Measurement of Haplotypic Variation in *Xanthomonas oryzae* pv. *oryzae* Within a Single Field by rep-PCR and RFLP Analyses

C. M. Vera Cruz, E. Y. Ardales, D. Z. Skinner, J. Talag, R. J. Nelson, F. J. Louws,
H. Leung, T. W. Mew, and J. E. Leach

First and ninth authors: Department of Plant Pathology, 4024 Throckmorton Plant Sciences Center, Kansas State University, Manhattan 66506-5502; second, fourth, fifth, and eighth authors: Division of Entomology and Plant Pathology, International Rice Research Institute, P.O. Box 933, 1099 Manila, the Philippines; third author: USDA-ARS and Department of Agronomy, Kansas State University, Manhattan 66506-5502; sixth author: MSU-DOE Plant Research Lab, Michigan State University, East Lansing 48824; seventh author: Department of Plant Pathology, Washington State University, Pullman 99164-6430.

ABSTRACT

Vera Cruz, C. M., Ardales, E. Y., Skinner, D. Z., Talag, J., Nelson, R. J., Louws, F. J., Leung, H., Mew, T. W., and Leach, J. E. 1996. Measurement of haplotypic variation in *Xanthomonas oryzae* pv. *oryzae* within a single field by rep-PCR and RFLP analyses. Phytopathology 86:1352-1359.

The haplotypic variation of *Xanthomonas oryzae* pv. *oryzae* in a farmer's field that had endemic bacterial blight in the Philippines was evaluated at a single time. The genomic structure of the field population was analyzed by repetitive sequence-based polymerase chain reaction with oligonucleotide primers corresponding to interspersed repeated sequences in prokaryotic genomes and restriction fragment length polymorphism (RFLP) with the insertion sequence IS*1113*. The techniques and specific probes and primers were selected because they grouped consistently into the same lineages a set of 30 selected *X. oryzae* pv. *oryzae* strains that represented the four distinct RFLP lineages found in the Philippines did. Strains (155) were systematically collected from a field planted to rice

cv. Sinandomeng, which is susceptible to the indigenous pathogen population. Two of the four Philippine lineages, B and C, which included race 2 and races 3 and 9, respectively, were detected in the field. Lineage C was the predominant population (74.8%). The haplotypic diversities of 10 of the 25 blocks were significantly greater than the total haplotypic diversity of the collection in the entire field; however, between individual blocks the haplotypic diversities were not significantly different. Haplotypes from both lineages were distributed randomly across the field. Analysis of genetic diversity at the microgeographic scale provided insights into the finer scale of variation of *X. oryzae* pv. *oryzae*, which are useful in designing experiments to study effects of host resistance on the population structure of the bacterial blight pathogen.

Additional keywords: genetic variation, population genetics, rice bacterial blight.

Bacterial blight of rice, caused by *Xanthomonas oryzae* pv. *oryzae* (40), has been managed primarily by host resistance. Although resistance has proven to be a successful management practice for the disease in a number of cases, virulent races or pathotypes of *X. oryzae* pv. *oryzae* have emerged to overcome deployed resistance (7,23,25). This change in the race structure of the pathogen population may have resulted from several factors, including genetic change leading to a virulent race (mutation or recombination), the introduction of a new race from another geographic area (migration), or the build-up of a preexisting, but minor, component of the population (selection). To guide the selection and deployment of bacterial blight-resistance genes in the future, we have undertaken a hierarchical analysis to evaluate the structure of *X. oryzae* pv. *oryzae* populations. "Population structure" includes diversity, phylogeny, and partitioning in time, space, host, and environment (15,18).

Studies of populations between countries and regions within countries have indicated that regionally defined pathogen populations are distinct, which could be due to either slow pathogen

migration/dispersal or spatial partitioning of host genotypes (1,16, 31). A comprehensive study of the extent and distribution of pathogen diversity among agroecosystems, sites, and individual fields on the island of Luzon in the Philippines was conducted by Ardales et al. (2). In their study (2), over 1,200 strains of *X. oryzae* pv. *oryzae* were systematically collected from 13 sites along a 310-km transect spanning an indigenous rice-growing area and a modern agroecosystem on Luzon. The study (2) revealed population substructuring at the levels of agroecosystems, sites, fields within sites, and sampling areas within fields. Whether the geographic differentiation detected by Ardales et al. (2) was due to varietal selection or a low rate of migration is unknown.

Our goal was to continue the hierarchical analysis of *X. oryzae* pv. *oryzae* population structures in the Philippines by examining the variation and microgeographic distribution of the pathogen in a single field at a single time. However, in designing these and future studies, it became clear that improved techniques for monitoring genomic diversity were needed to process the large number of required samples. In previous studies, restriction fragment length polymorphism (RFLP) analysis was documented as a reliable tool for understanding the population biology and structure of *X. oryzae* pv. *oryzae* (1,2,15,16,18). Although useful, the RFLP procedure is time-consuming and expensive. These problems are exacerbated in population studies due the large number of samples required to obtain the desired level of confidence in the data.

Recently, the utility of the repetitive sequence-based polymerase chain reaction (rep-PCR) technique was demonstrated for epidemiological studies of several mammalian pathogens (3,10,35,39). Louws et al. (19,20) used the technique to differentiate pathovars and strains of plant-pathogenic xanthomonads and pseudomonads. The technique is based on the amplification of DNA with oligonucleotide primers from three families of unrelated repetitive DNA sequences corresponding to repetitive extragenic palindromic (REP [11,43]), enterobacterial repetitive intergenic consensus (ERIC [13]), and BOX elements (21). We reasoned that if the dispersion of REP, ERIC, and BOX sequences is indicative of the structure and evolution of the bacterial genome, then the phylogenetic lineages of X. oryzae pv. oryzae previously inferred from RFLP data also should be predicted by rep-PCR data (15). Also, because the primers amplify DNA between the dispersed elements and these fragments are likely to be different sizes, differences between closely related strains might be detected. If so, then rep-PCR, which is simpler, faster, and cheaper than RFLP, should be useful in the analysis of genomic variation within pathogen populations. Thus, as part of this study and to aid in future studies requiring analysis of large numbers of samples, we compared the efficiency of RFLP and rep-PCR for measuring genomic diversity.

In the second part of the study, we used rep-PCR and RFLP to evaluate the genomic variation and distribution of X. oryzae pv. oryzae strains obtained after systematic sampling of a single field. This field, which had been planted for at least 3 years (six consecutive crops) to rice cv. Sinandomeng, was selected because a previous survey in the area indicated that more than one race of X. oryzae pv. oryzae was present. Also, cv. Sinandomeng is susceptible to the races of X. oryzae pv. oryzae indigenous to the area, and thus, major genes would not be expected to have exerted selection pressure on the pathogen populations. At the time of our sampling, the field was heavily blighted (estimated 100% disease incidence).

MATERIALS AND METHODS

Bacterial isolates and collection of infected leaf samples. Thirty strains of X. oryzae pv. oryzae were selected for comparison between RFLP and rep-PCR. These strains represented four distinct X. oryzae pv. oryzae genetic lineages identified by RFLP, utilizing insertion element (IS) probes IS*1112* and IS*1113* (16,31). All X. oryzae pv. oryzae isolates were maintained in 15% glycerol at –80°C or were lyophilized for long-term storage.

Severely blighted leaves of cv. Sinandomeng, which is susceptible to indigenous races of X. oryzae pv. oryzae, were collected on 7 September 1993 from a farmer's field located at Calauan, Laguna, the Philippines. The 37 × 53-m field plot (0.2 ha) was divided equally into 25 blocks of approximately 7.4 × 10.6 m (37 × 53 hills) each. Blocks located diagonally across the field were sampled more intensively, i.e., within each block along the diagonal one infected leaf from eight sampling areas, each having 4 × 4 hills spaced equally within the block, were collected, whereas in the middle sampling area five infected leaves were sampled (n = 13 samples per block). At blocks not located on the diagonals, five infected leaves were obtained from the middle 4 × 4 hills. Based on the equation $N = \log(1 - P)/\log(1 - F)$, where N is the number of samples ($N = 197$) required to detect a given haplotype present at a frequency (F) at least once with a probability (P) of 95%, this sampling strategy should detect variants present at a frequency of 1.5% in the population (18,22).

The leaves were held at –20°C until bacterial isolations were conducted. Bacteria were isolated from the advancing lesion as described previously (24). The bacterial suspension was streaked on modified Wakimoto's medium (14), and a single colony was selected after 96 h of incubation at 28°C. Isolates were maintained in the same medium at 4°C for routine use or were stored in 15% glycerol at –80°C. Although 197 leaves were sampled, only 155

yielded X. oryzae pv. oryzae isolates, because isolation was attempted only once for each leaf and some leaves also contained the related pathogen, X. oryzae pv. oryzicola.

RFLP analysis. Genomic DNA from each strain was extracted by a modification of the procedure of Murray and Thompson (27) or by the lysozyme-sodium dodecyl sulfate lysis method (34) as modified by Leach et al. (17). Bacterial DNA (~5 to 10 µg) was digested to completion with *Eco*RI (2 units/µg of DNA) at 37°C for 2 h as recommended by the enzyme manufacturer (Bethesda Research Labs, Gaithersburg, MD, or Boehringer Mannheim, Far East, Singapore). DNA fragments were separated by electrophoresis in 0.7% agarose in 0.5× TBE buffer (89 mM Tris base, pH 7.8; 89 mM boric acid; 2 mM EDTA, pH 8.0) (37) at 10 V for 69 to 72 h and blotted onto Hybond N or Hybond N+ (Amersham, Far East, Hong Kong) by alkali transfer as described by the manufacturer.

Probes IS*1112* (17,46) and IS*1113* (31) were selected to detect polymorphism by RFLP because they gave the most robust grouping of X. oryzae pv. oryzae when used to determine genetic relationships among 155 strains of X. oryzae pv. oryzae (31). Both probes were used to detect polymorphism among the 30 selected X. oryzae pv. oryzae strains; only IS*1113* was used to analyze the field isolates from cv. Sinandomeng. The probes were labeled by random priming with digoxigenin-labeled deoxyuridine triphosphate (Boehringer Mannheim), and hybridized bands were detected directly on the blot by a chromogenic method with nitro blue tetrazolium as chloride and 5-bromo-4-chloro-3-indolylphosphate substrates for alkaline phosphatase (Boehringer Mannheim).

rep-PCR analysis. Genomic DNAs from the selected X. oryzae pv. oryzae set were quantified by spectrofluorimetry with a model TKO-100 minifluorometer (Hoefer Scientific Instruments, San Francisco). Genomic DNA concentrations from field samples were estimated in an agarose gel by comparing them to reference X. oryzae pv. oryzae DNA adjusted to 50 ng/µl.

The primers used for PCR amplification were BOX (BOXA1R [5′-CTACGGCAAGGCGACGCTGACG-3′]), ERIC (ERIC1R [5′-ATGTAAGCTCCTGGGGATTCAC-3′] and ERIC2 [5′-AAG-TAAGTGACTGGGGTGAGCG-3′]), and REP (REP1R-I [5′-IIII-CGICGICATCIGGC-3′] and REP2-I [5′-ICGICTTATCIGGCCT-AC-3′]) (4,19,20,43). Collectively, the PCR protocols with these primers are referred to as rep-PCR. The PCR conditions were as described previously (4,41), except the final dNTP concentration was adjusted to 625 µM. PCR amplifications were performed in an automated thermal cycler (Perkin-Elmer DNA thermal cycler [Foster City, CA] or MJ Research, Inc., Watertown, MA). The initial denaturation at 95°C for 7 min was followed by 30 (for BOX primer) or 35 cycles (for ERIC and REP primers) that included denaturation at 94°C for 1 min, annealing at 53, 52, or 44°C for 1 min with BOX, ERIC, and REP primers, respectively, and extension at 65°C for 8 min. The final extension cycle was at 65°C for 15 min followed by final incubation at 4°C. A 10-µl portion of each amplified PCR product was separated at 4°C on a gel mixture of 0.75% agarose and 0.75% Synergel (Diversified Biotech, Inc., Boston) in 0.75× TAE (4 mM Tris-acetate, 2 mM EDTA) or 0.5× TBE buffer, stained with ethidium bromide, and photographed on a UV transilluminator. Throughout the study, PCR amplifications with DNA or cell lysates from randomly selected strains were done 2 to 3 times for confirmation of banding patterns. With field samples, only ERIC and REP primers were used for fingerprinting, because the BOX primer did not provide useful additional information for differentiating closely related strains of X. oryzae pv. oryzae.

Virulence analysis. Field isolates (a total of 32: 19 from lineage C and 13 from lineage B), representing nine distinct banding patterns based on combined data from RFLP and rep-PCR analyses, were tested for virulence on the bacterial blight differential cultivars and near-isolines containing specific resistance genes: IR24 (*Xa*18), IRBB4 (*Xa*4), IR1545 (*xa*5), IRBB5 (*xa*5), CAS209
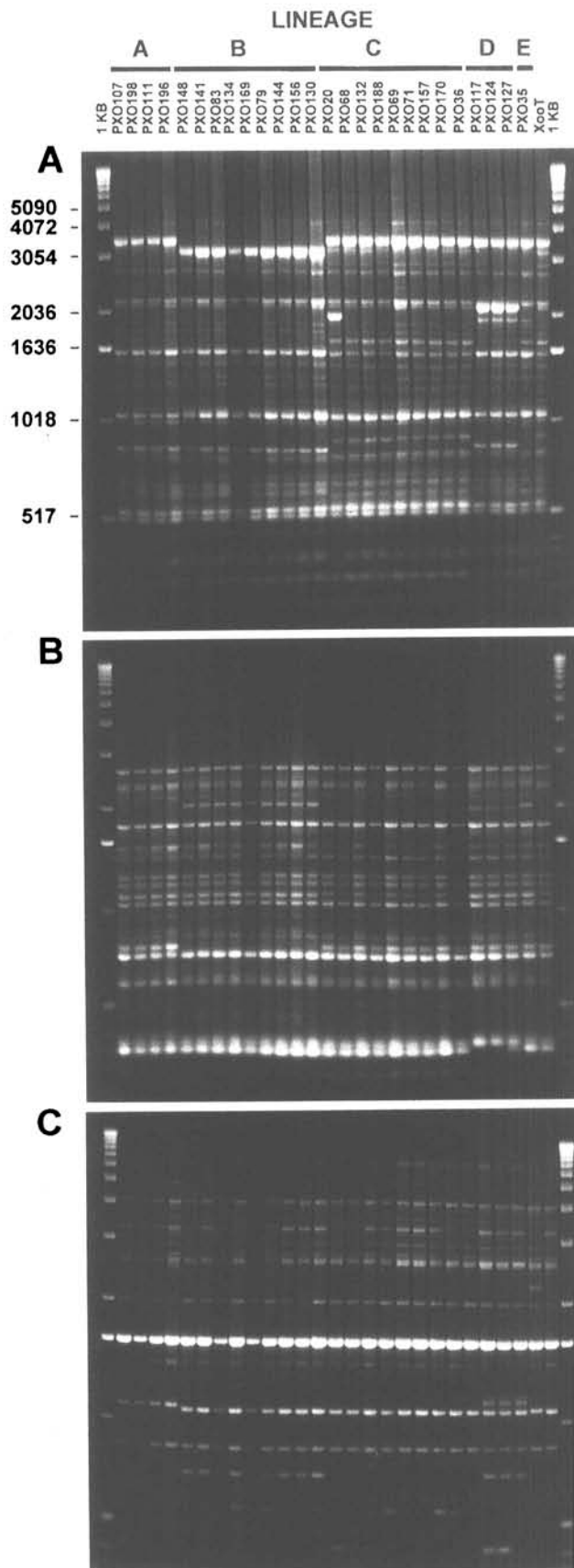
Fig. 1. Agarose gel showing representative patterns of each Philippine lineage of *Xanthomonas oryzae* pv. *oryzae* generated by repetitive sequence-based polymerase chain reaction (rep-PCR) with **A,** enterobacterial repetitive intergenic consensus (ERIC); **B,** BOX; and **C,** repetitive extragenic palindromic (REP) primers. After amplification by PCR, the DNA fragments were separated in agarose gels and visualized by staining with ethidium bromide. Strain designations are indicated above each lane. Numbers to the left indicate molecular weight in base pairs.

(Xa10), IRBB10 (Xa10), IRBB7 (Xa7), IRBB13t (Xa13), and DV85 (*xa5* and *Xa7*) (23,24,26,32). Five pregerminated seeds were grown in plastic pots arranged in plastic flats in a greenhouse as described by Adhikari et al. (1). Bacterial inoculum was prepared and inoculated in fully expanded leaves following the procedure of Mew and Vera Cruz (24). Lesion length from five leaves at similar positions from each main or secondary tiller were measured at 14 and 21 days after inoculation. Lesions ranging from 0 to 3 cm were rated as resistant, and those longer than 3 cm were considered susceptible.

Data analysis. *Multiple correspondence and cluster analyses.* Fingerprints generated for each strain were compared visually and grouped according to unique banding patterns, based on restriction fragment profile alone or rep-PCR amplification alone. To detect diversity among *X. oryzae* pv. *oryzae* strains from a single field, the composite data, i.e., the combined patterns from restriction fragment profiles and rep-PCR amplifications, were used. Each unique banding pattern generated by a probe or primer set alone or as a composite for the two techniques was regarded as a haplotype. For each technique, representative strains of a haplotype were analyzed in the same gel to confirm band positions. Binary data, coded as 1 for presence or 0 for absence of band, were scored at each position along a lane. For the 30 *X. oryzae* pv. *oryzae* strains, a total of 79 and 38 band positions were obtained for IS*1112* and IS*1113* haplotypes, respectively, whereas the composite rep-PCR (REP and ERIC) data gave a total of 42 positions. For the field samples, a total of 18 band positions were distinguished in IS*1113* haplotypes, whereas the composite rep-PCR haplotypes gave a total of 37 positions. The combined RFLP and rep-PCR data for the field analysis represented the total band positions, 55, generated for the two techniques.

A pairwise comparison of strains was generated by NTSYS-pc (version 1.80, Exeter Biological Software, Setauket, NY) with Jaccard's coefficient of similarity (38). The resulting data were converted to coefficients of dissimilarity (1 − Jaccard coefficient) and were used for multiple correspondence analysis (version 6, SAS/IML Software: Usage and Reference; SAS/STAT User's Guide, SAS Institute, Inc., Cary, NC), which determined the positions of the strains on a three-dimensional graph (1) that was drawn by the program JMP (version 3.1, SAS). The number of clusters based on a consensus among three clustering statistics (local peaks of the cubic clustering criterion and pseudo-F statistic combined with a small value of the pseudo-t^2 statistic and a larger t^2 for the next cluster fusion) were assigned by the average linkage method (SAS). The genetic distances among clusters were obtained to determine the relative distance of the clusters from each other. The cluster consistency index (CCI), which represents percent matches for each cluster, was obtained with 1,000 resampling cycles (6; D. Z. Skinner, M. C. Duque, C. M. Vera Cruz, and S. Kelemu, *unpublished data*). For field samples, a phenogram generated by the unweighted pair group arithmetic mean (UPGMA) (NTSYS-pc, version 1.80) was used to represent the relationships among strains. Bootstrap analysis (8,9) was performed with 1,000 iterations by the program Winboot (31,45) to determine the reproducibility of the phenograms.

Genetic diversity and analysis of molecular variance of field samples. Frequencies of haplotypes were tabulated for each block in the field and for the total population. The genetic diversity of *X. oryzae* pv. *oryzae* subpopulations from each block and of the total field population was estimated by Nei's haplotypic diversity index (30,29). The genetic diversity for the subpopulation from each block (H_B) and of the entire field population (H_T) were estimated based on the equation: $H = [n/(n − 1)][1 − \Sigma X_i^2]$, where X_i is the frequency of the ith haplotype at each block or the frequency of the ith haplotype in the entire population and n is the number of haplotypes from each block or total number of haplotypes for the entire population.

To determine whether the subpopulations were more diverse than the entire field population, the standard error (σ) was estimated

from the sampling variance for each subpopulation and total population. The sampling variance of H was modified for haploid organisms as

$$V_S(h_k) = 2/\{n[2(n-1)]\} \{\Sigma p_i^2 - (\Sigma p_i^2)^2 + 2(n-1) [\Sigma p_i^3 - (\Sigma p_i^2)^2]\}$$

where p_i and n correspond to X_i and n, respectively (28). Significance was declared if $H_B - 2\sigma$ was greater than H_T. Using haplotypic diversity indices from all blocks, the ability of rep-PCR and RFLP to detect diversity in this field plot was compared. The Wilcoxon signed-rank test (SAS) was used to analyze the magnitude of the rank of differences between pairs of haplotypic diversity indices from 25 blocks. The differentiation in the distribution of composite IS*1113*–rep-PCR haplotypes was determined by chi-square (12) and Fisher's exact tests by the program StatXact (version 1.00, Cytel Software Corp., Cambridge, MA).

RESULTS

Comparison of RFLP and rep-PCR analyses on selected *X. oryzae* pv. *oryzae* strains. Patterns generated by PCR with ERIC, BOX, and REP primers were compared among *X. oryzae* pv. *oryzae* strains from the Philippines (Fig. 1). Amplification with the BOX primer (Fig. 1B) showed the least variation between strains of *X. oryzae* pv. *oryzae* within a lineage, defined previously (31). For example, amplification with the BOX primer yielded only one haplotype for the 12 strains of lineage C, whereas the ERIC and REP primers yielded five and seven haplotypes, respectively (patterns for 9 of the 12 strains are shown in Figure 1). UPGMA with data from the ERIC primers alone resulted in the most consistent grouping of strains into lineages (data not shown), and ERIC patterns defining lineage were the easiest to recognize visually (Fig. 1A). For subsequent analysis, we combined data generated from the ERIC and REP primers because the combination of the two sets of primers differentiated lineages and distinguished more strains within a lineage (detected diversity) than either primer set alone.

Clusters generated from RFLP patterns after hybridization of digested genomic DNA with *X. oryzae* pv. *oryzae* IS elements were compared to clusters generated from rep-PCR patterns for the 30 *X. oryzae* pv. *oryzae* strains. Strains were assigned to clusters by the average linkage method after determining consensus among three clustering statistics ([1] SAS). On this basis, the strains were grouped initially into four clusters with similar members by both RFLP and rep-PCR. However, in both cases, the clusters containing one race 1 strain (PXO35) were unstable due to the high genetic distance within those clusters. For example, resampling analysis with 1,000 iterations indicated CCI for clusters A through D of 0.78, 0.50, 0.95, and 0.54 for RFLP data and 0.61, 0.32, 0.67, and 0.23 for rep-PCR data, respectively ([6] D. Z. Skinner, M. C. Duque, C. M. Vera Cruz, and S. Kelemu, *unpublished data*). Examination of the cooccurrence matrix revealed that isolate PXO35 was assigned inconsistently to clusters B and D in both cases. After manual separation of PXO35 into its own cluster, the remaining 29 isolates formed four clusters with CCI values of 0.78 or higher for RFLP and 0.50 or higher for rep-PCR. Thus, five clusters are reported for each technique (Fig. 2). The five clusters of strains derived by both techniques represent lineages that conformed to the lineages of *X. oryzae* pv. *oryzae* previously defined by Nelson et al. (31).

Comparison of RFLP and rep-PCR analyses on field isolates. RFLP with the IS*1113* probe and rep-PCR with the ERIC and REP primers were used to determine the level of diversity in a collection of *X. oryzae* pv. *oryzae* from a single farmer's field planted to rice cv. Sinandomeng. The 155 *X. oryzae* pv. *oryzae* strains collected from the field were distinctly separated into two lineages, with only 20% similarity in RFLP analysis (Fig. 3A) and with 55% similarity in rep-PCR (Fig. 3B). The two lineages discerned by rep-PCR contained the same strains as the lineages defined by RFLP analysis. Lineage C as detected by RFLP with

IS*1113* consisted of one haplotype (designated C-002), whereas lineage B was separated into three haplotypes (designated B-002, B-004, and B-005; Fig. 3A). In the combined RFLP and rep-PCR data, lineage C was separated into three haplotypes, and lineage B was separated into four haplotypes (Fig. 3C). Bootstrap values (100% for each lineage) indicated highly reliable groups.

Haplotypic diversity of the field population. RFLP and rep-PCR analyses detected the same levels of haplotypic diversity throughout the field (Wilcoxon signed-rank test $P = 0.69$; Table 1). Therefore, either method can be used with essentially equal efficiency. For further analysis of field data in this study, we used a combination of RFLP and rep-PCR data. The haplotypic diversities of 10 of the 25 blocks (H_B) were significantly greater than the total haplotypic diversity of the entire population ($H_T = 0.43$; Table 1). This suggests that small patches of the field contained
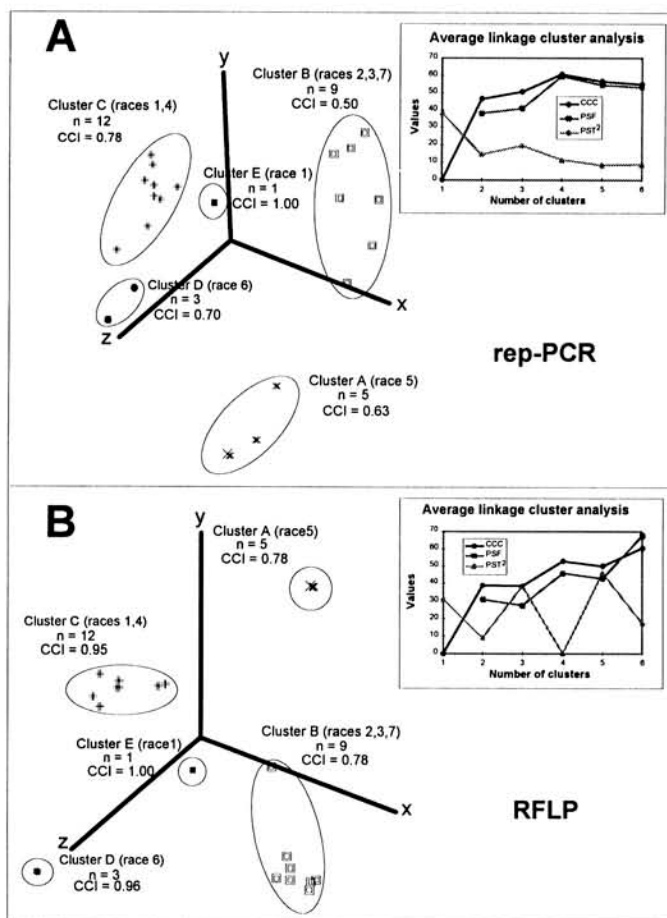

Fig. 2. Comparison of *Xanthomonas oryzae* pv. *oryzae* clusters derived by **A,** repetitive sequence-based polymerase chain reaction (rep-PCR) with repetitive extragenic palindromic (REP) and enterobacterial repetitive intergenic consensus (ERIC) primers and **B,** restriction fragment length polymorphism (RFLP) analyses with probes IS*1112* and IS*1113*. Positions on the three-dimensional graph were determined by multiple correspondence analysis. Clusters originally were assigned by the average linkage method (inset). The consensus among three clustering statistics, i.e., local peaks of the cubic clustering criterion (CCC) and pseudo-F (PSF) statistic combined with a small value of the pseudo-t^2 (PST^2) statistic and a larger t^2 for the next cluster fusion, indicated four clusters. The final adjustments to five clusters were made based on cluster stability, as indicated by cluster consistency index (CCI) analysis ([6]; D. Z. Skinner, M. jC. Duque, C. M. Vera Cruz, and S. Kelemu, *unpublished data*). For the 30 *X. oryzae* pv. *oryzae* strains, the first three dimensions (x, y, and z) accounted for **A,** 69.3, 16.9, and 7.5% (total = 93.7%) of the variation for rep-PCR analysis and **B,** 62.7, 17.6, and 9.2% (total = 89.5%) for RFLP analysis. Identical symbols within **A and B** indicate positions of strains within a cluster as grouped after analysis by both RFLP and rep-PCR. The total number of strains within each cluster is indicated by n; patterns for some strains are identical and, thus, are represented by a single point.

more diverse *X. oryzae* pv. *oryzae* populations than did the rest of the field.

The distribution of seven composite haplotypes from 25 blocks is shown in Figure 4. Haplotype C-1 (74.8% of the total strains) was the predominant population in the field and was found in all blocks, except in block IV-c (Fig. 4). The frequency of the six other haplotypes or subpopulations was low, ranging from 1.3 to 10.3%. Haplotype B-1 (10.3%) was found in 11 of 25 blocks, and the frequency ranged from 10 to 40% among the 11 blocks. Haplotype B-3 was present in 9 of 25 blocks. Haplotypes C-2, C-3, and B-2 were sparsely distributed across the 25 blocks. Haplotype B-4 was found only in the eastern blocks (I-e, II-d, IV-e, and V-e) and only at a very low frequency (2.6%) (Fig. 4).

Genetic differentiation in a single field as revealed by RFLP and rep-PCR. Differences in the distribution of the haplotypes among the 25 blocks were evaluated by chi-square and Fisher's exact statistics. In comparisons of the distribution of all haplotypes in all blocks, no significant differences were observed (data not shown), indicating the distribution of the seven haplotypes across the entire field was random.

In this single field (37 × 53 m), variance in haplotype frequencies was partitioned into variation among blocks (7.4 × 10.6 m), among groups within blocks (e.g., squares versus circles; Fig. 4),

and within isolates in each sampling area (e.g., isolates in a circle or a set of squares). About 83.8% of the variation was attributed to isolates within each sampling area (Table 2). None of the components was a source of significant variation, however, probably due to the low number of haplotypes detected.

Race structure of the field population. To determine the race structure of the field population from cv. Sinandomeng, a subset (32 isolates) of the population representing each of the seven composite haplotypes from the two lineages (19 isolates from lineage C and 13 from lineage B) was inoculated to a bacterial blight differential set of rice cultivars and near-isolines IRBB13 and IRBB7. IRBB7 is highly resistant to races 2 and 3 and allows distinction of race 9 from race 3 (M. R. Finckh V. M. Luman-ag, R. J. Nelson, *unpublished data*). Under our greenhouse conditions, race 3 strains induced shorter lesions (0 to 2 cm) on IRBB7 than were caused by race 9 strains. Race 9 caused lesions ranging from 5.8 to 11.6 cm. Three races, 2, 3, and 9, were detected in the field samples (Fig. 3). Haplotypes of lineage C consisted of both race 3 (11 of 19 isolates) and race 9 (8 of 19 isolates) (Fig. 3C). All four lineage B haplotypes contained only race 2 strains (Fig. 3C).

DISCUSSION

The efficiency of two techniques for measuring genomic diversity in *X. oryzae* pv. *oryzae* was compared: RFLP analysis, which has been used in previous studies of *X. oryzae* pv. *oryzae* diversity (1,2,16,31), and rep-PCR (4,19,20,43,44). When a selected set of 30 strains representing the extremes in diversity previously de-
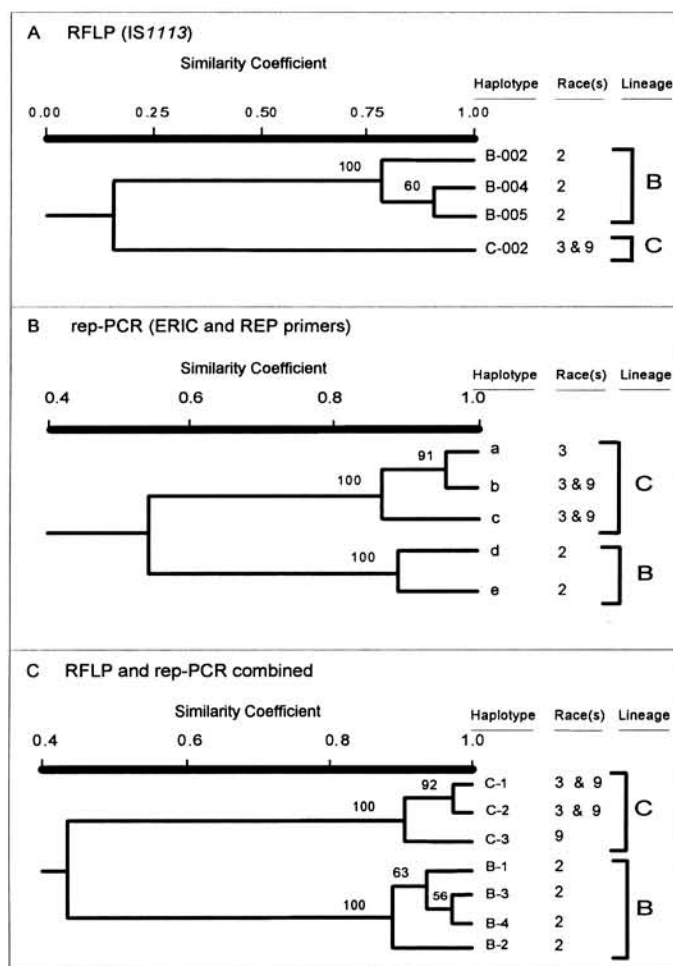
TABLE 1. Haplotypic diversity analysis of a *Xanthomonas oryzae* pv. *oryzae* population from 25 blocks in a single field planted to rice cv. Sinandomeng by restriction fragment length polymorphism (RFLP) and repetitive sequence-based polymerase chain reaction (rep-PCR)[a]

Block	RFLP (IS1113)	rep-PCR	Composite[b]
I-a	0	0	0
I-b	0.40	0.40	0.40
I-c	0	0	0
I-d	0.40	0.40	0.40
I-e	0.64	0.64	0.64*
II-a	0.60	0.60	0.60*
II-b	0.53	0.48	0.53
II-c	0.67	0.67	0.67*
II-d	0.25	0.25	0.25
II-e	0.50	0.50	0.50
III-a	0.60	0.60	0.60*
III-b	0	0	0
III-c	0.20	0.38	0.38
III-d	0.67	0.67	0.67*
III-e	0	0	0
IV-a	0	0	0
IV-b	0.35	0.47	0.49
IV-c	1	0.67	1*
IV-d	0.17	0.17	0.17
IV-e	0.70	0.60	0.70*
V-a	0.20	0.20	0.20
V-b	0	0	0
V-c	0.70	0.70	0.70*
V-d	0.67	0.67	0.67*
V-e	0.69	0.67	0.81*
H_S[c]	0.39	0.38	0.41
H_T[d]	0.38	0.40	0.43

[a] Haplotypic diversities were calculated as described in text. Levels of diversity detected by rep-PCR and RFLP were compared by the Wilcoxon signed-rank test with the haplotypic diversities of each block (H_B). There were no differences between techniques at $P = 0.69$.

[b] Data are the composites of both RFLP and rep-PCR patterns. Diversity in the individual blocks (H_B) for the composite data was considered significantly different (*) from the total diversity (H_T) if $H_B − 2\sigma$ was greater than H_T. σ = standard error.

[c] H_S is the average of the estimated haplotype diversities in the blocks.

[d] H_T is the estimate of the haplotype diversity of the total population.

Fig. 3. A, Restriction fragment length polymorphism (RFLP); **B,** repetitive sequence-based polymerase chain reaction (rep-PCR); and **C,** combined RFLP–rep-PCR analyses of *Xanthomonas oryzae* pv. *oryzae* field populations distinguished two lineages, B and C, that contain strains of race 2 and races 3 and 9, respectively. Similarity between pairs of strains was determined by Jaccard's method, and clusters were generated by the unweighted pair group arithmetic mean by NTSYS-pc, version 1.80 (Exeter Biological Software). The numbers at the base of each node are bootstrap values analyzed by the program Winboot (45) that represent the statistical reliability of the nodes. Nomenclature of RFLP-derived haplotypes is based on Nelson et al. (31).

tected in the Philippines (i.e., four distinct RFLP-derived lineages [16,31]) was used for comparison, both RFLP and rep-PCR consistently grouped the same subsets of strains together into genetic groups or lineages (Fig. 2). Similarly, phylogenetic groups for strains of *Bradyrhizobium japonicum* obtained by RFLP also were correlated with groups defined by REP- and ERIC-PCR (5). Both techniques detected distinct haplotypes or subgroups within each lineage, indicating they would be useful in measuring and monitoring spatial distribution of pathogen haplotypes.

For analysis of large populations of bacteria, rep-PCR has several distinct advantages over RFLP. The technique is simpler technically, less expensive, and less time-consuming. In both cases, the resolving power, i.e., the number of haplotypes detected, is dependent on the probe or primers used. Nelson et al. (31) demonstrated that although the lineages revealed were consistent, the use of different probes with RFLP revealed different levels of population diversity. In this study, BOX primers used in rep-PCR detected the least polymorphism in *X. oryzae* pv. *oryzae* populations, and REP primers detected the most. Thus, as Louws et al. (20) found for *X. campestris* pv. *vesicatoria*, each primer set gave unique information, and the choice of primer combinations depends on the level of separation or grouping desired. For analysis of the field population in the Philippines, we used data from rep-PCR with two primer sets (ERIC and REP) and RFLP with one probe (IS*1113*) because the combination of results allowed detection of a higher level diversity within the field than either primer/probe set or technique alone.

The variation and microgeographic distribution of *X. oryzae* pv. *oryzae* in a rice field in the Philippines was estimated by comparison of the genomic structure of strains systematically sampled at a single time. Based on genotypic analysis, the field population grouped into two highly robust clusters (bootstrap value = 100%) that correspond to two of the four genetic lineages previously described in the Philippines (2,16,31). Of the two, strains in lineage C were detected more frequently than strains in lineage B (77 versus 23%), and of the lineage C strains, 97% were of one haplotype (C-1). Lineage B was composed only of race 2, whereas lineage C contained strains of races 3 and 9. The observation that the relative proportion of the lineage C (races 3 and 9) population is higher than that of lineage B (race 2) in this field in southern Luzon (Laguna) is consistent with the findings of Ardales et al. (2) for populations in three provinces in central Luzon sampled in 1990 and 1991. The two studies together confirm that there has been a shift in the race structure of the population in Luzon from a prevalent population of race 2 (lineage B, originally designated as cluster I [16]), as observed by Leach et al. (16), Mew et al. (25), and Vera Cruz and Mew (42) during 1985 to 1988, to a prevalent population of a lineage (lineage C) that contains races 3 and 9.

This is intriguing because within this field, which has been planted to cv. Sinandomeng for more than 3 years, there has been no apparent host selection for either lineage of the bacterial blight pathogen. Because there was a predominance of lineage C (races 3 and 9) in the field, we speculate that this pathogen lineage is better adapted to cv. Sinandomeng in the absence of selection by major resistance genes. In a field study in which a race 1 (lineage C) strain and a race 2 (lineage B) strain were inoculated in a cultivar without *Xa-4* (susceptible to both), Roberts (36) observed that the race 1 strain was detected more frequently than the race 2 strain over the growing season. Thus, lineage C may be better adapted than lineage B in the absence of resistance gene selection under similar environmental conditions (the field sites were within 15 km of each other in Laguna province).

The total haplotypic diversity in the field planted to the improved cv. Sinandomeng was $H_T = 0.43 \pm 0.07$. In the study by Ardales et al. (2), genetic diversity values from pooled pathogen collections across ecosystems were $H_{TV} = 0.77$ for fields planted to traditional cultivars and $H_{IV} = 0.63$ for those planted to improved cultivars at 5 and 10 sites, respectively. They tentatively concluded that host diversity did not affect pathogen diversity. The haplotypic diversity we report for a single field is lower than the average reported for improved cultivars in 10 sites (2), even though the combination of rep-PCR primers and the RFLP-probe (IS*1113*) used in our study would have detected more diversity overall for both lineages C (the prevalent population for all sites) and B (Fig. 2) than the RFLP-probe (IS*1113*) and restriction enzyme digestion (*Pst*I) combination used by Ardales et al. (2). These findings emphasize that diversity values depend on the tools used for measurement and that the prevalent pathogen population should be considered in the selection of primers/probes for future studies to evaluate the effects of host diversity on pathogen diversity. Studies in progress will give insights into the effects of continuous cultivation of a host genotype on pathogen diversity.

Based on Fisher's exact test, strains of lineages B (race 2) and C (races 3 and 9) were distributed randomly in the field. However, because visual assessment of the pattern of distribution in the field suggested there was nonrandom distribution of some minor haplotypes, we also compared the distribution of individual haplotypes against all other haplotypes in blocks situated along (i) rows, (ii) columns, and (iii) across the diagonal blocks of the entire field. In general, these groupings, which increased the number of isolates in the comparisons, also did not show significant differences in haplotype distribution. Except for haplotype B-4, which was found only in the eastern blocks of the field (Fig. 4, open boxes), all haplotypes exhibited random distribution after these combined analyses. Our sampling strategy was designed to detect one isolate of a haplotype at a frequency of 1.5% in the population ($P = 0.95$); a
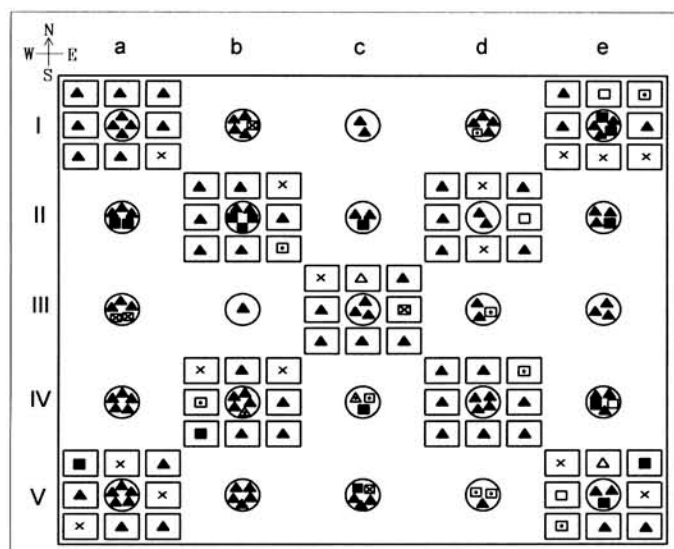

Fig. 4. Distribution of two lineages of *Xanthomonas oryzae* pv. *oryzae* by composite insertion element IS*1113* and enterobacterial repetitive intergenic consensus–repetitive extragenic palindromic haplotypes among the 25 blocks in a farmer's field planted to rice cv. Sinandomeng. Lineage B: ■ = haplotype B-1; ⊠ = haplotype B-2; ⊡ = haplotype B-3; and □ = haplotype B-4. Lineage C: ▲ = haplotype C-1; △ = haplotype C-2; ⟁ = haplotype C-3; and x = no isolate recovered.

TABLE 2. Hierarchical analysis of a *Xanthomonas oryzae* pv. *oryzae* population in a single field planted to susceptible rice cv. Sinandomeng

| Variance component | Observed partition | | P^a |
	Variance	% Total	
Among blocks	0.184	4.42	0.8515
Among populations/block	0.142	11.78	0.5149
Within population	2.803	83.80	0.5545

[a] Probability of having a more extreme variance component than the observed values by chance alone; none of the values were significant.

more intensive sampling might have shown random distribution of B-4 (frequency = 2.6%).

Ardales et al. (2) examined 39 fields across a transect that included many cultivars and diverse environments on the island of Luzon, the Philippines, and found that *X. oryzae* pv. *oryzae* was nonrandomly distributed within the majority of the fields. Information on the number of years a given cultivar had been planted within a field were not presented. It is possible that the continuous cropping of cv. Sinandomeng in our field contributed to the random distribution of the pathogen. Alternatively, the random distribution pattern in this field might indicate well-dispersed initial inoculum, perhaps through irrigation waters. In the tropics, high bacterial populations are maintained year round in irrigation waters (33). It is important to note that the sampling in this study was performed only once and this was late in the season. Thus, conclusions about mechanisms of dispersal or the influence of host are not possible. However, using rep-PCR, which allows rapid evaluation of sufficient numbers of samples to detect low-frequency variants, hypotheses concerning the epidemiology of this disease as well as the impact of host selection can be tested. Information on the microgeographic distribution of pathogens in fields can provide useful information for guiding deployment strategies and designing studies of the epidemiology of the disease.

LITERATURE CITED

1. Adhikari, T. B., Vera Cruz, C. M., Zhang, Q., Nelson, R. J., Skinner, D. Z., Mew, T. W., and Leach, J. E. 1995. Genetic diversity of *Xanthomonas oryzae* pv. *oryzae* in Asia. Appl. Environ. Microbiol. 61:966-971.
2. Ardales, E. Y., Leung, H., Vera Cruz, C. M., Mew, T. W., Leach, J. E., and Nelson, R. J. 1996. A hierarchical analysis of spatial variation of the rice bacterial blight pathogen across diverse ecosystems in the Philippines. Phytopathology 86:241-252.
3. Chatelut, M., Dournes, J. L., Chabanon, G., and Marty, N. 1995. Epidemiological typing of *Stenotrophomonas* (*Xanthomonas*) *maltophilia* by PCR. J. Clin. Microbiol. 33:912-914.
4. de Bruijn, F. J. 1992. Use of repetitive (repetitive extragenic palindromic and enterobacterial repetitive intergenic consensus) sequences and the polymerase chain reaction to fingerprint the genomes of *Rhizobium meliloti* isolates and other soil bacteria. Appl. Environ. Microbiol. 58:2180-2187.
5. Dudd, A. K., Schneider, M., Sadowsky, M. J., and de Bruijn, F. J. 1993. Use of repetitive sequences and the polymerase chain reaction to classify genetically related *Bradyrhizobium japonicum* serocluster 123 strains. Appl. Environ. Microbiol. 59:1702-1708.
6. Efron, B., and Tibshirani, R. 1993. An introduction to the bootstrap. Chapman and Hall, London.
7. Ezuka, A., and Sakaguchi, S. 1978. Host-parasite relationship in bacterial leaf blight of rice caused by *Xanthomonas oryzae*. Rev. Plant Prot. Res. 11:93-118.
8. Felsenstein, J. 1985. Confidence limits on phylogenies: An approach using the bootstrap. Evolution 39:783-791.
9. Felsenstein, J. 1988. Phylogenies from molecular sequences: Interference and reliability. Annu. Rev. Genet. 22:521-565.
10. Georghiou, P. R., Doggett, A. M., Kielhofner, M. A., Stout, J. E., Watson, D. A., Lupski, J. R., and Hamill, R. J. 1994. Molecular fingerprinting of *Legionella* species by repetitive element PCR. J. Clin. Microbiol. 32:2989-2994.
11. Higgins, C. F., Ames, G. F.-L., Barnes, W. M., Clement, J. M., and Hofnung, M. 1982. A novel intercistronic regulatory element of prokaryotic operons. Nature (Lond.) 298:760-762.
12. Hudson, R. R., Boos, D. D., and Kaplan, N. L. 1992. A statistical test for detecting geographic subdivision. Mol. Biol. Evol. 9:138-151.
13. Hulton, C. S. J., Higgins, C. F., and Sharp, P. M. 1991. ERIC sequences: A novel family of repetitive elements in the genomes of *Escherichia coli*, *Salmonella typhimurium* and other enterobacteria. Mol. Microbiol. 5:825-834.
14. Karganilla, A. D., and Natural, M. P. 1973. A comparative study of culture media for *Xanthomonas oryzae*. Philipp. Agric. 57:141-152.
15. Leach, J. E., Leung, H., Nelson, R. J., and Mew, T. W. 1995. Population biology of *Xanthomonas oryzae* pv. *oryzae* and approaches to its control. Curr. Opin. Biotechnol. 6:298-304.
16. Leach, J. E., Rhoads, M. L., Vera Cruz, C. M., White, F. F., Mew, T. W., and Leung, H. 1992. Assessment of genetic diversity and population structure of *Xanthomonas oryzae* pv. *oryzae* with a repetitive DNA ele-

ment. Appl. Environ. Microbiol. 58:2188-2195.
17. Leach, J. E., White, F. F., Rhoads, M. L., and Leung, H. 1990. A repetitive DNA sequence differentiates *Xanthomonas campestris* pv. *oryzae* from other pathovars of *X. campestris*. Mol. Plant-Microbe Interact. 3:238-246.
18. Leung, H., Nelson, R. J., and Leach, J. E. 1993. Population structure of plant pathogenic fungi and bacteria. Adv. Plant Pathol. 10:157-205.
19. Louws, F. J., Fulbright, D. W., Stephens, C. T., and de Bruijn, F. J. 1994. Specific genomic fingerprints of phytopathogenic *Xanthomonas* and *Pseudomonas* pathovars and strains generated with repetitive sequences and PCR. Appl. Environ. Microbiol. 60:2286-2295.
20. Louws, F. J., Fulbright, D. W., Stephens, C. T., and de Bruijn, F. J. 1995. Differentiation of genomic structure by rep-PCR fingerprinting to rapidly classify *Xanthomonas campestris* pv. *vesicatoria*. Phytopathology 85:528-536.
21. Martin, B., Humbert, O., Camara, M., Guenzi, E., Walker, J., Mitchell, T., Andrew, P., Prudhomme, M., Alloing, G., Hakenbeck, R., Morrison, D. A., Boulnois, G. J., and Claverys, J.-P. 1992. A highly conserved repeated DNA element located in the chromosome of *Streptococcus pneumoniae*. Nucleic Acids Res. 20:3479-3483.
22. McDonald, B. A., and Martinez, J. P. 1990. DNA restriction fragment length polymorphisms among *Mycosphaerella graminicola* (anamorph *Septoria tritici*) isolates collected from a single wheat field. Phytopathology 80:1368-1373.
23. Mew, T. W. 1987. Current status and future prospects on research on bacterial blight of rice. Annu. Rev. Phytopathol. 25:359-382.
24. Mew, T. W., and Vera Cruz, C. M. 1979. Variability of *Xanthomonas oryzae*: Specificity in infection of rice differentials. Phytopathology 69:152-155.
25. Mew, T. W., Vera Cruz, C. M., and Medalla, E. S. 1992. Changes in the race frequency of *Xanthomonas oryzae* pv. *oryzae* in response to rice cultivars planted in the Philippines. Plant Dis. 76:1029-1032.
26. Mew, T. W., Vera Cruz, C. M., and Reyes, R. C. 1982. Interaction of *Xanthomonas campestris* pv. *oryzae* and a resistant rice cultivar. Phytopathology 72:786-789.
27. Murray, M. G., and Thompson, W. F. 1980. Rapid isolation of high molecular weight plant DNA. Nucleic Acids Res. 8:4321-4325.
28. Nei, M. 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics 89:583-590.
29. Nei, M. 1987. Molecular Evolutionary Genetics. Columbia University Press, New York.
30. Nei, M., and Tajima, F. 1981. DNA polymorphism detectable by restriction endonucleases. Genetics 97:145-163.
31. Nelson, R. J., Baraoidan, M. R., Vera Cruz, C. M., Yap, I. V., Leach, J. E., Mew, T. W., and Leung, H. 1994. Relationship between phylogeny and pathotype for the bacterial blight pathogen of rice. Appl. Environ. Microbiol. 60:3275-3283.
32. Ogawa, T., Tabien, R. E., Yamamoto, T., Busto, G. A., and Ikeda, R. 1990. Breeding for near-isogenic lines resistant to bacterial blight in rice. Rice Genet. Newsl. 7:10.
33. Ou, S. H. 1985. Rice Diseases. 2nd ed. Commonwealth Mycological Institute, Kew, England.
34. Owen, R. J., and Borman, P. 1987. A rapid biochemical method for purifying high molecular weight bacterial chromosomal DNA for restriction enzyme analysis. Nucleic Acids Res. 15:3631.
35. Reboli, A. C., Houston, E. D., Monteforte, J. S., Wood, C. A., and Hamill, R. J. 1995. Discrimination of epidemic and sporadic isolates of *Acinetobacter baumannii* by repetitive element PCR-mediated DNA fingerprinting. J. Clin. Microbiol. 32:2635-2640.
36. Roberts, P. D. 1991. Evaluation and application of monoclonal antibodies specific to *Xanthomonas oryzae* pv. *oryzae* as epidemiological tools. M.S. thesis. University of Hawaii, Honolulu.
37. Sambrook, J., Fritsch, E. F., and Maniatis, T. 1989. Molecular cloning: A Laboratory Manual. 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
38. Sneath, P. H. A., and Sokal, R. R. 1973. Numerical taxonomy. W.H. Freeman and Co., San Francisco.
39. Snelling, A. M., Gerner-Smidt, P., Hawkey, P. M., Hertiage, J., Parnell, P., Porter, C., Bodenham, A. R., and Inglis, T. 1996. Validation of use of whole-cell repetitive extragenic palindromic sequence-based PCR (REP-PCR) for typing strains belonging to the *Acinetobacter calcoaceticus-Acenitobacter baumannii* complex and application of the method to the investigation of a hospital outbreak. J. Clin. Microbiol. 34:1193-1202.
40. Swings, J., Van der Mooter, M., Vauterin, L., Hoste, B., Gillis, M., Mew, T. W., and Kersters, K. 1990. Reclassification of the causal agents of bacterial blight (*Xanthomonas campestris* pv. *oryzae*) and bacterial leaf streak (*Xanthomonas campestris* pv. *oryzicola*) of rice as pathovars of *Xanthomonas oryzae* (ex Ishiyama 1922) sp. nov., nom. ret. Int. J. Syst. Bacteriol. 40:302-311.

41. Vera Cruz, C. M., Halda-Alija, L., Louws, F. J., Skinner, D. Z., George, M. L., Nelson, R. J., de Bruijn, F. J., Rice, C. W., and Leach, J. E. 1995. Repetitive sequence-based polymerase chain reaction of *Xanthomonas oryzae* pv. *oryzae* and *Pseudomonas* species. Int. Rice Res. Note 20:23-24.

42. Vera Cruz, C. M., and Mew, T. W. 1989. How variable is *Xanthomonas campestris* pv. *oryzae*? Pages 153-166 in: Bacterial Blight of Rice. International Rice Research Institute, Manila, the Philippines.

43. Versalovic, J., Koeuth, T., and Lupski, J. R. 1991. Distribution of repetitive DNA sequences in eubacteria and application to fingerprinting of bacterial genomes. Nucleic Acids Res. 19:6823-6831.

44. Versalovic, J., Schneider, M., de Bruijn, F. J., and Lupski, J. R. 1994. Genomic fingerprinting of bacteria using repetitive sequence-based polymerase chain reaction. Methods Mol. Cell. Biol. 5:25-40.

45. Yap, I. V., and Nelson, R. J. 1996. WinBoot: A program for performing bootstrap analysis of binary data to determine the confidence limits of UPGMA-based dendrograms. IRRI Discussion Paper Series 14. International Rice Research Institute, Manila, the Philippines.

46. Yun, C. H. 1991. Molecular characterization of a repetitive element of *Xanthomonas oryzae* pv. *oryzae*. Ph.D. dissertation. Kansas State University, Manhattan.