the entire training set in the BPN training process is called an epoch, and in the batch mode, weights updating is performed after every epoch instead of after every pattern (Haykin, 1994). The use of batch mode training allows the true gradient to be calculated (Demuth and Beale, 1994).

The stopping criterion is critical in the training. A network trained too specifically to the training data may not generalize well to untrained data sets. Cross-validation is a common scheme that avoids the problem of overtraining (Rumelhart et al., 1994). In cross-validation, a set of untrained testing data is presented to the network, following each training iteration, to evaluate the generalization performance. The training is stopped when an increase in generalization error is found (Schittenkopf et al., 1997).

Learning rate and momentum are two parameters that control BPN learning. The learning rate determines the amount of weight adjustment. A high learning rate may accelerate the learning, but it may also cause instability. The momentum is added to stabilize the training and to keep the weight adjustment moving in a steady direction

(Haykin, 1994; Freeman, 1994). Vogl et al. (1988) introduced the method of adaptive learning rate and momentum to accelerate the BPN convergence. Starting with a small learning rate and large momentum, the learning rate is increased continuously by a small factor when the error is decreasing smoothly. However, when there is a rise in the error, the learning rate is decreased by a large factor and the momentum is set to zero. A faster convergence and more stable learning can be achieved with this method.

The knowledge of a BPN is stored in the final connection weights obtained through the training. With the final connection weights, the BPN can be used for pattern recognition and classification. In the BPN application stage, only the forward pass is required to generate the corresponding outputs.

## 1.6 Performance Evaluation

Evaluation of classifier performance is needed to estimate the classification accuracy and expected error rates (Castleman, 1996). Classification errors include misclassification error, false rejection error and false acceptance error. In addition to evaluating the performance of the classifier, the quality of the training data is also being evaluated in the process (Skapura, 1996). A testing set is needed for the evaluation. The partitioning of the available data into two equal halves, one for training and one for testing is the most common method. To improve classification accuracy, classifiers trained on different features can be combined to yield better results (Cho, 1997).

## 1.7 Systems Analysis

Whitten and Bentley (1998) introduced systems development life cycle, a systematic and orderly approach to solving system problems. A diagram of the systems development life cycle is shown in Figure 1.4 (Whitten and Bentley, 1998).

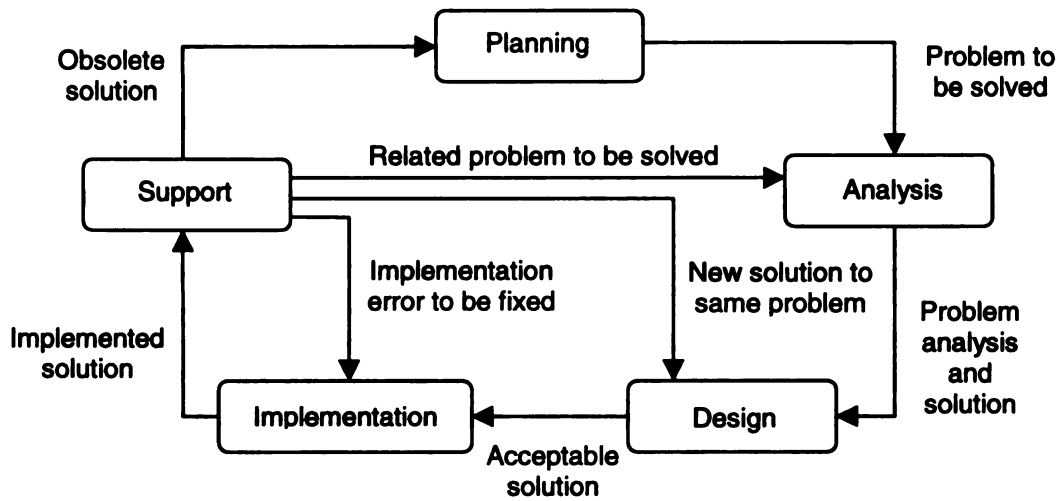



**Figure 1.4** Systems Development Life Cycle

The five general problem solving steps are planning, analysis, design, implementation and support. Planning involves identifying the scope and boundary of the problem and formulating the development strategy and goals. Analysis involves studying and analyzing the problems, causes and effects, as well as identifying and analyzing the requirements that must be fulfilled by any successful solution. Design involves designing the solution. Implementation involves implementing the solution. And lastly, support involves analyzing the implemented solution, refining the design, and implementing improvements to the solution (Whitten and Bentley, 1998).

Wetherbe and Vitalari (1994) developed a problem-solving framework, PIECES, consisting of six categories:

*P – the need to improve Performance*

*I – the need to improve Information and data*

*E – the need to improve Economics and control costs*

*C – the need to improve Control or security*

*E – the need to improve Efficiency of people and processes*

*S – the need to improve Service*

The PIECES framework outlined some of the underlying needs and challenges in system development.

Structured analysis was one of the first formal strategies developed for systems analysis of information systems and computer applications (Whitten and Bentley, 1998). Modern structured analysis is a process-centered technique, using a series of process models to represent the essential processes of a system along with inputs, outputs and storage (Yourdon, 1989). Another widely used approach for systems analysis is Information Engineering (IE). It evolved from structured analysis. IE is a data-centered, but process-sensitive technique (Whitten and Bentley, 1998). Since information is a product of data, the study and definition of data requirements are emphasized before those of process and interface requirements. Both data models and process models similar to those in structured analysis are used in IE (Whitten and Bentley, 1988).

Data modeling is a technique for organizing and documenting the data of a system. It defines the requirements for a database. An entity relationship diagram (ERD) is used to depict data in terms of the entities and relationships described by the data. Some concepts and definitions related to data modeling are summarized in Table 1.1:

**Table 1.1** Data Modeling Concepts and Definitions (Whitten and Bentley, 1998)

| Term | Definition |
|---|---|
| Entity | a class of objects, events or concepts about which we need to capture and store data. |
| Attribute | a descriptive property or characteristic of an entity. |
| Relationship | an association that exists between one or more entities. |
| Cardinality | a definition of the minimum and maximum number of occurrences of one entity for a single occurrence of the related entity. Cardinality must be defined in both directions for every relationship, since all relationships are bidirectional. |
| Primary key | an attribute that assumes a unique value for each entity instance. |
| Foreign key | a primary key duplicated in another entity to identify instances of a relationship. |

Processes transform data into useful information. A process model depicts the structure and flow of data through a system and the work performed by the system. It is made up of the following four components, namely: external agent, process, data store, and data flow. External agents provide the net input to a system and receive net outputs from a system. A complex system can be decomposed into smaller subsystems for improved communication, analysis and design. Through process modeling, the interactions of a system with its environment and the interactions among processes within the system are described (Whitten and Bentley, 1998).

The systems analysis, data modeling and process modeling methods can be applied to the design of a genomic fingerprint classification system.

# Chapter II

# MATERIALS AND METHODS

## 2.1 Data Collection

The rep-PCR genomic fingerprints data used in this research were obtained from MSU DOE-Plant Research Laboratory. There were three sets of 752 fingerprints from 376 bacterial strains, and they were generated using BOX, ERIC and REP primers in the rep-PCR reactions. The data represented 89 *Xanthomonas* pathovars. Each bacterial strain in the BOX, REP and ERIC data sets had a duplicate set of fingerprints generated from two different assays. A list of the data is shown in Table A.1 of Appendix A.

The fingerprint gel images were digitized using a flatbed scanner and pre-processed in GelCompar (version 3.1 Applied Maths, Kortrijk, Belgium, no endorsement implied), a window-based gel image processing software. Individual tracks of bacterial fingerprints, each representing a bacterial strain, were segmented from the gel image. The individual fingerprints were normalized with respect to two reference tracks, and background subtraction was performed to enhance image quality (Rademaker et al., 1997). The normalized rep-PCR genomic fingerprints were exported from GelCompar. Each fingerprint was represented by 400 data points, with each point having a gray level ranging from 0 to 255. A fingerprint can be viewed as a densitometric curve. A sample BOX-PCR genomic fingerprint is shown in Figure 2.1.

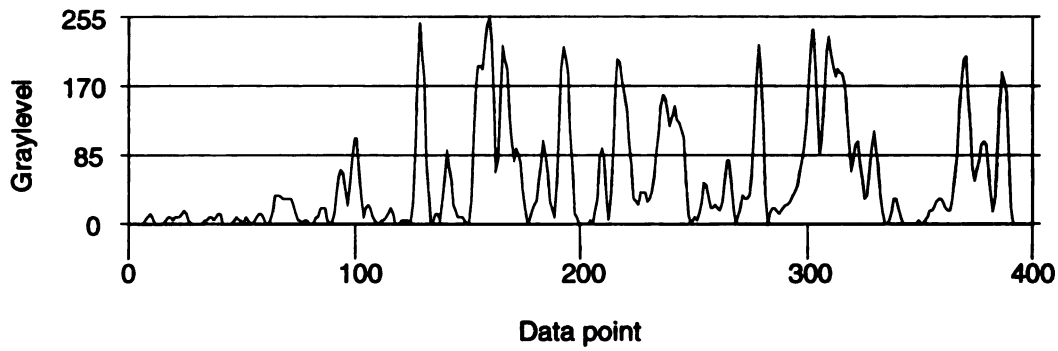



**Figure 2.1** BOX-PCR Genomic Fingerprint (LMG 12141)

Furthermore, the BOX, ERIC and REP PCR genomic fingerprints of the same bacterial strain were linearly combined to form the BER fingerprints. The BER fingerprint therefore, of one bacterial strain was composed of 1200 points. Then, the 1200 points were compressed to 400 points by replacing every three neighboring points with their arithmetic mean. A sample BER fingerprint is shown in Figure 2.2.

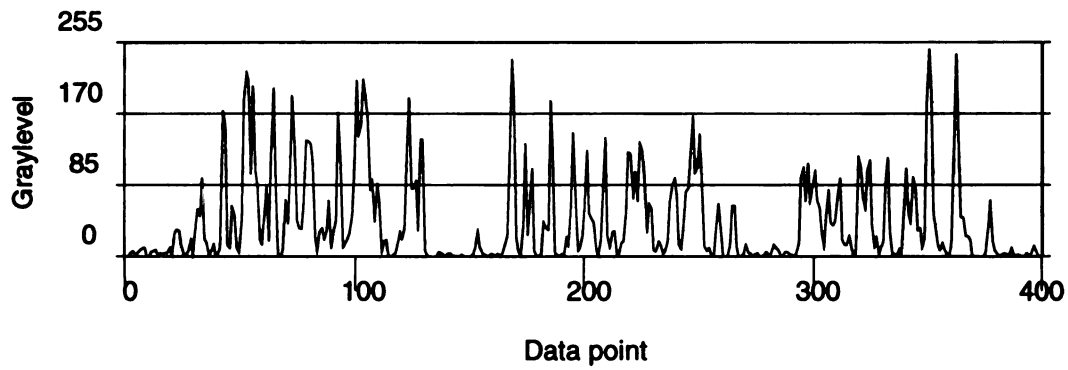



**Figure 2.2** BER Fingerprint (LMG 12141)

The rep-PCR genomic fingerprint data were in ASCII format and the structure of the data files can be found in Figure A.1 of Appendix A.

## 2.2 Systems Analysis

A systematic approach is needed to design and develop a rep-PCR genomic fingerprint classification system. The development of the rep-PCR genomic fingerprint classification system described here follows the systems development life cycle shown in Figure 1.4. The ultimate goal of the rep-PCR genomic fingerprint classification system is to identify bacterial pathovars and assign them to the corresponding pathovar classifications defined in the system. In the planning stage, PIECES problem-solving framework (Wetherbe, 1994; Whitten and Bentley, 1998) was used for problem identification, and the needs are summarized below:

(1) The need to improve performance

- The genomic fingerprint classification system should be fast and accurate. The current system using cluster analysis method classifies 50 bacterial strains in 14 seconds on a Pentium II 266 PC with 96MB of RAM.
- The need to improve information and data
- The inputs including pathovar names, strains, fingerprints, indexes and others should be captured accurately and not redundantly.
- The outputs of the system should be in a useful format, with accurate, necessary and relevant information.
- Data should be well organized and accessible.
- Fingerprints graphical visualization capability should be available.

(2) The need to improve economics and control costs

- The costs for building the system and implementing the system should not be too high. The prices of the GelCompare basic software, the cluster analysis

module and the identification module (developed by Applied Maths, Kortrijk, Belgium) are $2600, $1800 and $1400, respectively.

(3) The need to improve control or security

- Redundantly stored data should be consistent in different files or databases.
- Processing errors should be controlled.

(4) The need to improve efficiency of people and processes

- Routine data processing procedures should be automated.

(5) The need to improve service

- The system should produce accurate, consistent and reliable classifications.
- The system operating procedures should be easy to learn and use.
- The system should be flexible to adapt to new situations and change.
- The system should coordinate well with external packages.
- The system should be compatible with other systems.

In the development process, systems planning was followed by systems analysis. The problems identified in the planning stage were studied and analyzed, and the data and process requirements were defined with the use of data modeling and process modeling.

### 2.2.1 Data modeling

An entity relationship diagram (ERD) of the rep-PCR genomic fingerprint data is shown in Figure 2.3. The data entities are pathovar name, strain, BOX fingerprint, ERIC fingerprint, REP fingerprint, target pathovar, non-target pathovar, BOX classification, ERIC classification and REP classification. The attributes of each entity are listed in the respective boxes in the diagram. PK denotes primary key and FK denotes foreign key.

The line connecting two entities asserts the existence of a relationship. The definitions of entities, attributes, primary key and foreign key can be found in Table 1.1.

As more than one strain could belong to the same pathovar, the pathovar name entity has a one-to-many (zero or more) relationship with the strain entity. Since the BOX, ERIC and REP fingerprints were generated from the same bacterial strains by using different primers in the PCR reactions and each strain could have more than one fingerprint, the strain entity has one-to-many relationships with the BOX, ERIC and REP fingerprint entities. The target pathovar entity represents a collection of pathovar classifications that the classifier is trained to recognize. The non-target pathovar entity represents a collection of pathovar classifications that the classifier is not trained to recognize. The pathovar name entity has one-to-zero or one-to-one relationship with both target pathovar and non-target pathovar entities. The BOX, ERIC and REP classification entities represent selected subsets of fingerprints for BPN training and classification. The attributes include filtered fingerprint and classification results. Each of the BOX, ERIC and REP fingerprint entities has one-to-zero or one-to-one relationship with the respective BOX, ERIC or REP classification entity.

A separate ERD for BER fingerprint data is shown in Figure 2.4. It was separated from the BOX, ERIC and REP fingerprint data so that it could be analyzed independently. The relationships among the different entities are similar to those in Figure 2.3. The data model in Figure 2.4 can also be adapted to analyze BOX, ERIC and REP fingerprint data independently.

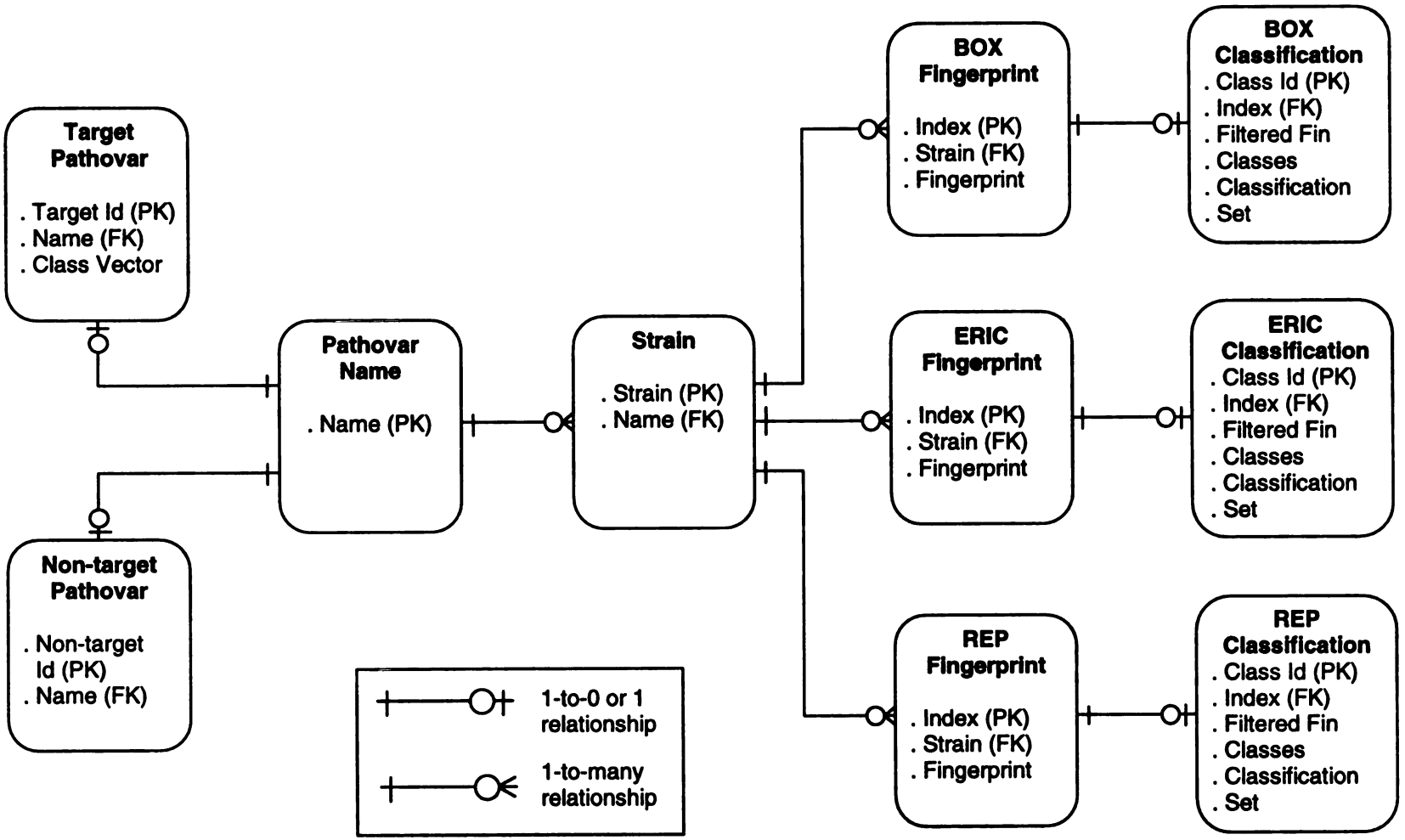



**Figure 2.3** Entity Relationship Diagram of rep-PCR Fingerprints

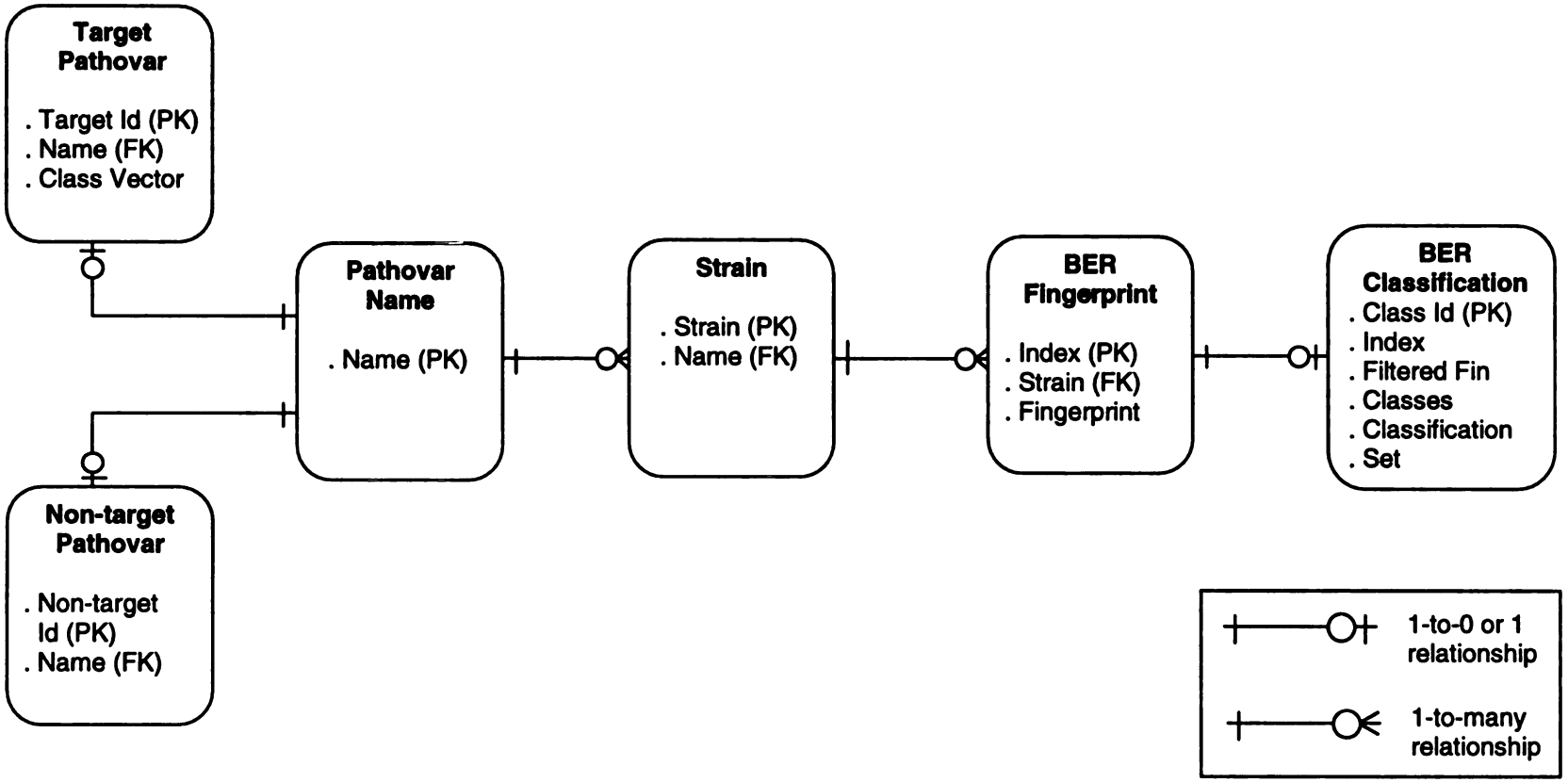
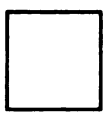
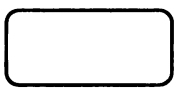
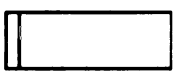
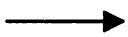



**Figure 2.4** Entity Relationship Diagram of BER Fingerprints

### 2.2.2 Process Modeling

Data flow diagrams (DFD) were used to depict the processes and data flows in the rep-PCR genomic fingerprint classification system. The notations of DFD are shown in Table 2.1. The DFD of the whole rep-PCR genomic fingerprint classification system is depicted in Figure 2.5. The main processes in the system were data input, image filtering, data partitioning, BPN training, BPN classifier testing and performance evaluation. Detailed DFDs of each processing components are shown in the respective sections following. All BOX, ERIC, REP and BER data went through the same processes in the system.

**Table 2.1** Notations of Data Flow Diagram (Whitten and Bently,1998)

| Symbol | Description |
|---|---|
|  | External Agent |
|  | Process |
|  | Data Store (Database or Files) |
|  | Data Flow |
|  | Exclusive Or Junction |
|  | And Junction |

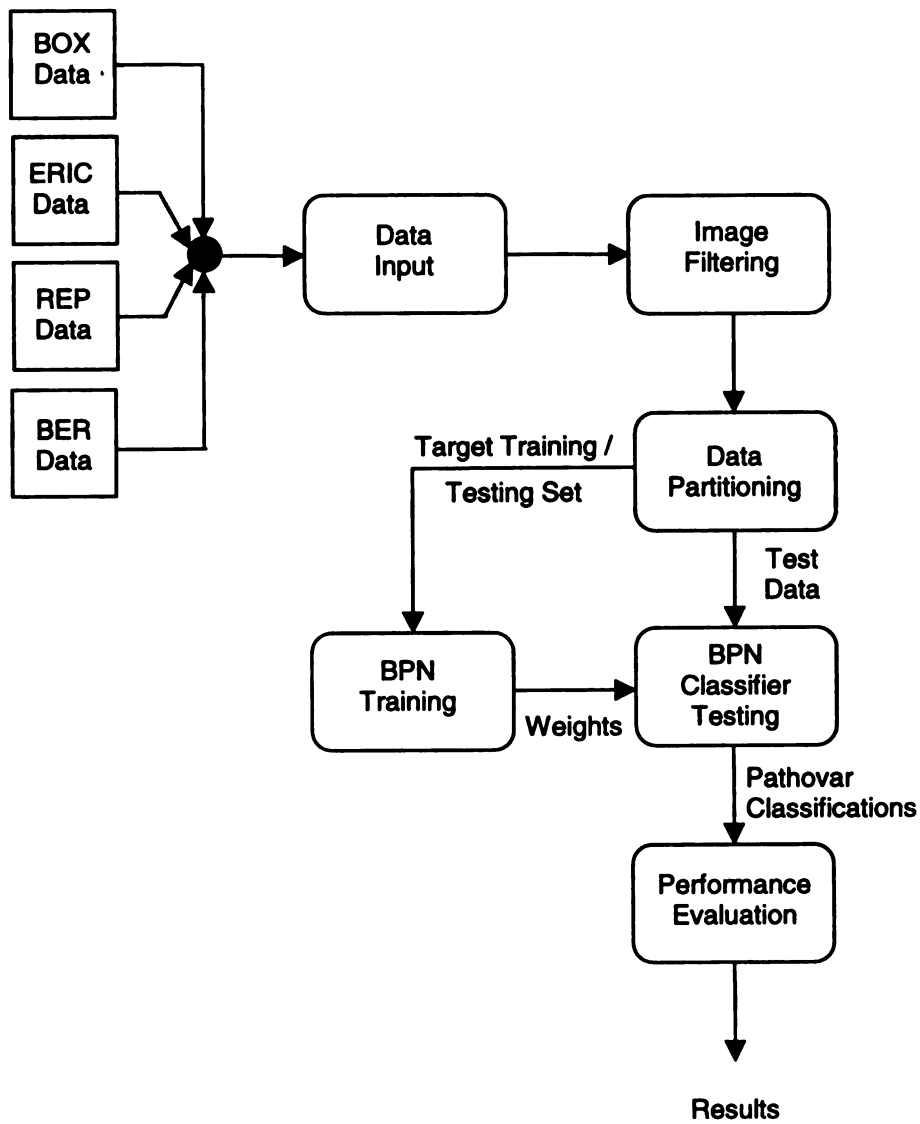



**Figure 2.5** DFD of rep-PCR Genomic Fingerprint Classification System

## 2.3 Data Processing

Data processing involved the creation of the rep-PCR genomic fingerprint database and noise reduction in the fingerprint images.

### 2.3.1 Data Input

A database in Microsoft Access format was created for the rep-PCR genomic fingerprints based on the data model in Figure 2.3, using Visual Basic 5.0 (Microsoft, Redmond, WA, no endorsement implied). The structure of the database is specified in Table B.1 of Appendix B. The BOX, ERIC, and REP fingerprint data were read into the database, and data were stored in the *Name*, *Strain*, *BFingerprint*, *EFingerprint* and *RFingerprint* tables, respectively. Similarly, a database was created for BER data based on the data model in Figure 2.4. A model of the data input process representing any one of BOX, ERIC, REP or BER data is shown in Figure 2.6.

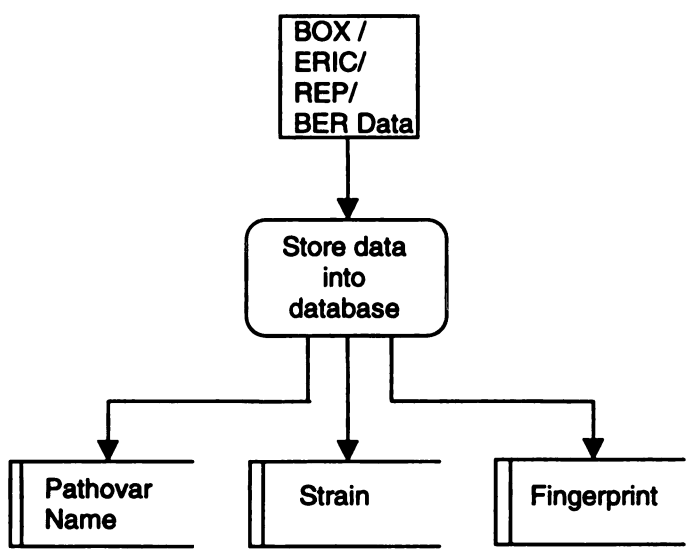



**Figure 2.6** DFD of Data Input

### 2.3.2 Image Filtering

Some noise was present in the fingerprint images. Although BPNs are quite tolerant of noise, noise reduction might still help to improve classification accuracy. Two filters, namely mean and wavelet, were used. Mean filter was implemented using equation 1.1, at a neighborhood size of 5. Wavelet filtering was implemented with Matlab Wavelet ToolBox (The Math Works Inc, Natick, MA, no endorsement implied), using the Daubechies db8 wavelet with decompositions at levels 3 and 4. The threshold for wavelet coefficient thresholding was selected automatically using the minimax criterion (Misiti et al., 1996). The scaling function and mother wavelet of db8 are shown in Figure 2.7.

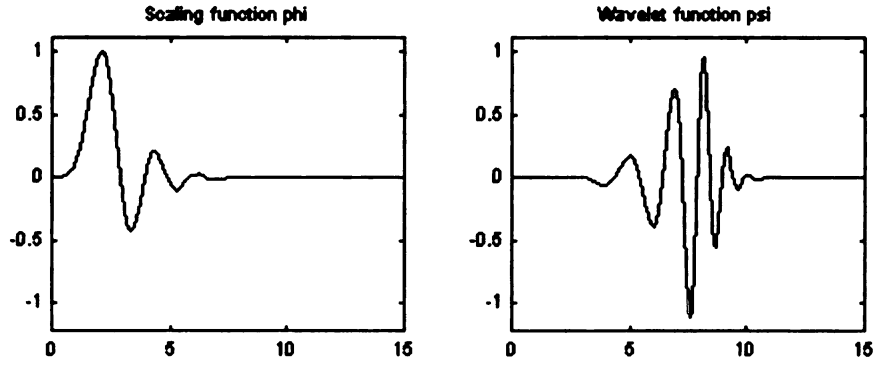



**Figure 2.7** Scaling Function and Wavelet Function of db8

A DFD of image filtering is shown in Figure 2.8. Data for BPN training and testing were selected from the fingerprint database. The selected data were stored in a text file. The filtering programs would read the file and generate a new file with filtered fingerprints. The filtered fingerprints were then stored in the field *Filtered* under the respective *B/E/RClassification* tables of the database. These filtered fingerprints were used for BPN training and classification.

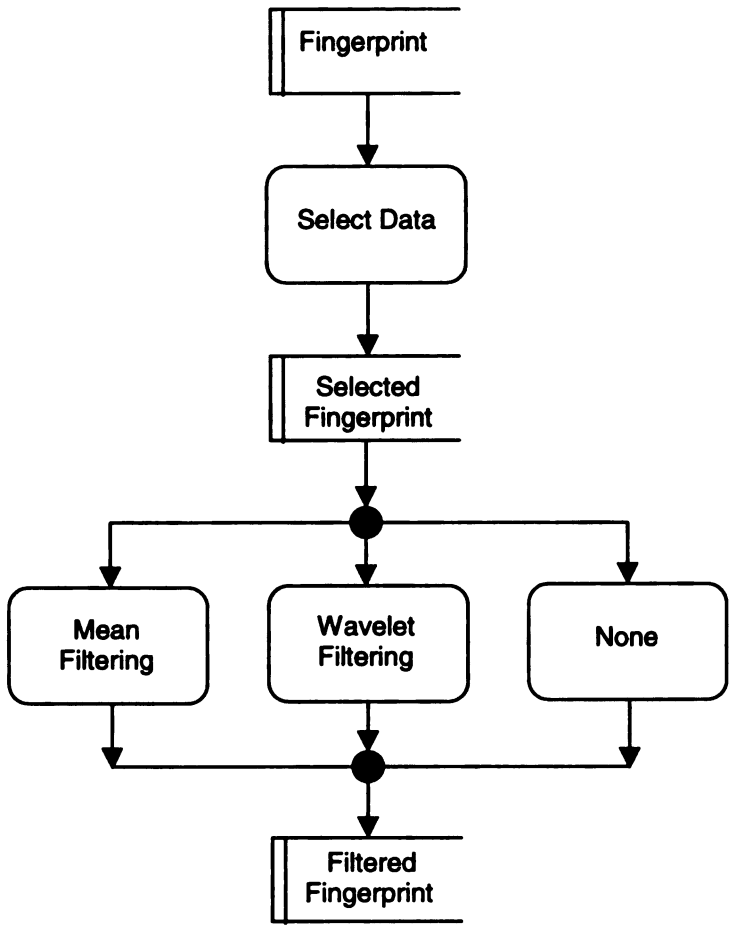



**Figure 2.8** DFD of Image Filtering

## 2.4 BPN Classifier Design

The BPN classifier design involved the following procedures: (1) Determination of pathovar classifications to be included in the classifier (target classifications), (2) the division of data into the training set and the testing set, and (3) BPN training.

### 2.4.1 Data Partitioning

The target pathovar classifications were determined from the data. To account for possible variations in the fingerprints of the same pathovar, only pathovars with more than four strains were selected as target pathovars. Those with fewer than four bacterial

strains were set aside as non-target pathovars. There were a total of 63 target pathovars and 26 non-target pathovars. The target and non-target pathovar names were stored in the *Target* and *Non-target* tables in the database, respectively. Each of the fingerprints that corresponded to the target pathovars was assigned a target vector represented in binary form. The $n^{th}$ target classification vector was expressed in terms of one 0.8 and nine 0.2s with the 0.8 in the $n^{th}$ position. Next, the target data were further divided into two equal halves: Set 1 and Set 2. Set 1 and Set 2 were used alternatively as training set and testing set for cross-validation.

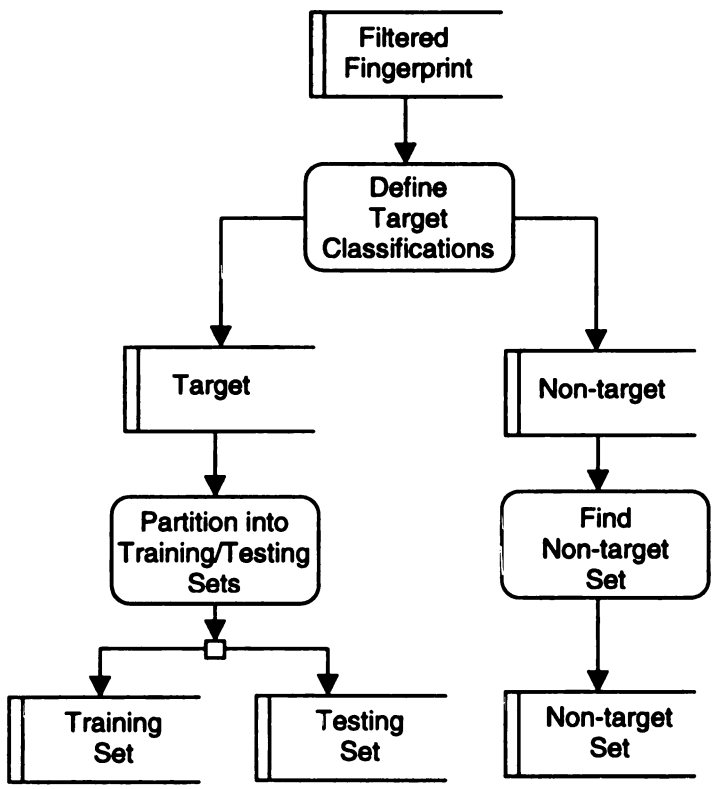



**Figure 2.9** DFD of Data Partitioning

The DFD of data partitioning is shown in Figure 2.9. The 752 bacterial strains in each of BOX, ERIC, REP and BER data types were partitioned as follows: 349 strains as training set, 349 strains as testing set and 54 strains as non-target set. The training set,

testing set and non-target set were stored in three separate files for the BPN. The non-target set could be used to test the true rejection capability of the BPN classifier.

### 2.4.2 BPN Training

BPN Classifiers for BOX, ERIC, REP and BER fingerprints were trained independently using the respective data sets. The fingerprints were linearly normalized to the scale of [0,2] from [0, 255] by dividing each data point by 127.5. The BPN performed better with input data in this scale. The BPN was a multilayer feed-forward network composed of an input layer, a hidden layer and an output layer. A diagram of the BPN architecture is shown in Figure 2.10.

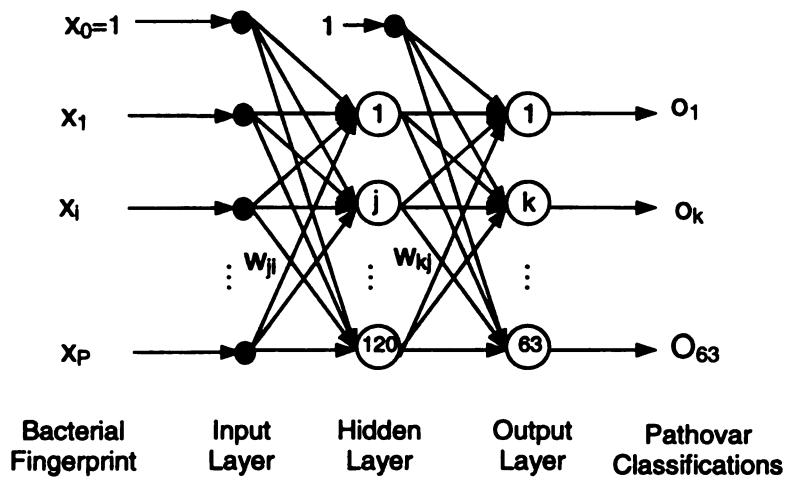



**Figure 2.10** The Architecture of a BPN Classifier

For the BOX, ERIC and REP BPNs, the input layer contained 360 input units which corresponded to the 31$^{st}$ to 390$^{th}$ data points of the 400-point rep-PCR genomic fingerprint. This segment represented DNA fragment sizes between 200 to 12,000 base pairs. The first thirty and last ten points were omitted, as they contained no relevant information. For the BER BPN, 387 input units corresponding to the 11$^{th}$ to 397$^{th}$ data

points were used, since the BER fingerprints were compressed versions of linearly combined BOX, ERIC and REP fingerprints. The output layer was made up of 63 output units, and each unit corresponded to a pathovar classification. The hidden layer had 120 hidden units. The number was determined by trial and error. The size of the hidden layer is important as it determines the capacity of the BPN to learn the different rep-PCR genomic fingerprint patterns. Although the learning capacity increases with the number of hidden units, too many hidden units may actually cause "overfitting". A network learned too specific to the training data will not perform well when presented with unfamiliar data sets. Moreover, a larger network will require more computations and training time. Thus, it is desirable to keep the network simple, while maintaining the capacity to learn the correct patterns. The nonlinear sigmoidal logistic function was used as the activation function for all hidden and output units.

Matlab Neural Network Toolbox (The Math Works Inc, Natick, MA, no endorsement implied) was used to simulate the BPNs. The supervised BPN training followed the backpropagation algorithm using the gradient descent method (Demuth and Beale, 1994). Batch mode training was applied, and the connection weights were updated after every epoch to minimize the errors between actual network outputs and desired outputs. Adaptive momentum and learning rate were used in the training. The system would increase the learning rate by a factor of 1.05 when the new training sum square error (SSE) was found lower than the training SSE in the previous epoch. If, however, the new training SSE was found greater or equal to the previous training SSE, the learning rate would be decreased using a factor of 0.7 and the momentum would be set to 0. Using an adaptive learning rate and momentum, the BPN converged fast. The status of the

training was determined by analyzing the training and generalization SSEs. The generalization SSE was found by presenting the testing data to the BPN, after every 1000 epochs. The number of 1000 was chosen so that the decision interval was large enough to avoid any local minimum. To avoid overtraining, the BPN training was terminated when the new generalization SSE was found greater than the previous generalization SSE. The final connection weights of each trained BPN were stored in a Matlab binary file. They were used to set up the respective BPN classifiers.

## 2.5 BPN Classifier Testing

The trained BPN classifiers were validated with the training data and tested with the testing data. The BPN classifiers were implemented with Matlab Neural Network ToolBox. The DFD of BPN classifier testing is shown in Figure 2.11.

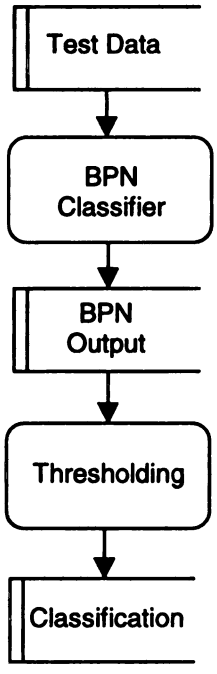



**Figure 2.11** DFD of BPN Classifier Testing

When presented with a test fingerprint as input, the BPN classifier assigned a corresponding pathovar classification. Each output unit of the BPN yielded an output score for the identification it represented, ranging from 0 to 1, with 0.8 and above for perfect match and 0.2 and below for no match at all. A threshold was selected for making classification decisions. Test bacterial strains were assigned pathovar classifications represented by output units that had output scores above the classification threshold. The classification threshold used had an effect on the classification accuracy. A high threshold would lead to low misclassification error rate and high rejection rate, while a low threshold would lead to high misclassification error rate and low rejection rate. The BPN classifiers were also tested for true rejection using fingerprints of non-target pathovars. A fingerprint would be rejected if none of the output units gave an output score greater than the threshold. The pathovar classifications identified by the BPNs were stored in the field *Classifications* under the respective *B/E/RClassification* tables of the database.

## 2.6 Performance Evaluation

The BPN classifications were verified with the desired classifications, and the performance of the BPN was evaluated based on top-2 recognition rate, misclassification error rate, false rejection error rate and false acceptance error rate. A tree representation of the BPN performance analysis is shown in Figure 2.12. A top-2 recognition rate was used for the evaluation since the top 2 identifications were important for the biological studies, and some pathovars of the same species had very similar fingerprints.

Furthermore, results from BOX, ERIC and REP BPN classifiers were combined to improve accuracy. The final classifications were based on majority votes on the

classifications identified by BOX, ERIC and REP BPN classifiers. No pathovar classification was assigned to a bacterial strain when there was a tie in the votes. Times required for the major processes were also recorded.

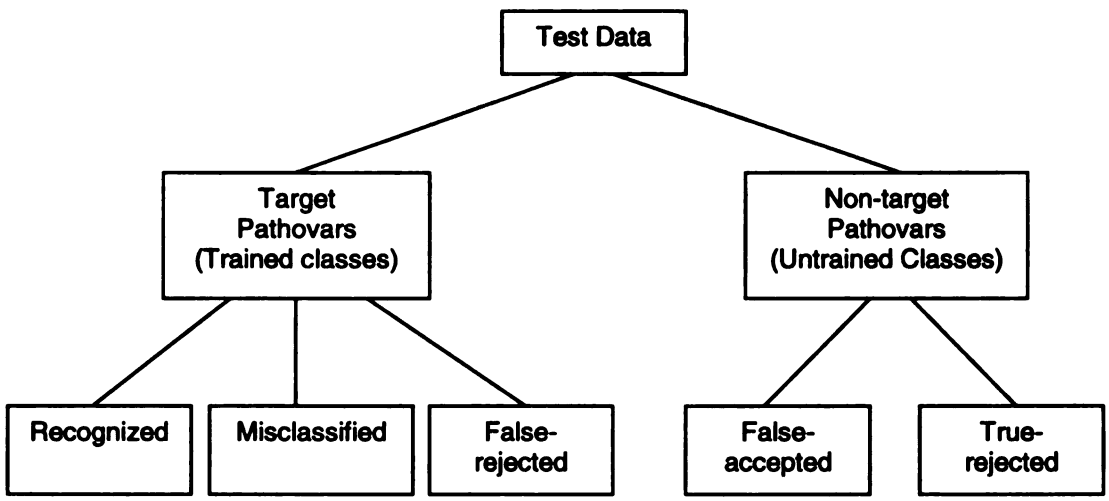



**Figure 2.12** A Tree Representation of the BPN Classification Performance Analysis

## 2.7 Graphical User Interface

A graphical user interface was developed using Microsoft Visual Basic 5.0 based on the data models and process models. The system responded to user-triggered events at the interface. The interface provided linkages to databases, files, process modules, and the external Matlab computation engine. Data could be searched and retrieved with the use of Structured Query Language (SQL). Information could be displayed on the interface in data tables or charts. Fingerprints could also be compared visually.

The rep-PCR genomic classification system was divided into four main components, namely Data Processing, BPN Training, BPN Classification, and

Comparison and Analysis. The Data Processing component consisted of data input and image filtering modules. The BPN Training component consisted of data partitioning and BPN training modules. The BPN Classification component consisted of BPN classifier testing and performance evaluation modules. And lastly, in the Comparison and Analysis component, a module was developed so that results of different BPN classifiers could be compared. With the use of Multiple-Document Interface (MDI), multiple instances of these components could operate simultaneously.

# Chapter III

# RESULTS AND DISCUSSION

## 3.1 Image Filtering

Fingerprint noise reduction was accomplished using the mean filter and wavelet filters methods. An example of a BOX fingerprint smoothed with mean filter is shown in Figure 3.1. The filtered fingerprint is compared with the original fingerprint in the figure.

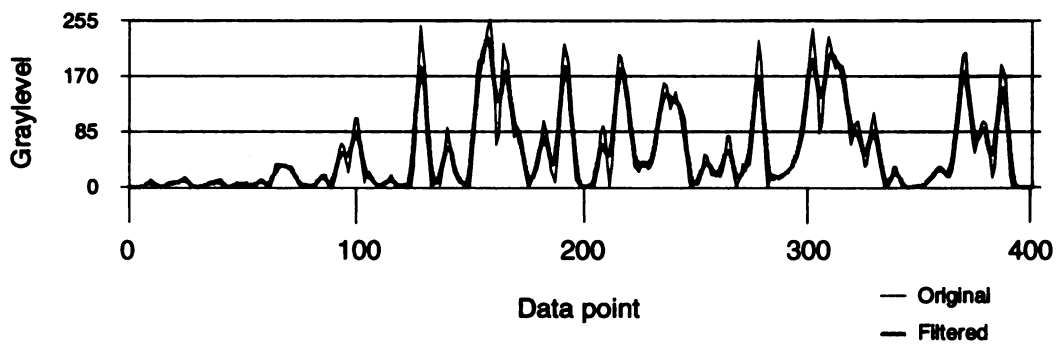



**Figure 3.1** BOX Fingerprint (LMG 12141) Filtered with Mean Filter

From Figure 3.1, we can see that the sharp edges in the peaks were rounded as a result of mean filtering. Examples of mean filtered REP and ERIC fingerprints are shown in Figure 3.2 and Figure 3.3, respectively.

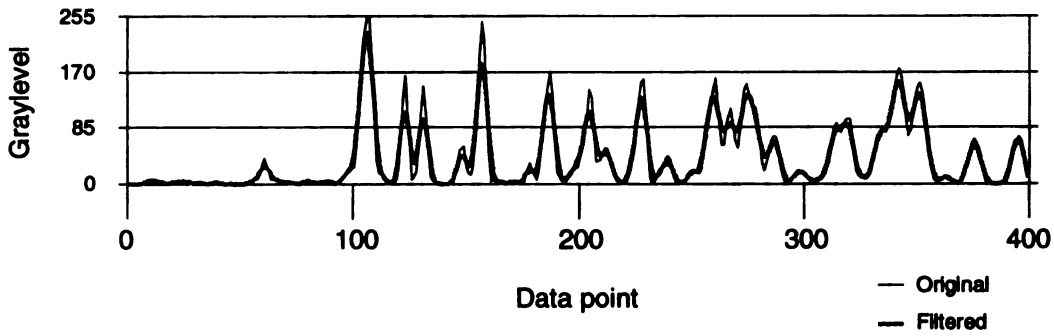



**Figure 3.2** ERIC Fingerprint (LMG 12141) Filtered with Mean Filter

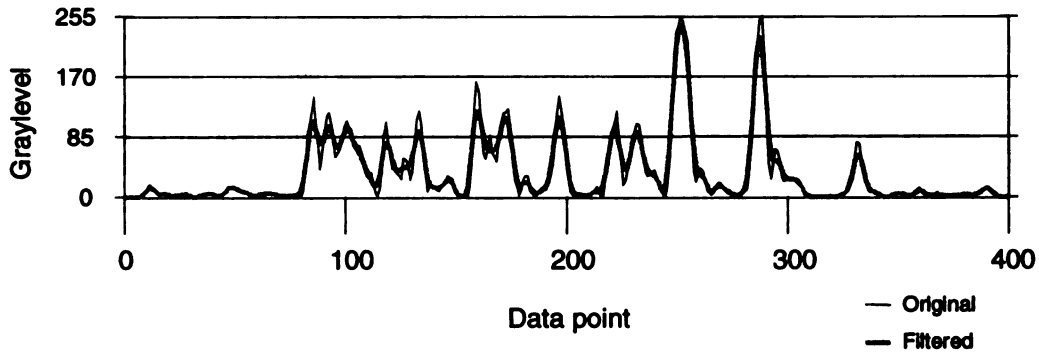



**Figure 3.3** REP Fingerprint (LMG 12141) Filtered with Mean Filter

The results of wavelet filtering are shown in figures 3.4-3.9. Figure 3.4 depicts the reconstructed original image and its approximations and details at each level before detail-coefficient-thresholding. The original image can be seen as a linear combination of a4, d4, d3, d2, and d1. Figure 3.5 depicts the detail wavelet transform coefficients of a BOX fingerprint (LMG 12141) and the associated thresholds at each of the four levels. Most of the noise filtered was from the details at levels one and two.

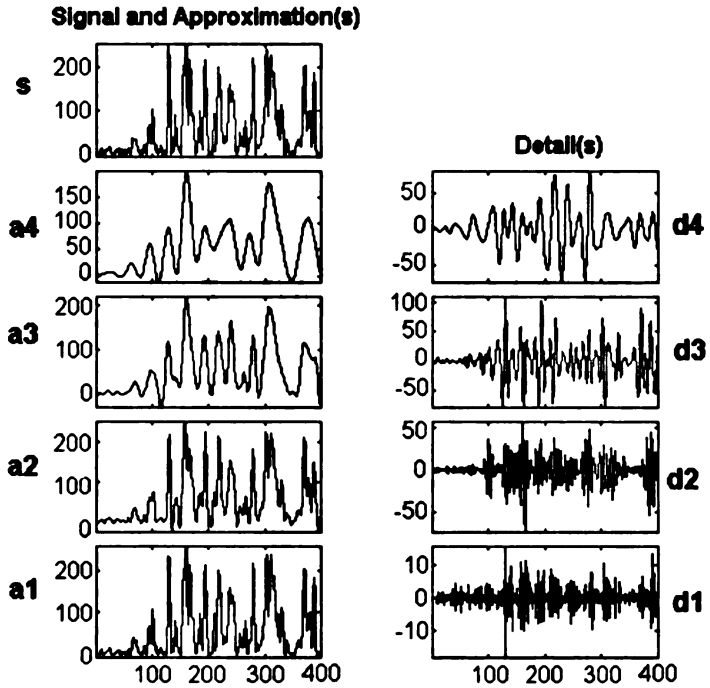



**Figure 3.4** Reconstructed Image, Approximation and Details before Thresholding (BOX Fingerprint – LMG 12141with db8 Wavelet at Level 4 Decomposition)

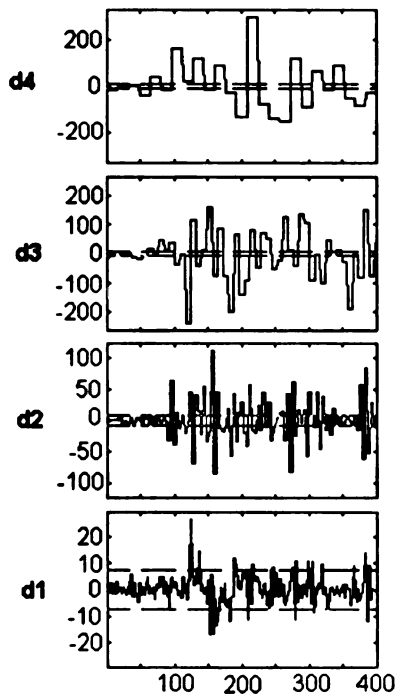



**Figure 3.5** Detail Coefficients and The Associated Threshold for Noise Reduction (BOX Fingerprint – LMG 12141with db8 Wavelet at Level 4 Decomposition)

Examples of BOX, ERIC, REP and BER fingerprints filtered with db8 wavelet at level three decomposition are depicted in Figures 3.6, 3.7, 3.8 and 3.9, respectively.

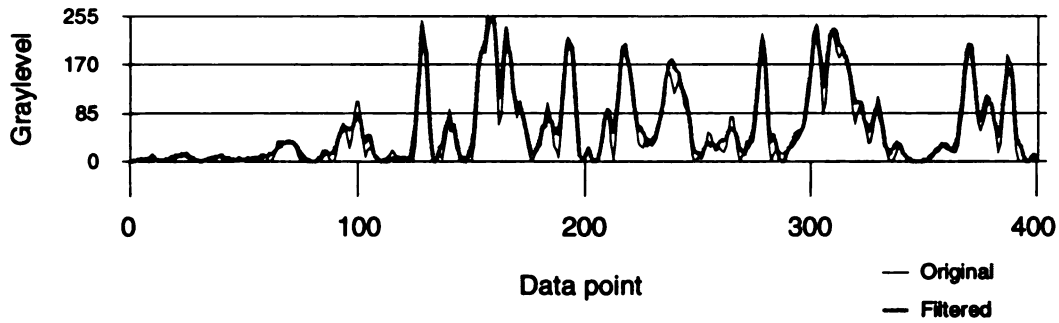



**Figure 3.6** BOX Fingerprint (LMG 12141) Filtered with db8 Wavelet at Level 3 Decomposition

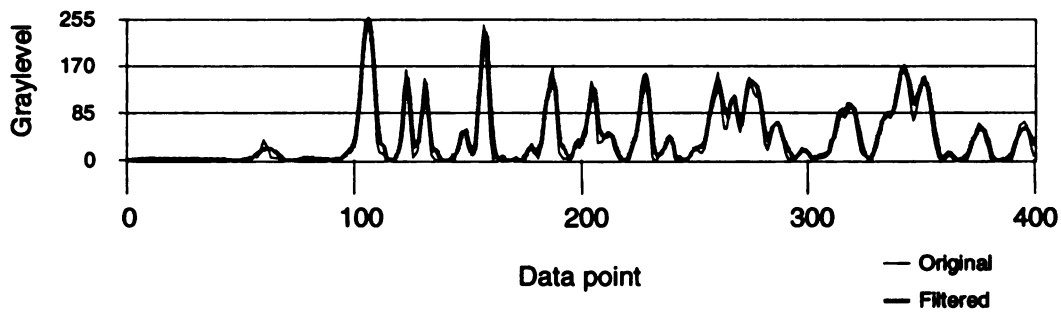



**Figure 3.7** ERIC Fingerprint (LMG 12141) Filtered with db8 Wavelet at Level 3 Decomposition

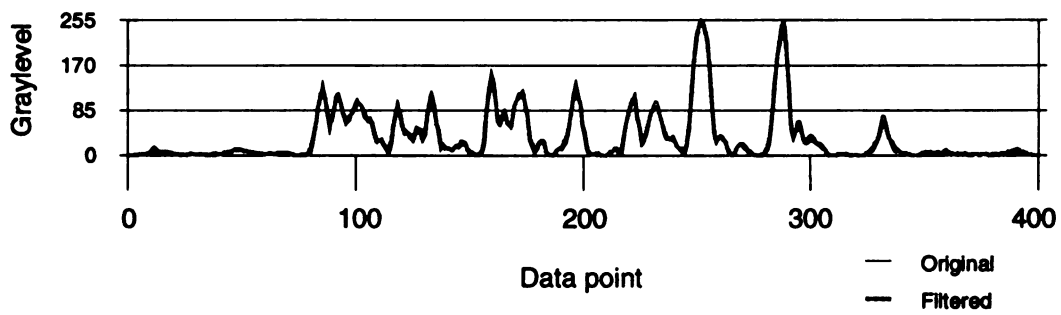



**Figure 3.8** REP Fingerprint (LMG 12141) Filtered with db8 Wavelet at Level 3 Decomposition

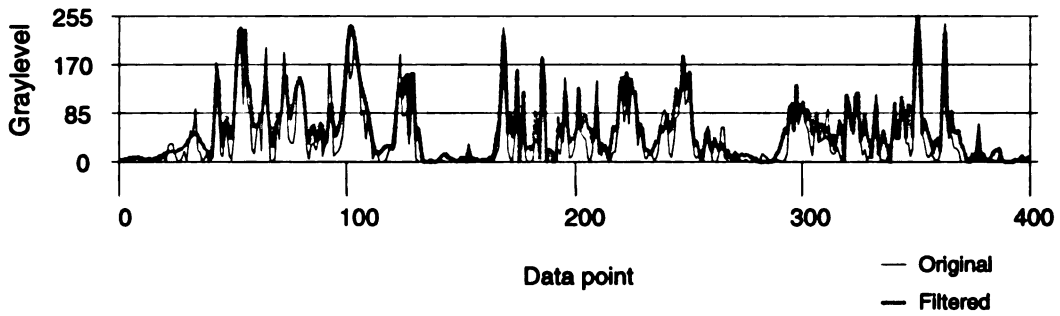



**Figure 3.9** BER Fingerprint (LMG 12141) Filtered with db8 Wavelet at Level 3 Decomposition

Examples of BOX, ERIC and REP fingerprints filtered with db8 wavelet at level four decomposition are depicted in Figures 3.10, 3.11 and 3.12, respectively. Wavelet filtering at level four was more rigorous since more detail coefficients were thresholded as compared with level three.

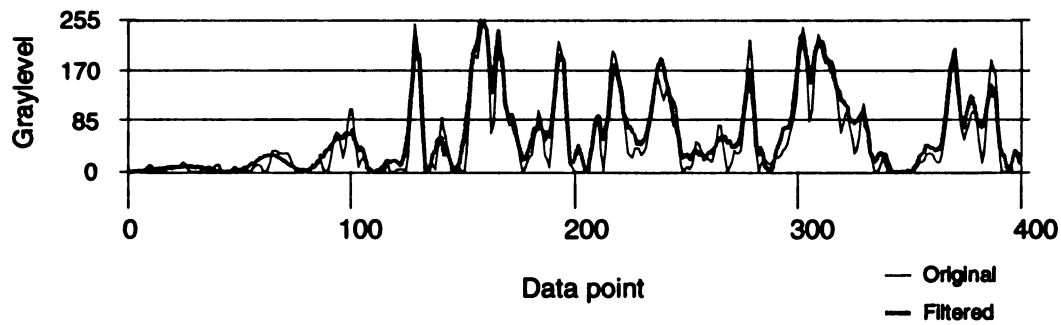



**Figure 3.10** BOX Fingerprint (LMG 12141) Filtered with db8 Wavelet at Level 4 Decomposition

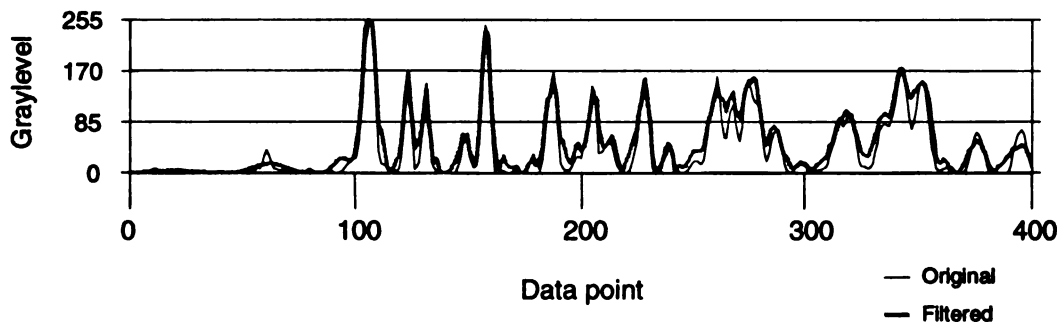



**Figure 3.11** ERIC Fingerprint (LMG 12141) Filtered with db8 Wavelet at Level 4 Decomposition

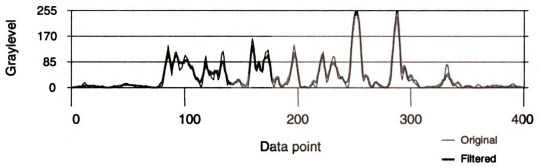



**Figure 3.12** REP Fingerprint (LMG 12141) Filtered with db8 Wavelet at level 4 Decomposition

## 3.2 BPN Training

Data available for training and testing were divided equally into two halves, Set 1 and Set 2. A summary of the results using data Set 1 for training and data Set 2 for testing is shown in Table 3.1. A summary of results using data Set 2 for training and data Set 1 for testing is shown in Table 3.2.

**Table 3.1** BPN Training Results (Set 1)

| Data Index | Filtering | Starting Learning rate | Number of Epochs | Training SSE | Generalization SSE |
|---|---|---|---|---|---|
| **BOX BPN** | | | | | |
| DAT00B-1 | None | 0.005 | 3000 | 1.85 | 28.29 |
| DAT01B-1 | Mean Filter | 0.005 | 4000 | 2.01 | 25.82 |
| DAT02B-1 | Wavelet db8, level 3 | 0.005 | 4000 | 1.58 | 27.12 |
| DAT03B-1 | Wavelet db8, level 4 | 0.005 | 3000 | 2.53 | 28.43 |
| **ERIC BPN** | | | | | |
| **DAT00E-1** | **None** | **0.005** | **3000** | **2.86** | **35.43** |
| **DAT01E-1** | **Mean Filter** | **0.005** | **3000** | **4.25** | **33.98** |
| **DAT02E-1** | **Wavelet db8, level 3** | **0.005** | **3000** | **3.00** | **34.86** |
| DAT03E-1 | Wavelet db8, level 4 | 0.005 | 4000 | 2.36 | 35.52 |
| **REP BPN** | | | | | |
| DAT00R-1 | None | 0.005 | 3000 | 2.29 | 35.22 |
| DAT01R-1 | Mean Filter | 0.005 | 3000 | 3.37 | 32.23 |
| DAT02R-1 | Wavelet db8, level 3 | 0.005 | 3000 | 2.40 | 34.38 |
| DAT03R-1 | Wavelet db8, level 4 | 0.005 | 3000 | 2.65 | 34.89 |
| **BER BPN** | | | | | |
| **BER00-1** | **None** | **0.004** | **7000** | **0.57** | **20.89** |
| **BER01-1** | **Wavelet db8, level 3** | **0.004** | **6000** | **1.03** | **25.41** |

**Table 3.2** BPN Training Results (Set 2)

| Data Index | Filtering | Starting Learning Rate | Number of Epochs | Training SSE | Generalization SSE |
|---|---|---|---|---|---|
| **BOX BPN** | | | | | |
| DAT00B-2 | None | 0.005 | 3000 | 1.70 | 28.65 |
| DAT01B-2 | Mean Filter | 0.005 | 5000 | 1.52 | 26.47 |
| DAT02B-2 | Wavelet db8, level 3 | 0.005 | 3000 | 2.05 | 27.91 |
| DAT03B-2 | Wavelet db8, level 4 | 0.005 | 3000 | 2.49 | 29.42 |
| **ERIC BPN** | | | | | |
| DAT00E-2 | **None** | **0.005** | **3000** | **2.87** | **36.64** |
| DAT01E-2 | **Mean Filter** | **0.005** | **3000** | **4.11** | **34.67** |
| DAT02E-2 | Wavelet db8, level 3 | 0.005 | 3000 | 2.91 | 35.37 |
| DAT03E-2 | Wavelet db8, level 4 | 0.005 | 3000 | 3.13 | 36.61 |
| **REP BPN** | | | | | |
| DAT00R-2 | None | 0.005 | 3000 | 2.29 | 33.33 |
| DAT01R-2 | Mean Filter | 0.005 | 4000 | 2.55 | 30.21 |
| DAT02R-2 | Wavelet db8, level 3 | 0.005 | 4000 | 1.78 | 31.15 |
| DAT03R-2 | Wavelet db8, level 4 | 0.005 | 4000 | 1.88 | 31.37 |
| **BER BPN** | | | | | |
| BER00-2 | **None** | **0.004** | **9000** | **0.36** | **19.89** |
| BER01-2 | **Wavelet db8, level 3** | **0.004** | **6000** | **0.92** | **25.03** |

A smaller starting learning rate of 0.004 was used for the BER BPNs. The BER BPNs training did not converge at a starting leaning rate of 0.005. Because the BER BPNs had a larger input layer and the networks were more complex compared with BOX, ERIC and REP BPNs, a lower learning rate was employed to avoid settling into local minima and to ensure convergence.

The number of training epochs was determined by the generalization sum square error (SSE). The BPN trainings were terminated when the generalization SSEs started to increase. Among the BOX, ERIC and REP BPNs, the number of training epochs was 3000 to 4000. The BER BPNs had the largest number of training epochs (6000 to 9000) and the lowest training and generalization SSEs among the BPNs, suggesting that there were fewer inconsistencies in the BER fingerprints. This was expected since the BER fingerprints contained all the features in BOX, ERIC and REP fingerprints and had a

higher information content. Any deficiency that existed in BOX, ERIC or REP features was offset by the contribution of other good features. A plot of DAT00B-1 BPN training and generalization SSE curves is shown in Figure 3.13. The gap between training and generalization errors started to increase after about 200 epochs. The training of a BPN took about four hours on a UNIX mainframe computer.

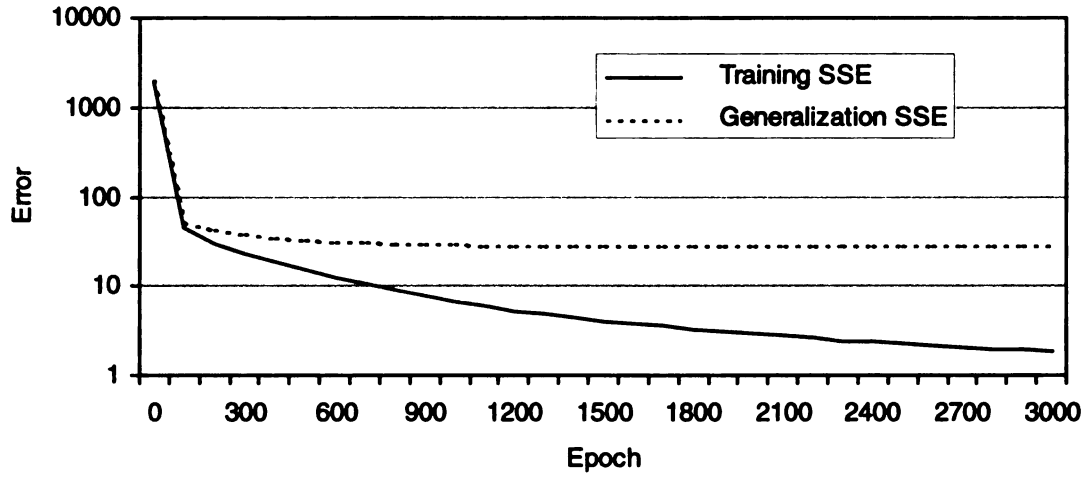



**Figure 3.13** BOX00-1 BPN Training and Generalization Sum Square Errors

## 3.3 BPN Classifier Testing

All BPN classifiers were validated with the training data. One hundred percent top-1 recognition rates were achieved when tested with data on which the BPN classifiers were trained. These results showed that the classifiers were capable of classifying bacterial fingerprints. Next, the BPN classifiers were tested with the testing data to evaluate the performance of the BPN classifiers on untrained data sets and estimate the true recognition rate and classification error rates. The classification threshold was found by trying different values. It was found that at a classification threshold of 0.45, a high

recognition rate could be achieved with some small penalty in misclassification and rejection errors.

An example of BPN output scores for a correctly identified bacterial strain is shown in Figure 3.14. It had a high output score at output unit 49, which represented the 49th bacterial pathovar classification. An example of BPN output scores for an unrecognized bacterial strain is shown in Figure 3.15. As shown in the chart, no output scores were greater than the classification threshold of 0.45. An example of BPN output scores for a bacterial strain with two identifications is shown in Figure 3.16. As shown in the chart, there were two output units with output scores greater than the threshold of 0.45.

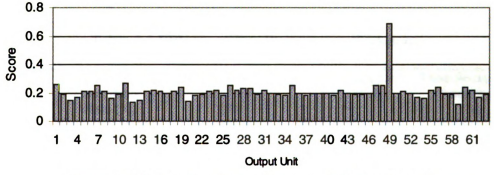



**Figure 3.14** BPN Output Scores of a Bacterial Strain with an Unique Identification (DAT00B-1 - LMG 471)

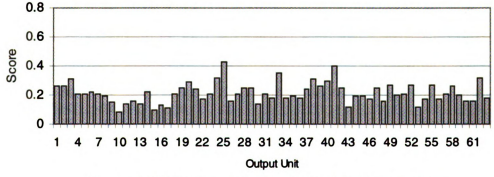



**Figure 3.15** BPN Output Scores of a Bacterial Strain with No Identification (DAT00B-1 - LMG 12141)

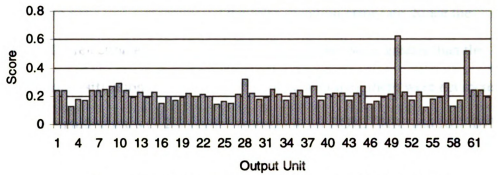



**Figure 3.16** BPN Output Scores of a Bacterial Strain with Two Identifications (DAT00B-1 - LMG 8684)

## 3.4 Performance Evaluation

The results of the BPN classifier performance evaluation are shown in Tables 3.3.

**Table 3.3** BPN Classifiers Performance Evaluation

| Data Index | Top-2 Recognition Rate (%) (Over 349 strains) | | | Misclassification Error Rate (%) (Over 349 strains) | | | False Rejection Error Rate (%) (Over 349 strains) | | | False Acceptance Error Rate (%) (Over 54 strains) | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | Mean | 1 | 2 | Mean | 1 | 2 | Mean | 1 | 2 | Mean |
| **BOX BPN** | | | | | | | | | | | | |
| DAT00B | 94 | 93 | 94 | 2.0 | 2.3 | 2.2 | 4.0 | 4.3 | 4.2 | 57 | 65 | 61 |
| DAT01B | 95 | 96 | 96 | 2.0 | 2.6 | 2.3 | 2.6 | 1.7 | 2.2 | 67 | 61 | 64 |
| DAT02B | 95 | 95 | 95 | 2.0 | 2.3 | 2.2 | 2.6 | 2.3 | 2.4 | 59 | 61 | 60 |
| DAT03B | 94 | 94 | 94 | 2.6 | 3.2 | 2.9 | 3.2 | 2.6 | 2.9 | 63 | 63 | 63 |
| **ERIC BPN** | | | | | | | | | | | | |
| DAT00E | 92 | 93 | 92 | 2.3 | 2.9 | 2.6 | 6.0 | 4.3 | 5.2 | 52 | 65 | 58 |
| DAT01E | 93 | 93 | 93 | 3.2 | 2.9 | 3.0 | 4.0 | 4.3 | 4.2 | 52 | 65 | 58 |
| DAT02E | 91 | 93 | 92 | 2.9 | 3.2 | 3.0 | 5.7 | 3.7 | 4.7 | 61 | 65 | 63 |
| DAT03E | 93 | 93 | 93 | 2.0 | 3.2 | 2.6 | 5.2 | 3.4 | 4.3 | 57 | 72 | 65 |
| **REP BPN** | | | | | | | | | | | | |
| DAT00R | 92 | 91 | 92 | 3.2 | 2.6 | 2.9 | 4.6 | 6.0 | 5.3 | 67 | 65 | 66 |
| DAT01R | 93 | 93 | 93 | 2.9 | 3.2 | 3.0 | 4.0 | 4.0 | 4.0 | 67 | 65 | 66 |
| DAT02R | 93 | 92 | 92 | 2.6 | 2.6 | 2.6 | 4.9 | 5.7 | 5.3 | 65 | 80 | 72 |
| DAT03R | 92 | 92 | 92 | 3.2 | 4.6 | 3.9 | 4.9 | 3.4 | 4.2 | 63 | 65 | 64 |
| **BER BPN** | | | | | | | | | | | | |
| BER00 | 99 | 98 | 98 | 1.4 | 0.86 | 1.1 | 0.0 | 0.86 | 0.43 | 50 | 61 | 56 |
| BER01 | 98 | 98 | 98 | 0.86 | 0.57 | 0.72 | 1.2 | 1.7 | 1.4 | 46 | 61 | 54 |

Table 3.3 shows that mean filtering (DAT01B/E/R) increased the misclassification error rate and decreased the false rejection error rate. Since the decreases in false rejection error rates (1.4% on the average) were greater than the increases in misclassification error rates (0.2% on the average), there were net increases in the top-2 recognition rates (1.2% on the average) in all BOX, ERIC and REP BPNs trained with mean filtered data. Also, wavelet filtering (DAT02B/E/R and DAT03B/E/R), when optimized at a specific level of decomposition, either decreased or maintained the misclassification error rates and decreased the false rejection error rates. Wavelet decomposition at level 3 using db8 worked best for BOX (DAT02B) and REP (DAT02R) data, while wavelet decomposition at level 4 using db8 worked best for ERIC (DAT03E) data. For the BPNs trained with the best wavelet-filtered data, the average decreases in misclassification and false rejection error rates were 0.10% and 0.86%, and the average increase in top-2 recognition rate was 0.96%.

Among the BPNs, the BER BPNs performed the best with the highest average top-2 recognition rate of 98% and lowest average misclassification error rate of 0.93%. The BOX BPNs performed better than the ERIC and REP BPNs with an average top-2 recognition rate of 95% and average misclassification error rate of 2.4%. The average top-2 recognition rates of ERIC and REP BPNs were 93% and 92%, respectively. The average misclassification error rates of ERIC and REP BPNs were 2.8% and 3.1%, respectively. The results showed that the BOX features and data were better than those of ERIC and REP. About 3% of the test strains were assigned two pathovar classifications by the BPNs.

The false acceptance error rates for all BPNs were high (54-72%) when tested with fingerprints of non-target pathovars. Ninety six percent of the misclassified and false accepted strains were assigned to some closely related pathovars of the same species. This was due to the fact that bacteria of the same species, but of different pathovar, still shared some common features in the fingerprints.

To enhance classification accuracy, the classification results from BOX, ERIC and REP BPNs were combined to give final classifications. Different combinations of BOX, ERIC and REP BPN classifiers were tried and the performances of the combined classifiers are shown in Table 3.4.

**Table 3.4** Performance of Combined BPN Classifiers

| Classifiers | Top 1 Recognition Rate (%) | Misclassification Error Rate (%) | False Rejection Error Rate (%) | False Acceptance Error Rate (%) |
|---|---|---|---|---|
| DAT00B-1<br>DAT00E-1<br>DAT00R-1 | 95 | 1.2 | 3.7 | 43 |
| DAT01B-1<br>DAT01E-1<br>DAT01R-1 | 96 | 1.2 | 2.6 | 44 |
| DAT02B-1<br>DAT02E-1<br>DAT02R-1 | 95 | 0.86 | 4.0 | 43 |
| DAT03B-1<br>DAT03E-1<br>DAT03R-1 | 95 | 0.57 | 4.6 | 43 |
| DAT02B-1<br>DAT03E-1<br>DAT02R-1 | 95 | 0.57 | 4.3 | 41 |

Since the final classifications were assigned based on majority votes, the recognition rates represented top-1 recognition rates. Recognition rates increased and both of the misclassification and false acceptance error rates decreased. The misclassification error rates of the combined classifiers were lower than those of the BER

BPN classifier. When the BOX, ERIC and REP BPNs with the smallest misclassification errors were combined, a top-1 recognition rate of 95% was achieved together with a lower misclassification error of 0.57%.

Bacterial strains wrongly identified or rejected by the BER01-1 BPN are shown in Table 3.5. Bacterial strains wrongly identified or rejected by the DAT02B-1, DAT03E-1 and DAT02R-1 BPNs are shown in Tables 3.6, 3.7 and 3.8, respectively. Bacterial strains wrongly identified or rejected by the combined DAT02B-1, DAT03E-1 and DAT02R-1 BPN classifiers are shown in Table 3.9. Bacterial strains listed in Table 3.5 were found in Table 3.6, 3.7 and 3.8 as well. Most wrongly-identified bacterial strains were assigned to a closely related pathovar of the same species. Bacterial strains listed in these tables need further examinations. The errors were due to variations caused by biological and physical factors in the fingerprinting process. The variations may also be inherent in characteristic of the bacterial species.

**Table 3.5** Bacterial Strains Wrongly Identified or Rejected by BER01-1 BPN

| Strain | Target Classification | BER BPN Classification | |
|---|---|---|---|
| LMG 539 | *X. axonopodis axonopodis* | *X. axonopodis vasculorum* | * |
| LMG 558 | *X. axonopodis cajani* | - | + |
| LMG 7473 | *X. axonopodis cajani* | *X. axonopodis glycines* | |
| LMG 8689 | *X. axonopodis vitians* | - | |
| LMG 937 | *X. axonopodis vitians* | *X. axonopodis malvacearum* | |
| LMG 7392 | *X. translucens cerealis* | - | |
| LMG 883 | *X. translucens secalis* | - | |

* misclassification error
+ false rejection error

**Table 3.6** Bacterial Strains Wrongly Identified or Rejected by DAT02B-1 BPN

| Strain | Target Classification | BOX BPN Classification |
|---|---|---|
| LMG 558 | *X. axonopodis cajani* | *X. axonopodis malvacearum* |
| LMG 7473 | *X. axonopodis cajani* | *X. axonopodis glycines* |
| LMG 9175 | *X. axonopodis citri E* | *X. axonopodis alfalfae* |
| LMG 8128 | *X. axonopodis glycines* | - |
| LMG 7427 | *X. axonopodis malvacearum* | *X. axonopodis vitians* |
| LMG 7443 | *X. axonopodis ricini* | - |
| LMG 937 | *X. axonopodis vitians* | - |
| LMG 7516 | *X. campestris campestris* | - |
| LMG 8134 | *X. campestris raphani* | - |
| LMG 7392 | *X. translucens cerealis* | - |
| LMG 891 | *X. translucens cerealis* | *X. translucens hordei* |
| LMG 8279 | *X. translucens hordei* | *X. translucens undulosa* |
| LMG 7507 | *X. translucens secalis* | - |
| LMG 883 | *X. translucens secalis* | - |
| LMG 875 | *X. translucens translucens* | *X. translucens hordei* |
| LMG 876 | *X. translucens translucens* | - |

**Table 3.7** Bacterial Strains Wrongly Identified or Rejected by DAT03E-1 BPN

| Strain | Target Classification | ERIC BPN Classification |
|---|---|---|
| LMG 488 | *X. albilineans* | - |
| LMG 8020 | *X. axonopodis alfalfae** | - |
| LMG 539 | *X. axonopodis axonopodis* | - |
| LMG 558 | *X. axonopodis cajani* | - |
| LMG 7473 | *X. axonopodis cajani* | - |
| LMG 9172 | *X. axonopodis citri E* | - |
| LMG 8128 | *X. axonopodis glycines* | - |
| LMG 761 | *X. axonopodis malvacearum* | *X. axonopodis glycines* |
| LMG 780t2 | *X. axonopodis manihotis* | *X. axonopodis phaseoli* |
| LMG 7455 | *X. axonopodis phaseoli* | - |
| LMG 821 | *X. axonopodis phaseoli* | *X. axonopodis manihotis* |
| LMG 8689 | *X. axonopodis vitians* | - |
| LMG 937 | *X. axonopodis vitians* | *X. axonopodis malvacearum* |
| LMG 7516 | *X. campestris campestris* | - |
| LMG 6518 | *X. oryzae oryzae* | *X. oryzae oryzicola* |
| LMG 595 | *X. translucens graminis* | - |
| LMG 726 | *X. translucens graminis* | - |
| LMG 8279t1 | *X. translucens hordei* | - |
| LMG 882 | *X. translucens hordei* | - |
| LMG 716 | *X. translucens phlei* | - |
| LMG 5260t2 | *X. translucens translucens* | *X. translucens undulosa* |
| LMG 875 | *X. translucens translucens* | - |
| LMG 885 | *X. translucens undulosa* | *X. translucens translucens* |
| LMG 8276 | *X. vasicola holcicola* | - |
| LMG 902 | *X. vasicola vasculorum* | - |

**Table 3.8** Bacterial Strains Wrongly Identified or Rejected by DAT02R-1 BPN

| Strain | Target Classification | REP BPN Classification |
|---|---|---|
| LMG 5402 | *X. arboricola poinsettiicola* | - |
| LMG 539 | *X. axonopodis axonopodis* | *X. axonopodis vasculorum* |
| LMG 7194 | *X. axonopodis begoniae* | - |
| LMG 558 | *X. axonopodis cajani* | - |
| LMG 7387t1 | *X. axonopodis cajani* | - |
| LMG 7473 | *X. axonopodis cajani* | *X. axonopodis glycines* |
| LMG 9176 | *X. axonopodis citri A* | - |
| LMG 9181 | *X. axonopodis citri C* | *X. axonopodis phaseoli(fusc)* |
| LMG 9182 | *X. axonopodis citri D* | *X. axonopodis citri C* |
| LMG 7429 | *X. axonopodis malvacearum* | *X. axonopodis vitians* |
| LMG 761 | *X. axonopodis malvacearum* | *X. axonopodis vitians* |
| LMG 937 | *X. axonopodis vitians* | *X. axonopodis malvacearum* |
| LMG 6518 | *X. oryzae oryzae* | - |
| LMG 471 | *X. sacchari* | - |
| LMG 8684 | *X. theicola* | - |
| LMG 8685 | *X. theicola* | - |
| LMG 880 | *X. translucens cerealis** | - |
| LMG 595 | *X. translucens graminis* | - |
| LMG 882 | *X. translucens hordei* | - |
| LMG 728 | *X. translucens poae* | - |
| LMG 7507 | *X. translucens secalis* | - |
| LMG 883 | *X. translucens secalis* | *X. oryzae oryzae* |
| LMG 876 | *X. translucens translucens* | - |
| LMG 885 | *X. translucens undulosa* | - |
| LMG 892 | *X. translucens undulosa* | *X. translucens translucens* |
| LMG 736t2 | *X. vasicola holcicola* | - |

**Table 3.9** Bacterial Strains Wrongly Identified or Rejected by Combined BOX, ERIC and REP BPN Classifiers

| Strain | Target Classification | DAT02B-1 Classification | DAT03E-1 Classification | DAT02R-1 Classification | Final Classification (by majority vote) |
|---|---|---|---|---|---|
| LMG 539 | *X. a. axonopodis* | *X. a. axonopodis* | - | *X. a. vasculorum* | - |
| LMG 558 | *X. a. cajani* | *X. a. malvacearum* | - | - | - |
| LMG 7473 | *X. a. cajani* | *X. a. glycines* | - | *X. a. glycines* | *X. a. glycines* |
| LMG 8128 | *X. a. glycines* | - | - | *X. a. glycines* | - |
| LMG 761 | *X. a. malvacearum* | *X. a. malvacearum* | *X. a. glycines* | *X. a. vitians* | - |
| LMG 937 | *X. a. vitians* | - | *X. a. malvacearum* | *X. a. malvacearum* | *X. a. malvacearum* |
| LMG 7516 | *X. c. campestris* | - | - | *X. c. campestris* | - |
| LMG 6518 | *X. o. oryzae* | *X. o. oryzae* | *X. o. oryzicola* | - | - |
| LMG 876 | *X. t. translucens* | - | *X. t. translucens* | - | - |
| LMG 891 | *X. t. cerealis* | *X. t. hordei* | *X. t. cerealis* | *X. t. cerealis*<br>*X. translucens hordei* | - |
| LMG 595 | *X. t. graminis* | *X. t. graminis* | - | - | - |
| LMG 882 | *X. t. hordei* | *X. t.hordei* | - | - | - |
| LMG 7507 | *X. t. secalis* | - | *X. t.secalis* | - | - |
| LMG 883 | *X. t. secalis* | - | *X. t. hordei*<br>*X. t. secalis* | *X. o. oryzae*<br>*X. t.undulosa* | - |
| LMG 875 | *X. t. translucens* | *X. t. hordei* | - | *X. t. translucens* | - |
| LMG 885 | *X. t. undulosa* | *X. t. undulosa* | *X. t.translucens* | - | - |
| LMG 892 | *X. t. undulosa* | *X. t. translucens*<br>*X. t. undulosa* | *X. t. undulosa* | *X. t. translucens* | - |

Times required for the major processes in the rep-PCR genomic fingerprint classification system are shown in Table 3.10. Data processing, BPN training and testing can be done in nine hours for 752 strains of bacteria. Fingerprint filtering and classification of 752 bacterial strains can be done in about 15 minutes with a trained BPN classifier. The BPN classification speed (0.2 seconds per strain on Pentium 133 MHz with 32MB RAM) was faster than the GelCompare software using cluster analysis (0.3 seconds per strain on Pentium II 266 MHz with 96MB RAM).

**Table 3.10** Speed Performance of rep-PCR Genomic Fingerprint Classification System

| Process | Number of Bacterial Strains | Time Required (P-133 32MB RAM Win 95) |
|---|---|---|
| Data Input | 752 | 35 seconds |
| Mean Filtering | 752 | 4 minutes 21 seconds |
| Wavelet Filtering | 752 | 11 minutes 7 seconds |
| Target Classifications Determination | 752 | 1 minute 38 seconds |
| Training and Testing Sets Partition | 752 | 50 seconds |
| BPN Training | 349 | 4 to 8 hours (UNIX) |
| Classification | 349 | 1 minute |
| Verification | 349 | 22 seconds |

The size of the training data file (349 strains) was 620 kilobytes, and the size of the BPN connection weights file was 408 kilobytes, amounting to a 34% saving in the storage space. The saving will be more significant when the number of training data is increased. The size of the BPN connection weights will remain the same as long as the network structure is kept unchanged. This is an advantage as compared to the cluster analysis method since increased number of training data per classification will improve recognition rate, but will not affect the storage space size and the time required for a trained BPN to perform the classification.

### 3.5 User Scenario

A snapshot of the graphical user interface of rep-PCR Genomic Fingerprint Classification System is shown in Figure 3.17. Names of databases created were shown on the upper part of the left panel. Opened databases and their associated components were shown on the lower part of the left panel. A tree view structure was used to represent the organization. Processes could be activated by clicking on the nodes of the tree. Activated components were shown on the right side of the main window. Several components could be opened at the same time.

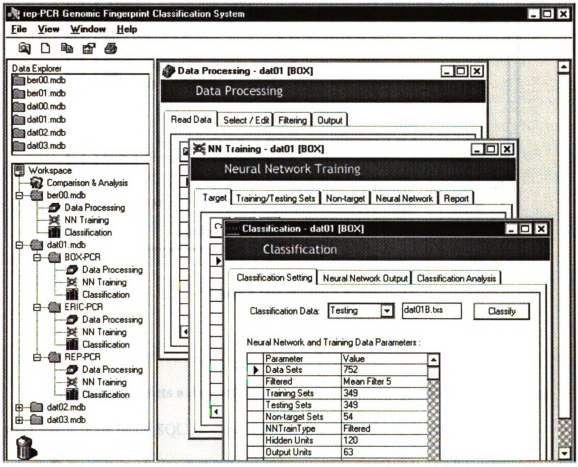



**Figure 3.17** A Snapshot of rep-PCR Genomic Fingerprint Classification System

The data input screen is shown in Figure 3.18. Selected fingerprints data file, in the format shown in Figure A.1 of Appendix A, could be read into the database.

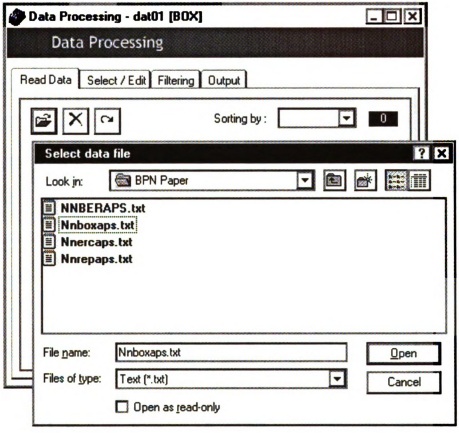



**Figure 3.18** Data Input Screen

Figure 3.19 depicts a data selection screen. Data could be selected for filtering and classification using SQL queries. Figure 3.20 depicts an image filtering settings screen. Different filtering techniques could be implemented.

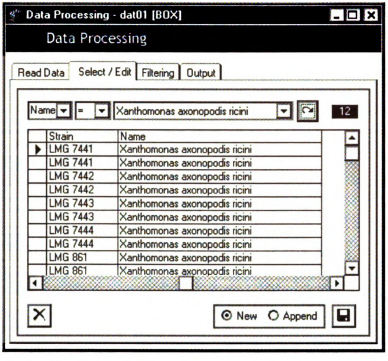



**Figure 3.19** Data Selection Screen

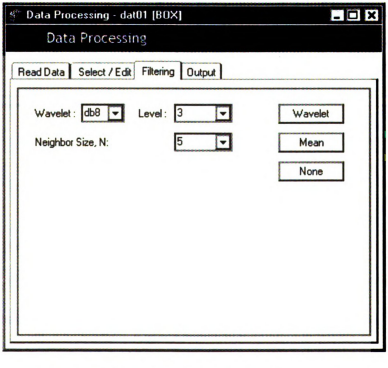



**Figure 3.20** Image Filtering Settings Screen

The filtering results were shown in the screen depicted in Figure 3.21. Fingerprints could also be visualized as densitometric curves.

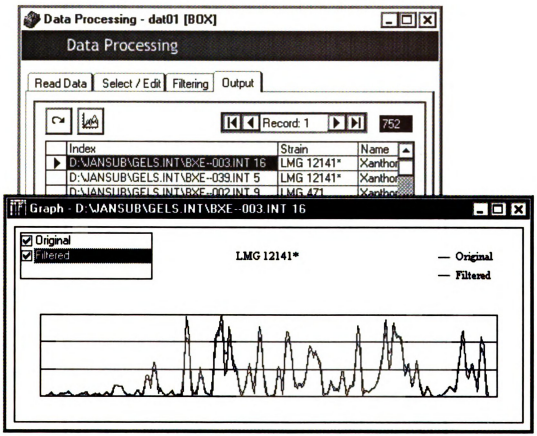



**Figure 3.21** Filtering Output Screen

Target pathovar classifications were defined in the target pathovar screen shown in Figure 3.22. Fingerprints belong to target pathovar classifications were divided into training and testing sets as shown in Figure 3.23. Non-target pathovars are defined in the non-target pathovar screen shown Figure 3.24. Figure 3.25 depicts the BPN settings screen. BPN training parameters could be set, and BPN training could be initiated.

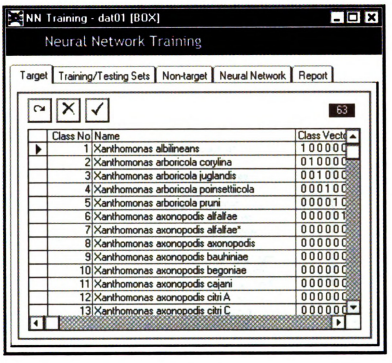



**Figure 3.22** Target Pathovar Screen

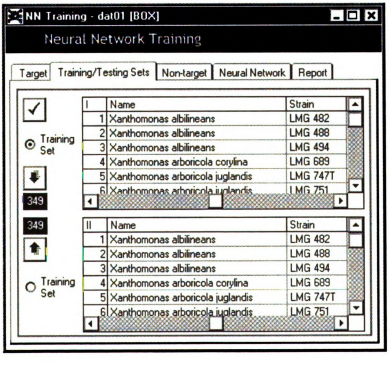



**Figure 3.23** Training / Testing Sets Screen

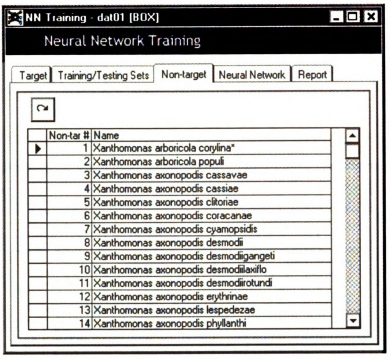



**Figure 3.24** Non-target Pathovar Screen

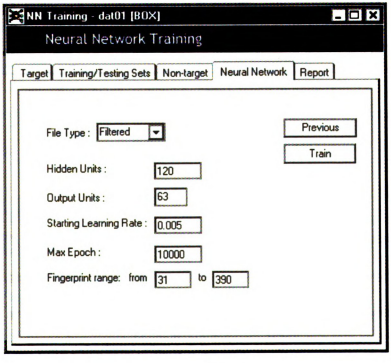



**Figure 3.25** BPN Settings Screen

BPN classifier setting screen is shown in Figure 3.26. Classification data could be selected. Fingerprints were classified using Matlab 5.1 as a computation engine.

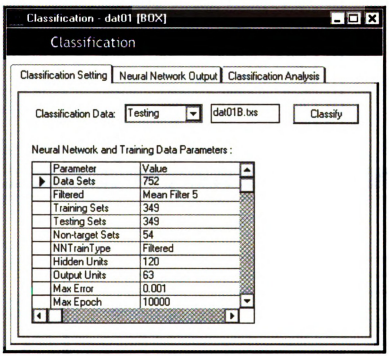



**Figure 3.26** Classification Setting Screen

The BPN outputs were shown on the BPN output screen depicted in Figure 3.27. The output scores could be shown on bar chart. Figure 3.28 depicts the pathovar classifications analysis screen. Incorrect classifications were identified, and recognition rate and error rates were calculated. Figure 3.29 depicts the classification comparison screen. Results form different classifiers could be loaded and compared.

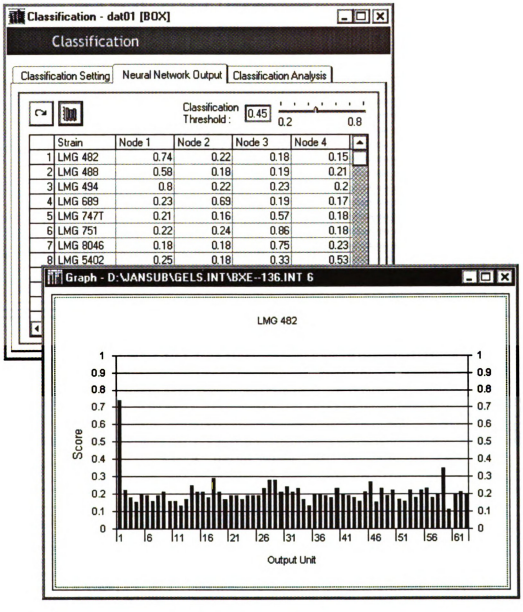



**Figure 3.27** BPN Output Screen

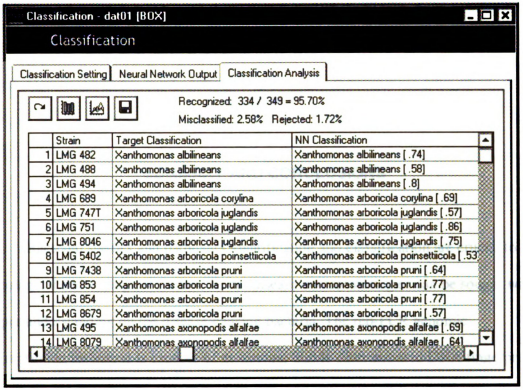



**Figure 3.28** BPN Classification Analysis Screen

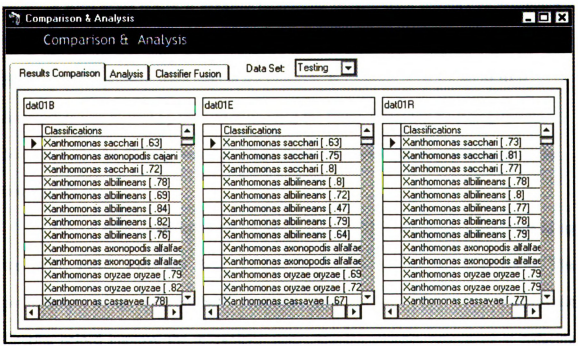



**Figure 3.29** Classification Comparison Screen

# Chapter IV

## SUMMARY AND CONCLUSIONS

Bacterial species and pathovars identification and classification are crucial to pathological and ecological studies. Better disease control strategies can be sought with deeper understanding on the disease-causing bacteria. The rep-PCR genomic fingerprinting technique is capable of generating highly discriminated DNA fingerprints for bacterial identification. An efficient classification system will better facilitate the DNA fingerprint classification and pathovar identification process.

A BPN based rep-PCR genomic fingerprint classification system was developed. Three sets of *Xanthomonas* genomic fingerprints generated from rep-PCRs, using BOX, ERIC and REP primers, were used in the research. In addition, a fourth set of BER fingerprints was formed by linearly combining the BOX, ERIC and REP fingerprints. Mean and wavelet filtering techniques were used to reduce high-frequency noise on the fingerprints.

Several BPN classifiers were trained to identify 63 *Xanthomonas* pathovars using the BOX, ERIC, REP and BER original fingerprints and filtered fingerprints. The performances of the different BPN classifiers were evaluated. Both mean and wavelet filtering helped improve the recognition rates. Wavelet filtering was better at reducing misclassification error rates, and mean filtering was better at reducing false rejection

error rates. The average top-2 recognition rates of BOX, ERIC, REP and BER BPN classifiers were 95%, 93%, 92% and 98%, respectively. Furthermore, classification results of selected BOX, ERIC and REP BPNs were combined to yield more accurate results. By combining the results of three BPN classifiers with the lowest misclassification error rates, a top-1 recognition rate of 95% was achieved together with a misclassification error rate of 0.57% and a false rejection rate of 4.3%.

Based on the results, the following conclusions are stated:

1. Bacterial rep-PCR genomic fingerprints could be classified effectively and efficiently at the pathovar level using BPN classifiers. The BPN-based genomic fingerprint classification system developed could be used to identify 63 *Xanthomonas* pathovars.
2. Classification accuracy improved with both mean filtering and wavelet denoising. Wavelet filtering was better at reducing misclassification error rates, and mean filtering was better at reducing false rejection error rates.
3. The BER BPNs performed the best, followed by BOX BPNs, REP BPNs and ERIC BPNs with the average top-2 recognition rates of 98%, 95%, 93% and 92%, respectively. Combined BOX, ERIC and REP classifiers improved classification accuracy. Ninety six percent of the wrongly identified bacterial strains were assigned to a closely related pathovar of the same species.

# Chapter V

## RECOMMENDATIONS

Listed below are some of the recommendations for future research and application:

1. Since the trained BPN classifier is very data dependent, the training data should cover all the possible variations in the fingerprint patterns. More training patterns per classification should be included in the training set.
2. The target classifications should be well-defined based on genomic features.
3. Since the cost of misidentification is higher than false rejection, the classifier design should be aimed at minimizing misclassification error rate while maintaining a low false rejection rate.
4. Analyze problems behind the wrongly identified or rejected bacterial strains.
5. Try other wavelets for wavelet filtering.
6. The methodology can be applied to other bacterial genus and species.
7. Improve the user interface to include more functionality for data analysis.
8. Make the BPN classifier available on-line so that it can be used by a larger group of researchers.

# APPENDICES

# APPENDIX A

## rep-PCR GENOMIC FINGERPRINT DATA

**Table A.1** List of rep-PCR Genomic Fingerprint Data

| Pathovar | Strain | BOX Index | ERIC Index | REP Index | BER Index |
|---|---|---|---|---|---|
| **Training/Testing Set 1** | | | | | |
| *X. albilineans* | LMG 482 | BXE-136.INT 6 | ECE-226.INT 22 | RPE-135.INT 6 | XBER4C1.INT 6 |
| *X. albilineans* | LMG 487 | BXE-237.INT 2 | ECE-229.INT 2 | RPE-233.INT 2 | XBER5C2.INT 2 |
| *X. albilineans* | LMG 488 | BXE-198.INT 27 | ECE-202.INT 27 | RPE-200.INT 27 | XBER13C1.INT 27 |
| *X. albilineans* | LMG 490 | BXE-258.INT 19 | ECE-262.INT 19 | RPE-266.INT 20 | XBER13C2.INT 19 |
| *X. albilineans* | LMG 494 | BXE-002.INT 10 | ECE-006.INT 10 | RPE-047.INT 13 | XBER3C1.INT 10 |
| *X. arboricola corylina* | LMG 688 | BXE-039.INT 2 | ECE-014.INT 9 | RPE-017.INT 9 | XBER1C2.INT 9 |
| *X. arboricola corylina* | LMG 689 | BXE-003.INT 10 | ECE-006.INT 10 | RPE-009.INT 10 | XBER1C1.INT 10 |
| *X. arboricola corylina* | LMG 8660 | BXE-012.INT 13 | ECE-014.INT 13 | RPE-017.INT 13 | XBER1C2.INT 13 |
| *X. arboricola juglandis* | LMG 747T | BXE-136.INT 25 | ECE-226.INT 4 | RPE-135.INT 25 | XBER4C1.INT 25 |
| *X. arboricola juglandis* | LMG 750 | BXE-239.INT 17 | ECE-231.INT 17 | RPE-235.INT 17 | XBER8C1.INT 3 |
| *X. arboricola juglandis* | LMG 751 | BXE-160.INT 23 | ECE-177.INT 21 | RPE-162.INT 25 | XBER7C1.INT 23 |
| *X. arboricola juglandis* | LMG 8045 | BXE-239.INT 29 | ECE-231.INT 29 | RPE-235.INT 29 | XBER8C1.INT 15 |
| *X. arboricola juglandis* | LMG 8046 | BXE-207.INT 22 | ECE-206.INT 22 | RPE-208.INT 22 | XBER15C1.INT 22 |
| *X. arboricola juglandis* | LMG 8047 | BXE-012.INT 14 | ECE-014.INT 14 | RPE-017.INT 14 | XBER1C2.INT 14 |
| *X. arboricola poinsettiicola* | LMG 5402 | BXE-207.INT 14 | ECE-206.INT 14 | RPE-208.INT 14 | XBER15C1.INT 14 |
| *X. arboricola poinsettiicola* | LMG 5403 | BXE-012.INT 15 | ECE-014.INT 15 | RPE-017.INT 15 | XBER1C2.INT 15 |
| *X. arboricola pruni* | LMG 7438 | BXE-199.INT 6 | ECE-203.INT 6 | RPE-201.INT 6 | XBER14C1.INT 6 |
| *X. arboricola pruni* | LMG 852 | BXE-012.INT 17 | ECE-014.INT 17 | RPE-017.INT 17 | XBER1C2.INT 17 |
| *X. arboricola pruni* | LMG 853 | BXE-207.INT 26 | ECE-206.INT 26 | RPE-208.INT 26 | XBER15C1.INT 26 |
| *X. arboricola pruni* | LMG 853R | BXE-272.INT 4 | ECE-272.INT 4 | RPE-272.INT 4 | XBER18C2.INT 4 |
| *X. arboricola pruni* | LMG 854 | BXE-207.INT 27 | ECE-206.INT 27 | RPE-208.INT 27 | XBER15C1.INT 27 |
| *X. arboricola pruni* | LMG 855 | BXE-260.INT 19 | ECE-264.INT 19 | RPE-268.INT 19 | XBER15C2.INT 19 |
| *X. arboricola pruni* | LMG 8679 | BXE-182.INT 7 | ECE-179.INT 5 | RPE-174.INT 9 | XBER8C2.INT 22 |
| *X. arboricola pruni* | LMG 8680 | BXE-012.INT 18 | ECE-014.INT 18 | RPE-017.INT 18 | XBER1C2.INT 18 |
| *X. axonopodis alfalfae* | LMG 495 | BXE-136.INT 4 | ECE-226.INT 24 | RPE-135.INT 4 | XBER4C1.INT 4 |
| *X. axonopodis alfalfae* | LMG 497 | BXE-225.INT 23 | ECE-255.INT 23 | RPE-227.INT 23 | XBER4C2.INT 23 |
| *X. axonopodis alfalfae* | LMG 8018t1-1 | BXE-237.INT 16 | ECE-175.INT 3 | RPE-161.INT 3 | XBER5C1.INT 8 |
| *X. axonopodis alfalfae* | LMG 8018t1-2 | BXE-139.INT 7 | ECE-229.INT 16 | RPE-233.INT 16 | XBER5C2.INT 17 |
| *X. axonopodis alfalfae* | LMG 8079 | BXE-139.INT 25 | ECE-175.INT 21 | RPE-161.INT 21 | XBER4C2.INT 27 |
| *X. axonopodis alfalfae** | LMG 8019 | BXE-243.INT 14 | ECE-246.INT 14 | RPE-249.INT 14 | XBER11C2.INT 14 |
| *X. axonopodis alfalfae** | LMG 8020 | BXE-192.INT 26 | ECE-194.INT 26 | RPE-204.INT 26 | XBER11C1.INT 26 |
| *X. axonopodis alfalfae** | LMG 8020R | BXE-258.INT 25 | ECE-360.INT 26 | RPE-266.INT 26 | XBER13C2.INT 25 |
| *X. axonopodis alfalfae** | LMG 8020RR | BXE-222.INT 24 | ECE-228.INT 6 | RPE-224.INT 24 | XBER18C1.INT 24 |
| *X. axonopodis axonopodis* | LMG 538t1T | BXE-039.INT 7 | ECE-014.INT 26 | RPE-017.INT 26 | XBER1C2.INT 26 |
| *X. axonopodis axonopodis* | LMG 539 | BXE-160.INT 7 | ECE-177.INT 5 | RPE-162.INT 9 | XBER6C2.INT 26 |
| *X. axonopodis axonopodis* | LMG 539R | BXE-241.INT 16 | ECE-244.INT 16 | RPE-247.INT 16 | XBER9C2.INT 16 |
| *X. axonopodis axonopodis* | LMG 540 | BXE-159.INT 17 | ECE-176.INT 16 | RPE-142.INT 17 | XBER8C1.INT 17 |
| *X. axonopodis bauhiniae* | LMG 548 | BXE-243.INT 18 | ECE-246.INT 18 | RPE-249.INT 18 | XBER11C2.INT 18 |
| *X. axonopodis bauhiniae* | LMG 548R | BXE-222.INT 3 | ECE-223.INT 3 | RPE-224.INT 3 | XBER17C2.INT 15 |
| *X. axonopodis begoniae* | LMG 551 | BXE-243.INT 4 | ECE-246.INT 4 | RPE-249.INT 4 | XBER11C2.INT 4 |
| *X. axonopodis begoniae* | LMG 552 | BXE-139.INT 3 | ECE-175.INT 25 | RPE-161.INT 25 | XBER5C1.INT 3 |
| *X. axonopodis begoniae* | LMG 7178 | BXE-257.INT 3 | ECE-300.INT 20 | RPE-265.INT 3 | XBER12C2.INT 3 |
| *X. axonopodis begoniae* | LMG 7188 | BXE-192.INT 27 | ECE-194.INT 27 | RPE-204.INT 27 | XBER11C1.INT 27 |
| *X. axonopodis begoniae* | LMG 7194 | BXE-237.INT 10 | ECE-229.INT 10 | RPE-233.INT 10 | XBER5C2.INT 10 |
| *X. axonopodis begoniae* | LMG 7194R | BXE-193.INT 12 | ECE-195.INT 12 | RPE-205.INT 12 | XBER12C1.INT 12 |
| *X. axonopodis begoniae* | LMG 7196 | BXE-225.INT 28 | ECE-226.INT 28 | RPE-227.INT 28 | XBER5C1.INT 24 |
| *X. axonopodis begoniae* | LMG 7196RA | BXE-193.INT 4 | ECE-195.INT 4 | RPE-205.INT 4 | XBER12C1.INT 4 |
| *X. axonopodis begoniae* | LMG 7226 | BXE-305.INT 2 | ECE-360.INT 2 | RPE-342.INT 2 | XBER12C2.INT 2 |
| *X. axonopodis begoniae* | LMG 7303 | BXE-139.INT 5 | ECE-175.INT 23 | RPE-161.INT 23 | XBER5C1.INT 5 |

| | | | | | |
|---|---|---|---|---|---|
| *X. axonopodis begoniae* | LMG 7304 | BXE--243.INT 17 | ECE--246.INT 17 | RPE--249.INT 17 | XBER11C2.INT 17 |
| *X. axonopodis begoniae* | LMG 7595 | BXE--139.INT 16 | ECE--175.INT 12 | RPE--161.INT 12 | XBER5C1.INT 16 |
| *X. axonopodis begoniae* | LMG 7601 | BXE--243.INT 12 | ECE--246.INT 12 | RPE--249.INT 12 | XBER11C2.INT 12 |
| *X. axonopodis begoniae* | LMG 7601R | BXE--198.INT 16 | ECE--202.INT 16 | RPE--200.INT 16 | XBER13C1.INT 16 |
| *X. axonopodis cajani* | LMG 558 | BXE--237.INT 18 | ECE--229.INT 18 | RPE--233.INT 18 | XBER5C2.INT 18 |
| *X. axonopodis cajani* | LMG 7387t1 | BXE--136.INT 8 | ECE--226.INT 20 | RPE--135.INT 8 | XBER4C1.INT 8 |
| *X. axonopodis cajani* | LMG 7473 | BXE--237.INT 13 | ECE--229.INT 13 | RPE--233.INT 13 | XBER5C2.INT 13 |
| *X. axonopodis citri A* | LMG 681 | BXE--193.INT 26 | ECE--360.INT 12 | RPE--307.INT 10 | XBER12C1.INT 26 |
| *X. axonopodis citri A* | LMG 681R | BXE--259.INT 12 | ECE--361.INT 13 | RPE--267.INT 13 | XBER14C2.INT 12 |
| *X. axonopodis citri A* | LMG 8650 | BXE--139.INT 13 | ECE--175.INT 15 | RPE--161.INT 15 | XBER5C1.INT 13 |
| *X. axonopodis citri A* | LMG 8654 | BXE--258.INT 10 | ECE--360.INT 18 | RPE--266.INT 12 | XBER13C2.INT 10 |
| *X. axonopodis citri A* | LMG 8657 | BXE--139.INT 15 | ECE--175.INT 13 | RPE--161.INT 13 | XBER5C1.INT 15 |
| *X. axonopodis citri A* | LMG 9176 | BXE--238.INT 3 | ECE--230.INT 3 | RPE--254.INT 4 | XBER6C2.INT 3 |
| *X. axonopodis citri A* | LMG 9321 | BXE--198.INT 13 | ECE--202.INT 13 | RPE--200.INT 13 | XBER13C1.INT 13 |
| *X. axonopodis citri A* | LMG 9665 | BXE--305.INT 14 | ECE--360.INT 23 | RPE--365.INT 7 | XBER13C2.INT 22 |
| *X. axonopodis citri A* | LMG 9669 | BXE--198.INT 23 | ECE--202.INT 23 | RPE--200.INT 23 | XBER13C1.INT 23 |
| *X. axonopodis citri A* | LMG 9671 | BXE--258.INT 27 | ECE--360.INT 28 | RPE--307.INT 19 | XBER13C2.INT 27 |
| *X. axonopodis citri C* | LMG 8655 | BXE--159.INT 7 | ECE--176.INT 6 | RPE--142.INT 7 | XBER5C2.INT 29 |
| *X. axonopodis citri C* | LMG 8656 | BXE--259.INT 2 | ECE--361.INT 3 | RPE--267.INT 3 | XBER14C2.INT 2 |
| *X. axonopodis citri C* | LMG 9181 | BXE--159.INT 22 | ECE--176.INT 21 | RPE--142.INT 22 | XBER6C1.INT 22 |
| *X. axonopodis citri C* | LMG 9654 | BXE--258.INT 12 | ECE--360.INT 19 | RPE--266.INT 13 | XBER13C2.INT 12 |
| *X. axonopodis citri C* | LMG 9658 | BXE--198.INT 25 | ECE--202.INT 25 | RPE--200.INT 25 | XBER13C1.INT 25 |
| *X. axonopodis citri D* | LMG 9182 | BXE--238.INT 6 | ECE--230.INT 6 | RPE--234.INT 6 | XBER6C2.INT 6 |
| *X. axonopodis citri D* | LMG 9182R | BXE--198.INT 5 | ECE--202.INT 5 | RPE--200.INT 5 | XBER13C1.INT 5 |
| *X. axonopodis citri E* | LMG 9160col | BXE--238.INT 22 | ECE--230.INT 22 | RPE--234.INT 22 | XBER6C2.INT 22 |
| *X. axonopodis citri E* | LMG 9160dro | BXE--160.INT 4 | ECE--177.INT 2 | RPE--162.INT 6 | XBER6C2.INT 23 |
| *X. axonopodis citri E* | LMG 9162 | BXE--238.INT 7 | ECE--230.INT 7 | RPE--234.INT 7 | XBER6C2.INT 7 |
| *X. axonopodis citri E* | LMG 9172 | BXE--193.INT 27 | ECE--301.INT 7 | RPE--205.INT 27 | XBER12C1.INT 27 |
| *X. axonopodis citri E* | LMG 9174 | BXE--257.INT 27 | ECE--261.INT 27 | RPE--265.INT 28 | XBER12C2.INT 27 |
| *X. axonopodis citri E* | LMG 9175 | BXE--207.INT 2 | ECE--206.INT 2 | RPE--208.INT 2 | XBER14C2.INT 22 |
| *X. axonopodis citri E* | LMG 9175R | BXE--259.INT 23 | ECE--361.INT 19 | RPE--267.INT 24 | XBER15C1.INT 3 |
| *X. axonopodis citri E* | LMG 9325 | BXE--198.INT 2 | ECE--202.INT 2 | RPE--205.INT 29 | XBER13C1.INT 2 |
| *X. axonopodis dieffenbachiae* | LMG 695 | BXE--238.INT 18 | ECE--230.INT 18 | RPE--234.INT 19 | XBER6C2.INT 18 |
| *X. axonopodis dieffenbachiae* | LMG 7484 | BXE--207.INT 18 | ECE--206.INT 18 | RPE--208.INT 18 | XBER15C1.INT 18 |
| *X. axonopodis dieffenbachiae* | LMG 8664 | BXE--238.INT 12 | ECE--230.INT 12 | RPE--234.INT 12 | XBER6C2.INT 12 |
| *X. axonopodis dieffenbachiae** | LMG 7399 | BXE--159.INT 20 | ECE--176.INT 18 | RPE--142.INT 20 | XBER6C1.INT 20 |
| *X. axonopodis dieffenbachiae** | LMG 7400 | BXE--238.INT 17 | ECE--230.INT 17 | RPE--234.INT 18 | XBER6C2.INT 17 |
| *X. axonopodis glycines* | LMG 712 | BXE--159.INT 5 | ECE--176.INT 4 | RPE--142.INT 5 | XBER5C2.INT 27 |
| *X. axonopodis glycines* | LMG 8026 | BXE--238.INT 8 | ECE--230.INT 8 | RPE--234.INT 8 | XBER6C2.INT 8 |
| *X. axonopodis glycines* | LMG 8126 | BXE--207.INT 16 | ECE--206.INT 16 | RPE--208.INT 16 | XBER15C1.INT 16 |
| *X. axonopodis glycines* | LMG 8128 | BXE--260.INT 16 | ECE--264.INT 16 | RPE--268.INT 16 | XBER15C2.INT 16 |
| *X. axonopodis malvacearum* | LMG 7427 | BXE--160.INT 17 | ECE--177.INT 15 | RPE--162.INT 20 | XBER7C1.INT 17 |
| *X. axonopodis malvacearum* | LMG 7429 | BXE--270.INT 4 | ECE--273.INT 4 | RPE--275.INT 4 | XBER16C3.INT 4 |
| *X. axonopodis malvacearum* | LMG 7430 | BXE--210.INT 25 | ECE--212.INT 25 | RPE--217.INT 24 | XBER16C1.INT 25 |
| *X. axonopodis malvacearum* | LMG 761 | BXE--239.INT 25 | ECE--231.INT 25 | RPE--235.INT 25 | XBER8C1.INT 10 |
| *X. axonopodis malvacearum* | LMG 762 | BXE--181.INT 4 | ECE--177.INT 26 | RPE--173.INT 4 | XBER7C2.INT 18 |
| *X. axonopodis malvacearum* | LMG 764 | BXE--239.INT 27 | ECE--256.INT 12 | RPE--235.INT 27 | XBER8C1.INT 13 |
| *X. axonopodis manihotis* | LMG 768 | BXE--181.INT 8 | ECE--178.INT 4 | RPE--173.INT 8 | XBER7C2.INT 23 |
| *X. axonopodis manihotis* | LMG 769 | BXE--241.INT 9 | ECE--244.INT 9 | RPE--247.INT 9 | XBER9C2.INT 9 |
| *X. axonopodis manihotis* | LMG 771 | BXE--210.INT 15 | ECE--212.INT 15 | RPE--217.INT 14 | XBER16C1.INT 15 |
| *X. axonopodis manihotis* | LMG 773 | BXE--239.INT 16 | ECE--231.INT 16 | RPE--235.INT 16 | XBER8C1.INT 2 |
| *X. axonopodis manihotis* | LMG 778 | BXE--210.INT 16 | ECE--212.INT 16 | RPE--217.INT 15 | XBER16C1.INT 16 |
| *X. axonopodis manihotis* | LMG 780t1 | BXE--270.INT 12 | ECE--273.INT 12 | RPE--306.INT 10 | XBER16C3.INT 12 |
| *X. axonopodis manihotis* | LMG 780t2 | BXE--207.INT 6 | ECE--206.INT 6 | RPE--208.INT 6 | XBER14C2.INT 26 |
| *X. axonopodis manihotis* | LMG 782 | BXE--270.INT 2 | ECE--273.INT 2 | RPE--275.INT 2 | XBER16C3.INT 2 |
| *X. axonopodis manihotis* | LMG 784 | BXE--210.INT 24 | ECE--212.INT 24 | RPE--217.INT 23 | XBER16C1.INT 24 |
| *X. axonopodis patelii* | LMG 811t2 | BXE--269.INT 4 | ECE--269.INT 4 | RPE--365.INT 8 | XBER16C2.INT 4 |
| *X. axonopodis patelii* | LMG 811t2R | BXE--210.INT 26 | ECE--212.INT 26 | RPE--217.INT 25 | XBER16C1.INT 26 |
| *X. axonopodis patelii* | LMG 811t3 | BXE--270.INT 7 | ECE--273.INT 7 | RPE--275.INT 7 | XBER16C3.INT 7 |
| *X. axonopodis phaseoli* | LMG 7455 | BXE--003.INT 28 | ECE--006.INT 28 | RPE--009.INT 28 | XBER1C1.INT 28 |
| *X. axonopodis phaseoli* | LMG 8014 | BXE--259.INT 24 | ECE--263.INT 24 | RPE--267.INT 25 | XBER15C1.INT 4 |

| | | | | | |
|---|---|---|---|---|---|
| *X. axonopodis phaseoli* | LMG 821 | BXE--181.INT 5 | ECE--177.INT 27 | RPE--173.INT 5 | XBER7C2.INT 19 |
| *X. axonopodis phaseoli* | LMG 821t1R | BXE--257.INT 26 | ECE--301.INT 4 | RPE--265.INT 27 | XBER12C2.INT 26 |
| *X. axonopodis phaseoli* | LMG 821t2R | BXE--305.INT 6 | ECE--261.INT 26 | RPE--307.INT 8 | XBER12C1.INT 23 |
| *X. axonopodis phaseoli* | LMG 823 | BXE--239.INT 9 | ECE--231.INT 9 | RPE--235.INT 9 | XBER7C2.INT 9 |
| *X. axonopodis phaseoli* | LMG 842 | BXE--207.INT 5 | ECE--206.INT 5 | RPE--208.INT 5 | XBER14C2.INT 25 |
| *X. axonopodis phaseoli(fusc)* | LMG 7456 | BXE--260.INT 3 | ECE--264.INT 3 | RPE--268.INT 3 | XBER15C2.INT 3 |
| *X. axonopodis phaseoli(fusc)* | LMG 7511 | BXE--199.INT 13 | ECE--203.INT 13 | RPE--201.INT 13 | XBER14C1.INT 13 |
| *X. axonopodis phaseoli(fusc)* | LMG 8036 | BXE--259.INT 13 | ECE--361.INT 14 | RPE--267.INT 14 | XBER14C2.INT 13 |
| *X. axonopodis phaseoli(fusc)* | LMG 8038 | BXE--199.INT 8 | ECE--203.INT 8 | RPE--201.INT 8 | XBER14C1.INT 8 |
| *X. axonopodis phaseoli(fusc)* | LMG 8251 | BXE--238.INT 29 | ECE--230.INT 29 | RPE--234.INT 29 | XBER7C1.INT 11 |
| *X. axonopodis phaseoli(fusc)* | LMG 8371 | BXE--160.INT 13 | ECE--177.INT 11 | RPE--162.INT 15 | XBER7C1.INT 13 |
| *X. axonopodis phaseoli(fusc)* | LMG 8371R | BXE--305.INT 15 | ECE--360.INT 24 | RPE--342.INT 14 | XBER13C2.INT 23 |
| *X. axonopodis phaseoli(fusc)* | LMG 8372 | BXE--160.INT 14 | ECE--177.INT 12 | RPE--162.INT 16 | XBER7C1.INT 14 |
| *X. axonopodis phaseoli(fusc)* | LMG 841 | BXE--239.INT 2 | ECE--231.INT 2 | RPE--235.INT 2 | XBER7C2.INT 2 |
| *X. axonopodis ricini* | LMG 7441 | BXE--199.INT 15 | ECE--203.INT 15 | RPE--201.INT 15 | XBER14C1.INT 15 |
| *X. axonopodis ricini* | LMG 7442 | BXE--259.INT 10 | ECE--361.INT 12 | RPE--267.INT 12 | XBER14C2.INT 10 |
| *X. axonopodis ricini* | LMG 7443 | BXE--181.INT 29 | ECE--178.INT 25 | RPE--174.INT 3 | XBER8C1.INT 29 |
| *X. axonopodis ricini* | LMG 7444 | BXE--305.INT 20 | ECE--361.INT 22 | RPE--307.INT 24 | XBER15C2.INT 4 |
| *X. axonopodis ricini* | LMG 861 | BXE--181.INT 28 | ECE--178.INT 24 | RPE--174.INT 2 | XBER8C1.INT 28 |
| *X. axonopodis ricini* | LMG 862 | BXE--259.INT 6 | ECE--361.INT 7 | RPE--267.INT 7 | XBER14C2.INT 6 |
| *X. axonopodis tamarindi* | LMG 869 | BXE--218.INT 7 | ECE--219.INT 7 | RPE--220.INT 7 | XBER16C3.INT 19 |
| *X. axonopodis tamarindi* | LMG 955 | BXE--270.INT 17 | ECE--273.INT 17 | RPE--275.INT 17 | XBER17C1.INT 5 |
| *X. axonopodis vasculorum* | LMG 894 | BXE--222.INT 2 | ECE--223.INT 2 | RPE--224.INT 2 | XBER17C2.INT 14 |
| *X. axonopodis vasculorum* | LMG 895 | BXE--271.INT 9 | ECE--274.INT 9 | RPE--276.INT 9 | XBER17C2.INT 9 |
| *X. axonopodis vasculorum* | LMG 899 | BXE--182.INT 13 | ECE--179.INT 11 | RPE--174.INT 15 | XBER9C1.INT 13 |
| *X. axonopodis vasculorum* | LMG 901 | BXE--240.INT 6 | ECE--232.INT 6 | RPE--236.INT 6 | XBER8C2.INT 6 |
| *X. axonopodis vasculorum* | LMG 901R | BXE--193.INT 22 | ECE--209.INT 5 | RPE--205.INT 22 | XBER12C1.INT 22 |
| *X. axonopodis vasculorum* | LMG 903 | BXE--271.INT 24 | ECE--274.INT 24 | RPE--276.INT 24 | XBER18C1.INT 12 |
| *X. axonopodis vesicatoria* | LMG 667 | BXE--182.INT 20 | ECE--179.INT 17 | RPE--174.INT 22 | XBER9C1.INT 20 |
| *X. axonopodis vesicatoria* | LMG 668 | BXE--241.INT 14 | ECE--244.INT 14 | RPE--247.INT 14 | XBER9C2.INT 14 |
| *X. axonopodis vesicatoria* | LMG 905 | BXE--003.INT 29 | ECE--006.INT 29 | RPE--009.INT 29 | XBER1C1.INT 29 |
| *X. axonopodis vesicatoria* | LMG 906 | BXE--271.INT 10 | ECE--274.INT 10 | RPE--276.INT 10 | XBER17C2.INT 10 |
| *X. axonopodis vesicatoria* | LMG 909 | BXE--218.INT 23 | ECE--219.INT 23 | RPE--220.INT 23 | XBER17C1.INT 23 |
| *X. axonopodis vesicatoria* | LMG 910 | BXE--013.INT 2 | ECE--015.INT 2 | RPE--018.INT 2 | XBER2C2.INT 2 |
| *X. axonopodis vesicatoria* | LMG 913 | BXE--270.INT 23 | ECE--273.INT 23 | RPE--275.INT 22 | XBER16C3.INT 23 |
| *X. axonopodis vesicatoria* | LMG 914t2 | BXE--306.INT 3 | ECE--378.INT 21 | RPE--306.INT 15 | XBER17C1.INT 16 |
| *X. axonopodis vesicatoria* | LMG 922 | BXE--222.INT 15 | ECE--223.INT 14 | RPE--224.INT 15 | XBER17C2.INT 27 |
| *X. axonopodis vesicatoria* | LMG 929 | BXE--013.INT 3 | ECE--031.INT 12 | RPE--018.INT 3 | XBER2C2.INT 3 |
| *X. axonopodis vesicatoria* | LMG 932 | BXE--218.INT 22 | ECE--219.INT 22 | RPE--220.INT 22 | XBER17C1.INT 22 |
| *X. axonopodis vitians* | LMG 8689 | BXE--306.INT 5 | ECE--378.INT 23 | RPE--306.INT 17 | XBER18C1.INT 13 |
| *X. axonopodis vitians* | LMG 937 | BXE--182.INT 14 | ECE--179.INT 12 | RPE--174.INT 16 | XBER9C1.INT 14 |
| *X. bromi* | LMG 8267 | BXE--225.INT 9 | ECE--255.INT 9 | RPE--227.INT 9 | XBER4C2.INT 9 |
| *X. bromi* | LMG 8269 | BXE--136.INT 18 | ECE--226.INT 10 | RPE--135.INT 18 | XBER4C1.INT 18 |
| *X. bromi* | LMG 8269R | BXE--269.INT 2 | ECE--378.INT 14 | RPE--269.INT 2 | XBER16C2.INT 2 |
| *X. bromi* | LMG 8272 | BXE--136.INT 17 | ECE--226.INT 12 | RPE--135.INT 17 | XBER4C1.INT 17 |
| *X. bromi* | LMG 8272R | BXE--225.INT 26 | ECE--255.INT 26 | RPE--227.INT 26 | XBER4C2.INT 26 |
| *X. bromi* | LMG 8272RR | BXE--210.INT 12 | ECE--212.INT 12 | RPE--217.INT 12 | XBER16C1.INT 12 |
| *X. campestris barbarae* | LMG 547 | BXE--004.INT 15 | ECE--015.INT 15 | RPE--018.INT 15 | XBER2C2.INT 15 |
| *X. campestris barbarae* | LMG 7385 | BXE--001.INT 10 | ECE--007.INT 16 | RPE--010.INT 16 | XBER2C1.INT 16 |
| *X. campestris campestris* | LMG 567 | BXE--243.INT 9 | ECE--246.INT 9 | RPE--249.INT 9 | XBER11C2.INT 9 |
| *X. campestris campestris* | LMG 568 | BXE--001.INT 11 | ECE--007.INT 17 | RPE--010.INT 17 | XBER2C1.INT 17 |
| *X. campestris campestris* | LMG 571 | BXE--237.INT 15 | ECE--229.INT 15 | RPE--233.INT 15 | XBER5C2.INT 15 |
| *X. campestris campestris* | LMG 571R | BXE--193.INT 20 | ECE--221.INT 13 | RPE--342.INT 7 | XBER12C1.INT 20 |
| *X. campestris campestris* | LMG 573 | BXE--243.INT 19 | ECE--246.INT 19 | RPE--249.INT 19 | XBER11C2.INT 19 |
| *X. campestris campestris* | LMG 573R | BXE--222.INT 6 | ECE--223.INT 6 | RPE--224.INT 6 | XBER17C2.INT 18 |
| *X. campestris campestris* | LMG 583 | BXE--237.INT 4 | ECE--229.INT 4 | RPE--233.INT 4 | XBER5C2.INT 4 |
| *X. campestris campestris* | LMG 7516 | BXE--139.INT 17 | ECE--175.INT 11 | RPE--161.INT 11 | XBER5C1.INT 17 |
| *X. campestris campestris* | LMG 7516RA | BXE--340.INT 2 | ECE--261.INT 6 | RPE--265.INT 6 | XBER12C2.INT 6 |
| *X. campestris campestris* | LMG 7662 | BXE--192.INT 24 | ECE--194.INT 24 | RPE--204.INT 24 | XBER11C1.INT 24 |
| *X. campestris campestris* | LMG 8001 | BXE--243.INT 13 | ECE--246.INT 13 | RPE--249.INT 13 | XBER11C2.INT 13 |
| *X. campestris campestris* | LMG 8003 | BXE--192.INT 5 | ECE--194.INT 5 | RPE--204.INT 5 | XBER11C1.INT 5 |

| | | | | | |
|---|---|---|---|---|---|
| *X. campestris campestris* | LMG 8032 | BXE–243.INT 20 | ECE–246.INT 20 | RPE–249.INT 20 | XBER11C2.INT 20 |
| *X. campestris campestris* | LMG 8035 | BXE–192.INT 22 | ECE–194.INT 22 | RPE–204.INT 22 | XBER11C1.INT 22 |
| *X. campestris campestris* | LMG 8035R | BXE–306.INT 4 | ECE–380.INT 17 | RPE–306.INT 16 | XBER17C2.INT 12 |
| *X. campestris campestris* | LMG 8051 | BXE–192.INT 14 | ECE–194.INT 14 | RPE–204.INT 14 | XBER11C1.INT 14 |
| *X. campestris campestris* | LMG 8055 | BXE–243.INT 26 | ECE–246.INT 26 | RPE–249.INT 26 | XBER11C2.INT 26 |
| *X. campestris campestris* | LMG 8082 | BXE–192.INT 2 | ECE–194.INT 2 | RPE–204.INT 2 | XBER11C1.INT 2 |
| *X. campestris campestris* | LMG 8099 | BXE–243.INT 10 | ECE–246.INT 10 | RPE–249.INT 10 | XBER11C2.INT 10 |
| *X. campestris campestris* | LMG 8100 | BXE–192.INT 6 | ECE–194.INT 6 | RPE–204.INT 6 | XBER11C1.INT 6 |
| *X. campestris campestris* | LMG 8112 | BXE–243.INT 8 | ECE–246.INT 8 | RPE–249.INT 8 | XBER11C2.INT 8 |
| *X. campestris campestris* | LMG 8119 | BXE–192.INT 15 | ECE–194.INT 15 | RPE–204.INT 15 | XBER11C1.INT 15 |
| *X. campestris campestris* | LMG 8119R | BXE–271.INT 16 | ECE–274.INT 16 | RPE–276.INT 16 | XBER18C1.INT 4 |
| *X. campestris campestris* | LMG 8121 | BXE–192.INT 4 | ECE–194.INT 4 | RPE–204.INT 4 | XBER11C1.INT 4 |
| *X. campestris campestris* | LMG 8123 | BXE–243.INT 22 | ECE–246.INT 22 | RPE–249.INT 22 | XBER11C2.INT 22 |
| *X. campestris campestris** | LMG 575 | BXE–136.INT 9 | ECE–226.INT 19 | RPE–135.INT 9 | XBER4C1.INT 9 |
| *X. campestris campestris** | LMG 575RA | BXE–257.INT 9 | ECE–261.INT 9 | RPE–265.INT 9 | XBER12C2.INT 9 |
| *X. campestris campestris** | LMG 7514 | BXE–139.INT 22 | ECE–175.INT 6 | RPE–161.INT 6 | XBER5C1.INT 22 |
| *X. campestris campestris** | LMG 7514R | BXE–257.INT 22 | ECE–261.INT 21 | RPE–265.INT 23 | XBER12C2.INT 22 |
| *X. campestris incanae* | LMG 7421 | BXE–001.INT 12 | ECE–007.INT 18 | RPE–010.INT 18 | XBER2C1.INT 18 |
| *X. campestris incanae* | LMG 7490 | BXE–004.INT 19 | ECE–015.INT 19 | RPE–018.INT 19 | XBER2C2.INT 19 |
| *X. campestris raphani* | LMG 7505 | BXE–004.INT 21 | ECE–007.INT 21 | RPE–010.INT 21 | XBER2C1.INT 21 |
| *X. campestris raphani* | LMG 8134 | BXE–020.INT 3 | ECE–015.INT 22 | RPE–018.INT 22 | XBER2C2.INT 22 |
| *X. campestris raphani** | LMG 8601 | BXE–199.INT 19 | ECE–203.INT 19 | RPE–201.INT 19 | XBER14C1.INT 19 |
| *X. campestris raphani** | LMG 8602 | BXE–259.INT 5 | ECE–361.INT 6 | RPE–342.INT 15 | XBER14C2.INT 5 |
| *X. cassavae* | LMG 5246 | BXE–198.INT 8 | ECE–202.INT 8 | RPE–200.INT 8 | XBER13C1.INT 8 |
| *X. cassavae* | LMG 5264 | BXE–012.INT 22 | ECE–014.INT 22 | RPE–017.INT 22 | XBER1C2.INT 22 |
| *X. cassavae* | LMG 5270 | BXE–136.INT 3 | ECE–226.INT 25 | RPE–135.INT 3 | XBER4C1.INT 3 |
| *X. cassavae* | LMG 5270R | BXE–257.INT 15 | ECE–261.INT 14 | RPE–265.INT 15 | XBER12C2.INT 15 |
| *X. cassavae* | LMG 670 | BXE–003.INT 19 | ECE–006.INT 19 | RPE–009.INT 19 | XBER1C1.INT 19 |
| *X. cassavae* | LMG 670 | BXE–259.INT 19 | ECE–263.INT 19 | RPE–267.INT 20 | XBER14C2.INT 19 |
| *X. cassavae* | LMG 673 | BXE–003.INT 21 | ECE–006.INT 21 | RPE–017.INT 21 | XBER1C1.INT 21 |
| *X. codiaei* | LMG 8677 | BXE–240.INT 3 | ECE–232.INT 3 | RPE–236.INT 3 | XBER8C2.INT 3 |
| *X. codiaei* | LMG 8678 | BXE–003.INT 23 | ECE–006.INT 23 | RPE–009.INT 23 | XBER1C1.INT 23 |
| *X. cucurbitae* | LMG 690 | BXE–238.INT 9 | ECE–230.INT 9 | RPE–234.INT 9 | XBER6C2.INT 9 |
| *X. cucurbitae* | LMG 690R | BXE–193.INT 16 | ECE–195.INT 16 | RPE–205.INT 16 | XBER12C1.INT 16 |
| *X. cucurbitae* | LMG 7480 | BXE–238.INT 4 | ECE–230.INT 4 | RPE–234.INT 4 | XBER6C2.INT 4 |
| *X. cucurbitae* | LMG 7481 | BXE–198.INT 7 | ECE–202.INT 7 | RPE–200.INT 7 | XBER13C1.INT 7 |
| *X. cucurbitae* | LMG 8661 | BXE–259.INT 4 | ECE–361.INT 5 | RPE–267.INT 5 | XBER14C2.INT 4 |
| *X. cucurbitae* | LMG 8662 | BXE–003.INT 25 | ECE–006.INT 25 | RPE–009.INT 25 | XBER1C1.INT 25 |
| *X. fragariae* | LMG 706 | BXE–012.INT 2 | ECE–031.INT 3 | RPE–017.INT 2 | XBER1C2.INT 2 |
| *X. fragariae* | LMG 706R | BXE–241.INT 20 | ECE–244.INT 20 | RPE–247.INT 20 | XBER9-1.INT 4 |
| *X. fragariae* | LMG 708 | BXE–012.INT 3 | ECE–055.INT 2 | RPE–017.INT 3 | XBER1C2.INT 3 |
| *X. fragariae* | LMG 710 | BXE–136.INT 22 | ECE–226.INT 7 | RPE–135.INT 22 | XBER4C1.INT 22 |
| *X. hortorum hederae* | LMG 733T | BXE–225.INT 2 | ECE–255.INT 2 | RPE–227.INT 2 | XBER4C2.INT 2 |
| *X. hortorum hederae* | LMG 733TR | BXE–159.INT 21 | ECE–176.INT 20 | RPE–142.INT 21 | XBER6C1.INT 21 |
| *X. hortorum hederae* | LMG 734 | BXE–238.INT 16 | ECE–230.INT 16 | RPE–234.INT 16 | XBER6C2.INT 16 |
| *X. hortorum hederae* | LMG 7413 | BXE–160.INT 8 | ECE–230.INT 27 | RPE–162.INT 11 | XBER6C2.INT 27 |
| *X. hortorum hederae* | LMG 7414 | BXE–238.INT 5 | ECE–230.INT 5 | RPE–234.INT 5 | XBER6C2.INT 5 |
| *X. hortorum hederae* | LMG 8665 | BXE–207.INT 17 | ECE–206.INT 17 | RPE–208.INT 17 | XBER15C1.INT 17 |
| *X. hortorum pelargonii* | LMG 7312 | BXE–239.INT 24 | ECE–231.INT 24 | RPE–235.INT 24 | XBER8C1.INT 9 |
| *X. hortorum pelargonii* | LMG 7314 | BXE–003.INT 4 | ECE–006.INT 4 | RPE–009.INT 4 | XBER1C1.INT 4 |
| *X. hortorum pelargonii* | LMG 7314R | BXE–225.INT 5 | ECE–255.INT 5 | RPE–227.INT 5 | XBER4C2.INT 5 |
| *X. hortorum pelargonii* | LMG 7314RR | BXE–160.INT 16 | ECE–177.INT 14 | RPE–162.INT 18 | XBER7C1.INT 16 |
| *X. hortorum pelargonii* | LMG 7315 | BXE–239.INT 15 | ECE–231.INT 15 | RPE–235.INT 15 | XBER7C2.INT 15 |
| *X. hortorum pelargonii* | LMG 7316 | BXE–160.INT 22 | ECE–177.INT 20 | RPE–162.INT 24 | XBER7C1.INT 22 |
| *X. hortorum pelargonii* | LMG 7317 | BXE–239.INT 28 | ECE–231.INT 28 | RPE–235.INT 28 | XBER8C1.INT 14 |
| *X. hortorum pelargonii* | LMG 7354 | BXE–210.INT 18 | ECE–212.INT 18 | RPE–217.INT 17 | XBER16C1.INT 18 |
| *X. hortorum pelargonii* | LMG 7356 | BXE–239.INT 26 | ECE–231.INT 26 | RPE–235.INT 26 | XBER8C1.INT 12 |
| *X. hortorum pelargonii* | LMG 7585 | BXE–207.INT 20 | ECE–206.INT 20 | RPE–208.INT 20 | XBER15C1.INT 20 |
| *X. hortorum pelargonii* | LMG 7690 | BXE–272.INT 5 | ECE–272.INT 5 | RPE–272.INT 5 | XBER18C2.INT 5 |
| *X. hortorum pelargonii* | LMG 7690R | BXE–210.INT 3 | ECE–212.INT 3 | RPE–217.INT 3 | XBER15C2.INT 23 |
| *X. hortorum pelargonii* | LMG 7706 | BXE–259.INT 15 | ECE–263.INT 15 | RPE–267.INT 16 | XBER14C2.INT 15 |

| | | | | | |
|---|---|---|---|---|---|
| *X. hortorum pelargonii* | LMG 7708 | BXE--207.INT 9 | ECE--206.INT 9 | RPE--208.INT 9 | XBER14C2.INT 29 |
| *X. hortorum pelargonii* | LMG 7710 | BXE--259.INT 18 | ECE--263.INT 18 | RPE--267.INT 19 | XBER14C2.INT 18 |
| *X. hortorum pelargonii* | LMG 7710R | BXE--207.INT 23 | ECE--206.INT 23 | RPE--208.INT 23 | XBER15C1.INT 23 |
| *X. hortorum pelargonii* | LMG 7712 | BXE--305.INT 26 | ECE--361.INT 26 | RPE--307.INT 28 | XBER15C2.INT 20 |
| *X. hortorum pelargonii* | LMG 7712R | BXE--210.INT 2 | ECE--361.INT 27 | RPE--217.INT 2 | XBER15C2.INT 22 |
| *X. hortorum pelargonii* | LMG 7715 | BXE--305.INT 22 | ECE--264.INT 6 | RPE--268.INT 6 | XBER15C2.INT 6 |
| *X. hortorum pelargonii* | LMG 7763 | BXE--160.INT 18 | ECE--177.INT 16 | RPE--162.INT 21 | XBER7C1.INT 18 |
| *X. hortorum pelargonii* | LMG 7764 | BXE--239.INT 5 | ECE--231.INT 5 | RPE--235.INT 5 | XBER7C2.INT 5 |
| *X. hortorum pelargonii* | LMG 820 | BXE--181.INT 7 | ECE--178.INT 3 | RPE--173.INT 7 | XBER7C2.INT 22 |
| *X. hortorum pelargonii* | LMG 820R | BXE--241.INT 17 | ECE--244.INT 17 | RPE--247.INT 17 | XBER9C2.INT 17 |
| *X. hortorum vitians* | LMG 7508 | BXE--181.INT 22 | ECE--178.INT 17 | RPE--173.INT 22 | XBER8C1.INT 22 |
| *X. hortorum vitians* | LMG 7510 | BXE--240.INT 23 | ECE--232.INT 23 | RPE--236.INT 24 | XBER9C1.INT 8 |
| *X. hortorum vitians* | LMG 8688 | BXE--218.INT 12 | ECE--219.INT 12 | RPE--220.INT 12 | XBER16C3.INT 25 |
| *X. hortorum vitians* | LMG 8690t1 | BXE--271.INT 5 | ECE--274.INT 5 | RPE--276.INT 5 | XBER17C2.INT 5 |
| *X. hortorum vitians* | LMG 8690t2 | BXE--218.INT 9 | ECE--219.INT 9 | RPE--220.INT 9 | XBER16C3.INT 22 |
| *X. hortorum vitians* | LMG 938 | BXE--254.INT 2 | ECE--256.INT 14 | RPE--254.INT 6 | XBER9C2.INT 18 |
| *X. hyacinthi* | LMG 739* | BXE--136.INT 23 | ECE--226.INT 6 | RPE--135.INT 23 | XBER4C1.INT 23 |
| *X. hyacinthi* | LMG 739R | BXE--254.INT 3 | ECE--256.INT 15 | RPE--254.INT 7 | XBER9C2.INT 19 |
| *X. hyacinthi* | LMG 740 | BXE--160.INT 24 | ECE--177.INT 22 | RPE--162.INT 26 | XBER7C1.INT 24 |
| *X. hyacinthi* | LMG 7419 | BXE--260.INT 26 | ECE--362.INT 3 | RPE--268.INT 26 | XBER16C1.INT 6 |
| *X. hyacinthi* | LMG 742 | BXE--002.INT 6 | ECE--008.INT 6 | RPE--047.INT 8 | XBER3C1.INT 6 |
| *X. hyacinthi* | LMG 8041 | BXE--005.INT 7 | ECE--055.INT 9 | RPE--048.INT 7 | XBER3C2.INT 7 |
| *X. hyacinthi* | LMG 8042 | BXE--210.INT 20 | ECE--212.INT 20 | RPE--217.INT 19 | XBER16C1.INT 20 |
| *X. melonis* | LMG 8670 | BXE--305.INT 24 | ECE--264.INT 10 | RPE--268.INT 10 | XBER15C2.INT 10 |
| *X. melonis* | LMG 8671 | BXE--181.INT 6 | ECE--178.INT 2 | RPE--173.INT 6 | XBER7C2.INT 20 |
| *X. melonis* | LMG 8672 | BXE--013.INT 8 | ECE--013.INT 3 | RPE--018.INT 8 | XBER2C2.INT 8 |
| *X. melonis* | LMG 8674 | BXE--181.INT 16 | ECE--178.INT 12 | RPE--173.INT 16 | XBER8C1.INT 16 |
| *X. oryzae oryzae* | LMG 5047T | BXE--013.INT 4 | ECE--056.INT 3 | RPE--018.INT 4 | XBER2C2.INT 4 |
| *X. oryzae oryzae* | LMG 5047TR | BXE--222.INT 18 | ECE--223.INT 17 | RPE--224.INT 18 | XBER18C1.INT 18 |
| *X. oryzae oryzae* | LMG 641t1 | BXE--270.INT 24 | ECE--273.INT 24 | RPE--275.INT 23 | XBER17C1.INT 11 |
| *X. oryzae oryzae* | LMG 6518 | BXE--218.INT 18 | ECE--219.INT 19 | RPE--220.INT 18 | XBER17C1.INT 18 |
| *X. oryzae oryzae* | LMG 803 | BXE--240.INT 10 | ECE--232.INT 10 | RPE--236.INT 10 | XBER8C2.INT 10 |
| *X. oryzae oryzae* | LMG 803R | BXE--193.INT 6 | ECE--195.INT 6 | RPE--205.INT 6 | XBER12C1.INT 6 |
| *X. oryzae oryzae* | LMG 806 | BXE--241.INT 7 | ECE--244.INT 7 | RPE--247.INT 7 | XBER9C2.INT 7 |
| *X. oryzae oryzicola* | LMG 657 | BXE--182.INT 12 | ECE--179.INT 9 | RPE--174.INT 14 | XBER9C1.INT 12 |
| *X. oryzae oryzicola* | LMG 657R | BXE--257.INT 20 | ECE--360.INT 7 | RPE--342.INT 6 | XBER12C2.INT 20 |
| *X. oryzae oryzicola* | LMG 661 | BXE--181.INT 24 | ECE--178.INT 20 | RPE--173.INT 24 | XBER8C1.INT 24 |
| *X. oryzae oryzicola* | LMG 661R | BXE--257.INT 19 | ECE--360.INT 6 | RPE--342.INT 5 | XBER12C2.INT 19 |
| *X. oryzae oryzicola* | LMG 665 | BXE--004.INT 5 | ECE--007.INT 5 | RPE--010.INT 5 | XBER2C1.INT 5 |
| *X. oryzae oryzicola* | LMG 665R | BXE--240.INT 16 | ECE--232.INT 16 | RPE--236.INT 17 | XBER9C1.INT 2 |
| *X. oryzae oryzicola* | LMG 793 | BXE--182.INT 5 | ECE--179.INT 3 | RPE--174.INT 7 | XBER8C2.INT 19 |
| *X. oryzae oryzicola* | LMG 793R | BXE--305.INT 3 | ECE--360.INT 3 | RPE--342.INT 3 | XBER12C2.INT 8 |
| *X. oryzae oryzicola* | LMG 797 | BXE--222.INT 9 | ECE--223.INT 9 | RPE--224.INT 9 | XBER17C2.INT 22 |
| *X. pisi* | LMG 847t1* | BXE--039.INT 12 | ECE--015.INT 7 | RPE--018.INT 7 | XBER9C2.INT 2 |
| *X. pisi* | LMG 847t1R | BXE--218.INT 3 | ECE--219.INT 3 | RPE--220.INT 3 | XBER2C1.INT 7 |
| *X. pisi* | LMG 847t2 | BXE--241.INT 2 | ECE--244.INT 2 | RPE--247.INT 2 | XBER17C1.INT 3 |
| *X. populi* | LMG 5743 | BXE--003.INT 7 | ECE--006.INT 7 | RPE--009.INT 7 | XBER1C1.INT 7 |
| *X. populi* | LMG 5753 | BXE--025.INT 2 | ECE--014.INT 8 | RPE--017.INT 8 | XBER1C2.INT 8 |
| *X. populi* | LMG 974 | BXE--003.INT 6 | ECE--006.INT 6 | RPE--009.INT 6 | XBER1C1.INT 6 |
| *X. sacchari* | LMG 471 | BXE--039.INT 29 | ECE--055.INT 12 | RPE--048.INT 9 | XBER3C2.INT 9 |
| *X. sacchari* | LMG 476 | BXE--139.INT 20 | ECE--175.INT 8 | RPE--161.INT 8 | XBER5C1.INT 20 |
| *X. sacchari* | LMG 476R | BXE--257.INT 16 | ECE--360.INT 5 | RPE--265.INT 16 | XBER12C2.INT 16 |
| *X. theicola* | LMG 8684 | BXE--002.INT 8 | ECE--008.INT 8 | RPE--048.INT 8 | XBER3C1.INT 8 |
| *X. theicola* | LMG 8685 | BXE--225.INT 13 | ECE--255.INT 13 | RPE--227.INT 13 | XBER4C2.INT 13 |
| *X. theicola* | LMG 8686 | BXE--136.INT 15 | ECE--226.INT 14 | RPE--227.INT 14 | XBER4C1.INT 15 |
| *X. translucens arrhenatheri* | LMG 590 | BXE--257.INT 4 | ECE--300.INT 22 | RPE--265.INT 4 | XBER12C2.INT 4 |
| *X. translucens arrhenatheri* | LMG 591 | BXE--136.INT 13 | ECE--226.INT 15 | RPE--135.INT 13 | XBER4C1.INT 13 |
| *X. translucens arrhenatheri* | LMG 727R | BXE--254.INT 5 | ECE--256.INT 17 | RPE--254.INT 9 | XBER9C2.INT 22 |
| *X. translucens arrhenatheri* | LMG 727t1 | BXE--004.INT 23 | ECE--007.INT 23 | RPE--010.INT 23 | XBER2C1.INT 23 |
| *X. translucens arrhenatheri* | LMG 7384 | BXE--225.INT 16 | ECE--255.INT 16 | RPE--227.INT 16 | XBER4C2.INT 16 |
| *X. translucens cerealis* | LMG 7392 | BXE--136.INT 11 | ECE--226.INT 17 | RPE--135.INT 11 | XBER4C1.INT 11 |

| | | | | | |
|---|---|---|---|---|---|
| *X. translucens cerealis* | LMG 887 | BXE--237.INT 14 | ECE--229.INT 14 | RPE--233.INT 14 | XBER5C2.INT 14 |
| *X. translucens cerealis* | LMG 891 | BXE--139.INT 4 | ECE--175.INT 24 | RPE--161.INT 24 | XBER5C1.INT 4 |
| *X. translucens cerealis** | LMG 679 | BXE--020.INT 5 | ECE--015.INT 24 | RPE--018.INT 24 | XBER2C2.INT 24 |
| *X. translucens cerealis** | LMG 679R | BXE--241.INT 23 | ECE--244.INT 23 | RPE--247.INT 23 | XBER9-1.INT 6 |
| *X. translucens cerealis** | LMG 880 | BXE--225.INT 18 | ECE--255.INT 18 | RPE--227.INT 18 | XBER4C2.INT 18 |
| *X. translucens graminis* | LMG 595 | BXE--160.INT 5 | ECE--177.INT 3 | RPE--162.INT 7 | XBER6C2.INT 24 |
| *X. translucens graminis* | LMG 713 | BXE--260.INT 27 | ECE--380.INT 16 | RPE--306.INT 4 | XBER16C1.INT 7 |
| *X. translucens graminis* | LMG 726 | BXE--136.INT 26 | ECE--226.INT 3 | RPE--135.INT 26 | XBER4C1.INT 26 |
| *X. translucens graminis* | LMG 726R | BXE--238.INT 2 | ECE--256.INT 7 | RPE--234.INT 2 | XBER6C2.INT 2 |
| *X. translucens hordei* | LMG 737(8/7) | BXE--182.INT 18 | ECE--179.INT 16 | RPE--174.INT 21 | XBER9C1.INT 18 |
| *X. translucens hordei* | LMG 737(9/7) | BXE--241.INT 13 | ECE--244.INT 13 | RPE--247.INT 13 | XBER9C2.INT 13 |
| *X. translucens hordei* | LMG 8279 | BXE--004.INT 27 | ECE--007.INT 27 | RPE--018.INT 27 | XBER2C1.INT 27 |
| *X. translucens hordei* | LMG 8279t1 | BXE--271.INT 17 | ECE--274.INT 17 | RPE--276.INT 17 | XBER18C1.INT 5 |
| *X. translucens hordei* | LMG 8279t2 | BXE--210.INT 5 | ECE--362.INT 2 | RPE--217.INT 5 | XBER15C2.INT 25 |
| *X. translucens hordei* | LMG 882 | BXE--020.INT 7 | ECE--031.INT 24 | RPE--018.INT 26 | XBER2C2.INT 26 |
| *X. translucens hordei* | LMG 884 | BXE--210.INT 8 | ECE--303.INT 2 | RPE--217.INT 8 | XBER15C2.INT 28 |
| *X. translucens phlei* | LMG 716 | BXE--240.INT 5 | ECE--232.INT 5 | RPE--236.INT 5 | XBER8C2.INT 5 |
| *X. translucens phlei* | LMG 719 | BXE--181.INT 26 | ECE--232.INT 12 | RPE--173.INT 26 | XBER8C1.INT 26 |
| *X. translucens phlei* | LMG 723 | BXE--305.INT 17 | ECE--378.INT 6 | RPE--267.INT 9 | XBER14C2.INT 8 |
| *X. translucens phlei* | LMG 730 | BXE--004.INT 28 | ECE--015.INT 28 | RPE--018.INT 28 | XBER2C1.INT 28 |
| *X. translucens poae* | LMG 594 | BXE--260.INT 15 | ECE--361.INT 24 | RPE--268.INT 15 | XBER15C2.INT 15 |
| *X. translucens poae* | LMG 728 | BXE--001.INT 4 | ECE--008.INT 2 | RPE--048.INT 2 | XBER3C1.INT 2 |
| *X. translucens secalis* | LMG 7507 | BXE--270.INT 18 | ECE--273.INT 18 | RPE--275.INT 18 | XBER17C1.INT 6 |
| *X. translucens secalis* | LMG 883 | BXE--005.INT 3 | ECE--016.INT 3 | RPE--048.INT 3 | XBER3C1.INT 3 |
| *X. translucens translucens* | LMG 5259 | BXE--240.INT 18 | ECE--232.INT 18 | RPE--236.INT 19 | XBER9C1.INT 4 |
| *X. translucens translucens* | LMG 5260t1 | BXE--182.INT 6 | ECE--179.INT 4 | RPE--174.INT 8 | XBER8C2.INT 20 |
| *X. translucens translucens* | LMG 5260t2 | BXE--240.INT 8 | ECE--232.INT 8 | RPE--236.INT 8 | XBER8C2.INT 8 |
| *X. translucens translucens* | LMG 5262 | BXE--222.INT 8 | ECE--223.INT 8 | RPE--224.INT 8 | XBER17C2.INT 20 |
| *X. translucens translucens* | LMG 875 | BXE--271.INT 23 | ECE--362.INT 7 | RPE--365.INT 10 | XBER18C1.INT 11 |
| *X. translucens translucens* | LMG 876 | BXE--002.INT 4 | ECE--008.INT 4 | RPE--048.INT 4 | XBER3C1.INT 4 |
| *X. translucens undulosa* | LMG 8283 | BXE--272.INT 3 | ECE--272.INT 3 | RPE--272.INT 3 | XBER18C2.INT 3 |
| *X. translucens undulosa* | LMG 885 | BXE--181.INT 18 | ECE--178.INT 14 | RPE--173.INT 18 | XBER8C1.INT 18 |
| *X. translucens undulosa* | LMG 886 | BXE--270.INT 16 | ECE--273.INT 16 | RPE--275.INT 16 | XBER17C1.INT 4 |
| *X. translucens undulosa* | LMG 888 | BXE--218.INT 13 | ECE--219.INT 13 | RPE--220.INT 13 | XBER16C3.INT 26 |
| *X. translucens undulosa* | LMG 892 | BXE--005.INT 5 | ECE--016.INT 5 | RPE--047.INT 7 | XBER3C2.INT 5 |
| *X. vasicola holcicola* | LMG 736t1 | BXE--160.INT 9 | ECE--177.INT 6 | RPE--162.INT 12 | XBER6C2.INT 28 |
| *X. vasicola holcicola* | LMG 736t1RA | BXE--257.INT 14 | ECE--300.INT 26 | RPE--265.INT 14 | XBER12C2.INT 14 |
| *X. vasicola holcicola* | LMG 736t1RR | BXE--210.INT 9 | ECE--212.INT 9 | RPE--217.INT 9 | XBER15C2.INT 29 |
| *X. vasicola holcicola* | LMG 736t2 | BXE--237.INT 25 | ECE--229.INT 24 | RPE--233.INT 25 | XBER6C1.INT 3 |
| *X. vasicola holcicola* | LMG 736t2R | BXE--199.INT 2 | ECE--203.INT 2 | RPE--201.INT 2 | XBER14C1.INT 2 |
| *X. vasicola holcicola* | LMG 7416 | BXE--013.INT 6 | ECE--031.INT 14 | RPE--018.INT 6 | XBER2C2.INT 6 |
| *X. vasicola holcicola* | LMG 7489 | BXE--160.INT 2 | ECE--176.INT 26 | RPE--162.INT 4 | XBER6C2.INT 20 |
| *X. vasicola holcicola* | LMG 7489R | BXE--340.INT 3 | ECE--378.INT 2 | RPE--307.INT 3 | XBER12C2.INT 12 |
| *X. vasicola holcicola* | LMG 8276 | BXE--210.INT 22 | ECE--212.INT 22 | RPE--217.INT 21 | XBER16C1.INT 22 |
| *X. vasicola holcicola* | LMG 8276R | BXE--272.INT 6 | ECE--272.INT 6 | RPE--272.INT 6 | XBER18C2.INT 6 |
| *X. vasicola holcicola* | LMG 8277 | BXE--210.INT 28 | ECE--212.INT 28 | RPE--217.INT 27 | XBER16C1.INT 28 |
| *X. vasicola vasculorum* | LMG 8284 | BXE--270.INT 28 | ECE--273.INT 28 | RPE--275.INT 28 | XBER17C1.INT 15 |
| *X. vasicola vasculorum* | LMG 900 | BXE--182.INT 11 | ECE--179.INT 8 | RPE--174.INT 13 | XBER9C1.INT 11 |
| *X. vasicola vasculorum* | LMG 902 | BXE--241.INT 15 | ECE--244.INT 15 | RPE--247.INT 15 | XBER9C2.INT 15 |
| *X. vesicatoria* | LMG 911 | BXE--004.INT 9 | ECE--055.INT 4 | RPE--010.INT 9 | XBER2C1.INT 9 |
| *X. vesicatoria* | LMG 916 | BXE--240.INT 13 | ECE--232.INT 13 | RPE--236.INT 14 | XBER8C2.INT 13 |
| *X. vesicatoria* | LMG 917 | BXE--182.INT 3 | ECE--178.INT 27 | RPE--174.INT 5 | XBER8C2.INT 17 |
| *X. vesicatoria* | LMG 919 | BXE--271.INT 19 | ECE--274.INT 19 | RPE--276.INT 19 | XBER18C1.INT 7 |
| *X. vesicatoria* | LMG 920t1 | BXE--004.INT 10 | ECE--055.INT 5 | RPE--010.INT 10 | XBER2C1.INT 10 |
| *X. vesicatoria* | LMG 925 | BXE--270.INT 27 | ECE--273.INT 27 | RPE--275.INT 27 | XBER17C1.INT 14 |
| *X. vesicatoria* | LMG 935 | BXE--222.INT 17 | ECE--223.INT 16 | RPE--224.INT 17 | XBER17C2.INT 29 |
| **Training/Testing Set 2** | | | | | |
| *X. albilineans* | LMG 482 | BXE--225.INT 22 | ECE--255.INT 22 | RPE--227.INT 22 | XBER4C2.INT 22 |
| *X. albilineans* | LMG 487 | BXE--139.INT 23 | ECE--175.INT 5 | RPE--161.INT 5 | XBER5C1.INT 23 |
| *X. albilineans* | LMG 488 | BXE--258.INT 28 | ECE--360.INT 29 | RPE--266.INT 29 | XBER13C2.INT 28 |
| *X. albilineans* | LMG 490 | BXE--196.INT 18 | ECE--202.INT 18 | RPE--200.INT 18 | XBER13C1.INT 18 |

| | | | | | |
|---|---|---|---|---|---|
| *X. albilineans* | LMG 494 | BXE--005.INT 10 | ECE--016.INT 10 | RPE--048.INT 10 | XBER3C2.INT 10 |
| *X. arboricola corylina* | LMG 688 | BXE--025.INT 5 | ECE--006.INT 9 | RPE--009.INT 9 | XBER1C1.INT 9 |
| *X. arboricola corylina* | LMG 689 | BXE--012.INT 10 | ECE--014.INT 10 | RPE--017.INT 10 | XBER1C2.INT 10 |
| *X. arboricola corylina* | LMG 8660 | BXE--003.INT 13 | ECE--006.INT 13 | RPE--009.INT 13 | XBER1C1.INT 13 |
| *X. arboricola juglandis* | LMG 747T | BXE--225.INT 4 | ECE--255.INT 4 | RPE--227.INT 4 | XBER4C2.INT 4 |
| *X. arboricola juglandis* | LMG 750 | BXE--181.INT 3 | ECE--177.INT 25 | RPE--173.INT 3 | XBER7C2.INT 17 |
| *X. arboricola juglandis* | LMG 751 | BXE--239.INT 13 | ECE--231.INT 13 | RPE--235.INT 13 | XBER7C2.INT 13 |
| *X. arboricola juglandis* | LMG 8045 | BXE--181.INT 15 | ECE--178.INT 11 | RPE--173.INT 15 | XBER7C2.INT 29 |
| *X. arboricola juglandis* | LMG 8046 | BXE--305.INT 25 | ECE--361.INT 23 | RPE--307.INT 25 | XBER15C2.INT 13 |
| *X. arboricola juglandis* | LMG 8047 | BXE--003.INT 14 | ECE--006.INT 14 | RPE--009.INT 14 | XBER1C1.INT 14 |
| *X. arboricola poinsettiicola* | LMG 5402 | BXE--260.INT 5 | ECE--264.INT 5 | RPE--268.INT 5 | XBER15C2.INT 5 |
| *X. arboricola poinsettiicola* | LMG 5403 | BXE--003.INT 15 | ECE--006.INT 15 | RPE--009.INT 15 | XBER1C1.INT 15 |
| *X. arboricola pruni* | LMG 7438 | BXE--259.INT 7 | ECE--361.INT 8 | RPE--267.INT 8 | XBER14C2.INT 7 |
| *X. arboricola pruni* | LMG 852 | BXE--003.INT 17 | ECE--006.INT 17 | RPE--009.INT 17 | XBER1C1.INT 17 |
| *X. arboricola pruni* | LMG 853 | BXE--260.INT 17 | ECE--361.INT 25 | RPE--268.INT 17 | XBER15C2.INT 17 |
| *X. arboricola pruni* | LMG 853R | BXE--222.INT 21 | ECE--223.INT 20 | RPE--224.INT 21 | XBER18C1.INT 21 |
| *X. arboricola pruni* | LMG 854 | BXE--260.INT 18 | ECE--264.INT 18 | RPE--268.INT 18 | XBER15C2.INT 18 |
| *X. arboricola pruni* | LMG 855 | BXE--207.INT 28 | ECE--206.INT 28 | RPE--208.INT 28 | XBER15C1.INT 28 |
| *X. arboricola pruni* | LMG 8679 | BXE--240.INT 22 | ECE--232.INT 22 | RPE--236.INT 23 | XBER9C1.INT 7 |
| *X. arboricola pruni* | LMG 8680 | BXE--003.INT 18 | ECE--006.INT 18 | RPE--009.INT 18 | XBER1C1.INT 18 |
| *X. axonopodis alfalfae* | LMG 495 | BXE--225.INT 24 | ECE--255.INT 24 | RPE--227.INT 24 | XBER4C2.INT 24 |
| *X. axonopodis alfalfae* | LMG 497 | BXE--136.INT 5 | ECE--226.INT 23 | RPE--135.INT 5 | XBER4C1.INT 5 |
| *X. axonopodis alfalfae* | LMG 8018t1-1 | BXE--139.INT 8 | ECE--175.INT 20 | RPE--161.INT 20 | XBER5C1.INT 8 |
| *X. axonopodis alfalfae* | LMG 8018t1-2 | BXE--237.INT 17 | ECE--229.INT 17 | RPE--233.INT 17 | XBER5C2.INT 17 |
| *X. axonopodis alfalfae* | LMG 8079 | BXE--225.INT 27 | ECE--226.INT 27 | RPE--227.INT 27 | XBER5C1.INT 25 |
| *X. axonopodis alfalfae** | LMG 8019 | BXE--192.INT 13 | ECE--194.INT 13 | RPE--204.INT 13 | XBER11C1.INT 13 |
| *X. axonopodis alfalfae** | LMG 8020 | BXE--243.INT 27 | ECE--246.INT 27 | RPE--249.INT 27 | XBER11C2.INT 27 |
| *X. axonopodis alfalfae** | LMG 8020R | BXE--198.INT 24 | ECE--202.INT 24 | RPE--200.INT 24 | XBER13C1.INT 24 |
| *X. axonopodis alfalfae** | LMG 8020RR | BXE--272.INT 7 | ECE--272.INT 7 | RPE--272.INT 7 | XBER18C2.INT 7 |
| *X. axonopodis axonopodis* | LMG 538t1T | BXE--003.INT 26 | ECE--006.INT 26 | RPE--009.INT 26 | XBER1C1.INT 26 |
| *X. axonopodis axonopodis* | LMG 539 | BXE--238.INT 26 | ECE--230.INT 26 | RPE--234.INT 26 | XBER7C1.INT 7 |
| *X. axonopodis axonopodis* | LMG 539R | BXE--182.INT 24 | ECE--179.INT 22 | RPE--174.INT 26 | XBER9C1.INT 24 |
| *X. axonopodis axonopodis* | LMG 540 | BXE--238.INT 10 | ECE--230.INT 10 | RPE--234.INT 10 | XBER6C2.INT 10 |
| *X. axonopodis bauhiniae* | LMG 548 | BXE--192.INT 17 | ECE--194.INT 17 | RPE--204.INT 17 | XBER11C1.INT 17 |
| *X. axonopodis bauhiniae* | LMG 548R | BXE--271.INT 15 | ECE--274.INT 15 | RPE--276.INT 15 | XBER18C1.INT 3 |
| *X. axonopodis begoniae* | LMG 551 | BXE--192.INT 3 | ECE--194.INT 3 | RPE--204.INT 3 | XBER11C1.INT 3 |
| *X. axonopodis begoniae* | LMG 552 | BXE--237.INT 22 | ECE--229.INT 22 | RPE--233.INT 22 | XBER5C2.INT 22 |
| *X. axonopodis begoniae* | LMG 7178 | BXE--193.INT 2 | ECE--195.INT 2 | RPE--205.INT 2 | XBER12C1.INT 2 |
| *X. axonopodis begoniae* | LMG 7188 | BXE--254.INT 9 | ECE--246.INT 28 | RPE--249.INT 28 | XBER11C2.INT 28 |
| *X. axonopodis begoniae* | LMG 7194 | BXE--139.INT 14 | ECE--175.INT 14 | RPE--161.INT 14 | XBER5C1.INT 14 |
| *X. axonopodis begoniae* | LMG 7194R | BXE--305.INT 4 | ECE--380.INT 19 | RPE--342.INT 4 | XBER12C2.INT 13 |
| *X. axonopodis begoniae* | LMG 7196 | BXE--139.INT 24 | ECE--175.INT 4 | RPE--161.INT 4 | XBER4C2.INT 28 |
| *X. axonopodis begoniae* | LMG 7196RA | BXE--257.INT 5 | ECE--300.INT 23 | RPE--265.INT 5 | XBER12C2.INT 5 |
| *X. axonopodis begoniae* | LMG 7226 | BXE--192.INT 29 | ECE--194.INT 29 | RPE--204.INT 29 | XBER11C1.INT 29 |
| *X. axonopodis begoniae* | LMG 7303 | BXE--237.INT 19 | ECE--229.INT 19 | RPE--233.INT 19 | XBER5C2.INT 19 |
| *X. axonopodis begoniae* | LMG 7304 | BXE--192.INT 16 | ECE--194.INT 16 | RPE--204.INT 16 | XBER11C1.INT 16 |
| *X. axonopodis begoniae* | LMG 7595 | BXE--237.INT 8 | ECE--229.INT 8 | RPE--233.INT 8 | XBER5C2.INT 8 |
| *X. axonopodis begoniae* | LMG 7601 | BXE--192.INT 10 | ECE--194.INT 10 | RPE--204.INT 10 | XBER11C1.INT 10 |
| *X. axonopodis begoniae* | LMG 7601R | BXE--258.INT 17 | ECE--262.INT 17 | RPE--266.INT 18 | XBER13C2.INT 17 |
| *X. axonopodis cajani* | LMG 558 | BXE--139.INT 6 | ECE--175.INT 22 | RPE--161.INT 22 | XBER5C1.INT 6 |
| *X. axonopodis cajani* | LMG 7387t1 | BXE--225.INT 20 | ECE--255.INT 20 | RPE--227.INT 20 | XBER4C2.INT 20 |
| *X. axonopodis cajani* | LMG 7473 | BXE--139.INT 12 | ECE--175.INT 16 | RPE--161.INT 16 | XBER5C1.INT 12 |
| *X. axonopodis citri A* | LMG 681 | BXE--305.INT 7 | ECE--387.INT 2 | RPE--342.INT 11 | XBER12C2.INT 28 |
| *X. axonopodis citri A* | LMG 681R | BXE--199.INT 10 | ECE--203.INT 10 | RPE--201.INT 10 | XBER14C1.INT 10 |
| *X. axonopodis citri A* | LMG 8650 | BXE--237.INT 12 | ECE--229.INT 12 | RPE--233.INT 12 | XBER5C2.INT 12 |
| *X. axonopodis citri A* | LMG 8654 | BXE--198.INT 9 | ECE--202.INT 9 | RPE--200.INT 9 | XBER13C1.INT 9 |
| *X. axonopodis citri A* | LMG 8657 | BXE--237.INT 9 | ECE--229.INT 9 | RPE--233.INT 9 | XBER5C2.INT 9 |
| *X. axonopodis citri A* | LMG 9176 | BXE--159.INT 9 | ECE--176.INT 8 | RPE--234.INT 3 | XBER6C1.INT 9 |
| *X. axonopodis citri A* | LMG 9321 | BXE--305.INT 12 | ECE--360.INT 21 | RPE--307.INT 14 | XBER13C2.INT 14 |
| *X. axonopodis citri A* | LMG 9665 | BXE--198.INT 20 | ECE--202.INT 20 | RPE--200.INT 20 | XBER13C1.INT 20 |
| *X. axonopodis citri A* | LMG 9669 | BXE--258.INT 24 | ECE--360.INT 25 | RPE--266.INT 25 | XBER13C2.INT 24 |

| | | | | | |
|---|---|---|---|---|---|
| *X. axonopodis citri A* | LMG 9671 | BXE--198.INT 26 | ECE--202.INT 26 | RPE--200.INT 26 | XBER13C1.INT 26 |
| *X. axonopodis citri C* | LMG 8655 | BXE--237.INT 29 | ECE--229.INT 29 | RPE--233.INT 29 | XBER8C1.INT 7 |
| *X. axonopodis citri C* | LMG 8656 | BXE--198.INT 29 | ECE--202.INT 29 | RPE--200.INT 29 | XBER13C1.INT 29 |
| *X. axonopodis citri C* | LMG 9181 | BXE--238.INT 15 | ECE--230.INT 15 | RPE--234.INT 15 | XBER8C2.INT 15 |
| *X. axonopodis citri C* | LMG 9654 | BXE--198.INT 10 | ECE--202.INT 10 | RPE--200.INT 10 | XBER13C1.INT 10 |
| *X. axonopodis citri C* | LMG 9658 | BXE--258.INT 26 | ECE--360.INT 27 | RPE--266.INT 27 | XBER13C2.INT 26 |
| *X. axonopodis citri D* | LMG 9182 | BXE--159.INT 13 | ECE--176.INT 12 | RPE--142.INT 13 | XBER8C1.INT 13 |
| *X. axonopodis citri D* | LMG 9182R | BXE--305.INT 9 | ECE--262.INT 6 | RPE--266.INT 7 | XBER13C2.INT 6 |
| *X. axonopodis citri E* | LMG 9160col | BXE--160.INT 3 | ECE--176.INT 27 | RPE--162.INT 5 | XBER7C1.INT 4 |
| *X. axonopodis citri E* | LMG 9160dro | BXE--238.INT 23 | ECE--230.INT 23 | RPE--234.INT 23 | XBER8C2.INT 22 |
| *X. axonopodis citri E* | LMG 9162 | BXE--159.INT 14 | ECE--176.INT 13 | RPE--142.INT 14 | XBER8C2.INT 7 |
| *X. axonopodis citri E* | LMG 9172 | BXE--257.INT 29 | ECE--378.INT 4 | RPE--307.INT 12 | XBER12C1.INT 27 |
| *X. axonopodis citri E* | LMG 9174 | BXE--193.INT 25 | ECE--209.INT 8 | RPE--205.INT 25 | XBER12C2.INT 27 |
| *X. axonopodis citri E* | LMG 9175 | BXE--259.INT 22 | ECE--361.INT 18 | RPE--267.INT 23 | XBER14C2.INT 22 |
| *X. axonopodis citri E* | LMG 9175R | BXE--207.INT 3 | ECE--206.INT 3 | RPE--208.INT 3 | XBER15C1.INT 3 |
| *X. axonopodis citri E* | LMG 9325 | BXE--258.INT 3 | ECE--360.INT 14 | RPE--266.INT 4 | XBER13C1.INT 2 |
| *X. axonopodis dieffenbachiae* | LMG 695 | BXE--159.INT 25 | ECE--176.INT 24 | RPE--142.INT 25 | XBER8C1.INT 25 |
| *X. axonopodis dieffenbachiae* | LMG 7484 | BXE--305.INT 23 | ECE--264.INT 9 | RPE--268.INT 9 | XBER15C2.INT 9 |
| *X. axonopodis dieffenbachiae* | LMG 8664 | BXE--159.INT 18 | ECE--176.INT 17 | RPE--142.INT 18 | XBER8C1.INT 18 |
| *X. axonopodis dieffenbachiae** | LMG 7399 | BXE--238.INT 13 | ECE--230.INT 13 | RPE--234.INT 13 | XBER8C2.INT 13 |
| *X. axonopodis dieffenbachiae** | LMG 7400 | BXE--159.INT 24 | ECE--176.INT 23 | RPE--142.INT 24 | XBER8C1.INT 24 |
| *X. axonopodis glycines* | LMG 712 | BXE--237.INT 27 | ECE--229.INT 27 | RPE--233.INT 27 | XBER8C1.INT 5 |
| *X. axonopodis glycines* | LMG 8026 | BXE--159.INT 15 | ECE--176.INT 14 | RPE--142.INT 15 | XBER8C1.INT 15 |
| *X. axonopodis glycines* | LMG 8126 | BXE--260.INT 7 | ECE--264.INT 7 | RPE--268.INT 7 | XBER15C2.INT 7 |
| *X. axonopodis glycines* | LMG 8128 | BXE--207.INT 25 | ECE--206.INT 25 | RPE--208.INT 25 | XBER15C1.INT 25 |
| *X. axonopodis malvacearum* | LMG 7427 | BXE--239.INT 7 | ECE--231.INT 7 | RPE--235.INT 7 | XBER7C2.INT 7 |
| *X. axonopodis malvacearum* | LMG 7429 | BXE--210.INT 19 | ECE--212.INT 19 | RPE--217.INT 18 | XBER16C1.INT 19 |
| *X. axonopodis malvacearum* | LMG 7430 | BXE--270.INT 9 | ECE--273.INT 9 | RPE--275.INT 9 | XBER16C3.INT 9 |
| *X. axonopodis malvacearum* | LMG 761 | BXE--181.INT 10 | ECE--178.INT 6 | RPE--173.INT 11 | XBER7C2.INT 25 |
| *X. axonopodis malvacearum* | LMG 762 | BXE--239.INT 18 | ECE--231.INT 18 | RPE--235.INT 18 | XBER8C1.INT 4 |
| *X. axonopodis malvacearum* | LMG 764 | BXE--181.INT 13 | ECE--231.INT 27 | RPE--173.INT 13 | XBER7C2.INT 27 |
| *X. axonopodis manihotis* | LMG 768 | BXE--239.INT 23 | ECE--231.INT 23 | RPE--235.INT 23 | XBER8C1.INT 8 |
| *X. axonopodis manihotis* | LMG 769 | BXE--182.INT 17 | ECE--179.INT 15 | RPE--174.INT 20 | XBER9C1.INT 17 |
| *X. axonopodis manihotis* | LMG 771 | BXE--269.INT 6 | ECE--269.INT 6 | RPE--306.INT 8 | XBER16C2.INT 6 |
| *X. axonopodis manihotis* | LMG 773 | BXE--181.INT 2 | ECE--177.INT 24 | RPE--173.INT 2 | XBER7C2.INT 16 |
| *X. axonopodis manihotis* | LMG 778 | BXE--305.INT 28 | ECE--387.INT 6 | RPE--342.INT 16 | XBER16C2.INT 7 |
| *X. axonopodis manihotis* | LMG 780t1 | BXE--210.INT 27 | ECE--212.INT 27 | RPE--217.INT 26 | XBER16C1.INT 27 |
| *X. axonopodis manihotis* | LMG 780t2 | BXE--259.INT 26 | ECE--263.INT 26 | RPE--267.INT 27 | XBER15C1.INT 6 |
| *X. axonopodis manihotis* | LMG 782 | BXE--210.INT 17 | ECE--212.INT 17 | RPE--217.INT 16 | XBER16C1.INT 17 |
| *X. axonopodis manihotis* | LMG 784 | BXE--270.INT 8 | ECE--273.INT 8 | RPE--275.INT 8 | XBER16C3.INT 8 |
| *X. axonopodis patelli* | LMG 811t2 | BXE--210.INT 13 | ECE--212.INT 13 | RPE--306.INT 6 | XBER16C1.INT 13 |
| *X. axonopodis patelli* | LMG 811t2R | BXE--270.INT 10 | ECE--273.INT 10 | RPE--275.INT 10 | XBER16C3.INT 10 |
| *X. axonopodis patelli* | LMG 811t3 | BXE--210.INT 23 | ECE--212.INT 23 | RPE--217.INT 22 | XBER16C1.INT 23 |
| *X. axonopodis phaseoli* | LMG 7455 | BXE--012.INT 28 | ECE--031.INT 7 | RPE--017.INT 28 | XBER1C2.INT 28 |
| *X. axonopodis phaseoli* | LMG 8014 | BXE--207.INT 4 | ECE--206.INT 4 | RPE--208.INT 4 | XBER14C2.INT 24 |
| *X. axonopodis phaseoli* | LMG 821 | BXE--239.INT 19 | ECE--231.INT 19 | RPE--235.INT 19 | XBER8C1.INT 5 |
| *X. axonopodis phaseoli* | LMG 821t1R | BXE--193.INT 24 | ECE--209.INT 7 | RPE--205.INT 24 | XBER12C1.INT 24 |
| *X. axonopodis phaseoli* | LMG 821t2R | BXE--340.INT 4 | ECE--360.INT 11 | RPE--342.INT 9 | XBER12C2.INT 25 |
| *X. axonopodis phaseoli* | LMG 823 | BXE--160.INT 20 | ECE--177.INT 17 | RPE--162.INT 22 | XBER7C1.INT 20 |
| *X. axonopodis phaseoli* | LMG 842 | BXE--259.INT 25 | ECE--263.INT 25 | RPE--267.INT 26 | XBER15C1.INT 5 |
| *X. axonopodis phaseoli(fusc)* | LMG 7456 | BXE--207.INT 12 | ECE--206.INT 12 | RPE--208.INT 12 | XBER15C1.INT 12 |
| *X. axonopodis phaseoli(fusc)* | LMG 7511 | BXE--259.INT 14 | ECE--361.INT 15 | RPE--267.INT 15 | XBER14C2.INT 14 |
| *X. axonopodis phaseoli(fusc)* | LMG 8036 | BXE--199.INT 12 | ECE--203.INT 12 | RPE--201.INT 12 | XBER14C1.INT 12 |
| *X. axonopodis phaseoli(fusc)* | LMG 8038 | BXE--259.INT 9 | ECE--361.INT 11 | RPE--267.INT 10 | XBER14C2.INT 9 |
| *X. axonopodis phaseoli(fusc)* | LMG 825t1 | BXE--160.INT 11 | ECE--177.INT 7 | RPE--162.INT 13 | XBER8C2.INT 29 |
| *X. axonopodis phaseoli(fusc)* | LMG 837t1 | BXE--239.INT 3 | ECE--231.INT 3 | RPE--235.INT 3 | XBER7C2.INT 3 |
| *X. axonopodis phaseoli(fusc)* | LMG 837t1R | BXE--198.INT 22 | ECE--202.INT 22 | RPE--200.INT 22 | XBER13C1.INT 22 |
| *X. axonopodis phaseoli(fusc)* | LMG 837t2 | BXE--239.INT 4 | ECE--231.INT 4 | RPE--235.INT 4 | XBER7C2.INT 4 |
| *X. axonopodis phaseoli(fusc)* | LMG 841 | BXE--160.INT 12 | ECE--177.INT 9 | RPE--162.INT 14 | XBER7C1.INT 12 |
| *X. axonopodis ricini* | LMG 7441 | BXE--259.INT 16 | ECE--263.INT 16 | RPE--267.INT 17 | XBER14C2.INT 16 |
| *X. axonopodis ricini* | LMG 7442 | BXE--199.INT 9 | ECE--203.INT 9 | RPE--201.INT 9 | XBER14C1.INT 9 |

| | | | | | |
|---|---|---|---|---|---|
| *X. axonopodis ricini* | LMG 7443 | BXE--240.INT 15 | ECE--232.INT 15 | RPE--236.INT 16 | XBER8C2.INT 15 |
| *X. axonopodis ricini* | LMG 7444 | BXE--207.INT 13 | ECE--206.INT 13 | RPE--208.INT 13 | XBER15C1.INT 13 |
| *X. axonopodis ricini* | LMG 861 | BXE--240.INT 14 | ECE--232.INT 14 | RPE--236.INT 15 | XBER8C2.INT 14 |
| *X. axonopodis ricini* | LMG 862 | BXE--199.INT 5 | ECE--203.INT 5 | RPE--201.INT 5 | XBER14C1.INT 5 |
| *X. axonopodis tamarindi* | LMG 869 | BXE--270.INT 19 | ECE--273.INT 19 | RPE--275.INT 19 | XBER17C1.INT 7 |
| *X. axonopodis tamarindi* | LMG 955 | BXE--218.INT 5 | ECE--219.INT 5 | RPE--220.INT 5 | XBER16C3.INT 17 |
| *X. axonopodis vasculorum* | LMG 894 | BXE--271.INT 14 | ECE--274.INT 14 | RPE--276.INT 14 | XBER18C1.INT 2 |
| *X. axonopodis vasculorum* | LMG 895 | BXE--218.INT 25 | ECE--219.INT 25 | RPE--220.INT 25 | XBER17C1.INT 25 |
| *X. axonopodis vasculorum* | LMG 899 | BXE--241.INT 5 | ECE--244.INT 5 | RPE--247.INT 5 | XBER9C2.INT 5 |
| *X. axonopodis vasculorum* | LMG 901 | BXE--181.INT 21 | ECE--178.INT 16 | RPE--173.INT 21 | XBER8C1.INT 21 |
| *X. axonopodis vasculorum* | LMG 901R | BXE--257.INT 24 | ECE--360.INT 9 | RPE--307.INT 7 | XBER12C2.INT 24 |
| *X. axonopodis vasculorum* | LMG 903 | BXE--222.INT 12 | ECE--223.INT 11 | RPE--224.INT 12 | XBER17C2.INT 24 |
| *X. axonopodis vesicatoria* | LMG 667 | BXE--241.INT 12 | ECE--244.INT 12 | RPE--247.INT 12 | XBER9C2.INT 12 |
| *X. axonopodis vesicatoria* | LMG 668 | BXE--182.INT 22 | ECE--179.INT 20 | RPE--174.INT 24 | XBER9C1.INT 22 |
| *X. axonopodis vesicatoria* | LMG 905 | BXE--012.INT 29 | ECE--014.INT 29 | RPE--017.INT 29 | XBER1C2.INT 29 |
| *X. axonopodis vesicatoria* | LMG 906 | BXE--218.INT 26 | ECE--219.INT 26 | RPE--220.INT 26 | XBER17C1.INT 26 |
| *X. axonopodis vesicatoria* | LMG 909 | BXE--271.INT 7 | ECE--274.INT 7 | RPE--276.INT 7 | XBER17C2.INT 7 |
| *X. axonopodis vesicatoria* | LMG 910 | BXE--004.INT 2 | ECE--007.INT 2 | RPE--010.INT 2 | XBER2C1.INT 2 |
| *X. axonopodis vesicatoria* | LMG 913 | BXE--306.INT 2 | ECE--378.INT 19 | RPE--306.INT 12 | XBER17C1.INT 10 |
| *X. axonopodis vesicatoria* | LMG 914t2 | BXE--218.INT 16 | ECE--219.INT 16 | RPE--220.INT 16 | XBER16C3.INT 29 |
| *X. axonopodis vesicatoria* | LMG 922 | BXE--271.INT 27 | ECE--274.INT 27 | RPE--276.INT 27 | XBER18C1.INT 15 |
| *X. axonopodis vesicatoria* | LMG 929 | BXE--004.INT 3 | ECE--007.INT 3 | RPE--010.INT 3 | XBER2C1.INT 3 |
| *X. axonopodis vesicatoria* | LMG 932 | BXE--271.INT 6 | ECE--274.INT 6 | RPE--276.INT 6 | XBER17C2.INT 6 |
| *X. axonopodis vitians* | LMG 8689 | BXE--222.INT 13 | ECE--223.INT 12 | RPE--224.INT 13 | XBER17C2.INT 25 |
| *X. axonopodis vitians* | LMG 937 | BXE--241.INT 6 | ECE--244.INT 6 | RPE--247.INT 6 | XBER9C2.INT 6 |
| *X. bromi* | LMG 8267 | BXE--136.INT 19 | ECE--226.INT 9 | RPE--135.INT 19 | XBER4C1.INT 19 |
| *X. bromi* | LMG 8269 | BXE--225.INT 10 | ECE--255.INT 10 | RPE--227.INT 10 | XBER4C2.INT 10 |
| *X. bromi* | LMG 8269R | BXE--210.INT 10 | ECE--212.INT 10 | RPE--217.INT 10 | XBER16C1.INT 10 |
| *X. bromi* | LMG 8272 | BXE--225.INT 12 | ECE--255.INT 12 | RPE--227.INT 12 | XBER4C2.INT 12 |
| *X. bromi* | LMG 8272R | BXE--136.INT 2 | ECE--226.INT 26 | RPE--135.INT 2 | XBER4C1.INT 2 |
| *X. bromi* | LMG 8272RR | BXE--269.INT 3 | ECE--269.INT 3 | RPE--269.INT 3 | XBER16C2.INT 3 |
| *X. campestris barbareae* | LMG 547 | BXE--001.INT 9 | ECE--007.INT 15 | RPE--010.INT 15 | XBER2C1.INT 15 |
| *X. campestris barbareae* | LMG 7365 | BXE--004.INT 16 | ECE--015.INT 16 | RPE--018.INT 16 | XBER2C2.INT 16 |
| *X. campestris campestris* | LMG 567 | BXE--192.INT 8 | ECE--194.INT 8 | RPE--204.INT 8 | XBER11C1.INT 8 |
| *X. campestris campestris* | LMG 568 | BXE--004.INT 17 | ECE--015.INT 17 | RPE--018.INT 17 | XBER2C2.INT 17 |
| *X. campestris campestris* | LMG 571 | BXE--139.INT 10 | ECE--175.INT 18 | RPE--161.INT 18 | XBER5C1.INT 10 |
| *X. campestris campestris* | LMG 571R | BXE--305.INT 5 | ECE--360.INT 8 | RPE--342.INT 8 | XBER12C2.INT 23 |
| *X. campestris campestris* | LMG 573 | BXE--192.INT 18 | ECE--194.INT 18 | RPE--204.INT 18 | XBER11C1.INT 18 |
| *X. campestris campestris* | LMG 573R | BXE--271.INT 18 | ECE--274.INT 18 | RPE--276.INT 18 | XBER18C1.INT 6 |
| *X. campestris campestris* | LMG 583 | BXE--139.INT 21 | ECE--175.INT 7 | RPE--161.INT 7 | XBER5C1.INT 21 |
| *X. campestris campestris* | LMG 7516 | BXE--237.INT 7 | ECE--229.INT 7 | RPE--233.INT 7 | XBER5C2.INT 7 |
| *X. campestris campestris* | LMG 7516RA | BXE--193.INT 5 | ECE--195.INT 5 | RPE--205.INT 5 | XBER12C1.INT 5 |
| *X. campestris campestris* | LMG 7662 | BXE--243.INT 25 | ECE--246.INT 25 | RPE--249.INT 25 | XBER11C2.INT 25 |
| *X. campestris campestris* | LMG 8001 | BXE--192.INT 12 | ECE--194.INT 12 | RPE--204.INT 12 | XBER11C1.INT 12 |
| *X. campestris campestris* | LMG 8003 | BXE--243.INT 6 | ECE--246.INT 6 | RPE--249.INT 6 | XBER11C2.INT 6 |
| *X. campestris campestris* | LMG 8032 | BXE--192.INT 19 | ECE--194.INT 19 | RPE--204.INT 19 | XBER11C1.INT 19 |
| *X. campestris campestris* | LMG 8035 | BXE--254.INT 7 | ECE--246.INT 23 | RPE--220.INT 27 | XBER11C2.INT 23 |
| *X. campestris campestris* | LMG 8035R | BXE--218.INT 27 | ECE--219.INT 27 | RPE--249.INT 23 | XBER17C1.INT 27 |
| *X. campestris campestris* | LMG 8051 | BXE--243.INT 15 | ECE--246.INT 15 | RPE--249.INT 15 | XBER11C2.INT 15 |
| *X. campestris campestris* | LMG 8055 | BXE--192.INT 25 | ECE--194.INT 25 | RPE--204.INT 25 | XBER11C1.INT 25 |
| *X. campestris campestris* | LMG 8082 | BXE--243.INT 3 | ECE--246.INT 3 | RPE--249.INT 3 | XBER11C2.INT 3 |
| *X. campestris campestris* | LMG 8099 | BXE--192.INT 9 | ECE--194.INT 9 | RPE--204.INT 9 | XBER11C1.INT 9 |
| *X. campestris campestris* | LMG 8100 | BXE--243.INT 7 | ECE--246.INT 7 | RPE--249.INT 7 | XBER11C2.INT 7 |
| *X. campestris campestris* | LMG 8112 | BXE--192.INT 7 | ECE--194.INT 7 | RPE--204.INT 7 | XBER11C1.INT 7 |
| *X. campestris campestris* | LMG 8119 | BXE--243.INT 16 | ECE--246.INT 16 | RPE--249.INT 16 | XBER11C2.INT 16 |
| *X. campestris campestris* | LMG 8119R | BXE--222.INT 4 | ECE--223.INT 4 | RPE--224.INT 4 | XBER17C2.INT 16 |
| *X. campestris campestris* | LMG 8121 | BXE--243.INT 5 | ECE--246.INT 5 | RPE--249.INT 5 | XBER11C2.INT 5 |
| *X. campestris campestris* | LMG 8123 | BXE--192.INT 20 | ECE--194.INT 20 | RPE--204.INT 20 | XBER11C1.INT 20 |
| *X. campestris campestris** | LMG 575 | BXE--225.INT 19 | ECE--255.INT 19 | RPE--227.INT 19 | XBER4C2.INT 19 |
| *X. campestris campestris** | LMG 575RA | BXE--193.INT 8 | ECE--195.INT 8 | RPE--205.INT 8 | XBER12C1.INT 8 |
| *X. campestris campestris** | LMG 7514 | BXE--237.INT 3 | ECE--229.INT 3 | RPE--233.INT 3 | XBER5C2.INT 3 |

| | | | | | |
|---|---|---|---|---|---|
| *X. campestris campestris** | LMG 7514R | BXE--193.INT 19 | ECE--221.INT 12 | RPE--205.INT 19 | XBER12C1.INT 19 |
| *X. campestris incanae* | LMG 7421 | BXE--039.INT 13 | ECE--031.INT 22 | RPE--018.INT 18 | XBER2C2.INT 18 |
| *X. campestris incanae* | LMG 7490 | BXE--001.INT 13 | ECE--007.INT 19 | RPE--010.INT 19 | XBER2C1.INT 19 |
| *X. campestris raphani* | LMG 7505 | BXE--020.INT 2 | ECE--015.INT 21 | RPE--018.INT 21 | XBER2C2.INT 21 |
| *X. campestris raphani* | LMG 8134 | BXE--004.INT 22 | ECE--007.INT 22 | RPE--010.INT 22 | XBER2C1.INT 22 |
| *X. campestris raphani** | LMG 8601t1 | BXE--305.INT 18 | ECE--361.INT 17 | RPE--307.INT 22 | XBER14C2.INT 20 |
| *X. campestris raphani** | LMG 8601t2 | BXE--199.INT 4 | ECE--203.INT 4 | RPE--201.INT 4 | XBER14C1.INT 4 |
| *X. cassavae* | LMG 5246 | BXE--258.INT 9 | ECE--262.INT 9 | RPE--266.INT 10 | XBER13C2.INT 9 |
| *X. cassavae* | LMG 5264 | BXE--003.INT 22 | ECE--006.INT 22 | RPE--009.INT 22 | XBER1C1.INT 22 |
| *X. cassavae* | LMG 5270 | BXE--225.INT 25 | ECE--255.INT 25 | RPE--227.INT 25 | XBER4C2.INT 25 |
| *X. cassavae* | LMG 5270R | BXE--193.INT 14 | ECE--195.INT 14 | RPE--205.INT 14 | XBER12C1.INT 14 |
| *X. cassavae* | LMG 670 | BXE--012.INT 19 | ECE--014.INT 19 | RPE--017.INT 19 | XBER1C2.INT 19 |
| *X. cassavae* | LMG 670 | BXE--199.INT 18 | ECE--203.INT 18 | RPE--201.INT 18 | XBER14C1.INT 18 |
| *X. cassavae* | LMG 673 | BXE--012.INT 21 | ECE--014.INT 21 | RPE--047.INT 15 | XBER1C2.INT 21 |
| *X. codiaei* | LMG 8677 | BXE--181.INT 17 | ECE--178.INT 13 | RPE--173.INT 17 | XBER8C1.INT 17 |
| *X. codiaei* | LMG 8678 | BXE--012.INT 23 | ECE--014.INT 23 | RPE--017.INT 23 | XBER1C2.INT 23 |
| *X. cucurbitae* | LMG 690 | BXE--159.INT 16 | ECE--176.INT 15 | RPE--142.INT 16 | XBER6C1.INT 16 |
| *X. cucurbitae* | LMG 690R | BXE--257.INT 18 | ECE--261.INT 17 | RPE--265.INT 19 | XBER12C2.INT 18 |
| *X. cucurbitae* | LMG 7480 | BXE--159.INT 11 | ECE--176.INT 9 | RPE--142.INT 11 | XBER6C1.INT 11 |
| *X. cucurbitae* | LMG 7481 | BXE--258.INT 8 | ECE--360.INT 17 | RPE--307.INT 13 | XBER13C2.INT 8 |
| *X. cucurbitae* | LMG 8661 | BXE--199.INT 3 | ECE--203.INT 3 | RPE--201.INT 3 | XBER14C1.INT 3 |
| *X. cucurbitae* | LMG 8662 | BXE--012.INT 25 | ECE--014.INT 25 | RPE--017.INT 25 | XBER1C2.INT 25 |
| *X. fragariae* | LMG 708 | BXE--003.INT 2 | ECE--014.INT 2 | RPE--009.INT 2 | XBER1C1.INT 2 |
| *X. fragariae* | LMG 708R | BXE--254.INT 4 | ECE--256.INT 16 | RPE--254.INT 8 | XBER9C2.INT 20 |
| *X. fragariae* | LMG 708 | BXE--003.INT 3 | ECE--031.INT 4 | RPE--009.INT 3 | XBER1C1.INT 3 |
| *X. fragariae* | LMG 710 | BXE--225.INT 7 | ECE--255.INT 7 | RPE--227.INT 7 | XBER4C2.INT 7 |
| *X. hortorum hederae* | LMG 733T | BXE--136.INT 27 | ECE--226.INT 2 | RPE--135.INT 27 | XBER4C1.INT 27 |
| *X. hortorum hederae* | LMG 733TR | BXE--238.INT 14 | ECE--230.INT 14 | RPE--234.INT 14 | XBER6C2.INT 14 |
| *X. hortorum hederae* | LMG 734 | BXE--159.INT 23 | ECE--176.INT 22 | RPE--142.INT 23 | XBER6C1.INT 23 |
| *X. hortorum hederae* | LMG 7413 | BXE--238.INT 27 | ECE--256.INT 10 | RPE--234.INT 27 | XBER7C1.INT 8 |
| *X. hortorum hederae* | LMG 7414 | BXE--159.INT 12 | ECE--176.INT 11 | RPE--142.INT 12 | XBER6C1.INT 12 |
| *X. hortorum hederae* | LMG 8665 | BXE--260.INT 8 | ECE--264.INT 8 | RPE--268.INT 8 | XBER15C2.INT 8 |
| *X. hortorum pelargonii* | LMG 7312 | BXE--181.INT 9 | ECE--178.INT 5 | RPE--173.INT 9 | XBER7C2.INT 24 |
| *X. hortorum pelargonii* | LMG 7314 | BXE--012.INT 4 | ECE--014.INT 4 | RPE--017.INT 4 | XBER1C2.INT 4 |
| *X. hortorum pelargonii* | LMG 7314R | BXE--136.INT 24 | ECE--226.INT 5 | RPE--135.INT 24 | XBER4C1.INT 24 |
| *X. hortorum pelargonii* | LMG 7314RR | BXE--239.INT 6 | ECE--231.INT 6 | RPE--235.INT 6 | XBER7C2.INT 6 |
| *X. hortorum pelargonii* | LMG 7315 | BXE--160.INT 25 | ECE--177.INT 23 | RPE--162.INT 27 | XBER7C1.INT 25 |
| *X. hortorum pelargonii* | LMG 7316 | BXE--239.INT 12 | ECE--231.INT 12 | RPE--235.INT 12 | XBER7C2.INT 12 |
| *X. hortorum pelargonii* | LMG 7317 | BXE--181.INT 14 | ECE--178.INT 9 | RPE--173.INT 14 | XBER7C2.INT 28 |
| *X. hortorum pelargonii* | LMG 7354 | BXE--270.INT 3 | ECE--273.INT 3 | RPE--275.INT 3 | XBER16C3.INT 3 |
| *X. hortorum pelargonii* | LMG 7356 | BXE--181.INT 12 | ECE--178.INT 7 | RPE--173.INT 12 | XBER7C2.INT 26 |
| *X. hortorum pelargonii* | LMG 7585 | BXE--260.INT 12 | ECE--264.INT 12 | RPE--268.INT 12 | XBER15C2.INT 12 |
| *X. hortorum pelargonii* | LMG 7690 | BXE--222.INT 22 | ECE--223.INT 21 | RPE--224.INT 22 | XBER18C1.INT 22 |
| *X. hortorum pelargonii* | LMG 7690R | BXE--260.INT 23 | ECE--264.INT 23 | RPE--268.INT 24 | XBER16C1.INT 3 |
| *X. hortorum pelargonii* | LMG 7706 | BXE--199.INT 14 | ECE--203.INT 14 | RPE--201.INT 14 | XBER14C1.INT 14 |
| *X. hortorum pelargonii* | LMG 7708 | BXE--305.INT 19 | ECE--263.INT 29 | RPE--307.INT 23 | XBER15C1.INT 9 |
| *X. hortorum pelargonii* | LMG 7710 | BXE--199.INT 17 | ECE--203.INT 17 | RPE--201.INT 17 | XBER14C1.INT 17 |
| *X. hortorum pelargonii* | LMG 7710R | BXE--260.INT 14 | ECE--264.INT 14 | RPE--268.INT 14 | XBER15C2.INT 14 |
| *X. hortorum pelargonii* | LMG 7712 | BXE--207.INT 29 | ECE--206.INT 29 | RPE--208.INT 29 | XBER15C1.INT 29 |
| *X. hortorum pelargonii* | LMG 7712R | BXE--305.INT 27 | ECE--378.INT 9 | RPE--268.INT 23 | XBER16C1.INT 2 |
| *X. hortorum pelargonii* | LMG 7715 | BXE--207.INT 15 | ECE--206.INT 15 | RPE--208.INT 15 | XBER15C1.INT 15 |
| *X. hortorum pelargonii* | LMG 7763 | BXE--239.INT 8 | ECE--231.INT 8 | RPE--235.INT 8 | XBER7C2.INT 8 |
| *X. hortorum pelargonii* | LMG 7764 | BXE--160.INT 15 | ECE--177.INT 13 | RPE--162.INT 17 | XBER7C1.INT 15 |
| *X. hortorum pelargonii* | LMG 820 | BXE--239.INT 22 | ECE--231.INT 22 | RPE--235.INT 22 | XBER8C1.INT 7 |
| *X. hortorum pelargonii* | LMG 820R | BXE--182.INT 25 | ECE--179.INT 23 | RPE--174.INT 27 | XBER9C1.INT 25 |
| *X. hortorum vitians* | LMG 7508 | BXE--240.INT 7 | ECE--232.INT 7 | RPE--236.INT 7 | XBER8C2.INT 7 |
| *X. hortorum vitians* | LMG 7510 | BXE--182.INT 8 | ECE--179.INT 6 | RPE--174.INT 11 | XBER8C2.INT 23 |
| *X. hortorum vitians* | LMG 8688 | BXE--270.INT 25 | ECE--273.INT 25 | RPE--275.INT 25 | XBER17C1.INT 12 |
| *X. hortorum vitians* | LMG 8690t1 | BXE--218.INT 21 | ECE--219.INT 21 | RPE--220.INT 21 | XBER17C1.INT 21 |
| *X. hortorum vitians* | LMG 8690t2 | BXE--270.INT 22 | ECE--273.INT 22 | RPE--275.INT 21 | XBER17C1.INT 9 |
| *X. hortorum vitians* | LMG 938 | BXE--241.INT 18 | ECE--244.INT 18 | RPE--247.INT 18 | XBER9-1.INT 2 |

| | | | | | |
|---|---|---|---|---|---|
| *X. hyacinthi* | LMG 739* | BXE--225.INT 6 | ECE--255.INT 6 | RPE--227.INT 6 | XBER4C2.INT 6 |
| *X. hyacinthi* | LMG 739R | BXE--241.INT 19 | ECE--244.INT 19 | RPE--247.INT 19 | XBER9-1.INT 3 |
| *X. hyacinthi* | LMG 740 | BXE--239.INT 14 | ECE--231.INT 14 | RPE--235.INT 14 | XBER7C2.INT 14 |
| *X. hyacinthi* | LMG 7419 | BXE--210.INT 6 | ECE--212.INT 6 | RPE--217.INT 6 | XBER15C2.INT 26 |
| *X. hyacinthi* | LMG 742 | BXE--005.INT 6 | ECE--016.INT 6 | RPE--048.INT 6 | XBER3C2.INT 6 |
| *X. hyacinthi* | LMG 8041 | BXE--002.INT 7 | ECE--016.INT 7 | RPE--047.INT 9 | XBER3C1.INT 7 |
| *X. hyacinthi* | LMG 8042 | BXE--270.INT 5 | ECE--273.INT 5 | RPE--275.INT 5 | XBER16C3.INT 5 |
| *X. melonis* | LMG 8670 | BXE--207.INT 19 | ECE--206.INT 19 | RPE--208.INT 19 | XBER15C1.INT 19 |
| *X. melonis* | LMG 8671 | BXE--239.INT 20 | ECE--231.INT 20 | RPE--235.INT 20 | XBER8C1.INT 6 |
| *X. melonis* | LMG 8672 | BXE--004.INT 8 | ECE--007.INT 10 | RPE--010.INT 8 | XBER2C1.INT 8 |
| *X. melonis* | LMG 8674 | BXE--240.INT 2 | ECE--232.INT 2 | RPE--236.INT 2 | XBER8C2.INT 2 |
| *X. oryzae oryzae* | LMG 5047T | BXE--004.INT 4 | ECE--055.INT 3 | RPE--010.INT 4 | XBER2C1.INT 4 |
| *X. oryzae oryzae* | LMG 5047TR | BXE--272.INT 2 | ECE--272.INT 2 | RPE--272.INT 2 | XBER18C2.INT 2 |
| *X. oryzae oryzae* | LMG 6411 | BXE--218.INT 11 | ECE--219.INT 11 | RPE--220.INT 11 | XBER16C3.INT 24 |
| *X. oryzae oryzae* | LMG 6518 | BXE--271.INT 3 | ECE--274.INT 3 | RPE--276.INT 3 | XBER17C2.INT 3 |
| *X. oryzae oryzae* | LMG 803 | BXE--181.INT 25 | ECE--178.INT 21 | RPE--173.INT 25 | XBER8C1.INT 25 |
| *X. oryzae oryzae* | LMG 803R | BXE--257.INT 7 | ECE--261.INT 7 | RPE--265.INT 7 | XBER12C2.INT 7 |
| *X. oryzae oryzae* | LMG 806 | BXE--182.INT 15 | ECE--179.INT 13 | RPE--174.INT 17 | XBER9C1.INT 15 |
| *X. oryzae oryzicola* | LMG 657 | BXE--241.INT 4 | ECE--244.INT 4 | RPE--247.INT 4 | XBER9C2.INT 4 |
| *X. oryzae oryzicola* | LMG 657R | BXE--193.INT 18 | ECE--221.INT 11 | RPE--307.INT 5 | XBER12C1.INT 18 |
| *X. oryzae oryzicola* | LMG 661 | BXE--240.INT 9 | ECE--232.INT 9 | RPE--236.INT 9 | XBER8C2.INT 9 |
| *X. oryzae oryzicola* | LMG 661R | BXE--193.INT 17 | ECE--195.INT 17 | RPE--307.INT 4 | XBER12C1.INT 17 |
| *X. oryzae oryzicola* | LMG 665 | BXE--013.INT 5 | ECE--015.INT 5 | RPE--018.INT 5 | XBER2C2.INT 5 |
| *X. oryzae oryzicola* | LMG 665R | BXE--182.INT 2 | ECE--178.INT 26 | RPE--174.INT 4 | XBER8C2.INT 16 |
| *X. oryzae oryzicola* | LMG 793 | BXE--240.INT 19 | ECE--232.INT 19 | RPE--236.INT 20 | XBER9C1.INT 5 |
| *X. oryzae oryzicola* | LMG 793R | BXE--193.INT 7 | ECE--195.INT 7 | RPE--307.INT 2 | XBER12C1.INT 7 |
| *X. oryzae oryzicola* | LMG 797 | BXE--271.INT 22 | ECE--274.INT 22 | RPE--276.INT 22 | XBER18C1.INT 9 |
| *X. pisi* | LMG 847t1* | BXE--004.INT 7 | ECE--007.INT 7 | RPE--010.INT 7 | XBER9C1.INT 9 |
| *X. pisi* | LMG 847t1R | BXE--270.INT 15 | ECE--273.INT 15 | RPE--275.INT 15 | XBER2C2.INT 7 |
| *X. pisi* | LMG 847t2 | BXE--182.INT 9 | ECE--179.INT 7 | RPE--174.INT 12 | XBER16C3.INT 15 |
| *X. populi* | LMG 5743 | BXE--012.INT 7 | ECE--014.INT 7 | RPE--017.INT 7 | XBER1C2.INT 7 |
| *X. populi* | LMG 5753 | BXE--003.INT 8 | ECE--006.INT 8 | RPE--009.INT 8 | XBER1C1.INT 8 |
| *X. populi* | LMG 974 | BXE--012.INT 6 | ECE--014.INT 6 | RPE--017.INT 6 | XBER1C2.INT 6 |
| *X. sacchari* | LMG 471 | BXE--002.INT 9 | ECE--016.INT 9 | RPE--047.INT 12 | XBER3C1.INT 9 |
| *X. sacchari* | LMG 476 | BXE--237.INT 5 | ECE--229.INT 5 | RPE--233.INT 5 | XBER5C2.INT 5 |
| *X. sacchari* | LMG 476R | BXE--193.INT 15 | ECE--261.INT 15 | RPE--205.INT 15 | XBER12C1.INT 15 |
| *X. theicola* | LMG 8684 | BXE--005.INT 8 | ECE--016.INT 8 | RPE--047.INT 11 | XBER3C2.INT 8 |
| *X. theicola* | LMG 8685 | BXE--136.INT 16 | ECE--226.INT 13 | RPE--135.INT 16 | XBER4C1.INT 16 |
| *X. theicola* | LMG 8686 | BXE--225.INT 14 | ECE--255.INT 14 | RPE--254.INT 2 | XBER4C2.INT 14 |
| *X. translucens arrhenatheri* | LMG 590 | BXE--193.INT 3 | ECE--195.INT 3 | RPE--205.INT 3 | XBER12C1.INT 3 |
| *X. translucens arrhenatheri* | LMG 591 | BXE--225.INT 15 | ECE--255.INT 15 | RPE--227.INT 15 | XBER4C2.INT 15 |
| *X. translucens arrhenatheri* | LMG 727R | BXE--241.INT 22 | ECE--244.INT 22 | RPE--247.INT 22 | XBER9-1.INT 5 |
| *X. translucens arrhenatheri* | LMG 727t1 | BXE--020.INT 4 | ECE--015.INT 23 | RPE--018.INT 23 | XBER2C2.INT 23 |
| *X. translucens arrhenatheri* | LMG 7384 | BXE--136.INT 12 | ECE--226.INT 16 | RPE--135.INT 12 | XBER4C1.INT 12 |
| *X. translucens cerealis* | LMG 7392 | BXE--225.INT 17 | ECE--255.INT 17 | RPE--227.INT 17 | XBER4C2.INT 17 |
| *X. translucens cerealis* | LMG 887 | BXE--139.INT 11 | ECE--175.INT 17 | RPE--161.INT 17 | XBER5C1.INT 11 |
| *X. translucens cerealis* | LMG 891 | BXE--237.INT 20 | ECE--229.INT 20 | RPE--233.INT 20 | XBER5C2.INT 20 |
| *X. translucens cerealis** | LMG 679 | BXE--004.INT 24 | ECE--007.INT 24 | RPE--010.INT 24 | XBER2C1.INT 24 |
| *X. translucens cerealis** | LMG 679R | BXE--254.INT 6 | ECE--256.INT 18 | RPE--254.INT 10 | XBER9C2.INT 23 |
| *X. translucens cerealis** | LMG 880 | BXE--136.INT 10 | ECE--226.INT 18 | RPE--135.INT 10 | XBER4C1.INT 10 |
| *X. translucens graminis* | LMG 595 | BXE--238.INT 24 | ECE--230.INT 24 | RPE--234.INT 24 | XBER7C1.INT 5 |
| *X. translucens graminis* | LMG 713 | BXE--210.INT 7 | ECE--378.INT 13 | RPE--217.INT 7 | XBER15C2.INT 27 |
| *X. translucens graminis* | LMG 726 | BXE--225.INT 3 | ECE--255.INT 3 | RPE--227.INT 3 | XBER4C2.INT 3 |
| *X. translucens graminis* | LMG 726R | BXE--159.INT 8 | ECE--230.INT 2 | RPE--142.INT 8 | XBER6C1.INT 8 |
| *X. translucens hordei* | LMG 737(8/7) | BXE--241.INT 10 | ECE--244.INT 10 | RPE--247.INT 10 | XBER9C2.INT 10 |
| *X. translucens hordei* | LMG 737(9/7) | BXE--182.INT 21 | ECE--179.INT 18 | RPE--174.INT 23 | XBER9C1.INT 21 |
| *X. translucens hordei* | LMG 8279 | BXE--039.INT 20 | ECE--015.INT 27 | RPE--033.INT 8 | XBER2C2.INT 27 |
| *X. translucens hordei* | LMG 8279t1 | BXE--222.INT 5 | ECE--223.INT 5 | RPE--224.INT 5 | XBER17C2.INT 17 |
| *X. translucens hordei* | LMG 8279t2 | BXE--260.INT 25 | ECE--378.INT 11 | RPE--306.INT 2 | XBER16C1.INT 5 |
| *X. translucens hordei* | LMG 882 | BXE--004.INT 26 | ECE--015.INT 26 | RPE--010.INT 26 | XBER2C1.INT 26 |
| *X. translucens hordei* | LMG 884 | BXE--260.INT 28 | ECE--362.INT 5 | RPE--268.INT 28 | XBER16C1.INT 8 |

| | | | | | |
|---|---|---|---|---|---|
| *X. translucens phlei* | LMG 716 | BXE–181.INT 19 | ECE–178.INT 15 | RPE–173.INT 20 | XBER8C1.INT 19 |
| *X. translucens phlei* | LMG 719 | BXE–240.INT 11 | ECE–256.INT 13 | RPE–236.INT 11 | XBER8C2.INT 11 |
| *X. translucens phlei* | LMG 723 | BXE–199.INT 7 | ECE–361.INT 9 | RPE–201.INT 7 | XBER14C1.INT 7 |
| *X. translucens phlei* | LMG 730 | BXE–020.INT 8 | ECE–055.INT 6 | RPE–033.INT 12 | XBER2C2.INT 28 |
| *X. translucens posae* | LMG 594 | BXE–207.INT 24 | ECE–264.INT 15 | RPE–208.INT 24 | XBER15C1.INT 24 |
| *X. translucens posae* | LMG 728 | BXE–002.INT 2 | ECE–016.INT 2 | RPE–047.INT 4 | XBER3C2.INT 2 |
| *X. translucens secalis* | LMG 7507 | BXE–218.INT 6 | ECE–219.INT 6 | RPE–220.INT 6 | XBER16C3.INT 18 |
| *X. translucens secalis* | LMG 883 | BXE–045.INT 11 | ECE–031.INT 26 | RPE–047.INT 5 | XBER3C2.INT 3 |
| *X. translucens translucens* | LMG 5259 | BXE–182.INT 4 | ECE–179.INT 2 | RPE–174.INT 6 | XBER8C2.INT 18 |
| *X. translucens translucens* | LMG 5260t1 | BXE–240.INT 20 | ECE–232.INT 20 | RPE–236.INT 21 | XBER9C1.INT 6 |
| *X. translucens translucens* | LMG 5260t2 | BXE–181.INT 23 | ECE–178.INT 18 | RPE–173.INT 23 | XBER8C1.INT 23 |
| *X. translucens translucens* | LMG 5262 | BXE–271.INT 20 | ECE–274.INT 20 | RPE–276.INT 20 | XBER18C1.INT 8 |
| *X. translucens translucens* | LMG 875 | BXE–222.INT 11 | ECE–274.INT 23 | RPE–224.INT 11 | XBER17C2.INT 23 |
| *X. translucens translucens* | LMG 876 | BXE–005.INT 4 | ECE–031.INT 28 | RPE–047.INT 6 | XBER3C2.INT 4 |
| *X. translucens undulosa* | LMG 8283 | BXE–222.INT 20 | ECE–223.INT 19 | RPE–224.INT 20 | XBER18C1.INT 20 |
| *X. translucens undulosa* | LMG 885 | BXE–240.INT 4 | ECE–232.INT 4 | RPE–236.INT 4 | XBER8C2.INT 4 |
| *X. translucens undulosa* | LMG 886 | BXE–218.INT 4 | ECE–219.INT 4 | RPE–220.INT 4 | XBER16C3.INT 16 |
| *X. translucens undulosa* | LMG 888 | BXE–270.INT 26 | ECE–378.INT 20 | RPE–306.INT 13 | XBER17C1.INT 13 |
| *X. translucens undulosa* | LMG 892 | BXE–002.INT 5 | ECE–006.INT 5 | RPE–048.INT 5 | XBER3C1.INT 5 |
| *X. vasicola holcicola* | LMG 736t1 | BXE–238.INT 28 | ECE–230.INT 28 | RPE–234.INT 28 | XBER7C1.INT 9 |
| *X. vasicola holcicola* | LMG 736t1RA | BXE–193.INT 13 | ECE–195.INT 13 | RPE–205.INT 13 | XBER12C1.INT 13 |
| *X. vasicola holcicola* | LMG 736t1RR | BXE–260.INT 29 | ECE–362.INT 6 | RPE–306.INT 5 | XBER16C1.INT 9 |
| *X. vasicola holcicola* | LMG 736t2 | BXE–159.INT 3 | ECE–176.INT 2 | RPE–142.INT 3 | XBER5C2.INT 25 |
| *X. vasicola holcicola* | LMG 736t2R | BXE–305.INT 16 | ECE–301.INT 28 | RPE–307.INT 20 | XBER14C2.INT 3 |
| *X. vasicola holcicola* | LMG 7416 | BXE–004.INT 6 | ECE–007.INT 6 | RPE–010.INT 6 | XBER2C1.INT 6 |
| *X. vasicola holcicola* | LMG 7489 | BXE–238.INT 20 | ECE–230.INT 20 | RPE–234.INT 20 | XBER7C1.INT 2 |
| *X. vasicola holcicola* | LMG 7489R | BXE–193.INT 10 | ECE–195.INT 10 | RPE–205.INT 10 | XBER12C1.INT 10 |
| *X. vasicola holcicola* | LMG 8276 | BXE–270.INT 6 | ECE–273.INT 6 | RPE–275.INT 6 | XBER16C3.INT 6 |
| *X. vasicola holcicola* | LMG 8276R | BXE–222.INT 23 | ECE–223.INT 22 | RPE–224.INT 23 | XBER18C1.INT 23 |
| *X. vasicola holcicola* | LMG 8277 | BXE–270.INT 13 | ECE–273.INT 13 | RPE–275.INT 13 | XBER16C3.INT 13 |
| *X. vasicola vasculorum* | LMG 8284 | BXE–218.INT 15 | ECE–219.INT 15 | RPE–220.INT 15 | XBER16C3.INT 28 |
| *X. vasicola vasculorum* | LMG 900 | BXE–241.INT 3 | ECE–244.INT 3 | RPE–247.INT 3 | XBER9C2.INT 3 |
| *X. vasicola vasculorum* | LMG 902 | BXE–182.INT 23 | ECE–179.INT 21 | RPE–174.INT 25 | XBER9C1.INT 23 |
| *X. vesicatoria* | LMG 911 | BXE–013.INT 9 | ECE–056.INT 4 | RPE–018.INT 9 | XBER2C2.INT 9 |
| *X. vesicatoria* | LMG 916 | BXE–181.INT 27 | ECE–178.INT 23 | RPE–173.INT 27 | XBER8C1.INT 27 |
| *X. vesicatoria* | LMG 917 | BXE–240.INT 17 | ECE–232.INT 17 | RPE–236.INT 18 | XBER9C1.INT 3 |
| *X. vesicatoria* | LMG 919 | BXE–222.INT 7 | ECE–223.INT 7 | RPE–224.INT 7 | XBER17C2.INT 19 |
| *X. vesicatoria* | LMG 920t1 | BXE–013.INT 10 | ECE–056.INT 5 | RPE–018.INT 10 | XBER2C2.INT 10 |
| *X. vesicatoria* | LMG 925 | BXE–218.INT 14 | ECE–219.INT 14 | RPE–220.INT 14 | XBER16C3.INT 27 |
| *X. vesicatoria* | LMG 935 | BXE–271.INT 29 | ECE–274.INT 29 | RPE–276.INT 29 | XBER18C1.INT 17 |
| **Non-target Data** | | | | | |
| *X. arboricola corylina** | LMG 8658 | BXE–003.INT 12 | ECE–006.INT 12 | RPE–009.INT 12 | XBER1C1.INT 12 |
| *X. arboricola corylina** | LMG 8658 | BXE–012.INT 12 | ECE–014.INT 12 | RPE–017.INT 12 | XBER1C2.INT 12 |
| *X. arboricola populi* | LMG 12141 | BXE–003.INT 16 | ECE–006.INT 16 | RPE–009.INT 16 | XBER1C1.INT 16 |
| *X. arboricola populi* | LMG 12141 | BXE–039.INT 5 | ECE–014.INT 16 | RPE–017.INT 16 | XBER1C2.INT 16 |
| *X. axonopodis cassavae* | LMG 8049 | BXE–258.INT 13 | ECE–301.INT 15 | RPE–200.INT 12 | XBER13C1.INT 12 |
| *X. axonopodis cassavae* | LMG 8049 | BXE–305.INT 10 | ECE–378.INT 5 | RPE–266.INT 14 | XBER13C2.INT 13 |
| *X. axonopodis cassiae* | LMG 675 | BXE–198.INT 15 | ECE–202.INT 15 | RPE–200.INT 15 | XBER13C1.INT 15 |
| *X. axonopodis cassiae* | LMG 675 | BXE–305.INT 13 | ECE–360.INT 22 | RPE–307.INT 15 | XBER13C2.INT 16 |
| *X. axonopodis citriae* | LMG 9045 | BXE–198.INT 14 | ECE–202.INT 14 | RPE–200.INT 14 | XBER13C1.INT 14 |
| *X. axonopodis citriae* | LMG 9045 | BXE–258.INT 15 | ECE–262.INT 15 | RPE–266.INT 16 | XBER13C2.INT 15 |
| *X. axonopodis coracanae* | LMG 7476 | BXE–159.INT 6 | ECE–176.INT 5 | RPE–142.INT 6 | XBER5C2.INT 28 |
| *X. axonopodis coracanae* | LMG 7476 | BXE–237.INT 28 | ECE–256.INT 6 | RPE–233.INT 28 | XBER6C1.INT 6 |
| *X. axonopodis cyamopsidis* | LMG 691 | BXE–198.INT 6 | ECE–202.INT 6 | RPE–200.INT 6 | XBER13C1.INT 6 |
| *X. axonopodis cyamopsidis* | LMG 691 | BXE–258.INT 7 | ECE–262.INT 7 | RPE–266.INT 8 | XBER13C2.INT 7 |
| *X. axonopodis desmodii* | LMG 692 | BXE–258.INT 5 | ECE–202.INT 4 | RPE–200.INT 4 | XBER13C1.INT 4 |
| *X. axonopodis desmodii* | LMG 692 | BXE–305.INT 8 | ECE–360.INT 16 | RPE–266.INT 6 | XBER13C2.INT 5 |
| *X. axonopodis desmodiigangeti* | LMG 693 | BXE–198.INT 28 | ECE–202.INT 28 | RPE–200.INT 28 | XBER13C1.INT 28 |
| *X. axonopodis desmodiigangeti* | LMG 693 | BXE–258.INT 29 | ECE–361.INT 2 | RPE–267.INT 2 | XBER13C2.INT 29 |
| *X. axonopodis desmodiilaxiflo* | LMG 9046 | BXE–198.INT 19 | ECE–202.INT 19 | RPE–200.INT 19 | XBER13C1.INT 19 |
| *X. axonopodis desmodiilaxiflo* | LMG 9046 | BXE–258.INT 20 | ECE–262.INT 20 | RPE–266.INT 22 | XBER13C2.INT 20 |

| | | | | | |
|---|---|---|---|---|---|
| *X. axonopodis desmodiirotundi* | LMG 694 | BXE--198.INT 17 | ECE--202.INT 17 | RPE--266.INT 19 | XBER13C1.INT 17 |
| *X. axonopodis desmodiirotundi* | LMG 694 | BXE--258.INT 18 | ECE--262.INT 18 | RPE--307.INT 16 | XBER13C2.INT 18 |
| *X. axonopodis erythrinae* | LMG 698 | BXE--210.INT 14 | ECE--212.INT 14 | RPE--217.INT 13 | XBER16C1.INT 14 |
| *X. axonopodis erythrinae* | LMG 698 | BXE--269.INT 5 | ECE--378.INT 16 | RPE--306.INT 7 | XBER16C2.INT 5 |
| *X. axonopodis lespedezae* | LMG 757 | BXE--207.INT 7 | ECE--206.INT 7 | RPE--208.INT 7 | XBER14C2.INT 27 |
| *X. axonopodis lespedezae* | LMG 757 | BXE--259.INT 27 | ECE--263.INT 27 | RPE--267.INT 28 | XBER15C1.INT 7 |
| *X. axonopodis phyllanthi* | LMG 844R | BXE--210.INT 4 | ECE--361.INT 29 | RPE--217.INT 4 | XBER15C2.INT 24 |
| *X. axonopodis phyllanthi* | LMG 844R | BXE--260.INT 24 | ECE--378.INT 10 | RPE--268.INT 25 | XBER16C1.INT 4 |
| *X. axonopodis poinsettiicola* | LMG 849 | BXE--182.INT 16 | ECE--179.INT 14 | RPE--174.INT 18 | XBER9C1.INT 16 |
| *X. axonopodis poinsettiicola* | LMG 849 | BXE--241.INT 8 | ECE--244.INT 8 | RPE--247.INT 8 | XBER9C2.INT 8 |
| *X. axonopodis rhynchosiae* | LMG 8021 | BXE--207.INT 10 | ECE--206.INT 10 | RPE--208.INT 10 | XBER15C1.INT 10 |
| *X. axonopodis rhynchosiae* | LMG 8021 | BXE--260.INT 2 | ECE--264.INT 2 | RPE--268.INT 2 | XBER15C2.INT 2 |
| *X. axonopodis sesbaniae* | LMG 867 | BXE--218.INT 17 | ECE--219.INT 17 | RPE--220.INT 17 | XBER17C1.INT 17 |
| *X. axonopodis sesbaniae* | LMG 867 | BXE--271.INT 2 | ECE--378.INT 22 | RPE--306.INT 14 | XBER17C2.INT 2 |
| *X. axonopodis vignaeradiatae* | LMG 936 | BXE--218.INT 24 | ECE--219.INT 24 | RPE--220.INT 24 | XBER17C1.INT 24 |
| *X. axonopodis vignaeradiatae* | LMG 936 | BXE--271.INT 8 | ECE--274.INT 8 | RPE--276.INT 8 | XBER17C2.INT 8 |
| *X. axonopodis vignicola* | LMG 828 | BXE--222.INT 14 | ECE--223.INT 13 | RPE--224.INT 14 | XBER17C2.INT 26 |
| *X. axonopodis vignicola* | LMG 828 | BXE--271.INT 26 | ECE--274.INT 26 | RPE--276.INT 26 | XBER18C1.INT 14 |
| *X. bromi or graminis?* | LMG 8274t1 | BXE--139.INT 19 | ECE--175.INT 9 | RPE--161.INT 9 | XBER5C1.INT 19 |
| *X. bromi or graminis?* | LMG 8274t1 | BXE--237.INT 6 | ECE--229.INT 6 | RPE--233.INT 6 | XBER5C2.INT 6 |
| *X. bromi or graminis?* | LMG 8274t2 | BXE--159.INT 2 | ECE--175.INT 27 | RPE--233.INT 24 | XBER5C2.INT 24 |
| *X. bromi or graminis?* | LMG 8274t2 | BXE--237.INT 24 | ECE--229.INT 25 | RPE--254.INT 3 | XBER6C1.INT 2 |
| *X. campestris aberrans* | LMG 9037 | BXE--001.INT 6 | ECE--007.INT 14 | RPE--010.INT 12 | XBER2C1.INT 12 |
| *X. campestris aberrans* | LMG 9037 | BXE--004.INT 12 | ECE--015.INT 12 | RPE--018.INT 12 | XBER2C2.INT 12 |
| *X. campestris armoraciae* | LMG 7383t2 | BXE--001.INT 8 | ECE--015.INT 14 | RPE--010.INT 14 | XBER2C1.INT 14 |
| *X. campestris armoraciae* | LMG 7383t2 | BXE--004.INT 14 | ECE--031.INT 18 | RPE--018.INT 14 | XBER2C2.INT 14 |
| *X. campestris armoraciae** | LMG 535 | BXE--001.INT 7 | ECE--015.INT 13 | RPE--010.INT 13 | XBER2C1.INT 13 |
| *X. campestris armoraciae** | LMG 535 | BXE--004.INT 13 | ECE--031.INT 17 | RPE--018.INT 13 | XBER2C2.INT 13 |
| *X. translucens hordei** | LMG 720 | BXE--207.INT 8 | ECE--361.INT 21 | RPE--208.INT 8 | XBER14C2.INT 28 |
| *X. translucens hordei** | LMG 720 | BXE--259.INT 28 | ECE--378.INT 8 | RPE--267.INT 29 | XBER15C1.INT 8 |
| *X. translucens hordei*** | LMG 879 | BXE--159.INT 26 | ECE--176.INT 25 | RPE--162.INT 3 | XBER6C1.INT 26 |
| *X. translucens hordei*** | LMG 879 | BXE--238.INT 19 | ECE--230.INT 19 | RPE--254.INT 5 | XBER6C2.INT 19 |
| *X. translucens phleipratensis* | LMG 843 | BXE--004.INT 29 | ECE--055.INT 7 | RPE--018.INT 29 | XBER2C1.INT 29 |
| *X. translucens phleipratensis* | LMG 843 | BXE--020.INT 9 | ECE--056.INT 7 | RPE--033.INT 14 | XBER2C2.INT 29 |

**Figure A.1** Sample rep-PCR Genomic Fingerprint Data File (First Two Strains)

```
List: NNBOXAPS
Traces: 752
Standard: JANGEL02 12
Resolution: 400
D:\JANSUB\GELS.INT\BXE--004.INT 2
Xanthomonas axonopodis vesicatoria
LMG 910
9


  39  39  39  39  25  11   0   2   7   9  16  23  27  37  43  46  32  41  48
  53  32  16   9   7   0   0   0   0   0   5   5  21  25  25  11   0   2   2
   5   2   0   0   0   2   5   2   0   7  11  18  14  16  18  25  23  21  18
  16  18  21  25  25  25  27  30  30  34  64  91 158 199 190 137  80  73  71
  73  82  85  89  80  89 107 105  91  66  71  73  69  59  53  62  69  75  64
  50  41  34  30  23  16  14   5   2   0   2   5  11  14  11   9   5   7   7
   5   2   0   0   0   0   0   0  11  18  53  62  55  27  16  21  21  18  11
   9   5   5   0   0   0   0   0  14  23  39  43  46  46  55  66  69  98 153
 219 255 254 197 137  94 105 117 130 128 130 130 133 117 112 119 135 139 130
```

```
117  87  71  78 107 126 112  85  64  43  37  30  27  23  23  21  27  30  32
 27  30  34  41  25   9   0   9  37  57  59  37   7   0  16  41  91 105 114
 80  55  23  11   0   0   0   0   9  48  85 133 139 146  96  64  34  41  46
 46  43  59  80 114 130 144 121 114 137 181 222 235 249 210 137  73  48  50
 46  41  50  57  71  82 110 126 114  80  48  53  62  71  80  85  85  73  59
 55  48  32  14   2   2   0   0   0   0   0   5  25  57  80  78  53  23   7
  2   0   0   7  16  25  21  21  34  48  69  64  62  55  62  85 105 110  96
103 119 151 158 165 135  98  78  80 107 126 142 137 130 112 101  78  59  43
 34  27  16   7   0   0   5   9  21  32  62  98 155 192 222 233 245 231 208
187 174 155 137 105  71  39  23   9   0   0  11  27  43  57  66  71  80  85
103 126 151 158 135  96  62  37  21   7   0   2   2   0   0   0   0   0   9
 18  46  94 139 144  91  41   5   0   0   0   0   0   0   0   0   0   0   0
  0
--------------------------------------------------------------------------
D:\JANSUB\GELS.INT\BXE--004.INT 3
Xanthomonas axonopodis vesicatoria
LMG 929
9


  0   0   0   0   0   0   0   2   5   5   9  14  14  12   7   9   5   5   5
  2   2   2   2   5   2   0   0   5   5   7   2   2   2   0   0   5  12  14
  9   2   0   0   0   5   2   0   0   0   0   2   5   2   0   0   0   0   0
  0   0   2   5   5   7   9  14  18  37  97 157 191 138  88  48  46  46  44
 55  60  69  69  71  71  60  51  51  58  65  58  46  32  25  21  18  21  28
 30  30  23  14   7   2   0   0   0   0   0   5   5   0   7  14  14   2   0
  2   9   9   7   2   0   0   0  14  28  35  23  14   7   9   7   7   9  14
 12   5   0   0   0   0   0   0   9  16  21  21  28  28  30  42  55 106 166
233 255 214 150  81  60  69  88  99 120 113 106  83  83  95 106 120 115  99
 67  35  42  74 101 111  99  71  46  28  32  58  78  85  58  32  23  21  23
 23  32  30  28  14   9  12  42  69  74  44  14   0   2  30  58  97  85  74
 46  32  21   5   0   0  23  46  78  97 125 141 122  81  39  23  25  30  32
 32  30  46  65  81  69  53  37  67 127 196 233 228 164  99  44  39  30  30
 32  39  46  48  65  88 113 108  71  42  30  39  46  53  62  55  51  28  23
 18  21  16   9   0   0   2   5   7   9  16  23  48  69  74  62  35  21   7
  0   0   0   0   0   0   9  14  30  39  51  51  62  74 101 115 122 113  99
106 125 145 138 106  69  44  60  81 104 111 104  99  99 111 138 161 175 148
 99  48  18  16  16  18  21  21  23  30  55  99 150 180 180 168 145 138 129
136 143 131 115  85  55  25   7   5   5  12  21  35  55  69  78  81  83  88
106 134 155 161 145 131 104  71  32   9   0   2   5  14  23  32  35  39  46
 78 115 157 155  99  55   0   0   0   0   0   0   0   2  21  21  12   0   0
  0
--------------------------------------------------------------------------
```

# APPENDIX B

## DATABASE AND DATA FILES

**Table B.1** rep-PCR Genomic Fingerprint Database Properties

| Table | Column | Type | Size |
|---|---|---|---|
| **Name** | | | |
| Record Count: 89 | Name | Text | 50 |
| **Strain** | | | |
| Record Count: 376 | Strain | Text | 40 |
| | Name_ | Text | 50 |
| **BFingerprint** | | | |
| Record Count: 752 | Index | Text | 40 |
| | Strain_ | Text | 20 |
| | Fingerprint | Memo | - |
| **EFingerprint** | | | |
| Record Count: 752 | Index | Text | 40 |
| | Strain_ | Text | 20 |
| | Fingerprint | Memo | - |
| **RFingerprint** | | | |
| Record Count: 752 | Index | Text | 40 |
| | Strain_ | Text | 20 |
| | Fingerprint | Memo | - |
| **BClassification** | | | |
| Record Count: 752 | Class_Id | Number (Integer) | 2 |
| | Index_ | Text | 40 |
| | Filtered | Memo | - |
| | Classes | Number (Integer) | 2 |
| | Classifications | Text | 140 |
| | Set | Text | 10 |
| **EClassification** | | | |
| Record Count: 752 | Class_Id | Number (Integer) | 2 |
| | Index_ | Text | 40 |
| | Filtered | Memo | - |
| | Classes | Number (Integer) | 2 |

| | | | |
|---|---|---|---|
| | Classifications | Text | 140 |
| | Set | Text | 10 |
| **RClassification** | | | |
| Record Count: 752 | Class_Id | Number (Integer) | 2 |
| | Index_ | Text | 40 |
| | Filtered | Memo | - |
| | Classes | Number (Integer) | 2 |
| | Classifications | Text | 140 |
| | Set | Text | 10 |
| **Target** | | | |
| Record Count: 63 | Tar_Id | Number (Integer) | 2 |
| | Name_ | Text | 50 |
| | ClassVec | Memo | - |
| **Non-target** | | | |
| Record Count: 26 | Non_Id | Number (Integer) | 2 |
| | Name_ | Text | 50 |
| **BParameters** | | | |
| Record Count: 16 | Parameter | Text | 15 |
| | Value | Text | 40 |
| **EParameters** | | | |
| Record Count: 16 | Parameter | Text | 15 |
| | Value | Text | 40 |
| **RParameters** | | | |
| Record Count: 16 | Parameter | Text | 15 |
| | Value | Text | 40 |

**Table B.2** BER Fingerprint Database Properties

| Table | Column | Type | Size |
|---|---|---|---|
| **Name** | | | |
| Record Count: 91 | Name | Text | 50 |
| **Strain** | | | |
| Record Count: 378 | Strain | Text | 40 |
| | Name_ | Text | 50 |
| **BFingerprint** | | | |
| Record Count: 752 | Index | Text | 40 |
| | Strain_ | Text | 20 |
| | Fingerprint | Memo | - |
| **BClassification** | | | |
| Record Count: 752 | Class_Id | Number (Integer) | 2 |
| | Index_ | Text | 40 |
| | Filtered | Memo | - |
| | Classes | Number (Integer) | 2 |
| | Classifications | Text | 140 |
| | Set | Text | 10 |
| **Target** | | | |
| Record Count: 63 | Tar_Id | Number (Integer) | 2 |
| | Name_ | Text | 50 |
| | ClassVec | Memo | - |
| **Non-target** | | | |
| Record Count: 26 | Non_Id | Number (Integer) | 2 |
| | Name_ | Text | 50 |
| **BParameters** | | | |
| Record Count: 16 | Parameter | Text | 15 |
| | Value | Text | 40 |

Sample Data Files:

- Selected Data for Denoising – File Name: *.txo

```
D:\JANSUB\GELS.INT\BXE--003.INT 16 [Index]

LMG 12141 [Strain]
   0   0   0   0   0   0   0   4   8  12   8   0   0   0   0   0   4
8   8   4   8   8   8  12  16  12   4   0   0   0   0   0   0   4   4
8   8   4   8  12  12   0   0   0   0   0   4   8   4   4   0   8   4
0   0   4   8  12  12   8   0   0   4  16  35  35  35  31  31  31  31
31  23  12   4   4   0   4   4   0   0   0   8   8  19  19  19   4   0
0  12  35  58  66  62  39  23  43  78 105 105  70  35   8  19  23  19
8   4   0   0   4   4   8  12  19  12   0   0   0   4   4   4   4   0
27 101 175 248 213 178  78  35   0   4  12  12   4  31  54  89  70  54
23  16   8   8   8   4   0   0  31 101 171 194 194 190 225 244 255 225
147  66  81 147 217 202 190 128 101  78  93  85  74  43  19   0   4  16
23  27  50  74 101  85  66  27  19   8  54 120 190 217 202 186 116  70
12   8   0   0   0   0   0   4   0  16  31  81  93  81  31   4  35 101
167 202 198 178 159 140 101  70  31  27  23  39  39  39  27  31  39  58
85 112 144 159 155 136 120 132 144 128 124 116 105  58  12   0   4   8
4  12  27  50  47  35  19  19  23  19  16  27  50  78  78  47  19   0
8  16  35  31  31  35  62 120 175 221 190 120  35   0  12  19  19  16
12  16  19  23  23  27  35  39  47  62  78  93 124 171 221 240 202 147
85 101 155 213 229 213 194 182 190 186 186 171 136  93  66  81  97 101
74  54  31  35  58  93 112  93  74  35  16   0   0   4  16  31  31  19
8   0   0   0   0   0   0   0   4   0   0   4   8  16  16  19  27  31
31  27  19  16  16  27  58  85 136 178 202 206 159 116  70  54  66  78
97 101  97  70  43  16  31  89 144 186 175 163  81  19   0   0   0   0
0   0   0   0   0   [Fingerprint]
```

- Filtered Data – File Name: *.txd

```
D:\JANSUB\GELS.INT\BXE--003.INT 16 [Index]

LMG 12141 [Strain]
   0   0   0   0   0   1   2   5   6   6   6   4   2   0   1   2   4
5   6   7   7   8  10  11  10   9   6   3   1   0   0   1   2   3   5
6   6   8   9   7   6   5   2   0   1   2   3   4   4   5   4   3   2
3   3   5   7   9   8   6   5   6  11  18  25  30  33  33  32  31  29
26  20  15   9   5   3   2   2   2   2   3   7  11  15  14  12   8   7
10  21  34  47  52  50  47  49  58  71  80  79  65  47  31  21  15  15
11   6   3   2   3   6   9  11  10   9   6   3   2   2   3   3   8  27
61 110 153 183 178 150 101  59  26  13   6  13  23  38  50  60  58  50
34  22  13   9   6   4   9  27  61  99 138 170 195 209 222 228 219 187
155 133 132 143 167 177 168 140 118  97  86  75  63  44  28  16  12  14
24  38  55  67  75  71  60  41  35  46  78 118 157 183 182 158 117  78
41  18   4   2   0   1   1   4  10  26  44  60  63  58  49  50  68 102
141 169 181 175 155 130 100  74  50  38  32  33  33  35  35  39  48  65
88 112 131 141 143 140 137 132 130 129 123 106  83  58  36  16   6   6
11  20  28  34  36  34  29  23  19  21  27  38  50  56  54  44  30  18
```

```
16  18  24  30  39  56  85 123 154 165 148 113  71  37  17  13  16  16
16  17  19  22  25  29  34  42  52  64  81 106 137 170 192 196 179 155
138 140 157 182 201 206 202 193 188 183 174 154 130 109  95  88  84  81
71  59  50  54  66  78  86  81  66  44  25  11   7  10  16  20  21  18
12   5   2   0   0   0   1   1   1   2   3   6   9  13  17  22  25  27
27  25  22  21  27  40  64  97 132 161 176 172 151 121  93  77  73  79
88  89  82  65  51  50  65  93 125 151 150 125  88  53  20   4   0   0
0   0   0   0   0 [Filtered Fingerprint]
```

- Training Data – File Name: *.txr

```
D:\JANSUB\GELS.INT\BXE--136.INT 6 [Index]
 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 [Target Vector]
LMG 482 [Strain]
   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
1   2   4   7  10  13  15  17  17  17  17  17  16  17  17  19  21  23
25  25  25  24  22  21  19  17  16  15  15  15  14  13  11  10   8   8
8   7   5   4   2   1   1   2   4   6   9  10  12  13  13  14  14  13
11   9   7   4   2   1   1   1   2   4   5   5   6   5   4   3   2   1
1   1   2   3   4   5   5   6   7   9  10  10   8   7   5   3   2   2
1   1   2   3   5   6   7   7   6   4   3   2   1   1   0   0   1   1
2   2   2   2   1   1   2   3   4   4   4   3   3   3   3   3   2   1
1   0   0   2   4   5   5   5   3   2   1   0   0   0   0   1   1   1
1   1   1   3  12  27  40  47  45  36  23  12   8   9   9   8   7   5
3   2   1   1   1   2   2   1   2   2   3   4   6   9  11  13  13  13
10   8   5   2   1   0   1   1   1   1   1   1   2   3   5   8  12  18
27  47  77 118 158 192 198 182 142 101  59  34  19  15  13  10   9   8
8   9   9  10  10  10  10  10  10   9   8   7   5   3   1   0   0   0
0   1   3   8  16  27  40  53  66  77  80  75  64  50  34  23  18  20
27  40  60  88 123 155 173 174 158 130  98  73  57  49  43  38  33  29
24  20  17  15  13  11   9   8   7   6   6   5   4   4   4   5   7   9
11  10   8   5   3   1   1   2   4   7  11  16  24  39  61  91 128 170
197 201 182 151 111  81  69  74  87 104 123 139 144 139 126 107  87  77
73  73  69  62  50  35  20  10   5   2   3   6  11  17  25  31  36  41
45  47  51  51  49  43  37  29  24  21  20  22  25  31  36  38  36  30
22  13   7   3   3   3   3   3   3   1   1   0   0   0   0   0   0   0
0   0   0   0   0 [Filtered Fingerprint]
```

- Testing Data – File Name: *.txs

```
D:\JANSUB\GELS.INT\BXE--225.INT 22
 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0
LMG 482
   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0   0
0   1   1   3   5   7   7   7   6   5   5   6   8   8   7   5   5   4
```

```
5   7   9  12  14  15  15  14  12  11  13  18  22  28  31  32  31  30
29  27  24  20  16  12   9   7   5   3   3   3   4   8  15  24  32  41
49  55  59  61  63  62  60  56  50  43  35  28  22  19  18  17  15  12
9   6   3   2   1   1   2   3   5   6   6   5   5   5   5   7   8   7
5   4   5  10  20  32  45  51  51  46  35  24  13   9   9  15  23  33
36  34  26  17  11  12  19  27  36  40  45  47  50  51  48  41  31  20
11   6   2   2   2   3   5   7   7   7   6   7   9  13  17  20  20  18
14   9   7  11  24  44  64  78  84  80  73  67  66  67  69  67  64  57
49  38  28  17   8   3   1   1   3   9  16  24  29  31  28  22  14   8
4   1   2   5   9  13  15  15  13  11  11  12  15  18  21  25  32  43
60  80 101 122 142 159 166 164 147 118  80  48  25  15  15  23  34  43
51  56  58  55  51  44  38  30  22  14   7   3   1   0   0   1   2   3
7  14  22  31  39  47  55  66  80  97 113 124 124 115 102  89  79  77
82  91 106 127 150 170 178 171 144 108  70  40  18   9   8  11  14  17
17  16  14  16  21  30  40  49  55  56  55  52  50  49  51  52  51  45
36  25  14   6   1   0   0   1   1   2   6  13  28  51  87 129 173 206
215 192 147  96  55  36  41  61  82  97 101  96  85  72  56  45  41  44
48  51  45  35  21   9   1   0   1   4   9  17  27  36  41  41  38  30
22  15  11   7   4   3   2   2   2   2   2   3   9  16  23  29  30  23
16   9   3   1   1   1   1   2   4   6   8   8   7   5   3   1   0   0
0   0   0   0   0
```

- **Non-target Data – File Name: *.txj**

```
D:\JANSUB\GELS.INT\BXE--003.INT 16

LMG 12141
   0   0   0   0   0   1   2   5   6   6   6   4   2   0   1   2   4
5   6   7   7   8  10  11  10   9   6   3   1   0   0   1   2   3   5
6   6   8   9   7   6   5   2   0   1   2   3   4   4   5   4   3   2
3   3   5   7   9   8   6   5   6  11  18  25  30  33  33  32  31  29
26  20  15   9   5   3   2   2   2   2   3   7  11  15  14  12   8   7
10  21  34  47  52  50  47  49  58  71  80  79  65  47  31  21  15  15
11   6   3   2   3   6   9  11  10   9   6   3   2   2   3   3   8  27
61 110 153 183 178 150 101  59  26  13   6  13  23  38  50  60  58  50
34  22  13   9   6   4   9  27  61  99 138 170 195 209 222 228 219 187
155 133 132 143 167 177 168 140 118  97  86  75  63  44  28  16  12  14
24  38  55  67  75  71  60  41  35  46  78 118 157 183 182 158 117  78
41  18   4   2   0   1   1   4  10  26  44  60  63  58  49  50  68 102
141 169 181 175 155 130 100  74  50  38  32  33  33  35  35  39  48  65
88 112 131 141 143 140 137 132 130 129 123 106  83  58  36  16   6   6
11  20  28  34  36  34  29  23  19  21  27  38  50  56  54  44  30  18
16  18  24  30  39  56  85 123 154 165 148 113  71  37  17  13  16  16
16  17  19  22  25  29  34  42  52  64  81 106 137 170 192 196 179 155
138 140 157 182 201 206 202 193 188 183 174 154 130 109  95  88  84  81
71  59  50  54  66  78  86  81  66  44  25  11   7  10  16  20  21  18
12   5   2   0   0   0   1   1   1   2   3   6   9  13  17  22  25  27
27  25  22  21  27  40  64  97 132 161 176 172 151 121  93  77  73  79
88  89  82  65  51  50  65  93 125 151 150 125  88  53  20   4   0   0
0   0   0   0   0
```

- Training Record – File Name: *.txn

| Epoch | LR | TSSE | GSSE |
|---|---|---|---|
| 0 | 0.005 | 1971.001 | 1971.209 |
| 1000 | 0.007 | 7.682 | 28.687 |
| 2000 | 0.012 | 3.445 | 27.042 |
| 3000 | 0.008 | 2.179 | 26.956 |
| 4000 | 0.004 | 1.584 | 27.119 |

- BPN Output Score – File Name: *.txc

```
D:\JANSUB\GELS.INT\BXE--225.INT 22
LMG 482
0.78  0.19  0.17  0.21  0.22  0.22  0.23  0.25  0.24  0.19  0.19  0.19
0.21  0.23  0.17  0.23  0.21  0.22  0.20  0.21  0.22  0.18  0.21  0.18
0.18  0.20  0.15  0.21  0.20  0.18  0.22  0.22  0.19  0.17  0.24  0.18
0.22  0.20  0.23  0.20  0.18  0.21  0.20  0.23  0.19  0.18  0.21  0.21
0.20  0.20  0.18  0.25  0.22  0.21  0.18  0.19  0.22  0.21  0.21  0.20
0.19  0.21  0.19
```

# APPENDIX C

# PROGRAM LISTING

## Mean Filtering (Visual Basic)

```
Private Sub cmdFilter_Click()
Dim dataPoint(400) As Long, dataFilter(400) As Long
Dim i As Integer, j As Integer, k As Integer

Me.MousePointer = 11

With fMainForm.ProgressBar1
    .Visible = True
    .Min = 0
    .Max = txtCount1.Text
    .Value = .Min
End With

datafile1 = dataPath + "\" + lblDbName.Caption + lblType.Caption + ".txo"
FileNum1 = FreeFile
Open datafile1 For Input As FileNum1

datafile2 = dataPath + "\" + lblDbName.Caption + lblType.Caption + ".txd"
FN2 = FreeFile
Open datafile2 For Output As FN2

' Read Data
Do Until EOF(FileNum1)
    DoEvents

    Line Input #FileNum1, Index
    Line Input #FileNum1, Header
    Line Input #FileNum1, Strain
    Line Input #FileNum1, Fingerprint
    Line Input #FileNum1, Header

'Mean Filtering
    For i = 0 To 399
        j = 4 * i + 1
        dataPoint(i) = Val(Mid(Fingerprint, j, 4))
    Next i

    k = ((Val(cboNeighbor.Text) + 1) / 2 - 1)

    For i = 0 To k - 1
        dataFilter(i) = dataPoint(i)
    Next i

    For i = 0 To 400 - Val(cboNeighbor.Text)
        dataFilter(i + k) = 0
        For j = 0 To (Val(cboNeighbor.Text) - 1)
            dataFilter(i + k) = dataFilter(i + k) + dataPoint(i + j)
        Next j
        dataFilter(i + k) = dataFilter(i + k) / j
    Next i
```

```
    For i = (400 - k) To 399
        dataFilter(i) = dataPoint(i)
    Next i

    Fingerprint = ""
    For i = 0 To 399
        Data = str(dataFilter(i))
        If Len(Data) = 2 Then
            Data = "  " + Data
        ElseIf Len(Data) = 3 Then
            Data = " " + Data
        End If
        Fingerprint = Fingerprint + Data
    Next i

    'Generate Output
    Print #FN2, Index
    Print #FN2, ""
    Print #FN2, Strain
    Print #FN2, Fingerprint
    Print #FN2, ""

    fMainForm.ProgressBar1.Value = fMainForm.ProgressBar1.Value + 1

Loop

Close #FN2, #FileNum1
Call storeParameters("Filtered", "Mean Filter " + cboNeighbor.Text)
Me.MousePointer = 0
fMainForm.ProgressBar1.Visible = False

End Sub
```

## Wavelet Filtering (Matlab)

### Initialization File: wavinit.m

```
clear all
addpath D:\PROJECTS\DNAVB
filename = 'ber01B.txo' ;
outputf = 'ber01B.txd' ;
W = 'db8' ;
level = 3 ;
denoise
```

### Wavelet ToolBox Script File: denoise.m

```
% Wavelet Filtering

fidr = fopen(filename,'r');
fidw = fopen(outputf,'w');

while 1

% Read data

index = fgetl(fidr);
```

```
if ~isstr(index), break, end
blank = fgetl(fidr);
strain = fgetl(fidr);
data = fscanf(fidr,'%d');

% Wavelet Filtering

datad = wden(data,'minimaxi','s','sln',level,W);

maxdatad = max(datad);

% Normalized Output
for i = 1:length(datad);
       if datad(i) < 0
          datad(i) = 0;
       else
           datad(i) = datad(i)/maxdatad * 255;
       end;
end;


% Write Filtered data

fprintf(fidw,'%s \r\n',index);
fprintf(fidw,'%s \r\n',blank);
fprintf(fidw,'%s \r\n',strain);

for idx = 1:length(datad)
       fprintf(fidw,'%3.0f ',datad(idx));
end;
fprintf(fidw,'\r\n');
fprintf(fidw,'\r\n');

end;

fclose(fidr);
fclose(fidw);

fidf = fopen('endflag.txt','w');
fprintf(fidf,'\r\n');
fclose(fidf);


'End'
```

## Define Target Classification (Visual Basic)

```
Private Sub cmdTargVec_Click()
Dim SQL1 As String, SQL2 As String
Dim vector As String
Dim cell() As Long
Dim N As Long, i As Long, j As Long, count As Long
Dim TrainClass() As String

On Error GoTo ErrHandle
Me.MousePointer = 11
DBGrid1.Visible = False

'Search for unique name
```

```
    SQLB = "SELECT DISTINCT Strain.Name_ AS BacName FROM BClassification INNER
JOIN (Strain INNER JOIN BFingerprint ON Strain.Strain = BFingerprint.Strain_)ON
BClassification.Index_ = BFingerprint.Index ORDER BY Strain.Name_;"

    SQLE = "SELECT DISTINCT Strain.Name_ AS BacName FROM EClassification INNER
JOIN (Strain INNER JOIN EFingerprint ON Strain.Strain = EFingerprint.Strain_)ON
EClassification.Index_ = EFingerprint.Index ORDER BY Strain.Name_;"

    SQLR = "SELECT DISTINCT Strain.Name_ AS BacName FROM RClassification INNER
JOIN (Strain INNER JOIN RFingerprint ON Strain.Strain = RFingerprint.Strain_)ON
RClassification.Index_ = RFingerprint.Index ORDER BY Strain.Name_;"

    datUName.DatabaseName = dataPath + "\" + lblDbName.Caption + ".Mdb"

    Select Case lblType.Caption
    Case "B"
        datUName.RecordSource = SQLB
    Case "E"
        datUName.RecordSource = SQLE
    Case "R"
        datUName.RecordSource = SQLR
    End Select
    datUName.Refresh

'Partition into target and non-target classification

    SQLB = "SELECT Strain.Name_ AS Name FROM BClassification INNER JOIN (Strain
INNER JOIN BFingerprint ON Strain.Strain = BFingerprint.Strain_)ON
BClassification.Index_ = BFingerprint.Index ORDER BY Strain.Name_;"

    SQLE = "SELECT Strain.Name_ AS Name FROM EClassification INNER JOIN (Strain
INNER JOIN EFingerprint ON Strain.Strain = EFingerprint.Strain_)ON
EClassification.Index_ = EFingerprint.Index ORDER BY Strain.Name_;"

    SQLR = "SELECT Strain.Name_ AS Name FROM RClassification INNER JOIN (Strain
INNER JOIN RFingerprint ON Strain.Strain = RFingerprint.Strain_)ON
RClassification.Index_ = RFingerprint.Index ORDER BY Strain.Name_;"

    datAName.DatabaseName = datUName.DatabaseName
    Select Case lblType.Caption
    Case "B"
        datAName.RecordSource = SQLB
    Case "E"
        datAName.RecordSource = SQLE
    Case "R"
        datAName.RecordSource = SQLR
    End Select
    datAName.Refresh

    With fMainForm.ProgressBar1
        .Visible = True
        .Min = 0
        .Max = datUName.Recordset.RecordCount
        .Value = .Min
    End With

    i = 0
    ReDim TrainClass(0)
    datUName.Recordset.MoveFirst
    Do While datUName.Recordset.EOF = False
        count = 0
        datAName.Recordset.MoveFirst
        Do While datAName.Recordset.EOF = False
```

```
            If datUName.Recordset!BacName = datAName.Recordset!Name Then
               count = count + 1
            End If
            If count > 2 Then
               ReDim Preserve TrainClass(UBound(TrainClass) + 1)
               i = i + 1
               TrainClass(i) = datUName.Recordset!BacName
               Exit Do
            End If
        datAName.Recordset.MoveNext
        Loop

        datUName.Recordset.MoveNext
        fMainForm.ProgressBar1.Value = fMainForm.ProgressBar1.Value + 1
    Loop

DoEvents
Call TargetVector(i, TrainClass)

fMainForm.ProgressBar1.Visible = False
Me.MousePointer = 0

Exit Sub

ErrHandle:
MsgBox Err.Description
Me.MousePointer = 0
Exit Sub
End Sub
```

## Generate Target Vector (Visual Basic)

```
Private Sub TargetVector(N As Long, TrainClass() As String)
Me.MousePointer = 11

'Clear database
datTar.Refresh
Do While datTar.Recordset.EOF = False
    datTar.Recordset.Delete
    datTar.Recordset.MoveNext
Loop

'Unique classification count

    txtCount1.Text = N
    txtOutputU.Text = N

'Generate target vector

    ReDim cell(N)
    For i = 1 To N
          cell(i) = 0
    Next i

    j = 1
    For j = 1 To N
        cell(j) = 1
        vector = ""
        For i = 1 To N
            vector = vector + str(cell(i))
        Next i
```

```
        With datTar
            .Recordset.AddNew
            .Recordset!Tar_Id = j
            .Recordset!Name_ = TrainClass(j)
            .Recordset!ClassVec = vector
            .Recordset.Update
        End With

        cell(j) = 0
    Next j


datTar.Refresh
DBGrid1.Visible = True
DBGrid1.Refresh

Call TrainingTestingSets
Me.MousePointer = 0

End Sub
```

## BPN Training (Matlab)

### Initialization File: nninit.m

```
%% Neural Network Initialization File
clear all
filename1 = 'dat01B.txr';
filename2 = 'dat01B.txs';
hidden = 120 ;
lr = 0.005;
maxepoch = 0;
range1 = 31 ;
range2 = 390 ;
matfilename = 'dat01B';
train_re = 'dat01B.txn';
```

### BPN Training Script File: nntrain.m

```
% Neural Networks Training
% -------------------------
nninit

% Training Data

fidr1 = fopen(filename1,'r');
i=1;
while 1

% Read data

       index = fgetl(fidr1);
       if ~isstr(index), break, end
       traintarget(:,i) = fscanf(fidr1,'%d');
       strain = fgetl(fidr1);
       traindata(:,i) = fscanf(fidr1,'%d');
i=i+1;
```

```
end

fclose(fidr1);

P = 2*traindata(range1:range2,:)/255;
T = traintarget*0.6 + 0.2;


% Testing Data

fidr2 = fopen(filename2,'r');
i=1;
while 1

% Read data

       index = fgetl(fidr2);
       if ~isstr(index), break, end
       testtarget(:,i) = fscanf(fidr2,'%d');
       strain = fgetl(fidr2);
       testdata(:,i) = fscanf(fidr2,'%d');
i=i+1;


end

fclose(fidr2);

U = 2*testdata(range1:range2,:)/255;
TU = testtarget*0.6 + 0.2;

% ---- Weight Initialization with Nguyen-Widrow initial conditions ----

R = size(P,1);
L = size(P,2);
S1 = hidden;
S2 = size(traintarget,1);

p1 = zeros(R,2);
p1(:,2) = ones(R,1)*2;

[W1,B1] = nwlog(S1,p1);

W2 = rands(S2, S1)*0.01;
B2 = rands(S2, 1)*0.01;

% ---- BPN Training ----

disp_freq = 100;
max_epoch = maxepoch;
err_goal = err;
lr_inc = 1.05;
lr_dec = 0.7;
momentum = 0.9;
err_ratio = 1.04;

TP = [disp_freq max_epoch err_goal lr lr_inc lr_dec momentum err_ratio];
[W1,B1,W2,B2,epochs] =
tbpx2(W1,B1,'logsig',W2,B2,'logsig',P,T,U,TU,TP,train_re);

% --- saving ---

save(matfilename,'W1','B1','W2','B2','range1','range2')
```

## BPN Classification (Matlab)

### Initialization File: clinit.m

```
%% Classification Initialization File
clear all
addpath D:\PROJECTS\DNAVB
fname1 = 'dat01B.txs';
fname2 = 'dat01B.txc';
matfilename = 'dat01B';
classify
```

### BPN Classification Script File: classify.m

```
% Testing Data
% -------------

load(matfilename)

fidr = fopen(fname1,'r');
fidw = fopen(fname2,'w');

i=1;
while 1

% Read data

       index = fgetl(fidr);
       if ~isstr(index), break, end
       blank = fgetl(fidr);
       strain = fgetl(fidr);
       fingerprint = fscanf(fidr,'%d');

% Classification

       finp = 2*fingerprint(range1:range2)/255;
       [A1,A2] = simuff(finp,W1,B1,'logsig',W2,B2,'logsig');

       result = compet(A2);

       fprintf(fidw, '%s \r\n',index);
       fprintf(fidw, '%s \r\n',strain);
       fprintf(fidw, '%3.2f  ',A2);
       fprintf(fidw, '\r\n','');
       fprintf(fidw, '\r\n','');


i=i+1;

end

fclose(fidr);
fclose(fidw);

flag = 'ok.'
```

## BPN Output Score Thresholding and Pathovar Assignment (Visual Basic)

```
Private Sub cmdResult_Click()
Dim colNumberNNScore As Long, colNumberCheck As Long
Dim N As Long, Idx As Long, i As Long, j As Long, k As Long, l As Long, m As
Long
Dim threshold As Single, score As Single, vflag As Long

On Error GoTo ErrHandle

Me.MousePointer = 11

DataNNinfo.Recordset.FindFirst ("Parameter = 'Output Units'")
If Not DataNNinfo.Recordset.NoMatch Then
    N = DataNNinfo.Recordset!Value
End If

With DataClassification
    .DatabaseName = dataPath + "\" + lblDbName.Caption + ".mdb"
    .RecordSource = "Target"
    .Refresh
End With

With DataMain
    SQLB = "SELECT BFingerprint.Index, Strain.Strain, Strain.Name_ As Name,
BFingerprint.Fingerprint As Fingerprint, BClassification.Filtered As DenFin
FROM BClassification INNER JOIN (Strain INNER JOIN BFingerprint ON
Strain.Strain = BFingerprint.Strain_ ) ON BClassification.Index_ =
BFingerprint.Index ORDER BY  BFingerprint.Index;"

    SQLE = "SELECT EFingerprint.Index, Strain.Strain, Strain.Name_ As Name,
EFingerprint.Fingerprint As Fingerprint, EClassification.Filtered As DenFin
FROM EClassification INNER JOIN (Strain INNER JOIN EFingerprint ON
Strain.Strain = EFingerprint.Strain_ ) ON EClassification.Index_ =
EFingerprint.Index ORDER BY  EFingerprint.Index;"

    SQLR = "SELECT RFingerprint.Index, Strain.Strain, Strain.Name_ As Name,
RFingerprint.Fingerprint As Fingerprint, RClassification.Filtered As DenFin
FROM RClassification INNER JOIN (Strain INNER JOIN RFingerprint ON
Strain.Strain = RFingerprint.Strain_ ) ON RClassification.Index_ =
RFingerprint.Index ORDER BY  RFingerprint.Index;"

    .DatabaseName = dataPath + "\" + lblDbName.Caption + ".mdb"

    Select Case lblType.Caption
    Case "B"
        .RecordSource = SQLB
    Case "E"
        .RecordSource = SQLE
    Case "R"
        .RecordSource = SQLR
    End Select
    .Refresh
End With

colNumberNNScore = fgrNNScore.Cols
With fgrCheck
    .Visible = False
    .Clear
    .Cols = 7
    colNumberCheck = .Cols
    .Rows = 2
```

```
    .ColWidth(0) = 400
    .ColWidth(1) = 3200
    .TextArray(0 * colNumberCheck + 1) = "Index                "
    .ColWidth(2) = 1000
    .TextArray(0 * colNumberCheck + 2) = "Strain    "
    .ColWidth(3) = 3000
    .TextArray(0 * colNumberCheck + 3) = "Target Classification"
    .ColWidth(4) = 3600
    .TextArray(0 * colNumberCheck + 4) = "NN Classification"
    .ColWidth(5) = 660
    .TextArray(0 * colNumberCheck + 5) = "Classes"
    .ColWidth(6) = 1000
    .TextArray(0 * colNumberCheck + 6) = "Verification"
End With

With fMainForm.ProgressBar1
    .Visible = True
    .Min = 1
    .Max = fgrNNScore.Rows - 1
    .Value = .Min
End With

l = 0
m = 0
For i = 1 To (fgrNNScore.Rows - 1)
    fMainForm.ProgressBar1.Value = i

    Entry = i
    Entry = Entry & Chr(9) & fgrNNScore.TextArray(i * colNumberNNScore + 1)
    Entry = Entry & Chr(9) & fgrNNScore.TextArray(i * colNumberNNScore + 2)

'Search for Target Classification
    DataMain.Recordset.FindFirst ("Index = '" + fgrNNScore.TextArray(i *
colNumberNNScore + 1) + "' ")
    If Not DataMain.Recordset.NoMatch Then
        Entry = Entry & Chr(9) & DataMain.Recordset!Name
    Else
        Entry = Entry & Chr(9) & ""
    End If

'Determine NN Classification by checking with Target table
    threshold = Val(txtThreshold.Text)
    Entry = Entry & Chr(9) & ""

    vflag = 0
    k = 0
    For j = 3 To N + 2
        score = Val(fgrNNScore.TextArray(i * colNumberNNScore + j))
        If score >= threshold Then
            txtClass.Text = j - 2

            DataClassification.Recordset.FindFirst ("Tar_Id = " + txtClass.Text
+ " ")
            If Not DataClassification.Recordset.NoMatch Then
                Entry = Entry & DataClassification.Recordset!Name_ _
                & " [" + str(score) + "]" & "  "
            End If

'Check If NN Classification is Correct
            If DataMain.Recordset!Name = DataClassification.Recordset!Name_
Then
                vflag = 1
                l = l + 1
```

```
            End If

        k = k + 1
        End If
    Next j

'Count number of classes

    Entry = Entry & Chr(9) & k

    If k = 0 Then m = m + 1

' Verification
    If vflag = 1 Then
        Entry = Entry & Chr(9) & "T"
    Else
        Entry = Entry & Chr(9) & " "
    End If

    fgrCheck.AddItem Entry
Next i

fgrCheck.RemoveItem 1
fgrCheck.Refresh

'Error Analysis
fMainForm.ProgressBar1.Visible = False
subError = (fgrCheck.Rows - 1 - l - m) / (fgrCheck.Rows - 1) * 100
rejError = m / (fgrCheck.Rows - 1) * 100
lblRec.Caption = "Recognized: " + str(l) + " / " + str(fgrCheck.Rows - 1) + " =
" + Format((l / (fgrCheck.Rows - 1) * 100), "##0.00") + "%"
lblSub.Caption = "Misclassified: " + Format(subError, "##0.00") + "%"
lblRej.Caption = "Rejected: " + Format(rejError, "##0.00") + "%"

Call storeParameters("Threshold", txtThreshold.Text)
If cboClassData.Text = "Non-target" Then
    Call storeParameters("False Accept", Format(subError, "##0.00") + "%")
Else
    Call storeParameters("Recognition", Format((l / (fgrCheck.Rows - 1) * 100),
"##0.00") + "%")
    Call storeParameters("Misclassified", Format(subError, "##0.00") + "%")
    Call storeParameters("False Rejection", Format(rejError, "##0.00") + "%")
End If
Call storeClassification
fgrCheck.Visible = True

Me.MousePointer = 0
Exit Sub

ErrHandle:
Me.MousePointer = 0
MsgBox Err.Description
Exit Sub

End Sub
```

# LIST OF REFERENCES

## LIST OF REFERENCES


Antonini, M., M. Barlaud, P. Mathieu and I. Daubechies. 1992. Image Coding Using Wavelet Transform. *IEEE Transactions on Image Processing*, 1(2): 205-220.

Barclay, V.J. and R.F. Bonner. 1997. Application of Wavelet Transforms to Experimental Spectra: Smoothing, Denoising, and Data Set Compression. *Analytical Chemistry*, 69(1): 78-90.

Boddy, L., C.W. Morris, M.F. Wilkins, G.A. Tarran and P.H. Burkill. 1994. Neural Network Analysis of Flow Cytometric Data for 40 Marine Phytoplankton Species. *Cytometry*, 15: 283-293.

Castleman, K.R. 1996. *Digital Image Processing*. Englewood Cliffs: Prentice-Hall, Inc.

Cho, S. 1997. Neural-Network Classifiers for Recognizing Totally Unconstrained Handwritten Numerals. IEEE Transactions on Neural Networks, 8(1): 43-53.

Civerolo, E.L. and M.J. Hattingh. 1993. Xanthomonas campestris pv. Pruni: cause of prunus bacterial spot, in *Xanthomonas* (ed. J.G. Swings and E.L. Civerolo). London: Chapman & Hall.

Curry, B. and D.E. Rumelhart. 1990. Neural Network for Classification of Mass Spectra. *Tetrahedron Computing Methodology*, 3: 213-237.

Demuth, H. and M. Beale. 1994. Neural Network Toolbox User's Guide. Natick: The MathWorks, Inc.

Donoho, D.L. 1995. De-Noising by Soft-Thresholding. *IEEE Transactions on Information Theory*, 41(3): 613-627.

Donoho, D.L. and I.M. Johnstone. 1995. Adapting to Unknown Smoothness via Wavelet Shrinkage. *Journal of the American Statistical Association*, 90(432): 1200-1224.

Duveiller, E. and H. Maraite. 1993. Xanthomonas campestris pathovars on cereals: cause of leaf streak or black chaff diseases, in *Xanthomonas* (ed. J.G. Swings and E.L. Civerolo). London: Chapman & Hall.

Edwards, M. and D.R. Morse. 1995. The potential for computer-aided identification in biodiversity research. *Trends. Ecol. Evol.* 10(4): 153-158.

Freeman, J. A. 1994. *Simulating Neural Networks with Mathematica*. Reading: Addison-Wesley Publishing Company, Inc.

Goodfellow, M., G.P. Manfio and J. Chun. 1997. Towards a practical species concept for cultivable bacterial. In *Species: The units of Biodiversity* (ed. M.F. Claridge, H.A. Dawah and M.R. Wilson). London: Chapman & Hall.

Hartl, D.L. 1994. *Genetics*. Boston: Jones and Bartlett Publishers.

Haykin, S. 1994. *Neural networks: a comprehensive foundation*. New York: Macmillan College Publishing Company.

Hayward, A.C. 1993. The host of Xanthomonas, in *Xanthomonas* (ed. J.G. Swings and E.L. Civerolo). London: Chapman & Hall.

Hokawat, S. and K. Rudolph. 1993. Xanthomonas campestris pv. Glycines: cause of bacterial pustule of soybean, in *Xanthomonas* (ed. J.G. Swings and E.L. Civerolo). London: Chapman & Hall.

Jain, A.K., J. Mao and K.M. Mohiuddin. 1996. Artificial Neural Networks: A Tutorial. IEEE Computer, March, 31-44.

Jain, R., R. Kasturi and B.G. Schunck. 1995. *Machine Vision*. New York: McGraw-Hill, Inc.

King, R.C. and W.D. Stanfield. 1997. *A Dictionary of Genetics*. New York: Oxford University Press.

Leyns, F. 1993. Xanthomonas campestris pv. Graminis: cause of bacterial wilt of forage grases, in *Xanthomonas* (ed. J.G. Swings and E.L. Civerolo). London: Chapman & Hall.

Leyns, F., M.D. Cleene, J. Swings and J.D. Ley. 1984. The Host Range of the Genus Xanthomonas. *The Botanical Review*, 50(3): 308-353.

Liang, J. 1997. Highly Scalable Image Coding for Multimedia Applications. *Proceedings ACM Multimedia 97 Seattle Washington*, 11-16.

Macdonell, M.T. and R.R. Colwell. 1985. The contribution of numerical taxonomy to the systematics of gram-negative bacteria. In *Computer-assisted bacterial systematics*. (ed. by M. Goodfellow, D. Jones and F.G. Priest). London: Academic Press.

Margulis, L. and K.V. Schwartz. 1998. *Five Kingdoms*. New York: W. H. Freeman and Company.

Mew, T.W.1993. Xanthomonas oryzae pathovars on rice: cause of bacterial blight and bacterial leaf streak in *Xanthomonas* (ed. J.G. Swings and E.L. Civerolo). London: Chapman & Hall.

Misiti, M., Y. Misiti, G. Oppenheim and J. Poggi. 1996. *Wavelet Toolbox User's Guide*. Natick: The Math Works, Inc.

Pankhurst, Richard J. 1991. *Practical taxonomic computing*. Cambridge: Cambridge University Press.

Priest, F. and B. Austin. 1993. *Modern Bacterial Taxonomy*. London: Chapman & Hall.

Pruvost, O. and B.Q. Manicom. 1993. Xanthomonas campestris pv. Mangiferaeindicae: cause of bacterial black spot of mangoes, in *Xanthomonas* (ed. J.G. Swings and E.L. Civerolo). London: Chapman & Hall.

Rademaker, J.L.W. and F.J. de Bruijn. 1997. Characterization and Classification of Microbes by rep-PCR Genomic Fingerprinting and computer-assisted pattern analysis, in *DNA markers: Protocols, Applications and Overviews* (ed. Caetano-Anollés, G., Gresshoff, P.M.). Chapter 10, 151-171. J. Wiley & Sons, Inc.

Rademaker, J.L.W., F.J. Louws, M.H. Schultz, U. Rossbach, L. Vauterin, J. Swings and F.J. de Bruijn. 1997. Molecular Systematics of Xanthomonas by rep-PCR Genomic Fingerprinting and Computer-Assisted Pattern Analysis. Poster B6 presented at APS annual meeting.

Rat, 1993. Xanthomonas fragariae: cause of angular leaf spot of strawberry, in *Xanthomonas* (ed. J.G. Swings and E.L. Civerolo). London: Chapman & Hall.

Ride, M. 1993. Xanthomonas populi: cause of bacterial canker of poplar, in *Xanthomonas* (ed. J.G. Swings and E.L. Civerolo). London: Chapman & Hall.

Rott, P. 1993. Xanthomonas albilineans: cause of leaf scald of sugar cane, in *Xanthomonas* (ed. J.G. Swings and E.L. Civerolo). London: Chapman & Hall.

Rumelhart, D.E., B. Widrow and M.A. Lehr. 1994. The Basic Ideas in Neural Network. Communications of the ACM, 27(3): 87-92.

Schaad, N.W. and A. Alvarez. 1993 Xanthomonas campestris pv campestris: cause of black rot of crucifers, in *Xanthomonas* (ed. J.G. Swings and E.L. Civerolo). London: Chapman & Hall.

Schindler, J., P. Paryzek and J. Farmer III. 1994. Identification of Bacterial by Artificial Neural Networks. *Binary Computing in Microbiology*, 6:191-196.

Schittenkopf, C., G. Deco and W. Brauer. 1997. Two Strategies to Avoid Overfitting in Feedforward Networks. *Neural Networks*, 10(3): 505-516.

Skapura, D.M.1996. *Building neural networks*. New York: ACM Press.

Stackebrandt, E., R.G.E. Murray and H.G. Truper. 1988. Proteobacteria class nov., a name for the phylogenetic taxon that includes the "Pulple bacterial and their relatives". *International Journal of systematic bacteriology*, 38(3): 321-325.

Stall, R.E. 1993. Xanthomonas campestris pv. Vesicatoria: cause of bacterial spot of tomato and pepper, in *Xanthomonas* (ed. J.G. Swings and E.L. Civerolo). London: Chapman & Hall.

Stall, R.E. and Civero E.L. 1993. Xanthomonas campestris pv. Citri: cause of citrus canker, in *Xanthomonas* (ed. J.G. Swings and E.L. Civerolo). London: Chapman & Hall.

Strang, G. and T. Nguyen. 1996. *Wavelets and Filter Banks*. Wellesley: Wellesley-Cambridge Press.

Stryer, L. 1995. *Biochemistry*. New York: W.H. Freeman and Company.

Swings, J., L. Vauterin, and Kersters. 1993. The bacterium Xanthomonas, in *Xanthomonas* (ed. J.G. Swings and E.L. Civerolo). London: Chapman & Hall.

Towner, K.J. and A. Cockayne. 1993. *Molecular Methods for Microbial Identification and Typing*. London: Chapman & Hall.

Vauterin, L., B. Hoste, K. Kersters and J. Swings. 1995. Reclassification of Xanthomonas. *Inetrnational Journal of systematic bacteriology*, 45(3): 472-489.

Vauterin, L., B. Hoste, P. Yang, A. Alvarez, K. Kersters and J. Swings. 1993. Taxonomy of the genus Xanthomonas, in *Xanthomonas* (ed. J.G. Swings and E.L. Civerolo). London: Chapman & Hall.

Vidaver, A.K. 1993. Xanthomonas campestris pv. Phaseoli: cause of common bacterial blight on bean, in *Xanthomonas* (ed. J.G. Swings and E.L. Civerolo). London: Chapman & Hall.

Vogl, T. P., J. K. Mangis, A. K. Rigler, W. T. Zink and D. L. Alkon. 1988. Accelerating the convergence of the backpropagation method. *Biological Cybernetics*, 59: 257-263.

Watson, J.D., M. Gilman, J. Witkowski and M. Zoller. 1992. *Recombinant DNA*. Scientific American Books.

Weaver, J.B., Y. Xu, D.M. Healy Jr. and L.D. Cromwell. 1991. Filtering Noise from Images with Wavelet Transforms. Magnetic Resonance in Medicine, 21: 288-295.

Weeks, P.J.D. and K.J. Gaston. 1997. Image Analysis, Neural Networks and the Taxonomic Impediment to Biodiversity Studies. *Biosiversity and Conservation*, 6: 263-274.

Wetherbe, J. and N.P. Vitalari. 1994. *Systems Analysis and Design: Traditional, Best Practices*. St. Paul: West Publishing.

Whitten, J.L. and L.D. Bentley. 1998. Systems Analysis and Design Methods. Boston: Irwin McGraw-Hill.

Wilkins, M.F., L. Boddy, C.W. Morris and R. Jonker. 1996. A Comparison of Some Neural and Non-neural methods for identification of phytoplankton from flow cytometry data. *Computer Applications in the Biosciences*, 12(1): 9-18.

Wu, C., G. Whitson, J. McLarty, A. Ermongkonchai and T. Chang. 1992. Protein classification artificial neural system. *Protein Science*, 1: 667-667.

Wu, C., S. Shivakumar. 1994. Back-propagation neural networks for phylogenetic classification of ribosomal RNA sequences. *Nucleic Acids Research*, 22(20): 4291-4299.

Yourdon, E. 1989. *Modern Structured Analysis*. Englewood Cliffs: Yourdon Press.

Zomorodian, A. and K. Rudolph. 1993. Xanthomonas campestris pv. Malvacearum: cause of bacterial blight of cotton, in *Xanthomonas* (ed. J.G. Swings and E.L. Civerolo). London: Chapman & Hall.