Gene Flow Between Geographic Populations of *Mycosphaerella graminicola* (Anamorph *Septoria tritici*) Detected with Restriction Fragment Length Polymorphism Markers

J. M. Boeger, R. S. Chen, and B. A. McDonald

Department of Plant Pathology and Microbiology, Texas A&M University, College Station 77843-2132.

ABSTRACT

Boeger, J. M., Chen, R. S., and McDonald, B. A. 1993. Gene flow between geographic populations of *Mycosphaerella graminicola* (anamorph *Septoria tritici*) detected with restriction fragment length polymorphism markers. Phytopathology 83:1148-1154.

We used DNA restriction fragment length polymorphism (RFLP) markers to assess the potential for gene flow between geographically separated populations of *Mycosphaerella graminicola*. We used 13 RFLP loci and DNA fingerprints to compare the genetic structure of Oregon and California populations of *M. graminicola* separated by 750 km. The two populations shared the majority of frequent alleles at a number of loci. Many alleles were present at similar frequencies. The populations had similar levels of genetic diversity. Genetic and genotypic variation were distributed on a very small scale in both populations. An indirect measure of gene flow based on Wright's island model (Nm) estimated the number of individuals that have migrated between the populations as 12, which is sufficient to prevent the populations from differentiating significantly. Gene flow on this scale could have significant implications for plant breeding programs.

Additional keywords: epidemiology, migration.

Gene flow is the exchange or movement of gametes, individuals, and populations on a geographic scale (30). Although gene flow may have important implications for the control of plant disease, few studies have assessed the impact of gene flow on the population biology of plant pathogens. Gene flow affects the movement of all genes, including those involved in virulence or resistance to fungicides, among populations of plant pathogens. The movement of genes can either limit the ability of populations to adapt to local conditions or promote the spread of genes and genotypes throughout a large geographic area (29,30). Although a positive correlation has been demonstrated between the geographic and the genetic distance between populations (e.g., 28), this correspondence is by no means absolute, especially for species that have the potential for widespread dispersal (14). Species that disperse genes over large geographic distances may possess a high degree of genetic similarity throughout their range.

Gene flow may have important consequences for the deployment of resistance genes or fungicides in agricultural ecosystems (15). If gene flow between populations is limited, then regional resistance gene deployment strategies that focus on local populations are more likely to be successful. Conversely, if there is extensive gene flow between populations separated by long distances, then plant breeders conducting trials for resistance should consider virulence present in distant populations as well as local populations. These same principles apply to the screening of populations of fungal pathogens for resistance to fungicides.

Gene flow can be measured directly or indirectly. The movement of individuals between populations can be observed directly with the help of DNA fingerprinting methods that can distinguish between individuals without ambiguity (5,9,13,20). However, such direct observations, while possible, are impractical for many plant pathogens because it is unlikely that the same individual will be found in two populations unless 1) asexual reproduction is a significant component of the pathogen life cycle, 2) asexual progeny are widely dispersed over long distances, and 3) the clones that result from asexual reproduction occur at a high enough frequency to be detected in both populations.

Indirect measures of gene flow rely on the analysis of gene frequencies in populations to estimate the effects of gene flow averaged over time (31). Indirect methods have been used to estimate gene flow among populations of many eukaryotes (e.g., 10,11,14,25,34) but have only rarely been applied to plant pathogens.

Mycosphaerella graminicola (Fuckel) J. Schröt. in Cohn (anamorph *Septoria tritici* Roberge in Desmaz.) causes Septoria tritici leaf blotch, a significant disease of wheat worldwide (12). The most economical way to control the disease is to use resistant cultivars (7,33). Eyal et al (8) compared the virulence of 97 isolates originating from 22 countries. They concluded that populations of *M. graminicola* from different countries differ significantly in virulence, and they suggested that regional gene deployment strategies are likely to be successful for disease control. Information on the potential for gene flow between geographic populations could provide a more reliable indicator of the likelihood for success of regional resistance gene deployment strategies.

We have developed DNA markers to study the population genetics of *M. graminicola*. Using single-locus restriction fragment length polymorphisms (RFLPs) and DNA fingerprints, we showed that a California population of *M. graminicola* from a single field was arranged in a fine-scaled mosaic of overlapping clones (17,20). Our previous results suggested that the asexual spores do not travel long distances in the field and that the primary inoculum probably was ascospores of *M. graminicola*. Although the teleomorph has been identified in many locations around the world (12), the mating control system remains unknown. Analyses of associations among loci suggested that the sexual stage is an important component of the population biology of *M. graminicola* (16).

Our objectives in the experiment described here were to use RFLP markers to compare the genetic structure of the California population with that of a geographically distant population and to assess the potential for gene flow between geographically separated populations of *M. graminicola*. We present evidence that populations separated by 750 km have nearly identical genetic structures. We interpret this finding as evidence for extensive gene flow between these populations.

MATERIALS AND METHODS

***M. graminicola* isolates.** A hierarchical sampling method was used to collect *M. graminicola* isolates from a naturally infected

wheat field in Davis, CA, in 1989 (17). The 93 isolates in this collection originated from 35 lesions on 19 leaves sampled from seven locations in a single field. Each field location was approximately 1 m² in area. The genetic structure of this population was described previously (17).

A different strategy was used to sample a population of *M. graminicola* from a field experiment in Corvallis, OR, in 1990. Collections were made from a cultivar mixture experiment in which four wheat cultivars that differed in resistance to *M. graminicola* were planted in pure stands and in all possible two-, three-, and four-way mixtures (15 treatments total) in a randomized complete block design with three replications. The results that compared the effects of host genotype on *M. graminicola* popu-

lations will be published elsewhere. For the purposes of assessing gene flow between populations, all isolates collected from this field in 1990 were treated as a single population.

Wheat plots measuring 1.5×6.1 m were planted on 16 October 1989 in a field that had been fallow the previous year. Before that, this field was planted to annual crops that are not hosts of *M. graminicola*. The nearest wheat fields were breeders' seed increases 2 km away. All wheat plots showed symptoms of uniform, natural infection by *M. graminicola* within 6 wk of seedling emergence. Twenty infected leaves were randomly chosen from each plot midway through the growing season (6 March) and again near the end of the growing season but before leaf senescence (30 May). The infected leaf tissue was air-dried at room tem-

TABLE 1. Clone-corrected allele frequencies at restriction fragment length polymorphism (RFLP) loci in California and Oregon populations of *Mycosphaerella graminicola* [a]

RFLP locus	Allele [b]	Population CA	Population OR	RFLP locus	Allele [b]	Population CA	Population OR
pSTS192 *Xho*I-A	1	0.73	0.927		8	0	0.002
	2	0.27	0.045		9	0	0.002
(N = 179)	3	0	0.006		10	0	0.007
	4	0	0.006		11	0	0.005
	5	0	0.006		12	0	0.002
	6	0	0.006		13	0	0.005
	7	0	0.006		14	0	0.002
pSTS192 *Xho*I-B	1	1.00	0.995		15	0	0.002
(N = 183)	2	0	0.006	pSTL53 *Pst*I	1	0.50	0.475
pSTS192 *Pst*I-A	1	0.64	0.943	(N = 404)	2	0.18	0.119
(N = 401)	2	0.36	0.030		3	0.14	0.161
	3	0	0.005		4	0.05	0
	4	0	0.003		5	0.09	0.042
	5	0	0.008		6	0.05	0.119
	6	0	0.008		7	0	0.027
	7	0	0.003		8	0	0.042
	9	0	0.003		9	0	0.003
SS192 *Pst*I-B	1	1.00	0.980		12	0	0.007
(N = 406)	2	0	0.015		13	0	0.003
	3	0	0.003		14	0	0.003
	4	0	0.003	pSTL53 *Xho*I	1	0.50	0.759
SS14 *Pst*I	1	0.91	0.808	(N = 170)	2	0.14	0.159
(N = 407)	2	0.09	0.187		3	0.14	0.035
	3	0	0.003		4	0.05	0
	4	0	0.003		5	0.14	0
SS196 *Xho*I	1	0.82	0.715		6	0.05	0.006
(N = 179)	2	0.18	0.257		8	0	0.006
	3	0	0.022		9	0	0.006
	4	0	0.006		10	0	0.006
SS2 *Pst*I	1	0.86	0.648		14	0	0.012
(N = 400)	2	0.09	0.085		15	0	0.006
	3	0.05	0.203		16	0	0.006
	4	0	0.018	pSTS199 *Pst*I	1	0.45	0.327
	5	0	0.005	(N = 266)	2	0.41	0.451
	6	0	0.003		3	0.14	0.026
	7	0	0.010		4	0	0.004
	8	0	0.003		5	0	0.124
	9	0	0.015		6	0	0.045
	10	0	0.010		7	0	0.008
	11	0	0.003		8	0	0.008
SL2 *Xho*I	1	0.59	0.349		9	0	0.008
(N = 175)	2	0.36	0.394	pSTS197 *Pst*I	1	0.73	0.632
	3	0.05	0.074	(N = 269)	2	0.27	0.227
	4	0	0.177		3	0	0.108
	5	0	0.006		4	0	0.007
pSTL10 *Pst*I	1	0.91	0.680		5	0	0.004
(N = 409)	2	0.05	0.051		6	0	0.004
	3	0.05	0.220		7	0	0.007
	4	0	0.015		8	0	0.007
	6	0	0.002		9	0	0.004
	7	0	0.002				

[a] Each RFLP locus is defined by a probe-enzyme combination. The clone-corrected sample size used to calculate allele frequencies in the California population is $N = 22$ for each locus. Clone-corrected sample sizes used to calculate allele frequencies in the Oregon population are shown below each RFLP locus.

[b] RFLP allele refers to a DNA fragment or set of fragments of a particular size that hybridize with a probe. Each allele has been assigned a unique number. Only alleles present in the Oregon or California population are given. Several alleles (e.g., allele 8 at the pSTS192 *Pst*I-A locus) are unique to other geographic populations and so are missing here.

perature for 2 wk before isolations were made. One single-spore isolation of *M. graminicola* was made from each leaf that expressed a viable cirrus. Although many leaves did not produce viable cirri, 728 isolates originating from different leaves were collected in Corvallis in 1990. RFLP data were collected from 444 of these isolates.

In a separate experiment, multiple isolations of *M. graminicola* were made from 10 leaves in the Oregon sample to allow a comparison of the microgeographic distribution of genetic variation in the California and Oregon populations. To analyze genetic variation among lesions on a leaf, isolations were made from four different lesions on each of four leaves that had a number of discrete, nonoverlapping lesions. To analyze genetic variation among pycnidia within a lesion, single-spore isolations were made from four different pycnidia within each lesion from four leaves that had well-defined lesions. Finally, to analyze genetic variation within a pycnidium, 10 single-spore isolations were made from a single pycnidium on each of two leaves.

DNA extractions, probes, and hybridization. DNA was extracted from each isolate by a CTAB extraction protocol described previously (17). Purified DNA (5 μg) was digested individually with the restriction enzymes *Pst*I or *Xho*I. DNA fragments were separated on 0.75% agarose gels and then transferred to nylon membranes by the alkaline transfer method (24), as recommended by the manufacturer (BioRad, Hercules, CA).

Development of anonymous DNA probes used for RFLP analysis was described previously (18). These probes came from a partial *Sau*3A digestion of total DNA from *M. graminicola*. DNA fragments were size-fractionated and then cloned into a pGEM4 plasmid vector (Promega, Madison, WI). Four hundred randomly chosen clones were kept for use as probes. After an initial screen of 32 of these probes, we selected 10 probes that hybridized to single loci and three probes that hybridized to dispersed, repetitive DNA sequences for genetic analysis of *M. graminicola* populations. We showed previously that several of the repetitive probes are useful for DNA fingerprinting (20). We previously used the single-locus probes to determine the genetic structure of the *M. graminicola* population sampled in California (17).

Probes were labeled by nick translation following the manufacturer's recommendations (BRL, Gaithersburg, MD). The single-locus probes used for this experiment are shown in Table 1. We demonstrated previously that probe pSTS192 hybridizes simultaneously to two loci on different chromosomes (19,20). In this experiment, we treated each of these loci separately. DNA fingerprint probe pSTL70 was used in addition to pSTL40. Probe pSTL70 hybridizes to more fragments on more chromosomes and provides a more uniform profile of restriction fragments than pSTL40 with the restriction enzyme *Pst*I (Fig. 1).

Data analysis. Each probe-enzyme combination defined an RFLP locus. DNA fragments or combinations of fragments of different sizes were treated as alleles at each RFLP locus. The number of isolates used in each analysis varied because of differences in the sampling methods for the two populations. Sample sizes for each locus also varied because data from some isolates were incomplete as a result of partial digestions, differences in the amount of DNA loaded in each lane, and occasional nonspecific background hybridization. Only alleles that could be scored unambiguously were included in each analysis. Midway through the collection of RFLP data from the Oregon isolates, we stopped using *Xho*I to digest DNA from each isolate in order to lower the cost of data collection. As a result, sample sizes are larger for all RFLP loci that use *Pst*I than for those that use *Xho*I.

Most population genetic analyses use allele frequency data from each locus. In the California population, almost three isolations were made from each lesion, and nearly five isolates were collected from each leaf on average. As a result of this sampling strategy, many clones were resampled several times. In the Oregon population, only one isolation was made from each leaf, and few clones were present more than once in the sample. In each population, isolates with the same multilocus haplotype (i.e., with the same alleles at each of the single RFLP loci) were compared with DNA fingerprints. Isolates with the same DNA fingerprint and multi-

locus haplotype were assumed to be individual members of the same clone and were counted only once in the analysis. To compare allele frequencies in the two populations, we used only one representative of each clone to calculate a clone-corrected allele frequency. The 93 isolates in the California population comprised 22 different genotypes, so only these 22 genotypes were used to calculate allele frequencies. In the Oregon population, the clone-corrected sample size ranged from 170 to 409 individuals for each RFLP locus.

We used the same methods described previously (17) to measure genetic variation in the Oregon and California populations. Single-locus measures of genetic diversity included the number of alleles per locus and Nei's (22) measure of gene diversity. In addition, we used a measure of genotypic diversity based on the number
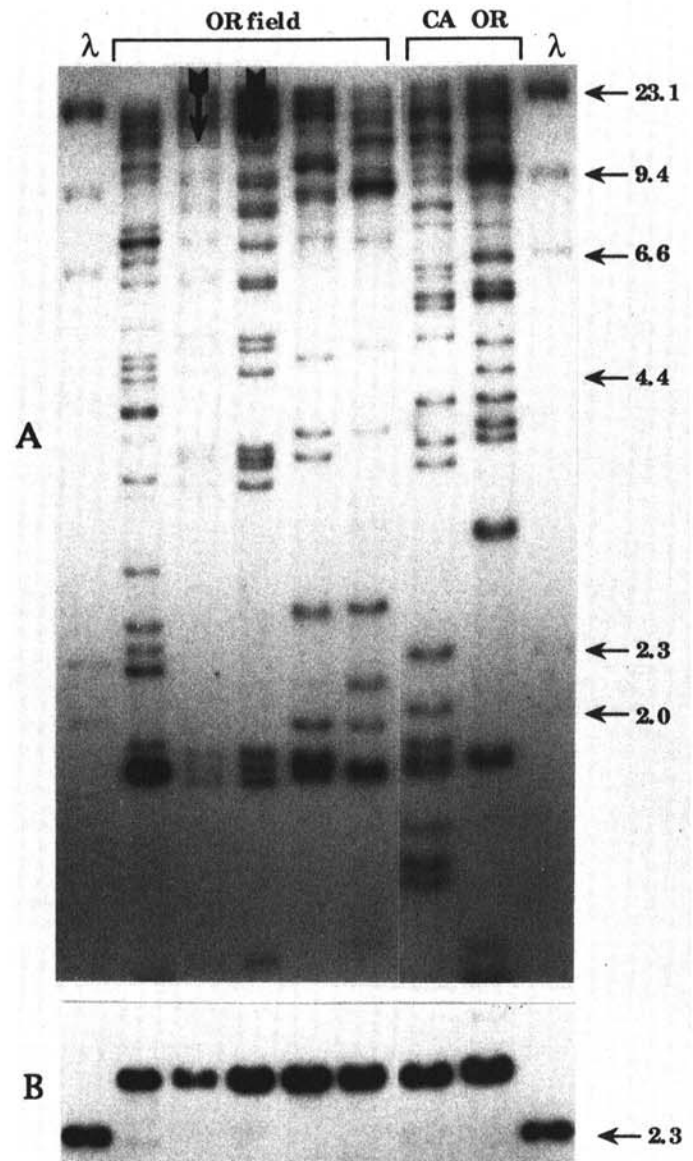

Fig. 1. DNA fingerprints of *Mycosphaerella graminicola* isolates that have the same multilocus haplotype within and between populations. Panel A: Probe pSTL70 was hybridized to DNA from isolates digested with *Pst*I. The first set of lanes shows the DNA fingerprints of isolates from Oregon (OR) that have the same alleles at 11 restriction fragment length polymorphism (RFLP) loci. Vertical arrows indicate isolates that have the same DNA fingerprint. These isolates were collected from the same plot. The second set of lanes shows isolates from California (CA) and Oregon that have the same alleles at 10 RFLP loci. Panel B: DNA from the same isolates hybridized with probe pSTL10. Each of the isolates has allele 1, the most common allele at the pSTL10 *Pst*I RFLP locus. The first and last lanes are λ *Hin*dIII size standards.

of multilocus haplotypes in each population. The multilocus haplotype is a summary of the alleles present at each RFLP locus in each individual (17). Genotypic diversity $\hat{G}$ in each population was calculated from the following formula:

$$\hat{G} = 1 \: / \: \Sigma_{x=0}^{N}[f_x \cdot (x/N)^2] \, ,$$

where N is the sample size and f_x is the number of multilocus haplotypes observed x times in the sample (32). The maximum possible value for $\hat{G}$, which occurs when each individual in the sample has a different genotype, is the number of individuals in the sample. To compare $\hat{G}$ in populations with different sample sizes, we divided $\hat{G}$ from each population by N to calculate the percentage of maximum diversity that was obtained. For this analysis, we used only isolates with complete data for each RFLP locus: complete haplotypes were available for 406 isolates in the Oregon population and for all 93 isolates in the California population.

We used hierarchical gene diversity analysis (1) to partition the distribution of gene diversity within and among plots in the Oregon population. This partitioning of genetic variation was analogous to the analysis used to measure variation within and among sampling locations in the California field (17). In the California population, seven sampling locations along two transects were used to partition variation. In the Oregon population, each of the 26 plots sampled in this experiment (from the 45 available) was treated as a separate location.

We assessed the genetic similarity of the Oregon and California populations in several ways. Individual RFLP loci were compared directly in terms of allele frequencies, number of alleles per locus, Nei's diversity (22), Nei's genetic distance and identity (21), and Nei's G_{st} (22). We also calculated average values across all 13 RFLP loci.

The amount of gene flow between populations—Nm, where N is the population size and m is the fraction of individuals in a population that are immigrants—was estimated by substituting Nei's G_{st} for F_{st} in Wright's island model of gene flow (35). According to Wright's model,

$$Nm = 1 \: / \: 4[(1/F_{st}) - 1] \, .$$

The formula we used to calculate Nm was

$$Nm = 1 \: / \: 2[(1/G_{st}) - 1] \, .$$

The 4 was replaced with a 2 because *M. graminicola* is haploid.

If two populations are similar in size, then Nm estimates the average number of individuals that migrate between the populations per generation. Wright (35) showed that movement of just one individual per generation between two populations was sufficient to prevent the fixation of different selectively neutral alleles

in those populations. This result is independent of population size, because the force of gene flow, which is measured by the fraction m of immigrants in a population, is counteracted by the force of genetic drift, which is proportional to the inverse of population size N. If $Nm < 1$, then local populations will differentiate; if $Nm \geq 1$, then there will be little differentiation among populations.

We used DNA fingerprints to determine the genetic identity or nonidentity of isolates to analyze the microgeographic distribution of genotypes among pycnidia within a lesion and among lesions on a leaf. Isolates with the same DNA fingerprints were assumed to be genetically identical, as we demonstrated in previous experiments (20).

RESULTS

Table 1 summarizes the clone-corrected allele frequencies at RFLP loci in the California and Oregon populations. Sixty-five RFLP alleles present in the Oregon population were not found in the smaller sample of isolates from California. On average, the Oregon population had 7.8 alleles per RFLP locus, compared with 2.8 alleles at the same loci in the California population. Despite the difference in number of alleles, Nei's measure of genetic diversity across all loci was similar for both populations ($H = 0.356$ for California and 0.389 for Oregon) (Table 2).

Among the 406 Oregon isolates with complete data from individual RFLP loci, 315 multilocus haplotypes were found. No haplotype was present more than six times, which represents less than 2% of the sample. In all cases where a haplotype was present four or more times, common alleles (usually the one at highest frequency) were present at all loci. Genotypic diversity was greater in the Oregon population ($\hat{G} = 253$, or 62% of the theoretical maximum of 406) than in the California population ($\hat{G} = 13$, or 14% of the theoretical maximum of 93).

Many isolates in the Oregon population with the same multilocus haplotypes had different DNA fingerprints (Fig. 1). Hence, our estimate of genotypic diversity, which was based on multilocus haplotypes, underestimates the actual level of genotypic diversity. Most of the isolates with the same haplotype had the most frequent allele at every locus, which is in accord with expectations for a random-mating population. In four cases, isolates in California and Oregon had the same alleles at 11 RFLP loci, but in every case the isolates had different DNA fingerprints (Fig. 1).

Estimates of Nei's measures of gene diversity, genetic identity, genetic distance, G_{st}, and Nm for all loci are summarized in Table 2. Hierarchical gene diversity analysis showed that on average 93% of the genetic variation in the Oregon population was distributed within plots (Table 3), compared with 57% within locations in the California population (17).

In tests for variation among lesions on a single leaf, we found four genotypes on one leaf, three on another leaf, and two on

TABLE 2. Nei's measures of genetic diversity, probability of identity between population (J_{ij}), normalized genetic identity (I), genetic distance (D), and estimates of population differentiation (G_{st}) and amount of gene flow (Nm) between Oregon and California populations of *Mycosphaerella graminicola*.

RFLP locus	Genetic diversity		J_{ij}	I	D	G_{st}	Nm
	California	Oregon					
SS192 *Xho*I-A	0.394	0.139	0.689	0.953	0.048	0.076	6.1
SS192 *Xho*I-B	0.000	0.011	0.995	1.000	0.000	0.000	—[a]
SS192 *Pst*I-A	0.461	0.110	0.614	0.887	0.120	0.150	2.8
SS192 *Pst*I-B	0.000	0.039	0.980	1.000	0.000	0.025	19.5
SS14 *Pst*I	0.164	0.312	0.752	0.992	0.008	0.021	23.3
SS196 *Xho*I	0.295	0.422	0.633	0.991	0.009	0.012	41.1
SS197 *Pst*I	0.394	0.537	0.523	0.987	0.014	0.014	35.2
SS2 *Pst*I	0.250	0.532	0.575	0.970	0.030	0.042	11.4
SL2 *Xho*I	0.520	0.686	0.351	0.905	0.100	0.037	13.0
SL10 *Pst*I	0.168	0.487	0.631	0.965	0.035	0.059	8.0
SL53 *Pst*I	0.686	0.716	0.291	0.973	0.027	0.006	82.8
SL53 *Xho*I	0.690	0.397	0.407	0.940	0.062	0.045	10.6
SS199 *Pst*I	0.610	0.671	0.336	0.938	0.064	0.018	27.3
Average	0.356	0.389	0.598	0.962	0.040	0.039	12.3

[a] This value approaches infinity but cannot be determined because it requires division by zero.

each of the other leaves. These results were similar to our findings of microgeographic variation in the California population (17). In tests for variation within a single lesion, we found that two lesions had three different genotypes, one lesion had two genotypes, and the fourth lesion had only one genotype (Fig. 2). Isolates originating from the same pycnidium had the same DNA fingerprint (Fig. 3).

DISCUSSION

Genetic diversity. The large difference in the number of alleles per locus suggests that the Oregon population had more genetic variation than the California population, but the difference in allelic diversity is probably due solely to differences in sample size and sampling strategy between the two populations. The complete California sample included only 93 isolates from 19 leaves, compared to 444 isolates, each from a different leaf, in the complete Oregon sample. All but three of the alleles found in the California population were also found in the Oregon population, and two of these private alleles were unusual duplications at the pSTL53 loci that we described previously (19). Most of the additional alleles present in the Oregon population occurred at a low frequency, often in only a single individual.

The genotypic diversity of the Oregon population, which was 62% of its maximum value, was much higher than that of the California population, which was only 14% of its maximum value.

TABLE 3. Distribution of genetic variation within and among plots in the Oregon population of *Mycosphaerella graminicola*[a]

RFLP locus	Total diversity	Within-plot diversity	Among-plot diversity
SS192 *Xho*I-A	0.139	0.94	0.06
SS192 *Xho*I-B	0.011	1.00	0.00
SS192 *Pst*I-A	0.110	0.93	0.07
SS192 *Pst*I-B	0.039	0.92	0.08
SS14 *Pst*I	0.312	0.92	0.08
SS196 *Xho*I	0.422	0.88	0.12
SS197 *Pst*I	0.537	0.94	0.06
SS2 *Pst*I	0.532	0.96	0.04
SL2 *Xho*I	0.686	0.92	0.08
SL10 *Pst*I	0.487	0.90	0.10
SL53 *Pst*I	0.716	0.95	0.05
SL53 *Xho*I	0.397	0.97	0.03
SS199 *Pst*I	0.671	0.92	0.08
Average	0.389	0.93	0.07

[a] Total diversity was partitioned into components within and among plots by the method of Beckwitt and Chakraborty (1). Diversity within and among plots is shown as a proportion of the total diversity at each RFLP locus.

Again, this difference between the populations can be accounted for by the differences in sampling methods. Most of the isolates in the California sample with the same multilocus haplotype were isolated from the same leaf. In the California field, five clones accounted for more than 50% of the 93 isolates (17). These few clones, which were sampled many times from the same location in the field, undoubtedly gave a substantial downward bias to the measure of genotypic diversity in the California population.

The measure of genotypic diversity in the Oregon population also underestimates the actual diversity because it is based on single-locus data. Most of the isolates with the same multilocus haplotype had different DNA fingerprints (Fig. 1) and hence were different genotypes. Isolates with the same DNA fingerprint were always sampled from the same plot in the Oregon field. In California, we found only one case where the same genotype was present in two different locations in the field (17). The results from Oregon provide additional evidence for our hypothesis that asexual spores usually are not dispersed more than a few meters over the course of a growing season.

Genetic structure. The California and Oregon populations had very similar genetic structures. The analysis of genetic diversity in the Oregon population illustrates that the arrangement of *M. graminicola* in a field consists of a fine-scaled mosaic of overlapping clusters of clones, where each small cluster results from splash dispersal of asexual pycnidiospores, which are the dominant secondary inoculum. This result supports previous epidemiological studies (26,27) that suggested that ascospores are much more important than pycnidiospores as a source of primary inoculum.

Considering the differences in sample size and sampling methods, the overall amount and distribution of genetic diversity were remarkably similar in the two populations. Averaged across 13 RFLP loci, Nei's diversity differed by only 3% in the two populations, with Oregon isolates showing more variability than California isolates. For both populations, most of the diversity was distributed within locations in the field, with a smaller proportion distributed between locations in the field. Although a larger proportion of the total diversity in the Oregon population was accounted for by differences within plots, this difference is most likely due to differences in sampling strategies.

The scale at which diversity could be detected was the same in both populations: different genotypes were often found within

Fig. 3. DNA fingerprints from 10 independent single-pycnidiospore isolations from two pycnidia of *Mycosphaerella graminicola*. The first, middle, and last lanes are λ*Hind*III size standards. Probe pSTL40 was hybridized to DNA digested with *Pst*I. All isolates from the same pycnidium have the same DNA fingerprint, demonstrating that DNA fingerprints are stable through mitosis.
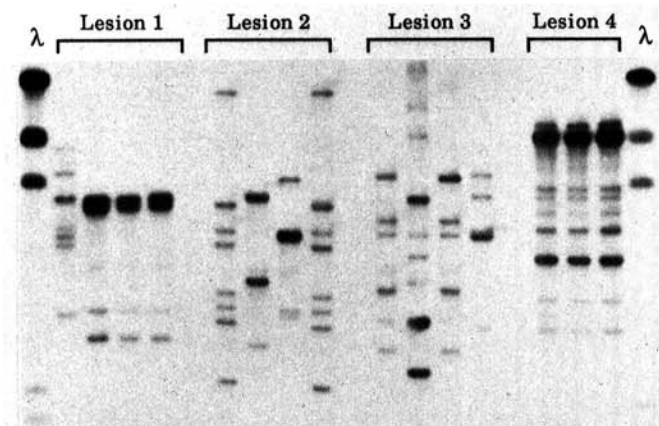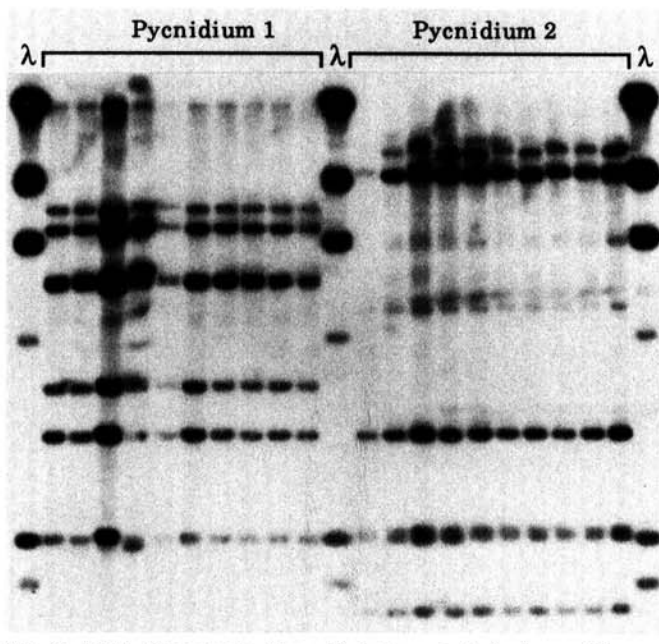

Fig. 2. DNA fingerprints of *Mycosphaerella graminicola* isolates collected from different pycnidia within each of four lesions from different leaves. Probe pSTL40 was hybridized to DNA digested with *Pst*I. The first and last lanes are λ *Hind*III size standards. Two or three genotypes were present in lesions 1–3; the fourth lesion had only one genotype.

a single lesion, and most lesions on the same leaf had different genotypes (Fig. 2). This result demonstrates that a lesion is often the result of coinfection by two or more genotypes, confirming our previous observations (17). The lack of variation among pycnidiospores originating from the same pycnidium (Fig. 3) shows that DNA fingerprints were stable through mitosis, also confirming our previous conclusions based on 10 sequential single-spore transfers to fresh growth media (20).

Genetic similarity. The Oregon and California populations were similar by all measures used for comparison. The same alleles were found in both populations, often at similar frequencies (Table 1). The average gene differentiation ($G_{st} = 0.039$) and the overall genetic distance ($D = 0.040$) between the populations were surprisingly small. Nei's normalized genetic identity (I) was 0.96 across all loci (Table 2), compared to the theoretical maximum of 1.00, which is reached when two populations have the same alleles present at the same frequencies. Nei's probability of gene identity (J_{ij}), which measures the probability of choosing the same allele at any locus in a random draw of two individuals from different populations, averaged 0.598 across all loci in the Oregon and California populations; this means that there is nearly a 60% chance that two randomly chosen individuals from these populations will have the same allele at any RFLP locus.

Despite the high degree of genetic similarity, there was no sharing of genotypes between the populations. Several individuals in the two populations had the same multilocus haplotypes but different DNA fingerprints (Fig. 1), indicating that they were not the same clone. In every case where the same haplotype occurred in both populations, the alleles in common at the RFLP loci were either the most frequent or the second most frequent alleles. The most frequent alleles at each locus are expected to occur together in the common multilocus haplotypes in random-mating populations.

What is a population? The term *population* cannot be strictly defined for many plant pathogens. Because pathogens in different geographic regions usually differ in allele or genotype frequencies, plant pathologists usually treat collections of isolates from different locations as separate populations. However, a high level of gene flow between two geographically distant populations may unite them into homogeneous genetic groups that are evolving together. This appears to be the case with the California and Oregon populations of *M. graminicola* described here.

The correlation between genetic distance and geographic distance depends largely on the dispersal ability of an organism. Organisms with the potential for widespread, long-distance dispersal will display greater genetic uniformity across local populations than organisms with very limited dispersal ability. We defined the Oregon and California isolates as distinct groups of individuals on the basis of their geographic separation of 750 km, but the results of our analyses suggest that these "geographic populations" are not evolving independently and hence may be considered part of the same "genetic population."

Natural selection or gene flow? We consider two hypotheses to explain the high degree of genetic similarity between these populations. One hypothesis is that natural selection has resulted in the same RFLP alleles achieving similar frequencies in both populations. This hypothesis supposes that selection for particular DNA sequences at the RFLP loci themselves, or for genes that are tightly linked to each RFLP locus, occurs similarly in both populations. In addition, balancing selection or frequency-dependent selection may be necessary to maintain the large number of alleles found for many of the RFLP loci in each population. The selection hypothesis also requires selection to operate at several loci independently, because most of these loci are unlinked. Furthermore, the same selection must be happening in the different climates of Corvallis, OR, and Davis, CA, and on different host genotypes. Selection operating on two isolated populations in this manner is unlikely.

We believe that the single-locus probes used in this experiment identify genetic variation that is on average selectively neutral. These probes were chosen randomly from a collection of anonymous *Sau*3A fragments. Except for loci such as pSTL53 *Xho*I

and pSTL53 *Pst*I, which assay different restriction sites along the same piece of chromosome, the majority of these RFLP loci are unlinked (19). A significant fraction of the genome of most eukaryotes is made up of noncoding DNA, which appears to be selectively neutral (4,6,23). It is possible that many of the probes represent DNA cloned from these noncoding regions.

We have found a large amount of genetic variation at most of the RFLP loci tested. Among 17 RFLP loci, we found an average of more than nine RFLP alleles per locus in a sample of more than 700 isolates (B. A. McDonald, *unpublished*). Several of these RFLP loci have 15 or more alleles. The finding of so many alleles, with many present at a frequency over 5% at the same locus in one population, is consistent with selective neutrality of these loci. We cannot exclude the possibility that one or more of the RFLP loci considered in this experiment are closely linked to genes that are under selection, but it seems unlikely that all nine of the probes hybridized to anonymous loci that are under selection.

An alternative hypothesis is that substantial gene flow has occurred between these geographic populations. Under this hypothesis, and assuming selective neutrality of our RFLP loci, we can use an island model (35) to estimate the number of individuals that would have to have successfully migrated between the populations to account for their degree of similarity. Our estimates of Nm under this model range from two to 82 for individual loci (Table 2). On the basis of an average G_{st} of 0.039 across all 13 loci, we estimate that movement of at least 12 individuals per generation would be necessary to account for the degree of genetic similarity between the populations. Considering that movement of only one individual per generation is adequate to prevent populations from diverging significantly by genetic drift, this level of gene flow is sufficient to make these geographically separated populations a coevolving unit.

Indirect measures of gene flow cannot be used to determine the time frame over which gene flow has occurred; rather, they estimate the average amount of movement between populations over many generations. Gene flow may be a continuous process, with a few individuals migrating every generation, or it may take place in short bursts, with many individuals moving in only one or a few generations. These data cannot tell us whether gene flow occurred over the course of decades, centuries, or millennia. Further studies using direct measures of gene flow based on DNA fingerprinting may allow us to determine the historical context of this gene flow.

Mechanisms for gene flow. Several mechanisms could facilitate gene flow between populations of *M. graminicola*. The most obvious is air dispersal of the ascospores. Shaw and Royle (27) demonstrated that ascospores have the potential to move at least several hundred meters, which would be adequate to move them between adjacent wheat fields. Thus, wind-mediated movement of ascospores is probably feasible over distances of at least tens of kilometers. But we believe that other, less obvious mechanisms may also facilitate gene flow.

The role of alternate hosts in the epidemiology of *M. graminicola* deserves further consideration. Brokenshire (2) showed that *M. graminicola* can infect many common grass species, and many of these alternate hosts produced pycnidia. It is possible that one or more of the alternate hosts for *M. graminicola* form a continuous host population throughout the wheat production areas along the West Coast of North America. If so, *M. graminicola* ascospores could move readily between local geographic populations of the alternate hosts, maintaining a uniform source population for the ascospore inoculum that infects wheat fields each autumn. Among the alternate hosts identified by Brokenshire (2) are annual bluegrass (*Poa annua*) and *Hordeum murinum*. *P. annua* is distributed throughout wheat-growing areas of North America and much of the rest of the world, and *H. murinum* is present in California and Oregon.

The potential role of infected seed also deserves additional research. Brokenshire (3) showed that *M. graminicola* can infect seed, which would give it the potential for dispersal over much longer distances via the transport of grain; however, Brokenshire

never demonstrated that infected seeds could give rise to infected seedlings. Because we have evidence that gene flow occurs over very long distances, we plan to conduct experiments to test the possibility of seedborne transmission.

We have other evidence based on the analysis of RFLPs in mitochondrial DNA (mtDNA) and nuclear DNA that supports our hypothesis of gene flow over very long distances. We found the common Oregon and California RFLP alleles present at similar frequencies in an international collection of isolates that represents populations sampled from four continents in 1991 and 1992 (B. A. McDonald and R. E. Pettway, *unpublished*). In addition, we found the same mtDNA haplotype present at a high frequency in all populations surveyed so far (B. A. McDonald and R. E. Pettway, *unpublished*). Considering that global wheat production is a relatively recent event in world history, the amount of gene flow that would be necessary to unify genetic variation in geographic populations around the world is likely to be substantial. We consider it unlikely that such a high level of gene flow could have taken place without human-mediated movement of *M. graminicola*, perhaps in infected grain or straw.

Our findings have significant implications for wheat breeding programs that seek to incorporate resistance to *M. graminicola*. Our evidence for gene flow between populations separated by hundreds or thousands of kilometers suggests that plant breeders should test the resistance of their cultivars at many locations away from the area of local adaptation. The fine-scaled pattern of genetic variability suggests that plant breeders should use a wide spectrum of *M. graminicola* genotypes when testing wheat cultivars for resistance to this pathogen in any location. Wheat breeding programs that are international in scope may find it useful to test the resistance of new cultivars in a location that is known to have the maximum level of pathogen diversity. The types of analyses described in this paper may be useful for identifying appropriate testing locations. Finally, the potential for seedborne transmission suggests that rigorous quarantine measures should be followed when plant breeding programs exchange germ plasm.

LITERATURE CITED

1. Beckwitt, R., and Chakraborty, R. 1980. Genetic structure of *Pileolaria pseudomilitaris* (Polychaeta: Spirobidae). Genetics 96:711-726.
2. Brokenshire, T. 1975. The role of graminaceous species in the epidemiology of *Septoria tritici* on wheat. Plant Pathol. 24:33-38.
3. Brokenshire, T. 1975. Wheat seed infection by *Septoria tritici*. Trans. Br. Mycol. Soc. 64:331-335.
4. Brosius, J. 1991. Retroposons—Seeds of evolution. Science 251:753.
5. Brown, J. K. M., Jessop, A. C., and Rezanoor, H. N. 1991. Genetic uniformity in barley and its powdery mildew pathogen. Proc. R. Soc. London B 246:83-90.
6. Doolittle, W. F., and Sapienza, C. 1980. Selfish genes, the phenotype paradigm and genome evolution. Nature (London) 284:601-603.
7. Eyal, Z. 1981. Integrated control of Septoria diseases of wheat. Plant Dis. 65:763-768.
8. Eyal, Z., Scharen, A. L., Huffman, M. D., and Prescott, J. M. 1985. Global insights into virulence frequencies of *Mycosphaerella graminicola*. Phytopathology 75:1456-1462.
9. Fry, W. E., Goodwin, S. B., Matuszak, J. M., Spielman, L. J., Milgroom, M. G., and Drenth, A. 1992. Population genetics and intercontinental migrations of *Phytophthora infestans*. Annu. Rev. Phytopathol. 30:107-129.
10. Hamrick, K. J., and Hamrick, J. L. 1989. Genetic variation within

and among populations of an alpine grasshopper, *Aeropedellus clavatus*. J. Hered. 80:186-192.
11. Karl, S. A., Bowen, B. W., and Avise, J. C. 1992. Global population genetic structure and male-mediated gene flow in the green turtle (*Chelonia mydas*): RFLP analyses of anonymous nuclear loci. Genetics 131:163-173.
12. King, J. E., Cook, R. J., and Melville, S. C. 1983. A review of Septoria diseases of wheat and barley. Ann. Appl. Biol. 103:345-373.
13. Kohli, Y., Morrall, R. A. A., Anderson, J. B., and Kohn, L. M. 1992. Local and trans-Canadian clonal distribution of *Sclerotinia sclerotiorum* on canola. Phytopathology 82:875-880.
14. Lehman, N., and Wayne, R. K. 1991. Analysis of coyote mitochondrial DNA genotype frequencies: Estimation of the effective number of alleles. Genetics 128:405-416.
15. McDermott, J. M., and McDonald, B. A. 1993. Gene flow in plant pathosystems. Annu. Rev. Phytopathol. 31:353-373.
16. McDonald, B. A. 1991. Gametic disequilibrium among anonymous, nuclear RFLP loci in a *Septoria tritici* population. (Abstr.) Phytopathology 81:1190.
17. 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.
18. McDonald, B. A., and Martinez, J. P. 1990. Restriction fragment length polymorphisms in *Septoria tritici* occur at a high frequency. Curr. Genet. 17:133-138.
19. McDonald, B. A., and Martinez, J. P. 1991. Chromosome length polymorphisms in a *Septoria tritici* population. Curr. Genet. 19:265-271.
20. McDonald, B. A., and Martinez, J. P. 1991. DNA fingerprinting of the plant pathogenic fungus *Mycosphaerella graminicola* (anamorph *Septoria tritici*). Exp. Mycol. 15:146-158.
21. Nei, M. 1972. Genetic distance between populations. Am. Nat. 106:283-292.
22. Nei, M. 1973. Analysis of gene diversity in subdivided populations. Proc. Natl. Acad. Sci. U.S.A. 70:3321-3323.
23. Orgel, L. E., and Crick, F. H. C. 1980. Selfish DNA: The ultimate parasite. Nature (London) 284:604-607.
24. Reed, K. C., and Mann, D. A. 1985. Rapid transfer of DNA from agarose gels to nylon membranes. Nucleic Acids Res. 13:7207-7221.
25. Schnabel, A., and Hamrick, J. L. 1990. Comparative analysis of population genetic structure in *Quercus macrocarpa* and *Q. gambelii* (Fagaceae). Syst. Bot. 15:240-251.
26. Schuh, W. 1990. Influence of tillage systems on disease intensity and spatial pattern of Septoria leaf blotch. Phytopathology 80:1337-1340.
27. Shaw, M. W., and Royle, D. J. 1989. Airborne inoculum as a major source of *Septoria tritici* (*Mycosphaerella graminicola*) infections in winter wheat crops in the UK. Plant Pathol. 38:35-43.
28. Singh, R. S., and Rhomberg, L. R. 1987. A comprehensive study of genic variation in natural populations of *Drosophila melanogaster*. II. Estimates of heterozygosity and patterns of geographic differentiation. Genetics 117:255-271.
29. Slatkin, M. 1985. Gene flow in natural populations. Annu. Rev. Ecol. Syst. 16:393-430.
30. Slatkin, M. 1987. Gene flow and the geographic structure of natural populations. Science 236:787-792.
31. Slatkin, M., and Barton, N. H. 1989. A comparison of three indirect methods for estimating average levels of gene flow. Evolution 43:1349-1367.
32. Stoddart, J. A., and Taylor, J. F. 1988. Genotypic diversity: Estimation and prediction in samples. Genetics 118:705-711.
33. Van Ginkel, M., and Scharen, A. L. 1988. Host-pathogen relationships of wheat and *Septoria tritici*. Phytopathology 78:762-766.
34. Wolf, P. G., and Soltis, P. S. 1992. Estimates of gene flow among populations, geographic races, and species in the *Ipomopsis aggregata* complex. Genetics 130:639-647.
35. Wright, S. 1951. The genetical structure of populations. Ann. Eugen. 15:323-354.