Hierarchical Analysis of Spatial Variation of the Rice Bacterial Blight Pathogen Across Diverse Agroecosystems in the Philippines

E. Y. Ardales, H. Leung, C. M. Vera Cruz, T. W. Mew, J. E. Leach, and R. J. Nelson

First, fourth, and sixth authors: Division of Entomology and Plant Pathology, International Rice Research Institute, P.O. Box 933, 1099 Manila, Philippines; second author: Department of Plant Pathology, Washington State University, Pullman 99164-6430; third and fifth authors: Department of Plant Pathology, Throckmorton Plant Sciences Center, Kansas State University, Manhattan 66506-5502.

This work was supported in part by the Rockefeller Foundation's International Program on Rice Biotechnology.

We thank C. G. McLaren and J. F. San Agustin (Biometrics, International Rice Research Institute [IRRI]) for statistical consultations, I. V. Yap (Entomology and Plant Pathology Division, IRRI) for writing the computer program for data subsampling, and D. Christian (Department of Plant Pathology, Washington State University) for assisting in data analysis. We also thank E. Medalla and C. Bautista (Entomology and Plant Pathology Division, IRRI) for helping in virulence testing of the strains; N. C. Exconde, M. R. Baraoidan, A. A. Bordeos (Entomology and Plant Pathology Division, IRRI), and C. M. Ramirez (Soil and Water Sciences Division, IRRI) for technical assistance.

Accepted for publication 15 November 1995.

ABSTRACT

Ardales, E. Y., Leung, H., Vera Cruz, C. M., Mew, T. W., Leach, J. E., and Nelson, R. J. 1996. Hierarchical analysis of spatial variation of the rice bacterial blight pathogen across diverse agroecosystems in the Philippines. Phytopathology 86:241-252.

A hierarchical analysis of variation was conducted for the rice bacterial blight pathogen in the Philippines. More than 1,200 strains were collected systematically from 13 sites along a 310-km transect spanning an indigenous rice-growing area in the mountainous region and the improved irrigated agroecosystem in the lowlands of the island of Luzon. Restriction fragment length polymorphism analysis using the transposable element IS*1113* and restriction enzyme analysis using *Pst*I were utilized to assess the genetic diversity and phylogenetic structure of the pathogen population. Nineteen haplotypes (distinct DNA fingerprints), grouping into three robust clusters, were defined using IS*1113*. *Pst*I fingerprinting revealed greater diversity (46 unique haplotypes defined), but the phenogram derived from *Pst*I data was not robust. Hierarchical analysis of genetic variation showed population substructuring within and between sites. For the entire transect, a high degree of genetic differentiation among sites was estimated, indicating geographic differentiation between pathogen populations in different agroecosystems. Reiterative subsampling of the *Pst*I data set suggested that pathogen populations could be sampled less intensively by decreasing the number of samples collected per field without resulting in considerable loss of detected diversity. Nearly 80% of the haplotypes found by sampling 35 samples per field could be detected by sampling 18 samples per field. The results of this study have provided a thorough understanding of the population structure of the bacterial blight pathogen in the Philippines, which is one of the important prerequisites for designing effective and long-term strategies for resistance-gene deployment.

Additional keywords: geographic subdivision, population genetics, *Xanthomonas oryzae* pv. *oryzae*.

Knowledge of pathogen population structure, coupled with an understanding of the mechanisms that drive genetic changes in the pathogen population, is essential for formulating long-term disease-management strategies. Information on pathogen diversity can be used in identifying and characterizing resistant germ plasm. Information on the extent and distribution of pathogen variation is needed to design effective deployment strategies for resistant germ plasm (10).

Bacterial blight of rice, caused by *Xanthomonas oryzae* pv. *oryzae* (24), is managed primarily through the use of resistant cultivars. Many sources of resistance to bacterial blight have been identified, but few have been used in commercial rice production. The *Xa-4* gene for resistance has been widely used in tropical Asia; however, pathotypes overcoming the gene have become prevalent (14). To guide the selection and deployment of resistance genes, information on the population structure of the pathogen is needed. Population structure refers to the amount of diversity in a population, the genetic relationships among types, and the partitioning of diversity in time and space (10).

For *X. oryzae* pv. *oryzae*, diversity has been analyzed based on virulence typing (12,13,14) and DNA fingerprinting (9,19,20). Mew et al. (14) detected six virulence types ("pathotypes" or "races") of the pathogen among 800 strains taken from around the Philippines over a 17-year period. The use of repetitive DNA elements allowed the detection of a higher level of diversity and the assessment of evolutionary relationships among strains and pathotypes (9,19). Four main phylogenetic lineages have been detected, some of which consist of multiple pathotypes. Two additional pathotypes have been defined based on the use of DNA typing data and further virulence testing (19).

Pathotype surveys indicate temporal and geographic differentiation of the rice bacterial blight pathogen population in the Philippines. Prior to the wide-scale release of rice cultivars carrying the *Xa-4* gene for bacterial blight resistance, pathotype 1 dominated the *X. oryzae* pv. *oryzae* population in central Luzon. Later, pathotype 2 became predominant in the population (14). In the mountainous area of northern Luzon, the pathogen population has been dominated by pathotype 5. DNA fingerprinting suggests that the pathotype 2 population was derived from pathotype 5 rather than from pathotype 1 (19).

Corresponding author: R. J. Nelson; E-mail address: r.j.nelson@cgnet.com

Publication no. P-1996-0130-01R

These studies suggest that diversity was nonrandomly distributed among different regions of the Philippines. Because previous collections were not sampled intensively or systematically, however, they were not suitable for a detailed analysis of the genetic differentiation of geographically defined *X. oryzae* pv. *oryzae* populations. An understanding of the spatial distribution of genetic variation would be useful for guiding the use of resistant rice cultivars and for designing optimal strategies for sampling the pathogen in other regions where bacterial blight is a significant constraint to rice production.

We conducted a hierarchical analysis to determine the distribution of genetic variation of pathogen populations across diverse ecosystems. In the Philippines, rice is produced in both indigenous and improved ecosystems. In the cool, mountainous highlands, one crop of diverse traditional rice cultivars is grown per year. In the tropical lowlands, two to three crops of semidwarf, early

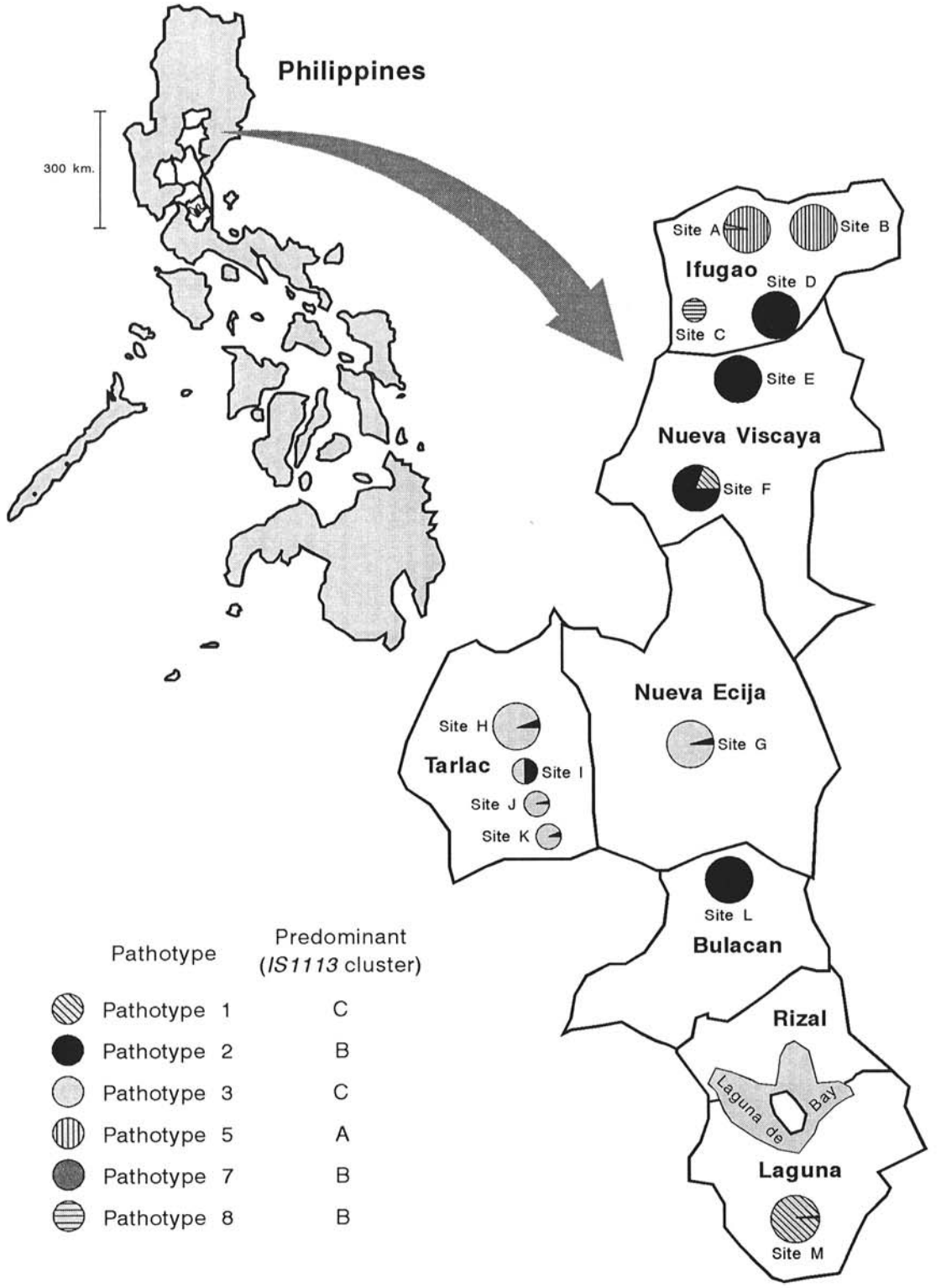

Fig. 1. Geographic location of 13 sites in Luzon, Philippines, from which strains of *Xanthomonas oryzae* pv. *oryzae* were collected. The distribution of pathotypes and predominant transposable element IS*1113* clusters is indicated. Table 1 describes the sites, and Table 2 lists the distances in kilometers between each pair of sites.

maturing cultivars are grown. Strains were sampled from farmers' fields at 13 sites across a 310-km transect spanning the two ecosystems on the island of Luzon. Strains were analyzed by DNA fingerprinting and by inoculation on a set of differential rice cultivars. Genetic differentiation was assessed at various levels to infer the influence of agroecosystems and host on population structure.

MATERIALS AND METHODS

Collection of infected leaf samples. Thirteen sites were selected from six provinces in northern (Ifugao and Nueva Vizcaya), central (Nueva Ecija, Tarlac, and Bulacan), and southern Luzon (Laguna) in the Philippines (Fig. 1; Table 1). At most of the sites, several rice fields showed symptoms of bacterial blight. Distances between the sites ranged from 1 to 310 km (Table 2). The six sites in northern Luzon, designated A through F, were sampled between 30 June and 2 July 1990; the six sites in central Luzon, designated G through L, were sampled between 27 and 28 September 1991; and site M in southern Luzon was sampled on 18 October 1991.

Four to five adjacent or neighboring fields were sampled for sites A, B, C, E, F, G, H, L, and M. Only one bacterial blight-infected field was sampled at sites C, I, J, and K. The single fields sampled at sites I, J, and K were 1 km apart. For each field at all sites, except those at site M, five symptomatic leaves (if present) were collected from each of seven 1-m^2 areas in a W-walk (Fig. 2). At site M, 10 leaves were collected from each 1-m^2 sampling area. Most of the sampling areas within the W-walk were 3 m apart. Distances between sampling areas varied depending on the incidence and severity of bacterial blight in the fields sampled.

Isolation, maintenance, and virulence analysis of bacteria. One colony was picked per leaf sample and streaked on modified Wakimoto's medium (20 g of sucrose, 5 g of peptone, 0.5 g of $Ca(NO_3)_2 \cdot 4H_2O$, 1.82 g of $Na_2HPO_4 \cdot 7H_2O$, 0.05 g of $FeSO_4 \cdot 7H_2O$, and 18 g of Bacto agar per liter [7]). For long-term storage, strains were lyophilized and/or maintained in 2% skim milk at –20°C.

Bacterial inoculum was prepared as described by Mew and Vera Cruz (13) and inoculated on a set of rice differential cultivars harboring specific resistance genes for bacterial blight: IR8 (*Xa-11*), IR24 (0), IR20 (*Xa-4*), IR1545-339 (*xa-5*), CAS209 (*Xa-10*), and DV85 (*xa-5* and *Xa-7*). Selected strains also were tested for virulence on IR-BB14 (*Xa-14*) to differentiate pathotypes 1 and 8 and pathotypes 5 and 7 (19). Three plants per cultivar were grown in the greenhouse, and the three youngest fully expanded leaves were clip-inoculated at maximum tillering stage (45 to 50 days after sowing). Disease reactions were scored by measuring lesion length 14 days after inoculation (8).

DNA typing. *DNA extraction.* Genomic DNA was extracted from 15 ml of overnight culture grown in nutrient broth following a modification of the method of Murray and Thompson (15) as described by Raymundo et al. (20). The DNA pellet was dissolved in 100 to 200 µl of TE (10 mM Tris, 1 mM EDTA, pH 8.0)

Restriction fragment length polymorphism (RFLP) analysis using IS1113. Approximately 5 to 10 µg of genomic DNA was digested overnight at 37°C with 50 units of *Eco*RI (Bethesda Research Laboratories [BRL], Gaithersburg, MD, or Boehringer Mannheim [Far East], Singapore) as described by the manufacturer. DNA fragments were separated by electrophoresis in 0.7% agarose-0.5× TBE (89 mM Tris [pH 7.8], 89 mM boric acid, 2 mM EDTA) gels at 10 V for 69 to 72 h prior to alkali transfer onto Hybond N (Amersham [Far East], Hong Kong) as directed by the membrane manufacturer.

The DNA probe IS*1113* (19) consisted of a 1.05-kb transposable element isolated from the genome of *X. oryzae* pv. *oryzae*. IS*1113* was propagated in *Escherichia coli* HB101 and later subcloned into pBluescript (N. Sakthivel and J. E. Leach, *unpublished*). The plasmid containing the IS*1113* fragment was isolated from overnight broth cultures by a modification of the alkaline lysis procedure of Birnboim and Doly (2). The IS*1113* fragment was excised from the vector by digestion with *Eco*RI and separated in 0.7% low melting-point agarose (BioRad Laboratories, Richmond, CA) by electrophoresis. The IS*1113* fragment was recovered from the gel using the Geneclean kit (BIO 101, La Jolla, CA). IS*1113* was labeled by random priming with digoxigenin-labeled dUTP (Boehringer Mannheim) and hybridized to bacterial DNA bound to nylon membranes (Hybond N, Amersham) according to the manufacturer's instructions. Hybridized bands were detected either on X-ray film (Kodak X-OMAT AR film [Eastman Kodak Co., Rochester, NY] or Amersham Hyperfilm) by chemiluminography using 3-(2′-spiroadamantane)-methoxy-4(3″phosphoryloxy)-1,2-dioxytane as chemiluminescent substrate for alkaline phosphatase (Boehringer Mannheim or Tropix, Bedford, MA) or directly on the membrane by the chromogenic method, using nitro blue

TABLE 1. Location of sampling sites on the island of Luzon, Philippines, agroecosystems, number of fields sampled, and rice cultivars planted at the 13 sites from which strains of *Xanthomonas oryzae* pv. *oryzae* were collected

Site	Location	Region	Type of agroecosystem[a]	No. of fields analyzed	Rice cultivar planted[b]
A	Banaue (near Batad), Ifugao	Northern Luzon	TH	4	Traditional
B	Banaue (near Taloc), Ifugao	Northern Luzon	TH	4	Traditional
C	Namulditan, Ifugao	Northern Luzon	TT	1	Traditional
D	Lamut, Ifugao	Northern Luzon	IT	4	IR60[c], IR66[c]
E	Bagabag, Nueva Vizcaya	Northern Luzon	IT	3	IR20[c], IR60
F	Bayombong, Nueva Vizcaya	Northern Luzon	I/TT	5	C4[d], IR[e], Traditional
G	Talavera, Nueva Ecija	Central Luzon	IIL	4	IR64[c], IR72[c]
H	La Paz, Tarlac	Central Luzon	IIL	4	IR72
I	La Paz, Tarlac	Central Luzon	IIL	1	IR72
J	La Paz, Tarlac	Central Luzon	IIL	1	IR72
K	La Paz, Tarlac	Central Luzon	IIL	1	IR72
L	San Miguel, Bulacan	Central Luzon	IIL	5	R10, IR66
M	Calauan, Laguna	Southern Luzon	I/TIL	2	Malagkit[f], IR74

[a] Sites were categorized into different agroecosystems based on the rice cultivar(s) planted at each of the sites and the elevation of the sites. TH = traditional highland; highland site commonly planted to a traditional rice cultivar. TT = traditional transition; transition zone site planted to a traditional rice cultivar. IT = improved transition; transition zone site planted to an improved rice cultivar. I/TT = improved and traditional transition; transition zone site planted to improved and traditional rice cultivars. IIL = improved irrigated lowland; irrigated lowland site planted to an improved rice cultivar. I/TIL = improved and traditional irrigated lowland; irrigated lowland site planted to improved and traditional rice cultivars.

[b] Information on cultivar(s) planted was obtained from farmers.

[c] Improved semidwarf cultivar with the *Xa-4* gene.

[d] Improved semidwarf cultivar without the *Xa-4* gene.

[e] Unidentified improved cultivar.

[f] Traditional glutinous rice cultivar.

tetrazolium chloride and 5-bromo-4-chloro-3-indolylphosphate as substrates.

Restriction analysis using Pst*I*. Approximately 5 to 8 µg of genomic DNA was digested overnight with a fivefold excess of *Pst*I (BRL or Boehringer Mannheim) at 37°C according to the manufacturer's instructions. DNA fragments were separated by electrophoresis in 0.7% agarose-1× TBE gels at 25 V for 16 h. The gels were stained in 0.5 µg of ethidium bromide per ml of H_2O and photographed under ultraviolet irradiation (302 nm) using type 55 or 57 Polaroid film (Polaroid Corp., Cambridge, MA).

Data analysis. *Cluster analysis.* DNA of strains with unique DNA profiles (haplotypes) was coelectrophoresced to confirm band identities and differences. To derive phenograms from the DNA band data, the DNA profiles were scored in binary form, i.e., the

TABLE 2. Geographic differentiation of *Xanthomonas oryzae* pv. *oryzae* among collection sites on the island of Luzon, Philippines, measured by frequency distributions of haplotypes defined by restriction fragment length polymorphism based on transposable element IS*1113* and on *Pst*I digestion

Site[a]	Data set and analysis[b]	Site[c–e]												
		A	B	C	D	E	F	G	H	I	J	K	L	M
A	IS*1113*, χ^2		6	17	38	44	55	153	172	173	174	175	200	310
	IS*1113*, FE													
	*Pst*I, χ^2													
	*Pst*I, FE													
B	IS*1113*, χ^2	135		15	32	38	51	149	169	170	171	172	196	305
	IS*1113*, FE	158												
	*Pst*I, χ^2	203												
	*Pst*I, FE	250												
C	IS*1113*, χ^2	153	135		27	34	42	138	156	157	158	159	186	296
	IS*1113*, FE	146	144											
	*Pst*I, χ^2	114	129											
	*Pst*I, FE	101	128											
D	IS*1113*, χ^2	236	267	160		7	20	117	139	140	141	142	164	272
	IS*1113*, FE	312	296	160										
	*Pst*I, χ^2	200	195	126										
	*Pst*I, FE	240	245	121										
E	IS*1113*, χ^2	176	162	95	133		18	114	136	137	138	139	159	267
	IS*1113*, FE	203	197	114	142									
	*Pst*I, χ^2	160	155	86	85									
	*Pst*I, FE	176	182	96	78									
F	IS*1113*, χ^2	251	219	167	69	143		98	119	120	121	122	145	255
	IS*1113*, FE	315	299	157	86	154								
	*Pst*I, χ^2	240	235	153	93	111								
	*Pst*I, FE	284	287	123	97	109								
G	IS*1113*, χ^2	253	298	174	232	173	154		27	28	29	30	49	164
	IS*1113*, FE	328	238	163	294	204	176							
	*Pst*I, χ^2	233	228	159	221	172	261							
	*Pst*I, FE	274	278	131	261	177	311							
H	IS*1113*, χ^2	254	240	219	216	166	140	3^{ns}		1	2	3	44	157
	IS*1113*, FE	326	309	163	271	198	158	3^{ns}						
	*Pst*I, χ^2	235	230	161	196	164	239	101						
	*Pst*I, FE	287	291	140	237	184	297	108						
I	IS*1113*, χ^2	117	132	67	156	89	84	71	72		1	2	43	156
	IS*1113*, FE	135	133	25	138	99	75	53	53					
	*Pst*I, χ^2	136	131	62	118	79	156	88	63					
	*Pst*I, FE	121	128	71	107	85	125	68	49					
J	IS*1113*, χ^2	248	132	66	144	83	66	0.62^{ns}	0.28^{ns}	19		1	42	155
	IS*1113*, FE	134	132	82	129	97	60	0.35^{ns}	0.14^{ns}	21				
	*Pst*I, χ^2	134	129	60	121	82	154	91	30	21				
	*Pst*I, FE	119	126	72	112	89	124	68	26	21				
K	IS*1113*, χ^2	152	139	70	127	77	51	5^{ns}	2^{ns}	23	2^{ns}		41	154
	IS*1113*, FE	144	142	87	115	93	49	5^{ns}	2^{ns}	24	2^{ns}			
	*Pst*I, χ^2	130	125	56	111	74	148	30	5^{ns}	24	24			
	*Pst*I, FE	109	116	68	96	80	110	23	5^{ns}	24	26			
L	IS*1113*, χ^2	289	273	206	40	165	107	276	260	114	185	166		116
	IS*1113*, FE	359	338	176	50	168	130	347	325	97	144	131		
	*Pst*I, χ^2	272	267	198	201	186	255	289	269	134	191	178		
	*Pst*I, FE	317	318	147	230	185	310	346	336	103	140	123		
M	IS*1113*, χ^2	247	233	166	251	186	168	8^{ns}	11	64	5^{ns}	19	291	
	IS*1113*, FE	318	302	162	342	224	206	12^{ns}	12	49	4^{ns}	14	376	
	*Pst*I, χ^2	228	223	154	220	180	169	253	255	156	154	150	288	
	*Pst*I, FE	281	284	138	273	200	198	313	326	139	136	124	356	

[a] Sites are described in Table 1.
[b] FE = Fisher's exact test.
[c] Values above the diagonal are distances in kilometers between pairs of sites.
[d] Values below the diagonal are chi-square and Fisher's exact values for each pairwise comparison. Unless marked ns, values are significant at $P = 0.05$.
[e] For the pairwise comparisons among the five sites at which IR72 was planted, the chi-square and Fisher's exact values are highlighted with a triangle.

presence or absence of bands were coded as 1 or 0, respectively. For the *Pst*I haplotypes, only bands larger than 7 kb were scored, giving a total of 27 band positions. In some cases, two *Pst*I haplotypes showed identical banding patterns, with one haplotype having one band that was denser than the corresponding band in the other haplotype. Although these were considered distinct haplotypes, only one was included in the analysis.

Similarity coefficients were calculated for all pairwise combinations of *Pst*I or IS*1113* haplotypes using Dice's coefficient (21). Cluster analysis was conducted based on similarity coefficients by the unweighted pair-group method, arithmetic mean (UPGMA [22]) using the computer program NTSYS-pc, version 1.70 (21).

Bootstrap analysis using the computer program "Winboot" (I. Yap and R. Nelson, *unpublished*) was used to assess the robustness of the groupings produced by cluster analysis of both IS*1113* and *Pst*I data sets. The IS*1113* and the *Pst*I phenograms were reconstructed 2,000 and 400 times, respectively, by repeated sampling with replacement, and the frequency with which a particular grouping was formed was used as a measure of the strength of the grouping (4,5,19).

Analysis of diversity. The genetic diversity of *X. oryzae* pv. *oryzae* populations was estimated using Nei's haplotypic diversity index, which takes into account both the number of types and their frequencies (17,18). The haplotypic diversity of the *X. oryzae* pv. *oryzae* subpopulations at each site (H_S), field (H_F), and agroecosystem/region (H_A) and the total haplotypic diversity of the entire *X. oryzae* pv. *oryzae* population (H_T) were estimated using the equation $H = [n/(n-1)](1 - \Sigma X_i^2)$, where X_i is the frequency of the *i*th haplotype (*Pst*I or IS*1113*) and n is the number of strains examined. The same equation was used to calculate the haplotypic diversities of the *X. oryzae* pv. *oryzae* subpopulations collected from traditional and improved rice cultivars.

Cluster analysis of sites based on haplotype distribution. To determine the relatedness among pathogen populations at the different sites, Pearson correlation analysis (Systat for the Macintosh, version 5.2, Systat, Evanston, IL) was conducted based on IS*1113* frequency distribution. The correlation matrix was used to construct a dendrogram depicting the relationships among pathogen populations at the 13 sites.

Analysis of differentiation. Geographic subdivision was analyzed at four levels: (i) between agroecosystems; (ii) between sites; (iii) among fields within a site; and (iv) among sampling areas within a field, using two approaches. Overall partitioning of variation was examined using the coefficient of genetic differentiation (3,6,16):

$$G_{ST} = (H_T - \bar{H}_S)/H_T$$

where $\bar{H}_S$ is the average of estimated haplotypic diversities of the subpopulations at the 13 sites and H_T is an estimate of the haplotypic diversity of the total population (6).

To determine the differentiation between individual regions and sites, pairwise comparisons of haplotype distribution were made using chi-square and Fisher's exact tests. Pairwise comparisons of the distribution of pathotypes and IS*1113* lineages were done between sites. Genetic differentiation of haplotype distribution within site among fields and within field among sampling areas also was assessed by chi-square and Fisher's exact tests. The computer programs Systat for the Macintosh (version 5.2) and StatXact (version 1.00, Cytel Software Corp., Cambridge, MA) were used to calculate chi-square and Fisher's exact statistics.

Data subsampling. To assess the relationship between within-field sampling intensity and detected pathogen diversity, the data sets for 39 fields were subjected to systematic subsampling. The *Pst*I data set was chosen for this subsampling exercise because of the higher discriminating power of this method, which allowed the detection of more than twice the number of DNA haplotypes compared with the IS*1113*-RFLP method. To simulate the result of different strategies for within-field sampling, reiterative subsampling was conducted using a program written in Borland Pascal 7.0 for this purpose (I. Yap, *unpublished*).

For each of the fields, the number of sampling areas and the number of samples per area were progressively reduced from seven to one and from five to one, respectively, using a random number generator. Subsampling was simulated using the actual frequencies of haplotypes detected in each area as probabilities for taking a random sample. Subsampling was conducted three times, with different random subsamples generated for each run. The proportion of *Pst*I haplotypes obtained for each random subsample was determined by dividing the number of haplotypes taken for each subsample by the actual number of haplotypes detected from the original (unsubsampled) sampling areas. The average across all fields and runs was taken for each area-sample combination. To illustrate the relationships between sampling scheme (number of sampling areas and number of samples per area) and detected diversity (proportion of *Pst*I haplotypes recovered), surface and contour plots, with distance-weighted least squares smoothing, were constructed using Systat version 5.03 for Windows.

RESULTS

Establishment and pathotypic analysis of the strain collection. More than 1,200 strains of *X. oryzae* pv. *oryzae* were collected from 13 sites along a 310-km transect between the traditional highland agroecosystem in northern Luzon and the improved irrigated lowland agroecosystem in central and southern Luzon in the Philippines. The sites were categorized into different agroecosystems on the basis of elevation and the rice cultivars planted at

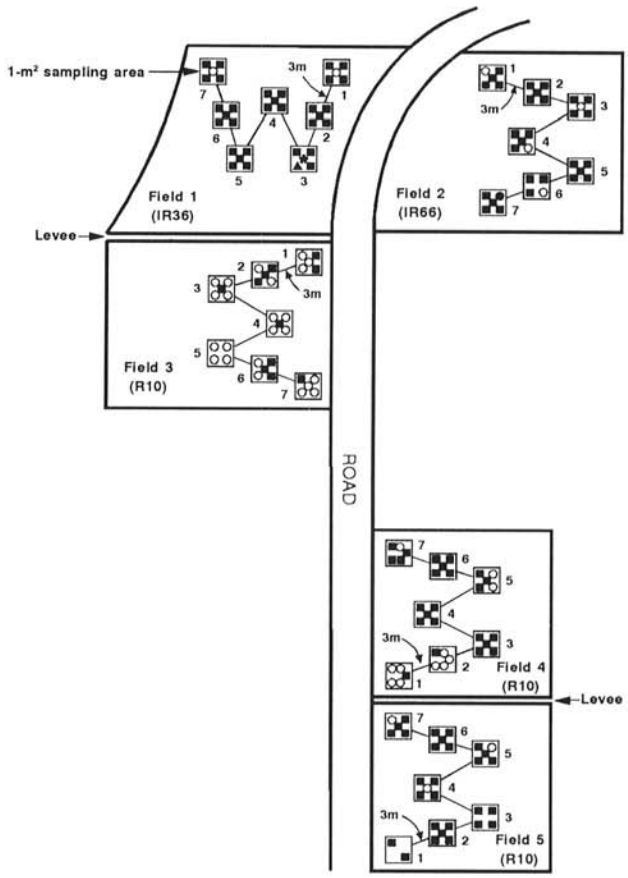

Fig. 2. Sketch of site L, located in San Miguel, Bulacan, Philippines, showing the W pattern used for sampling *Xanthomonas oryzae* pv. *oryzae* strains and the transposable element IS*1113* haplotypes detected. The seven 1-m² sampling areas were approximately 3 m apart. Field 4 showed significant within-field (between-sampling area) differentiation based on chi-square and Fisher's exact tests. Geometric shapes represent the IS*1113* haplotypes detected in the site: ● = B-002; ○ = B-004; ■ = B-005; ▲ = B-011; ✶ = B-012.

each of the sites. The locations and characteristics of the 13 sites are given in Table 1 and Figure 1.

Six pathotypes were detected, corresponding to a subset of those previously reported (14,19). Pathotypes 4 and 6, which were previously detected at sites not included in this study, were not found. Pathotype 2 was the most abundant in the collection (41%). The second most abundant was pathotype 3 (27%). Pathotypes 5 and 1 were represented at intermediate frequencies (17 and 13%, respectively), and pathotypes 8 and 7 were relatively rare (3 and 0.4%, respectively).

The distinct geographic distribution of *X. oryzae* pv. *oryzae* pathotypes reported in previous surveys in the Philippines (14) was apparent in this survey (Fig. 1). Pathotype 2 was found at all sites, except for A, B, and C, the three sites at which only traditional cultivars were grown (the two highland sites and the adjacent "transitional zone" site at Namulditan). Pathotype 3 was found at five sites in central Luzon (sites G, H, I, J, and K). Pathotype 1 was identified at sites F and M, the only two lowland sites at which traditional cultivars were grown.

Pathotype 5 was found only at the two highland areas of Ifugao Province (sites A and B). Pathotype 7 (previously classified as pathotype 5 but differentiated on rice cultivars carrying the *Xa-14* gene [19]) was found only at site A. Pathotype 8 (previously classified as pathotype 1 but also differentiated on rice cultivars carrying the *Xa-14* gene [19]) was detected only at site C.

A single pathotype was present at each of five sites (B, C, D, E, and L). At each of the other eight sites, however, two pathotypes were found. In central Luzon, strains of pathotypes 2 and 3 were found in the same fields. At highland site A, pathotypes 5 and 7 were found in the same fields. At sites F and M, both pathotypes 1 and 2 were collected from fields planted to cultivars expected not to carry the *Xa-4* gene for resistance to *X. oryzae* pv. *oryzae*, whereas only pathotype 2 was collected from the fields planted to cultivars carrying *Xa-4*.

Phylogenetic and bootstrap analyses. From a total of 1,242 strains examined for DNA polymorphism by RFLP analysis using the repetitive DNA probe IS*1113*, 19 unique RFLP patterns (IS*1113* haplotypes) were defined (Table 3; Figs. 3A and 4A). Between 9 and 12 hybridizing bands were scored per strain (average = 10.6 bands per strain), and 26 band positions were scored. Of these strains, 1,161 also were analyzed by *Pst*I digestion, and 46 unique *Pst*I digestion patterns (*Pst*I haplotypes) were identified (Table 4; Figs. 3B and 4B). For each strain, between 1 and 10 bands of greater than 7 kb were scored (average = 7.3 bands per strain). A total of 27 band positions was scored.

Based on UPGMA and bootstrap analyses, the 19 IS*1113* haplotypes formed three robust groupings (bootstrap values of >88%; Fig. 4A). The three IS*1113* lineages corresponded to those previously observed (19). Eight of the same haplotypes were seen in the previous study, and eleven "new" haplotypes were detected.

TABLE 3. Frequency distribution of transposable element IS*1113*-derived haplotypes and estimates of genetic diversity of *Xanthomonas oryzae* pv. *oryzae* at 13 sites on the island of Luzon, Philippines

IS*1113* haplotype[a]	Pathotype[b]	No. of strains per site[c]													Row total	%
		A	B	C	D	E	F	G	H	I	J	K	L	M		
A-002	5	65													65	5.23
A-006	5	1													1	0.08
A-007	5	1													1	0.08
A-003	5	26	93												119	9.58
A-004	5		4												4	0.32
A-008	5	13													13	1.05
A-009	5		3												3	0.24
B-002	2						32			2			1		35	2.82
B-004	2									13			41	1	55	4.43
B-008	2							2							2	0.16
B-005	2				124	13	75	6	11	1	2	6	125		363	29.23
	3							2	2						4	0.32
	1													1	1	0.08
B-009	2				1										1	0.08
B-011	2												1		1	0.08
B-010	5		1												1	0.08
	2					47									47	3.78
B-006	8			35											35	2.82
B-013	7	5													5	0.40
B-012	2												1		1	0.08
B-014	5	5													5	0.40
C-002	1						26							128	154	12.40
	2							1							1	0.08
	3							126	125	16	29	29			325	26.17
Total no. of isolates		116	101	35	125	60	133	137	138	32	31	35	169	130	1,242	
No. of fields sampled		4	4	1	4	3	5	4	4	1	1	1	5	2	39	
No. of IS*1113* haplotypes		7	4	1	2	2	3	3	2	4	2	2	5	3		
Proportion of variable areas[d]		20/26	3/24	0/7	4/28	1/13	18/27	5/28	9/28	6/7	2/7	5/7	24/35	2/14	99/251	
Genetic diversity		0.63	0.15	0	0.02	0.35	0.59	0.14	0.17	0.60	0.12	0.29	0.40	0.03	0.75	

[a] The letters A, B, and C in the haplotype designations refer to the groupings formed by cluster analysis (the unweighted pair-group method, arithmetic mean; Fig. 4A) that correspond to lineages of the pathogen (17). The numbers after each letter refer to a distinct haplotype (or hybridization profile) within the lineage.

[b] The majority of the isolates (n = 1,177; 95%) analyzed by IS*1113*-restriction fragment length polymorphism were subjected to pathotyping.

[c] Sites are described in Table 1.

[d] Number of 1-m^2 sampling areas with more than 1 haplotype/total number of areas sampled.

The association between IS*1113* lineage and pathotype was similar to that described by Nelson et al. (19).

The 46 *Pst*I haplotypes detected did not form robust groupings (Fig. 4B). Bootstrap values of the *PstI* clusters ranged from 0 to 77%. Although some of the groupings had low bootstrap values, there was a general correspondence between the groups formed by IS*1113* and *Pst*I analyses. The majority of strains in IS*1113* cluster A were found in *Pst*I cluster C, most of the strains of IS*1113* cluster B were found in the *Pst*I cluster D, and most of the strains of IS*1113* cluster C were found in *Pst*I clusters A and H (DNA from strains in cluster H were not digestible with *Pst*I; data not shown).

Genetic diversity. The total genetic (haplotypic) diversity (H_T) of the entire *X. oryzae* pv. *oryzae* collection analyzed by IS*1113* RFLP was 0.75 (Table 3). Diversity was highest at site A in the highlands, with seven IS*1113* haplotypes detected among 116 strains analyzed. The second most diverse populations were at site I in the lowlands and site F in the transition zone. Diversity was low at seven sites: B in the highlands; D in the transition zone; and G, H, J, K, and M in the lowlands. The population at site C was homogeneous with only one haplotype, B-006, detected among 35 strains analyzed. Based on IS*1113* haplotype frequency, the average genetic diversity ($\overline{H_S}$) of the *X. oryzae* pv. *oryzae* subpopulations at the 13 sites was estimated to be 0.27. The coefficient of genetic differentiation (G_{ST}) was 0.64.

A higher genetic diversity (H_T) estimate of 0.93 was obtained for the total population of *X. oryzae* pv. *oryzae* strains analyzed by *Pst*I digestion (Table 4). Diversity was highest at sites I and L, both lowland sites, with 6 and 11 *Pst*I haplotypes among 32 and 168 strains analyzed, respectively. Similar genetic diversity estimates were obtained from site F in the transition zone, from site G in the lowlands, and from site A in the highlands. Diversity was low at site B in the highlands and at site K in the lowlands. Only three haplotypes were detected from both sites out of 99 and 26 strains analyzed from sites B and K, respectively. The population in the single field sampled at site C was again homogeneous: all 30 strains analyzed from this site had the D-001 haplotype. The average genetic diversity ($\overline{H_S}$) of the *X. oryzae* pv. *oryzae* subpopulations at the 13 sites based on *Pst*I haplotype frequencies was 0.56. The coefficient of genetic differentiation (G_{ST}) was 0.40.

The genetic diversity of pathogen subpopulations collected from traditional cultivars was compared with the diversity of pathogen subpopulations collected from improved, short-stature cultivars bred at the International Rice Research Institute (IR cultivars) and at the University of the Philippines, Los Baños (cultivar C4). For 324 strains collected from traditional cultivars (from sites A, B, C, F, and M), 14 haplotypes were detected, and a genetic diversity estimate (H_{TV}) of 0.77 was obtained. For the 918 strains collected from improved, short-stature rice cultivars (from sites D through M), nine haplotypes were detected, and a genetic diversity estimate (H_{IV}) of 0.63 was obtained. The significance of the difference (0.77 versus 0.63), however, could not be tested statistically because the variances of the diversity estimates were not known.

Geographic differentiation. The relationships among the *X. oryzae* pv. *oryzae* populations at the 13 sites were assessed by Pearson correlation analysis. The IS*1113* frequency distribution was used to calculate the Pearson correlation between sites, and the correlation matrix obtained was used to construct a dendrogram depicting the relationship among the pathogen populations at different sites (Fig. 5). Four major clusters were obtained, corre-

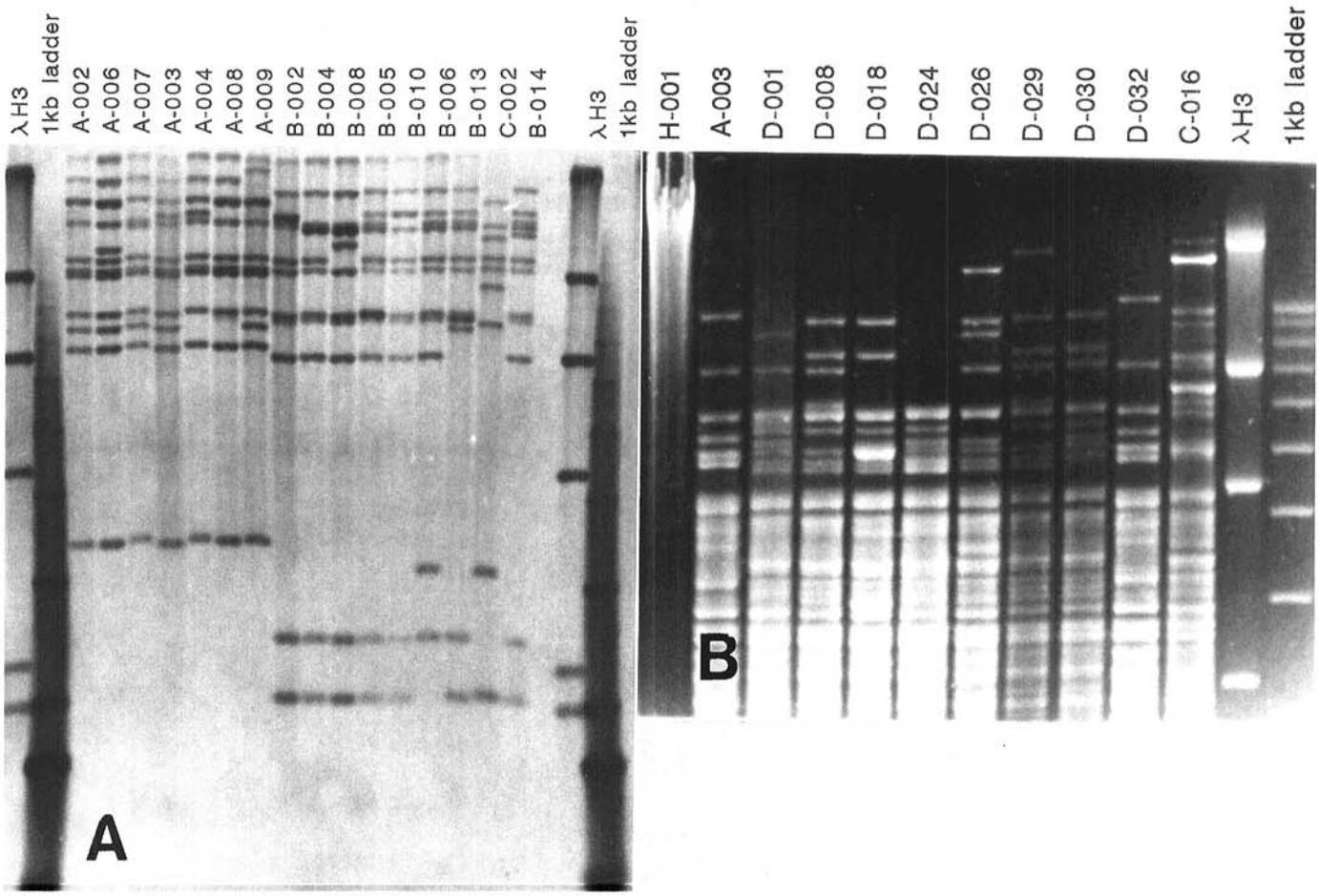

Fig. 3. Examples of DNA hybridization patterns and haplotype designations of *Xanthomonas oryzae* pv. *oryzae* strains obtained using **A,** transposable element IS*1113* as a probe and **B,** DNA banding patterns derived by *Pst*I digestion.

sponding closely to the agroecosystems along the transect. Sites A and B were in the traditional highland area and formed one cluster. Sites G, H, I, J, K, and M in the improved irrigated lowland area comprised another cluster. Two clusters were formed by sites C, D, E, and F, which were located in the transition zone between the traditional highland and the improved irrigated lowland agroecosystems. Site L, which was located between site G in Nueva Ecija and site M in Calauan, Laguna (Fig. 1), had a pathogen population closely resembling those in the transition zone and clustered with two transition zone sites, D and F.

Chi-square and Fisher's exact tests were used to evaluate the differentiation among the pathogen collections from the three major agroecological zones. The highlands were represented by sites A and B; the transition zone was represented by sites C, D, E, and F; and the lowlands were represented by sites G through M. Both IS*1113* and *Pst*I data were analyzed. Based on all tests, the three zones were significantly differentiated (all chi-square and Fisher's exact values exceeded 500; $P < 0.001$; data not shown).

The majority of the sites were strongly differentiated, based on the haplotypic composition of the pathogen populations. For each pair of sites, the distribution of haplotypes between sites was compared by the chi-square method and the Fisher's exact test, using both the IS*1113* and *Pst*I data sets. Of 78 site-by-site comparisons, 70 and 77 cases, respectively, showed significant differentiation of IS*1113* and *Pst*I haplotype distributions at $P < 0.01$ (Table 2). The eight pairs of sites that had similar IS*1113* haplotype distributions also showed similar distribution of pathotypes and IS*1113* lineages. Sites A and B differed in the distribution of IS*1113* haplotypes but were similar in the distribution of pathotypes and IS*1113* lineages.

Sampling of multiple fields at nine sites allowed us to assess genetic heterogeneity among fields within each of the sites. For each site, the two marker data sets (IS*1113* and *Pst*I) were each analyzed using chi-square and Fisher's exact tests (Table 5). The detection of differentiation was dependent on both the marker system used and the analytical method. The more discriminating

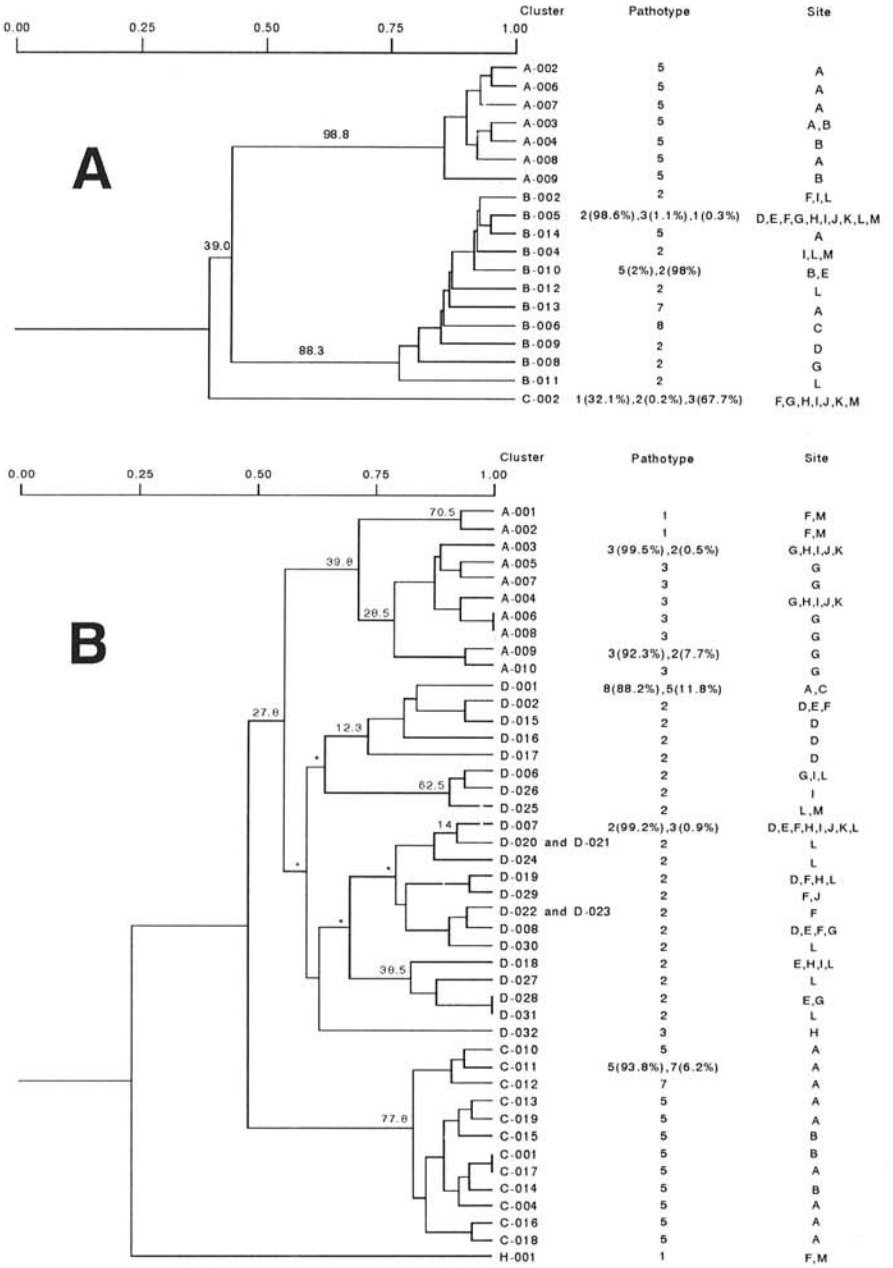

Fig. 4. Dendrogram derived by the unweighted pair group method, arithmetic mean, showing the similarity relationships among **A,** transposable element IS*1113* haplotypes of *Xanthomonas oryzae* pv. *oryzae* and **B,** *Pst*I haplotypes. Pathotype, site of origin, and bootstrap values (values on branches of the clusters, representing the percentage of times the group occurred out of 2,000 and 400 iterations for **A** and **B**, respectively) are indicated. An asterisk indicates the cluster is not in the bootstrap consensus tree.

TABLE 4. Frequency distribution of *Pst*I-derived haplotypes and estimates of genetic diversity of *Xanthomonas oryzae* pv. *oryzae* at 13 sites on the island of Luzon, Philippines

*Pst*I haplotype[a]	Pathotype[b]	No. of strains per site[c]													Row total	%
		A	B	C	D	E	F	G	H	I	J	K	L	M		
H-001	1						2							18	20	1.72
A-001	1						11							41	52	4.48
A-002	1						15							64	79	6.80
A-003	3							62	92	8	9	23			194	16.71
	2									1					1	0.09
A-004	3							3	25	7	19	1			55	4.74
A-005	3							8							8	0.69
A-006	3							7							7	0.60
A-007	3							15							15	1.29
A-008	3							10							10	0.86
A-009	3							12							12	1.03
	2							1							1	0.09
A-010	3							6							6	0.52
D-001	8			30											30	2.58
	5	4													4	0.34
D-002	2				15	4	2								21	1.81
D-015	2				1										1	0.09
D-016	2				11										11	0.95
D-017	2				2										2	0.17
D-007	2				48	9	28		7	2	1	2	20		117	10.08
	3								1						1	0.09
D-018	2					1			1	1			9		12	1.03
D-019	2				1		1		3				1		6	0.52
D-020	2												37		37	3.19
D-021	2												1		1	0.09
D-022	2						9								9	0.78
D-023	2						5								5	0.43
	1						1								1	0.09
D-024	2												2		2	0.17
D-006	2							2		8			26		36	3.10
D-025	2												15	1	16	1.38
D-026	2									5					5	0.43
D-027	2												1		1	0.09
D-008	2				18	10	59	1							88	7.58
D-028	2					32		2							34	2.93
D-029	2						3				1				4	0.34
D-030	2												54		54	4.65
D-031	2												2		2	0.17
D-032	3								2						2	0.17
C-010	5	4													4	0.34
C-011	5	15													15	1.29
	7	1													1	0.09
C-012	7	2													2	0.17
C-013	5	2													2	0.17
C-001	5		11												11	0.95
C-014	5		84												84	7.24
C-015	5		4												4	0.34
C-004	5	6													6	0.52
C-016	5	11													11	0.95
C-017	5	6													6	0.52
C-018	5	52													52	4.48
C-019	5	1													1	0.09
Total no. of isolates		104	99	30	96	56	136	129	131	32	30	26	168	124	1,161	
No. of fields sampled		4	4	1	4	3	5	4	4	1	1	1	5	2	39	
No. of *Pst*I types		10	3	1	7	5	10	11	6	6	4	3	11	4		
Proportion of variable areas[d]		19/26	6/24	0/7	15/24	10/14	24/29	27/28	11/28	6/7	5/7	4/7	32/35	13/14	172/250	
Genetic diversity		0.71	0.27	0	0.68	0.62	0.75	0.74	0.47	0.81	0.53	0.22	0.8	0.61	0.93	

[a] The letters A, C, D, and H in the haplotype designations refer to the groupings formed by cluster analysis (the unweighted pair-group method, arithmetic mean; Fig. 4B) at an arbitrary cut off point of 55% similarity. The numbers after each letter refer to a distinct haplotype (or restriction pattern) within the group.

[b] The majority of the isolates analyzed by *Pst*I digestion were subjected to pathotyping.

[c] The sites are described in Table 1.

[d] Number of 1-m^2 sampling areas with more than 1 haplotype/total number of areas sampled.

*Pst*I method revealed greater differentiation: using this data set, seven of the nine sites with multiple fields showed among-field differentiation. Sites A, F, and L were significantly ($P < 0.05$) differentiated based on both analyses for both data sets. Significant among-field heterogeneity also was obtained for sites B and G based on chi-square analysis of the IS*1113* data set. Sites D, E, H, and M also consistently showed significant differentiation in *Pst*I haplotype distribution among fields with both statistical tests.

Strain collections from 8 of the 13 sites contained more than one pathotype (Table 1). Pathotypic differentiation among fields within a site was evaluated by chi-square and Fisher's exact tests for five of the sites that had more than one field sampled (A, F, G, H, and M). Only site F showed highly significant differentiation ($P < 0.01$) in pathotypic composition among fields within the site. This site had the greatest number of different cultivars planted, with improved and traditional cultivars planted in adjacent fields.

Of the 39 fields sampled from 13 sites, 32 (82%) contained more than one IS*1113* haplotype. Differentiation in the distribution of IS*1113* haplotypes among the sampling areas within each of these 32 fields was assessed by chi-square and Fisher's exact tests. With the chi-square test, five fields showed significant differentiation of IS*1113* haplotype distribution among sampling areas, whereas with Fisher's exact test only two of these fields showed significant among-sampling area differentiation (data not shown). An example of the distribution of haplotypes in one of these fields is shown in Figure 2.

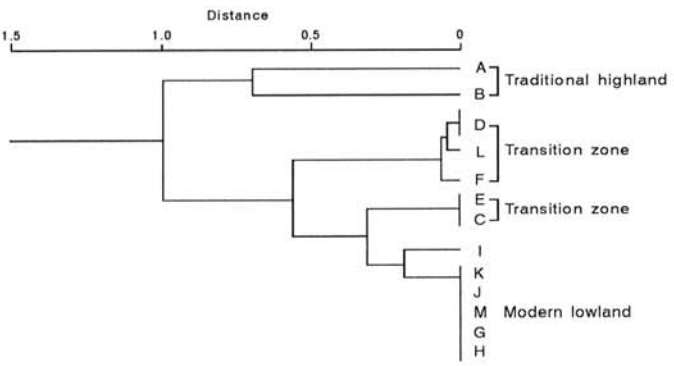

Fig. 5. Dendrogram depicting the relationships among *Xanthomonas oryzae* pv. *oryzae* subpopulations at different sites in Luzon, Philippines. Transposable element IS*1113* haplotype frequencies were used to calculate Pearson correlation between sites, and the correlation matrix was used to construct the dendrogram. The distance scale is $1 - r$, where r is the Pearson correlation coefficient.

TABLE 5. Within-site differentiation of *Xanthomonas oryzae* pv. *oryzae* populations based on frequencies of transposable element IS*1113* and *Pst*I haplotypes found at sampling sites on the island of Luzon, Philippines

Site[a]	Data set	No. of haplotypes (h)	No. of fields (f)	df $(h-1)(f-1)$	χ^2 value[b]	Fisher's exact value[b]
A	IS*1113*	7	4	18	53**	44**
	*Pst*I	10		27	56**	47*
B	IS*1113*	4	4	9	26**	16^{ns}
	*Pst*I	3		6	11^{ns}	10^{ns}
D	IS*1113*	2	4	3	4^{ns}	3^{ns}
	*Pst*I	7		18	107**	104**
E	IS*1113*	2	3	2	2^{ns}	2^{ns}
	*Pst*I	5		12	37**	33**
F	IS*1113*	3	5	8	111**	98**
	*Pst*I	10		36	140**	115**
G	IS*1113*	3	4	6	14*	10^{ns}
	*Pst*I	11		30	36^{ns}	35^{ns}
H	IS*1113*	2	4	3	2^{ns}	3^{ns}
	*Pst*I	6		15	116**	100**
L	IS*1113*	5	5	16	63**	58**
	*Pst*I	11		40	126**	105**
M	IS*1113*	3	2	2	2^{ns}	2^{ns}
	*Pst*I	4		3	15**	15**

[a] The nine sites with more than 1 field sampled were analyzed.
[b] ** = highly significant differentiation ($P < 0.01$); * = significant differentiation ($P < 0.05$); ns = no significant differentiation ($P > 0.05$).

Differentiation among fields planted to the same cultivar. To assess the role of host selection in structuring pathogen populations, the haplotypic compositions of pathogen subpopulations collected from individual fields planted to the same cultivar were compared. Fields planted to IR60, IR66, and IR72 were sampled at more than one site, allowing us to test the role of location versus host in determining differentiation of pathogen populations.

Strains were collected from IR60 from two fields at site D and from one field at site E. Although only lineage B was collected from the three fields of IR60, the predominant haplotypes present at the two sites were distinct. Therefore, based on both the chi-square and Fisher's exact tests, the subpopulations taken from the same host at the two sites were differentiated. Strains were collected from two fields each of IR66 at sites D and L. For both sites, no differentiation was seen between the two fields of IR66. There was also no significant differentiation between the pathogen subpopulations collected from this cultivar at the two sites.

Collections were made from cultivar IR72 at five sites (from three fields at site G, from four fields at site H, and from one field each at sites I, J, and K). Significant differentiation of pathogen subpopulations from IR72 was seen across fields because the subpopulation present at site I was distinct from those collected from the other sites. Within sites G and H, no differentiation was seen among the seven fields of IR72. No differentiation was seen for 6 of the 10 pairwise comparisons of sites planted to IR72 that were made using the chi-square and Fisher's exact tests. Only site I was significantly differentiated from the other four sites. The presence of IR72 as host cultivar can explain six of the eight pairs of sites that were not differentiated among the 78 pairwise comparisons of sites for the entire transect (Table 2). The two other pairs of sites showing nondifferentiation were planted to IR72 and IR74. This suggests a strong role played by host selection in causing differentiation of pathogen populations.

Data subsampling. Data subsampling was simulated to determine whether a lower sampling intensity could be used to detect the diversity observed in this study. Because differentiation was observed between fields and at higher geographic scales, subsampling was conducted using within-field data. With reiterative subsampling, decreases in both the number of sampling areas per field and the number of samples per area resulted in a reduction in the proportion of haplotypes detected. A greater number of haplotypes were lost by decreasing the number of areas than by decreasing the number of samples per area (Fig. 6A). Nearly 80% of the diversity found by sampling 35 samples per field could be detected by sampling 18 samples per field if 3 samples were taken in each of 6 sampling areas (Fig. 6B).

DISCUSSION

Hierarchical analysis of genetic differentiation. In this study, two DNA fingerprinting techniques—RFLP with the probe IS*1113* and *Pst*I digestion—were used to measure the diversity and differentiation of populations of the bacterial blight pathogen at various spatial scales in the Philippines. A large collection of strains ($n > 1{,}200$) obtained from intensive sampling of 13 sites along a 310-km transect spanning the highland agroecosystem, which is dominated by traditional cultivars, and the irrigated lowland agroecosystem, which is dominated by improved high yielding cultivars on the island of Luzon, was analyzed. The pathogen populations across the transect were strongly differentiated. For the entire transect, a high degree of genetic differentiation ($G_{ST} = 0.64$ based on the IS*1113* data set; $G_{ST} = 0.40$ based on the *Pst*I data set) was found. Genetic differentiation of the pathogen population was evaluated at the level of agroecosystems, sites, fields within sites, samp-

ling areas within fields, and cultivars. Partitioning of variation was seen at every level tested, although the degree of partitioning decreased with the size of the spatial units compared.

When the haplotypic composition of the collections from different sites was used as a basis for cluster analysis of the sites, the sites were grouped in a way that closely reflected the agroecosystems. Chi-square and Fisher's exact analyses showed that the pathogen populations found in the three major agroecosystems (highland, transition, and irrigated lowland) were distinct. Differentiation was further seen in most pairwise comparisons of sites. In the majority of cases, populations among fields at a given site were differentiated, and in some cases, differentiation was even detected within individual rice fields. In contrast, populations of *Rhizobium leguminosarum* collected from two sites 1 km apart, but from different plant communities, showed similar genetic structures (23).

Spatial differentiation of pathogen populations could result from selection or genetic drift. Selection could result from differences in the rice cultivars planted in different locations or other aspects of the environment and agricultural practice. We attempted to evaluate the role of selection by comparing differentiation between sites planted to the same or planted to different rice cultivars. Although 97% (66/68) of site-pairs planted to different hosts showed significantly differentiated pathogen populations, only 40% (4/10) of site-pairs planted to IR72 showed significantly differentiated populations of *X. oryzae* pv. *oryzae*. Differentiation attributable to host selection could result from differences in major genes for resistance or to minor genetic factors differentially affecting pathogen fitness.

Movement of pathogen subpopulations tends to counteract the forces leading to differentiation. Some haplotypes were widely distributed, suggesting the occurrence of migration. Fourteen *Pst*I haplotypes were found at more than one site. The most widely distributed among the *Pst*I haplotypes was D-007, which was found at eight sites located in four provinces. Five IS*1113* haplotypes were present at more than 1 site, and one of these haplotypes, B-005, was found at 10 sites located in all six provinces sampled in this study. However, the extent of dispersal along this transect appeared to be insufficient to counteract differentiation at all spatial scales examined. With a relatively low rate of migration, selection and genetic drift would lead to spatial differentiation among subpopulations.

Effect of host genetic diversity on pathogen diversity. We began with the hypothesis that pathogen diversity is affected by host diversity. We assumed that host diversity would be greater in the highland part of the transect, where many diverse traditional cultivars are still planted by the Ifugao tribes (25), rather than in the lowland portion of the transect, where the rice cultivars are more homogeneous. McArthur et al. (11) showed that the genetic diversity of natural populations of *Pseudomonas cepacia* was positively correlated with habitat variability (e.g., organic and inorganic nutrients or soil pH) across a landscape gradient.

Our results, however, do not strongly support the hypothesis that pathogen haplotype diversity was affected by host diversity. First, we compared the diversity of pathogen collections from different agroecosystems. Ten IS*1113* haplotypes were localized in the highland sites, six were present at the transition zone sites, and seven were detected at the lowland sites. The overall genetic diversity of the total collection from each of the agroecoystems was compared, but no difference in diversity was evident (H_S = 0.61 for sites A + B in the highlands; H_S = 0.60 for sites C + D + E + F in the transition zone; and H_S = 0.67 for sites G + H + I + J + K + L + M in the lowlands).

To compare the diversity of collections from traditional versus improved host cultivars, we pooled pathogen collections across ecosystems according to their host of origin. We detected a slight difference in the diversity of pathogen collections from different types of hosts (traditional versus improved). Strains from traditional cultivars (planted at sites A, B, C, F. and M; H_{TV} = 0.77) appeared to be more diverse than the subpopulations derived from the improved cultivars (planted at sites D through M; sites F and M had some fields planted to improved cultivars and other fields planted to traditional cultivars [Table 1]; H_{IV} = 0.63). However, because the statistical significance of this difference could not be tested, we could not draw firm conclusions regarding the effect of host types on pathogen diversity.

Based on our results, we tentatively conclude that host diversity does not strongly affect pathogen diversity. Two caveats, however, should be considered. First, our assumption regarding the functional diversity (in terms of disease resistance) of rice cultivars grown in the traditional and improved agroecosystems may be incorrect. Second, diversity at loci involved in host-pathogen interactions could be affected by differences in host population structure,

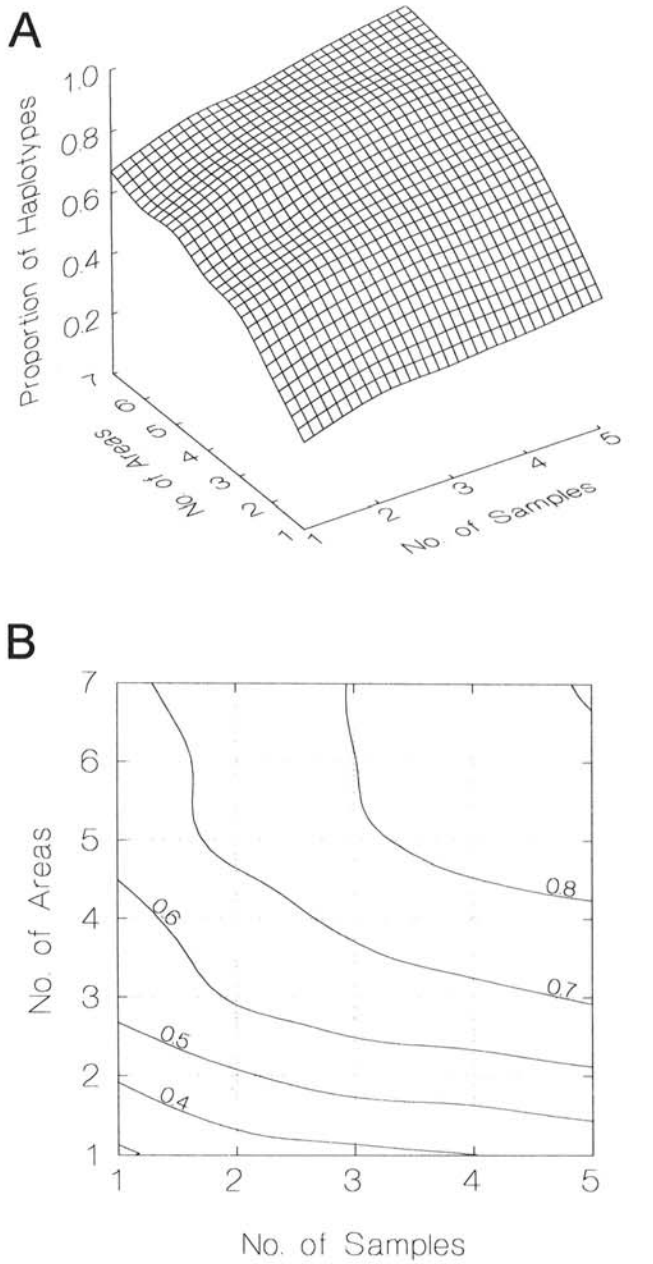

Fig. 6. Results of *Xanthomonas oryzae* pv. *oryzae* subsampling simulation of 39 within-field data sets. **A,** Surface plot and **B,** contour plot, showing the proportions of haplotypes detected as a function of the number of areas sampled per field and number of samples taken per area. Both plots were smoothed using distance-weighted least squares smoothing. The contour plot shows that nearly 80% of the *Pst*I haplotypes could be detected by sampling six areas and taking three samples per area.

whereas haplotypic diversity as defined by our DNA fingerprinting techniques are not. These issues could be addressed when allele-specific diagnostic tools for the relevant resistance and avirulence genes are available.

Effect of sampling intensity on diversity detected. In previous studies, a set of Philippine strains of *X. oryzae* pv. *oryzae* collected from around the Philippines over a 17-year period was analyzed with various molecular markers (9,19,20). Nearly the same set of approximately 150 strains was utilized in each of these studies to allow comparison of the different marker systems explored. These strains were mostly drawn from a larger collection of 960 strains that had been previously subjected to pathotypic analysis by Mew et al. (14) to represent each of the pathotypes detected. In these studies, five groups of strains (presumably representing lineages) were defined by cluster analysis of RFLP data.

Although the sampling strategy utilized in the present study was more systematic and intensive than those used in previous DNA fingerprinting studies and the number of strains analyzed was an order of magnitude larger, no "new" lineages of the pathogen were detected (19,20). Most of the haplotypes detected in this study had been detected previously. Thus, it seems that the bacterial blight pathogen population is not infinitely variable, and populations of the pathogen probably can be described with reasonable accuracy using a limited sampling strategy.

It should be noted that although the same lineages were detected in this study as in previous studies, the relative proportion of pathotypes (and presumably of lineages) was different from that previously reported. All of the 62 strains of *X. oryzae* pv. *oryzae* collected by Mew et al. (14) from the central Luzon provinces of Nueva Ecija, Tarlac, and Bulacan in 1985, 1986, and 1988 were of pathotype 2. In contrast, among the 542 strains analyzed from these provinces in the present study, only 213 (39%) were of pathotype 2, whereas the majority (329 or 61%) were of pathotype 3. There apparently has been a shift of pathotypes in central Luzon in recent years.

The results of this study provide a better understanding of the biology of the bacterial blight pathogen in the Philippines. The specific population structure we observed in Philippine populations of *X. oryzae* pv. *oryzae*, however, may not reflect the diversity and differentiation of pathogen populations in other countries where bacterial blight is an important constraint to rice production. Pathogen collections from various Asian countries varied substantially, and those from the Philippines were distinct from those from the other countries studied (1). The results of our data subsampling exercise could provide useful guidelines for designing optimal strategies for sampling *X. oryzae* pv. *oryzae* populations in other countries. Based on our results, less intensive sampling could be conducted per field. The differentiation detected in the present study suggests that it is important to sample many sites and that the strains collected should be well distributed per site.

Knowledge of pathogen population structure is one of the several prerequisites for designing deployment strategies. An understanding of pathogen population biology should have both spatial and temporal components. In this study, we have provided evidence for spatial differentiation and pathogen migration. It will be important to obtain further information on the rates of selection and migration, the genetics of host resistance, and farmers' preferences before specific recommendations on gene deployment can be made. A sequel to this study is underway to investigate the effect of host selection on the population structure of the bacterial blight pathogen.

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. Birnboim, H. C., and Doly, J. 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. 7: 1513-1523.
3. Crow, J. F. 1986. Basic Concepts in Population, Quantitative and Evolutionary Genetics. W.H. Freeman and Co., San Francisco.
4. Felsenstein, J. 1985. Confidence limits on phylogenies: An approach using the bootstrap. Evolution 39:783-791.
5. Hedges, S. B. 1992. The number of replications needed for accurate estimation of the bootstrap *P* value in phylogenetic studies. Mol. Biol. Evol. 9:366-369.
6. Hudson, R. R., Boos, D. D., and Kaplan, N. L. 1992. A statistical test for detecting geographic subdivision. Mol. Biol. Evol. 9:138-151.
7. Karganilla, A. D., and Natural, M. P. 1973. A comparative study of culture media for *Xanthomonas campestris* pv. *oryzae.* Philipp. Agric. 57: 141-152.
8. Kauffman, H. E., Reddy, A. P. K., Hsieh, S. P. Y., and Merca, S. D. 1973. An improved technique for evaluation of resistance of rice varieties to *Xanthomonas oryzae*. Plant Dis. Rep. 57:537-541.
9. 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 element. Appl. Environ. Microbiol. 58:2188-2195.
10. Leung, H., Nelson, R. J., and Leach, J. E. 1993. Population structure of plant pathogenic fungi and bacteria. Adv. Plant Pathol. 10:157-205.
11. McArthur, J. V., Kovacic, D. A., and Smith, M. H. 1988. Genetic diversity in natural populations of a soil bacterium across a landscape gradient. Proc. Natl. Acad. Sci. USA 85:9621-9624.
12. Mew, T. W. 1987. Current status and future prospects on research on bacterial blight of rice. Annu. Rev. Phytopathol. 25:359-382.
13. Mew, T. W., and Vera Cruz, C. M. 1979. Variability of *Xanthomonas oryzae*: Specificity in infection of rice differentials. Phytopathology 69: 152-155.
14. 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.
15. Murray, M. G., and Thompson, W. F. 1980. Rapid isolation of high-molecular weight plant DNA. Nucleic Acids Res. 8:4321-4325.
16. Nei, M. 1973. Analysis of gene diversity in subdivided populations. Proc. Natl. Acad. Sci. USA 70:3321-3323.
17. Nei, M. 1987. Molecular Evolutionary Genetics. Columbia University Press, New York.
18. Nei, M., and Tajima, F. 1981. DNA polymorphism detectable by restriction endonucleases. Genetics 97:145-163.
19. 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.
20. Raymundo, A. K., Ardales, E. Y., Yap, I. V., Baraoidan, M. R., Mew, T. W., and Nelson, R. J. 1993. Analysis of genetic variation in *X. oryzae* pv. *oryzae* using *Pst*I digestion. Philipp. J. Biotechnol. 4:9-27.
21. Rohlf, F. J. 1992. NTSYS-pc., version 1.7. Numerical Taxonomy and Multivariate Analysis System. Exeter Publishing, Ltd., Setauket, New York.
22. Sneath, P. H. A., and Sokal, R. R. 1973. Numerical Taxonomy: The Principles and Practice of Numerical Classification. W.H. Freeman and Co., San Francisco.
23. Strain, S. R., Leung, K., Whittam, T. S., De Bruijn, F. J., and Bottomley, P. J. 1994. Genetic structure of *Rhizobium leguminosarum* biovar *trifolii* and *viciae* populations found in two Oregon soils under different plant communities. Appl. Environ. Microbiol. 60:2772-2778.
24. 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. 40:309-311.
25. Vera Cruz, C. M., Nelson, R., Leung, H., Leach, J. E., and Mew, T. W. 1992. Reaction of rice cultivars from Ifugao Province, Philippines, to indigenous strains of the bacterial blight pathogen. Int. Rice. Res. Inst. Newsl. 17:8.