Variation at the *b* Mating Type Locus of *Ustilago maydis*

Paul Zambino, James V. Groth, Lewis Lukens, James R. Garton, and Georgiana May

First author: USDA Forest Service, North Central Forest Experiment Station, Forestry Sciences Laboratory, 5985 Highway K, Rhinelander, WI 54501-9128; second author: Department of Plant Pathology, University of Minnesota, St. Paul 55108; third, fourth, and fifth author: Department of Plant Biology, University of Minnesota, St. Paul 55108; and second, fourth, and fifth author: Center for Community Genetics, University of Minnesota, St. Paul.
Accepted for publication 30 August 1997.

ABSTRACT

Zambino, P., Groth, J. V., Lukens, L., Garton, J. R., and May, G. 1997. Variation at the *b* mating type locus of *Ustilago maydis*. *Phytopathology* 87:1233-1239.

Population level diversity at the *Ustilago maydis* *b* mating type locus was determined in samples from four Minnesota locations using a combination of plate mating techniques and a polymerase chain reaction (PCR)-based assay. The PCR method allows rapid identification of *b* types from samples of natural populations and utilizes the hypervariable regions of the *b* locus that determine mating type specificity. Results demonstrated high levels of *b* diversity within populations, with one population yielding 17 of the total 18 *b* types found in the study. Pairwise G_{ST} values were in the

range of 0.02 to 0.05, and common *b* mating types were found across broad geographic distances. These data demonstrated that very low levels of differentiation among *U. maydis* populations occur with respect to *b* locus variation. Consistent with frequency-dependent selection models, *b* types were represented at approximately equal frequencies within the entire Minnesota population. However, neutral evolutionary models for patterns of geographic distribution and variation at *b* cannot be entirely excluded. The importance to agricultural practices of understanding population genetic processes is discussed.

Additional keywords: barcode, basidiomycete, corn smut, DNA fingerprint, mating system.

The ability of *Ustilago maydis* to cause disease in maize is directly dependent on the availability of compatible mates. The interdependence of infection and sexual compatibility arises because the mating type loci that control sexual compatibility also control the subsequent transition from a saprophytic to a pathogenic phase of the life cycle. Specifically, we address the importance of variation at the *b* mating type locus that, when heterozygous, regulates the development of infectious filamentous cells (35,36) and of tumors (2,9). The requirement for heterozygosity at *b* and its further role in pathogenic development may have provided the strong selective force necessary to the evolution of the large number of *b* mating types observed for this species. To begin an investigation of the evolutionary forces maintaining *b* mating type variation, we determined the spatial scale and levels of variation at the *b* mating type locus in Minnesota field populations.

In *U. maydis*, there are two mating loci, *a* and *b*, that must both be heterozygous in mating haploids for compatibility. The *a* locus has but two types, *a1* and *a2*, and controls the signaling and cell fusion events early in compatibility (5). We focus on the *b* locus that is represented by up to 25 different compatibility types (31,38). The *b* locus acts as a genetic "master switch" and controls the transition from a harmless saprophytic yeast phase to a filamentous form that invades the plant (3,13,35,36). Each *b* locus is made up of two divergently transcribed genes, termed *bEast* (*bE*) and *bWest* (*bW*) (13). Surprisingly, it is the interaction of the gene product encoded by *bE* from one *b* mating type with the encoded protein by *bW* from a different *b* mating type that governs compatibility, e.g., presence of the *bE1* and *bW2* genes is sufficient to initiate developmental processes (13). The N-terminal regions of the *bE*-

and *bW*-encoded proteins are highly variable and include the molecular determinants of compatibility. The C-terminal regions are better conserved than the N-terminals and include a homeodomain-like motif in each protein (13,21,48). It is apparent that *b* is acting as a regulatory locus with the *bE* and *bW* gene products acting in combination to regulate additional genes involved in filamentous growth and pathogenicity (1,3,12,13).

Although variation at the *b* locus is of direct importance to the population genetic structure and pathogenicity of this organism, little is known of the evolutionary processes maintaining diversity at *b*. Recent molecular analyses of fungal and plant mating compatibility type genes demonstrate that different mating types are highly divergent in sequence and that these different mating types have likely been maintained in populations for very long time periods (7,13,19,21,33,34,43). Because sexual reproduction and subsequent development require a genetic difference at compatibility loci, frequency-dependent selection will maintain a large number of different mating types even in local populations (41,45-47,49). Arriving in the population by either mutation or migration, a haploid carrying a new mating compatibility type will be compatible with most other haploid cells in the population and increase in frequency. At equilibrium, different compatibility types will be present at equal frequencies (25). The model above generates a straightforward expectation; under frequency-dependent selection, different mating types will be present at equal frequencies.

To determine levels of variation, geographic distribution, and frequency distributions of alleles at the *b* mating type locus, we surveyed *U. maydis* mating types at four Minnesota localities. To rapidly assess *b* variation, we developed a polymerase chain reaction (PCR) technique, the "barcode" analysis, that uses restriction fragment length polymorphisms (RFLPs) of PCR products amplified from the hypervariable portions of *b* mating types. We demonstrate a 1-to-1 correspondence between RFLP patterns and the *b* mating types determined by plate mating methods. Our results demonstrate that different *b* mating types are widespread across Minnesota and that the 18 mating types occur at approximately equal frequencies.

Corresponding author: G. May; E-mail address: gmay@maroon.tc.umn.edu

Publication no. P-1997-0929-02R

This article is in the public domain and not copyrightable. It may be freely reprinted with customary crediting of the source. The American Phytopathological Society, 1997.

We see little evidence of association between specific *b* mating types recovered from the diploid teliospores, and data suggest that the organism is largely outcrossing.

MATERIALS AND METHODS

Collections and field plot locations. In September 1993, 10 galls were collected from each of three Minnesota fields (Becker, field corn; Lamberton, field corn at a site continuously planted to corn; and St. Paul [University of Minnesota], experimental breeding lines of sweet corn). Galls were collected from plants at least 10 m apart, and each were transported to the lab in a separate plastic bag, where galls were air-dried and stored at room temperature. One portion of each gall was then transferred to a 1.2-ml cryovial and stored with desiccant. In August 1995, galls were collected from a fourth field (Owatonna, sweet corn) in a defined sampling grid, with each gall collected at least 15 m from any other. Gall material was placed directly into cryovials in the field and subsequently dried in the lab as previously described. Teliospores subsequently stored over desiccant at 22°C remained viable for over 2 years.

Isolation of the haploid sporidial stage. Teliospores were vortexed with spore wash (0.5 g/liter, penicillin G [Sigma PEN-NA; Sigma Chemical Co., St. Louis]; 0.5 g/liter, streptomycin sulfate [Sigma S6501; Sigma Chemical Co.]; and 4 drops/liter, Tween 80) and plated at low densities on plates of PDA+ agar (24 g/liter, potato dextrose broth; 15 g/liter, agar [Sigma Chemical Co.]; 0.5 g/liter, penicillin G; and 0.5 g/liter, streptomycin sulfate). Incubation for 36 to 48 h at 25°C caused teliospores to undergo meiosis and produce probasidia and sporidia. The haploid sporidia were washed off plates with 2 ml of spore wash, vortexed to separate cells, and replated on PDA+ plates. Individual colonies representing single sporidia were then transferred to PDA plates lacking antibiotics, grown overnight at 25°C, and maintained at 4°C until crosses or DNA analyses were made. To ensure recovering all *b* mating types from each gall collection, colonies representing 15 to 20 haploid sporidia were isolated from each gall.

Determination of *b* mating types by plate matings. Plate matings to determine mating types were made on a rich agar medium derived and simplified from previous recipes (8,17,30). This pairing medium consisted of 48.0 g/liter, potato dextrose broth; 2.0 g/liter, yeast extract; 5.0 g/liter, casamino acids; 10 g/liter, powdered activated charcoal; 126.0 ml/liter, Holiday's (17) salt solution (KH₂PO₄, 16 g/liter; Na₂SO₄, 4.0 g/liter; KCl, 8.0 g/liter; MgSO₄, 2.0 g/liter; and CaCl₂, 1.0 g/liter); 874.0 ml of distilled water; 0.5 M NaOH to obtain pH 7.0; and 30 g/liter, agar (Sigma Chemical Co.). Mating types were determined by streaking aqueous suspensions of young, freshly harvested sporidia of unknown mating genotypes onto plates of pairing medium on which tester sporidia of known genotype had been streaked. A compatible plate mating was identified after 12 to 24 h of incubation at 25°C by the appearance of the "fuzz" reaction, indicating that filamentous growth occurred (30).

Mating types were identified in two screens. The first screen identified the *a* mating type of each sporidial strain by pairing them

with reference "tester" strains (listed below) carrying *a1 b1*, *a1 b2*, *a2 b1*, and *a2 b2* mating types. The second screen identified the *b* mating types. Strains were paired with testers carrying each of the 18 different *b* mating types, such that *a* types of the unknown and the tester were compatible. Identical *b* mating types were identified as noncompatible plate matings when the fuzz reaction was not observed. To ensure accurate scoring, any uncertain reactions were repeated with fresh sporidia. Tester strains carrying the *a* mating type allele *a1* and the *b* mating types, *b1* through *b18*, were as follows: 93-4-5, 93-4-2, 93-26-13, 93-41-2, 93-41-3, 93-59-1, 93-59-6, 93-30-1, 93-25-5, 93-33-10, 93-33-8, 93-5-4, 93-3-2, 93-7-7, 93-8-1, (no *a1 b16* tester), 93-9-3, and 93-11-4. Tester strains carrying *a2* and *b* mating types, *b1* through *b18*, were as follows: 93-4-1, 93-4-3, 93-26-1, 93-41-5, 93-41-4, 93-59-2, 93-59-5, 93-30-4, 93-25-2, 93-33-6, 93-33-7, 93-5-1, 93-3-1, 93-7-1, 93-8-8, 93-6-11, 93-9-16, and 93-11-13. All tester strains were kept in permanent storage at -80°C.

Molecular methods. Genomic DNA was prepared by miniprep methods (50) or the microwave method of Goodwin and Lee (14). PCR was carried out in a reaction cocktail of 20 ng of genomic DNA, 1.5 mM MgCl₂, 50 mM KCl, 10 mM Tris (pH 8.3), 200 nM dNTPs, 0.625 units of AmpliTaq polymerase (Perkin Elmer, Roche Molecular Systems, Branchburg, NJ), and 100 pmol each of oligonucleotide primers. The following primers allowed amplification of the hypervariable region for all 18 *b* types; primer be10 (5'-AAGGATCCATAGCGTGAGCTGATGA-3') (3) is within the region coding for the *bE* homeodomain. GenBank sequences M84179 to M84182 (13) were examined for nonvariable regions to construct the second primer, Umbw.qds (5'-CATCGAAGAAAGTAC-TGTCCTG-3'). The latter primer is 625 bases downstream from the end of the region coding for the *bW* homeodomain. The temperature cycle program for amplification of the 2.1-kb product was

TABLE 2. Occurrence and frequencies of *Ustilago maydis* *b* mating types sampled in galls from Minnesota fields of *Zea mays*

<i>b</i> allele	Occurrence among samples by field location				All samples	Frequencies
	Becker ^a	Lamberton ^a	St. Paul ^a	Owatonna ^b		
<i>b1</i>	0	0	2	3	5	0.06
<i>b2</i>	3	1	2	0	6	0.08
<i>b3</i>	3	1	0	0	4	0.04
<i>b4</i>	0	0	3	0	3	0.04
<i>b5</i>	0	3	2	3	8	0.10
<i>b6</i>	1	3	2	0	6	0.08
<i>b7</i>	3	0	3	2	8	0.10
<i>b8</i>	1	1	1	1	4	0.05
<i>b9</i>	1	0	1	0	2	0.03
<i>b10</i>	1	1	1	1	4	0.05
<i>b11</i>	3	0	1	1	5	0.06
<i>b12</i>	0	0	2	1	3	0.04
<i>b13</i>	0	1	1	0	2	0.03
<i>b14</i>	1	1	3	0	5	0.06
<i>b15</i>	1	1	2	0	4	0.05
<i>b16</i>	0	0	1	0	1	0.01
<i>b17</i>	0	2	1	3	6	0.08
<i>b18</i>	0	0	1	1	2	0.03

Total haploids tested	18	15	29	16	79
No. of <i>b</i> mating types	10	10	17	9	18
Gene diversity ^c	87.04	87.11	92.98	85.94	91.84
Mantel test					16.67
Statistic ^d					NS
Overall G _{ST} ^c					0.03
Overall N _m ^e					16.17

^a *b* mating type as determined by plate matings.

^b *b* mating types as determined by barcode.

^c Literature citation 27.

^d NS = not significant with a Mantel test statistic less than the chi-square critical value = 24.769 (df = 17 and α = 0.1) (22,26).

^e Literature citations 39 and 40.

TABLE 1. Dikaryons, haploids, and *b* mating types sampled in *Ustilago maydis* galls from *Zea mays* at four Minnesota localities

Sample	Galls tested	Dikaryons	Haploids	Number of <i>b</i> mating types
1993, Becker field corn	9	9	18	10 ^a
1993, Lamberton field corn	8	8	15	10 ^a
1993, St. Paul sweet corn	10	15	29	17 ^a
1995, Owatonna sweet corn	8	8	16	9 ^b
Total	35	40	78	18

^a *b* mating types determined by plate matings.

^b *b* mating types determined by the barcode.

40 cycles of 94°C for 30 s, 50°C for 60 s, and 72°C for 150 s; and one cycle of 72°C for 10 min.

After purification of the PCR product by ethanol precipitation, 300 to 500 ng of the PCR product was digested with either *TaqI* or *Sau96I* according to manufacturers instructions (Promega Corp., Madison, WI). Fragments were separated by size on a 1× Tris-borate-EDTA, 0.3% agarose gel (4:1 mixture of NuSieve GTG/SeaKem GTG agaroses; FMC BioProducts, Rockland, ME). DNA fragments were visualized with ethidium bromide staining and viewed and recorded on an UVP ImageStore 7500 system (UVP, Inc., Upland, CA). Results were saved in digital form for further analysis and printing. Image processing was carried out on a Macintosh Quadra 800 computer (Apple Computer, Inc., Cupertino, CA) using NIH Image 1.54 to label images and to invert images to represent black fragments on a white background. NIH Image is available online from the National Institute of Health.

Data analysis. The Mantel test was used to test the null hypothesis of equal frequencies among *b* mating types (22,26). The chi-square test statistic was calculated as follows.

$$\chi^2 = (n-1) \left(\sum_{j=1}^n C_j^2 - 4r^2/n \right) / (2r - 4r/n)$$

In which n = number of mating types recovered; r = number of diploids tested; and C_j = number of times each allele j appears in the sample. The test statistic approximates a chi-square distribution with $(n-1)$ degrees of freedom if r is large (6). Simulations carried out by Lawrence et al. (22) indicate that the Mantel test is conservative.

Gene diversity (H) and the coefficient of gene differentiation (G_{ST}) were calculated by the method of Nei (27,28). H is the gene diversity in which x_i is the frequency of the i th allele.

$$H = 1 - \sum_i x_i^2$$

H_T is the gene diversity for the entire population that can be broken down to H_S (the within population) and D_{ST} (the between population) components of gene diversity. G_{ST} compares differences between populations with the total diversity over the population.

$$H_T = H_S + D_{ST}$$

$$G_{ST} = D_{ST}/H_T$$

Estimates of N_m , the product of population size and migration rate, were calculated using G_{ST} for haploids (27,40).

$$N_m = (1/G_{ST} - 1)/2$$

RESULTS

Terminology and mating system. The mating system of *U. maydis* is controlled by two loci, *a* and *b*. Haploids must carry different *a* and different *b* mating types to be completely compatible and allow completion of the sexual cycle in corn. The *a* locus has but two types, *a1* and *a2*, and the *b* locus has an estimated 25 different mating types (38). We refer to different compatibility types controlled by the *b* locus as *b* mating types. For example, we designated *b1* through *b18* as different mating types, because they allow compatibility between mating haploids. We use the more inclusive term, *b* alleles, to designate any variants of *b*, whether a compatibility type or a difference in RFLP patterns.

Gene diversity and population subdivision. We chose sampling localities to broadly represent areas of Minnesota where the corn (*Zea mays*) host is grown, with distances of 44 to 115 miles (71 to 185 km) between localities. Sampling locations were close to the cities of Becker, Lamberton, St. Paul (University of Minnesota campus), and Owatonna, MN. Eighteen different *b* mating types were identified among the samples (Table 1). Samples from three of the four fields yielded nine or 10 of the 18 *b* mating types, whereas those from the St. Paul field yielded 17 mating types. Of

the total 18 *b* mating types identified, only two were found in all four locations. Becker and Lamberton samples shared seven mating types with each other and four mating types with samples from Owatonna. Following the method of Whitehouse (44), we calculated that we had recovered 94% of the total number of *b* types in each of the four populations and that the total number of different types in the combined population was approximately 19, with a range of 18 to 20 mating types using a 95% confidence interval. This value is comparable to Puhalla's (31) estimate of 21 mating types, with a range of 17 to 25 mating types using a 95% confidence interval.

Values obtained for H , the gene diversity (27), for each subpopulation and for the total population are given in Table 2 and were quite high in all sample locations. As the value of H increases, the chance of a sporidial strain encountering another that has the same *b* mating type decreases. We calculated G_{ST} , Nei's measure of gene diversity and population substructuring, as 0.03 over all populations, and the pairwise G_{ST} values were in the range of 0.02 to 0.05 (Table 3). Such low values of G_{ST} reflect low levels of population differentiation with respect to the *b* locus. As a measure of migration levels, we derived N_m from G_{ST} (40). Over all pairwise comparisons, the average N_m was quite high, 12.4, as were values for individual pairwise comparisons, ranging from 9.2 to 21.6 (Table 3). The lowest pairwise G_{ST} values and the highest pairwise values of N_m were found in comparisons of the St. Paul sample with each of the other three sampled localities. This was likely because the St. Paul samples yielded almost all of the *b* types found in the entire study. Most of the diversity at *b* was represented within each of the local samples, with little additional diversity attributed to differences between population.

Frequencies of *b* mating types. Across locations, the observed frequencies of specific mating types ranged from 1 to 10%, with no allele recovered at a frequency of more than 20% from any one location (Table 2). To test whether different *b* mating types occurred at equal frequencies, as predicted by the frequency-dependent selection hypothesis, we used the Mantel test (22,26) to analyze the relative frequencies of 18 different *b* types in the pooled sample of 78 haploid isolates from four sites. The null hypothesis for equal frequencies could not be rejected at $\alpha = 0.1$. We assumed that most galls resulted from infection by a single dikaryon, because we recovered only two *b* types. We scored these as having one diploid (r) for calculation of the Mantel test statistic. However, in the St. Paul samples, five of the 10 galls sampled were apparently the result of infection by two dikaryons, and three or four different *b* mating types were recovered. No galls contained more than four mating types. For these samples, we assumed that two dikaryons were sampled, bringing the total number of dikaryons sampled at the St. Paul site to 15 and the total number of diploids (r) in the test statistic to 40 (Table 1).

We asked whether specific combinations of *b* types occurred more often than predicted by chance association. However, with a total of 18 *b* types identified, there are 153 unique combinations of two different mating types and, with such a large number of unique combinations possible, our sample size was insufficient for statistical determination of associations. We note that, in 84% of galls analyzed (34 of the 40 total), we recovered a unique combination of *b* mating types for each gall. In three cases, the same

TABLE 3. Pairwise values of G_{ST} ^a (above diagonal) and N_m ^b (below diagonal) for *Ustilago maydis* *b* mating types in Minnesota sampling populations

	St. Paul	Lamberton	Becker	Owatonna
St. Paul	...	0.023	0.024	0.036
Lamberton	21.59	...	0.045	0.036
Becker	20.57	10.71	...	0.051
Owatonna	20.13	13.54	9.22	...

^a Literature citation 27.

^b Literature citations 39 and 40.

specific combination of two *b* mating types was recovered in two different galls and accounted for 16% of the sample (six of the 40 galls). In these three cases, the two galls with the same combination of two *b* types were not adjacent to each other. No specific combination of *b* types was recovered more than twice.

Assay for variation at the *b* mating type locus, the “barcodes.”

We developed a PCR-based method to rapidly determine *b* mating types of *U. maydis* strains based on the molecular structure of the *b* mating type locus. The *U. maydis* *b* mating type is made up of a pair of divergently transcribed genes, with variable N-terminal regions located next to each other and flanked by more conserved regions (13,37). Primers were located in conserved regions and amplified a 2.1-kb fragment containing the variable regions of both the *bE* and *bW* genes (discussed in Materials and Methods). Restriction endonuclease digestion with *TaqI* or *Sau96I* of the 2.1-kb fragment from each different *b* mating type resulted in distinct RFLP patterns that we call the “barcode” (Fig. 1).

To be useful, the RFLPs demonstrated by the PCR-based method must correlate with the *b* mating type observed in laboratory matings. Results of the barcode analysis are presented in Figures 1 and 2. When 2.1-kb PCR fragments amplified from each *b* locus were digested with *Sau96I*, numerous fragments ranging from 0.1 to 1.9 kb were detected (Fig. 1). Each *b* mating type, as determined by mating tests, yielded a unique barcode that could be easily distinguished from others.

To test the correlation between barcode and *b* mating type, we used *TaqI* and *Sau96I* to digest the 2.1-kb fragment from several representatives of the *b6*, *b8*, *b10*, and *b11* mating types that were isolated from different fields. We detected one *Sau96I* restriction site polymorphism among the *b10* mating types (Fig. 2). No other *b* locus had a barcode that was similar to either *b10* RFLP pattern. The barcodes, thus, allowed us to rapidly detect the diversity of *b* mating types occurring in different fields and to locate molecular variants of *b* mating types.

The utility of the barcode method in rapidly determining *b* mating types was tested on the samples from the Owatonna field. To empirically establish how many sporidial isolates from each gall must be assayed to determine two *b* types, we assayed four sporidial strains from each of nine galls. For seven of these nine, four sporidial isolates were sufficient to determine both *b* types. For two of the nine, we recovered only one *b* type in four sporidia and, for these, obtaining barcodes for four more sporidial isolates was sufficient to determine the second *b* type. For each strain, the RFLP pattern matched a known barcode for one of the 18 *b* types previously identified in the Becker, Lamberton, and St. Paul locations.

DISCUSSION

Our results demonstrate high variability at the *b* mating type locus in populations of *U. maydis*. Using plate matings, we identified 18 different *b* mating types that were represented in approximately equal frequencies across localities. Using RFLPs generated from PCR amplification of the hypervariable region of *b*, we demonstrated that each *b* locus is uniquely characterized by the restriction fragment fingerprint of the hypervariable region.

With the barcode, we took advantage of the two-gene structure of *b* to design PCR primers located in well-conserved regions flanking the hypervariable regions (13,37). Because the hypervariable region is known to determine mating type (48), we reasoned that the molecular variation found in this area should correspond to the mating type. The barcode patterns of restriction fragments showed an exact correspondence to different *b* mating types for samples from within and among different locations, a main criterion necessary to demonstrate the usefulness of the barcode for population studies. The barcode also detected molecular variation at the *b10* mating type that plate matings could not, and these alleles could be considered a neutral variation within the *b10* mating type.

The population level gene diversity values were very high, 0.86 to 0.93. This is not surprising given that heterozygosity at *b* is a

Fig. 1. Barcode analysis of 18 different *Ustilago maydis* *b* mating types. Individual lanes represent *Sau96I* restriction digests of the 2.1-kb polymerase chain reaction fragment amplified from the *b* mating type locus of *U. maydis* haploid strains carrying each different *b* type. Lanes 1 to 18 represent each of the 18 *b* mating types as determined by plate matings. Lanes L indicate size markers, 1-kb ladder (BRL Life Sciences, Grand Island, NY). The image is a negative digital image of an ethidium bromide-stained agarose gel.

prerequisite to completion of the life cycle and production of teliospores. Because we collected teliospores, our sampling method assured the heterozygous condition of *b* for each individual gall. Obtaining a high value for Nei's (27) measure of gene diversity reflects the recovery of a large number of distinct *b* mating types in approximately equal frequencies. It is a population level parameter, and high gene diversity values correspond to a low probability that two haploids chosen at random from a population will share the same *b* type. The values of *H* that we report are much higher than those reported for other loci in fungal populations. For example, in *Phytophthora infestans*, gene diversity values ranged from 0.126 to 0.258 for a DNA fingerprinting locus (15) and, in *Mycosphaerella graminicola*, values range from 0.230 to 0.750 (4). The numbers of different mating types we observed were a little lower than numbers of *S* alleles reported for plant gametophytic incompatibility systems (33) and were comparable to the numbers of *S* alleles reported for species with a sporophytic incompatibility system (20). As in plant *S* loci, common *b* mating types were recovered from more than one sampled location (29,32).

There is little question that frequency-dependent selection plays a major role in the evolution of mating compatibility genes. Wright's mutation/selection models estimated that a population size of 2,000 to 10,000 plants and a mutation rate of 10^{-6} would be sufficient to maintain the observed diversity at the *S* locus of *Oenothera organensis* (11). By direct observation, Emerson (10,11) had estimated population sizes at less than 500 individuals, but more recent estimates based on isozyme data placed the population size at approximately 5,000 to 10,000 (24). If so, the population size and the low F_{ST} values obtained in that study suggest that gene flow and population size are sufficient to prevent loss of *S* allele diversity by drift.

In *U. maydis*, mating types are at equal frequencies across the Minnesota populations, and common *b* mating types are distributed across several locations. We observed very low G_{ST} values for *b*, and little of the diversity at *b* could be attributed to differ-

ences between subpopulations. Either high levels of gene flow or strong selection for different *b* mating types will equalize *b* mating type frequencies across populations. However, under a neutral model, allele frequencies within subpopulations will not be equal. Our within population sample sizes were not large enough to determine if the *b* types were at equal frequencies. Certainly, not all of the 18 types were found in every location. Although the result for *b* frequencies across populations was consistent with the predictions of frequency-dependent selection, these data were not sufficient to distinguish a selection hypothesis from a neutral hypothesis for maintenance and spatial distribution of diversity at the *b* locus. The relative importance of selective and nonselective evolutionary forces could be analyzed by comparing G_{ST} values for *b* with those obtained for neutral loci (42) and by determining *b* mating type frequencies at different hierarchical spatial levels.

Recently, the use of *S* allele genealogies to estimate historic and current population sizes has allowed a different approach to understanding the maintenance of diversity (33). The genealogies for many genes under balancing selection demonstrate trans-specific polymorphism. In several plant and fungal systems, bursts of diversification are followed by long time periods when a large number of alleles are maintained, but few new alleles evolve (19,32,43; G. May, unpublished data). After diversification, nonselective forces could be of increasing importance to the maintenance of variation. In a fungus such as *U. maydis*, which produces a large number of airborne spores and has an extensive geographic distribution, population sizes might be quite large. If this model is correct, we would predict that historic population sizes of *U. maydis* were smaller and selective pressures higher than in the present.

To understand the maintenance of genetic variation in *U. maydis*, additional information for pathogen and host are required. We understand little of the life cycle of the fungus in field populations. Whereas in smuts such as *Tilletia* species, the primary sporidia mate immediately after they are produced on the promycelium and may,

Fig. 2. Barcode analysis of multiple strains from different Minnesota locations representing the *b6*, *b8*, *b10*, and *b11* mating types. Shown are two panels from separate gels (*b6/b8* and *b10/b11*) presented as a negative digital image of an ethidium bromide-stained agarose gel.

thus, lead to high levels of inbreeding (18); this does not seem to be common in *U. maydis*. If inbreeding was common, the diversity of *b* types would decline and repeated recovery of specific associations of *b* mating types should be common, especially in local regions. However, the likelihood of sibling matings will depend on agricultural practices and the location of teliospore germination, about which we know little. For example, if teliospore germination occurs in the soil before tilling and planting, cultivation could mix populations of fungi (16). In our analyses, we did not repeatedly recover the same two *b* mating types in association, suggesting that *U. maydis* populations are largely outbreeding.

On the level of the host, both the genetic interaction of *U. maydis* with maize and the history of the field site could affect the genetic structure of *U. maydis* populations. As an example, we recovered more *b* mating types at the St. Paul fields than at other locations. This site is used for breeding programs, and a mixture of many maize genotypes are maintained as sweet corn breeding lines. Because this field is used every year for this purpose, repeated cropping of susceptible genotypes in the same location could build up a large population of *U. maydis*, in contrast to the possible genetic bottleneck effect of rotation planting in a new locations. For those few pathogens for which data exist, information on pathogen migration rates, evolution of virulence, and response to variation in host resistance is invaluable to agricultural practices (23).

The advantages of using *U. maydis* to address fundamental questions in molecular evolution and in the evolution of plant pathogens are clear. The *b* mating type plays a central role in the evolution of *U. maydis* populations and will profoundly affect the population dynamics of the organism. Because this fungus is amenable to laboratory manipulations, mating types can be fully cross-referenced, and extensive genetic and molecular analyses on single haploid genotypes can be carried out. Since smut fungi are major pathogens on many cereal grain crops such as corn, barley, rye, and oats, our work with the genetically tractable *U. maydis* may provide a basis for understanding the evolution of many economically important plant-pathogen interactions.

ACKNOWLEDGMENTS

This research was supported from grants from the Minnesota Agriculture Experiment Station (to G. May and J. V. Groth), the Center for Community Genetics (to J. R. Garton), and NSF-REU (to G. May). We thank A. Gozdziołko and A. Lee (undergraduate research assistants) for their enthusiastic assistance in making isolations, D. Davis (Department of Horticulture, University of Minnesota, St. Paul), and personnel of the Lamberton and Becker Experiment Stations and of the Owatonna Canning Corp. for assistance in identifying suitable field locations or collecting galls of *Ustilago maydis*.

LITERATURE CITED

- Banuett, F. 1991. Identification of genes governing filamentous growth and tumor induction by the plant pathogen *Ustilago maydis*. *Proc. Natl. Acad. Sci. U.S.A.* 88:3922-3926.
- Banuett, F., and Herskowitz, I. 1994. Identification of Fuz7, a *U. maydis* MEK/MAPKK homologue required for α -dependent and independent processes in the fungal life cycle. *Genes Dev.* 8:1367-1378.
- Barrett, K. J., Gold, S. E., and Kronstad, J. W. 1993. Identification and complementation of a mutation to constitutive filamentous growth in *Ustilago maydis*. *Mol. Plant-Microbe Interact.* 6:274-283.
- 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.
- Bölker, M., Urban, M., and Kahmann, R. 1992. The α mating type locus of *U. maydis* specifies cell signaling components. *Cell* 68:441-450.
- Campbell, J. M., and Lawrence, M. J. 1981. The population genetics of the self-incompatibility polymorphism in *Papaver rhoeas*. II. The number and frequency of *S*-alleles in a natural population (R106). *Heredity* 46:81-90.
- Clark, A. G., and Kao, T.-H. 1991. Excess nonsynonymous substitution at shared polymorphic sites among self-incompatibility alleles of Solanaceae. *Proc. Natl. Acad. Sci. U.S.A.* 88:9823-9827.
- Day, P. R., and Anagnostakis, S. L. 1971. Corn smut dikaryons in culture. *Nature* 231:19-20.
- Day, P. R., Anagnostakis, S. L., and Puhalla, J. E. 1971. Pathogenicity resulting from mutation at the *b* locus of *Ustilago maydis*. *Proc. Natl. Acad. Sci. U.S.A.* 68:533-535.
- Emerson, S. 1938. The genetics of self-incompatibility in *Oenothera organensis*. *Genetics* 23:190-202.
- Emerson, S. 1939. A preliminary survey of the *Oenothera organensis* population. *Genetics* 24:524-536.
- Giasson, L., and Kronstad, J. W. 1995. Mutations in the *myl1* gene of *Ustilago maydis* attenuate mycelial growth and virulence. *Genetics* 141:491-501.
- Gillissen, B., Bergemann, J., Sandmann, C., Schroeder, B., Bölker, M., and Kahmann, R. 1992. A two-component regulatory system for self/non-self recognition in *Ustilago maydis*. *Cell* 68:647-657.
- Goodwin, D. C., and Lee, S. B. 1993. Microwave miniprep of total genomic DNA from fungi, plants, protists and animals for PCR. *Bio-Techniques* 15:438-444.
- Goodwin, S. B., Spielman, L. J., Matuszak, J. M., Bergeron, S. N., and Fry, W. E. 1992. Clonal diversity and genetic differentiation of *Phytophthora infestans* populations in northern and central Mexico. *Phytopathology* 82:955-961.
- Gordon, T. R., Okamoto, D., and Milgroom, M. G. 1992. The structure and interrelationship of fungal populations in native and cultivated soils. *Mol. Ecol.* 1:241-249.
- Holliday, R. 1974. *Ustilago maydis*. Pages 575-595 in: *Handbook of Genetics*, Vol. 1. R. C. King, ed. Plenum Press, New York.
- Holton, C. S., Hoffmann, J. A., and Duran, R. 1968. Variation in the smut fungi. *Annu. Rev. Phytopathol.* 6:213-242.
- Joerger, T. R., Clark, A. G., and Kao, T. 1990. Polymorphism at the self-incompatibility locus in Solanaceae predates speciation. *Proc. Natl. Acad. Sci. U.S.A.* 87:9732-9735.
- Karron, J. D., Marshall, D. L., and Oliveras, D. M. 1990. Numbers of sporophytic self-incompatibility alleles in populations of wild radish. *Theor. Appl. Genet.* 79:457-460.
- Kronstad, J. W., and Leong, S. A. 1990. The *b* mating type locus of *Ustilago maydis* contains variable and constant regions. *Genes Dev.* 4:1384-1395.
- Lawrence, M. J., O'Donnell, S., Lane, M. D., and Marshall, D. F. 1994. The population genetics of the self-incompatibility polymorphism in *Papaver rhoeas*. VIII. Sampling effects as a possible cause of unequal allele frequencies. *Heredity* 72:345-352.
- Leung, H., Nelson, R. J., and Leach, J. E. 1993. Population structure of plant pathogenic fungi and bacteria. Pages 157-205 in: *Advances in Plant Pathology*. J. H. Andrews and I. C. Tommerup, eds. Academic Press, New York.
- Levin, D. A., Ritter, K., and Ellstrand, N. C. 1979. Protein polymorphism in the narrow endemic *Oenothera organensis*. *Evolution* 33:534-542.
- Lewontin, R. C., Ginzburg, L. R., and Tuljapourkar, S. D. 1978. Heterosis as an explanation for large amounts of genic polymorphism. *Genetics* 88:149-170.
- Mantel, N. 1974. Approaches to a health research occupancy problem. *Biometrics* 30:355-362.
- Nei, M. 1973. Analysis of gene diversity in subdivided populations. *Proc. Natl. Acad. Sci. U.S.A.* 70:3321-3323.
- Nei, M. 1987. *Molecular Evolutionary Genetics*. Columbia University Press, New York.
- O'Donnell, S., Lane, M. D., and Lawrence, M. J. 1993. The population genetics of the self-incompatibility polymorphism in *Papaver rhoeas*. VI. Estimation of the overlap between the allelic complements of a pair of populations. *Heredity* 71:591-595.
- Puhalla, J. E. 1968. Compatibility reactions on solid medium and inter-strain inhibition in *Ustilago maydis*. *Genetics* 60:461-474.
- Puhalla, J. E. 1970. Genetic studies of the *b* incompatibility locus of *U. maydis*. *Genet. Res. Camb.* 16:229-232.
- Richman, A., Kao, T.-H., Schaeffer, S., and Uyenoyama, M. K. 1995. *S*-allele sequence diversity in natural populations of horsenettle, *Solanum carolinense* (horsenettle). *Heredity* 75:405-415.
- Richman, A. D., and Kohn, J. R. 1996. Learning from rejection: The evolutionary biology of single-locus incompatibility. *Trends Ecol. Evol.* 11:497-502.
- Richman, A. D., Uyenoyama, M. K., and Kohn, J. R. 1996. Allelic diversity and gene genealogy at the self-incompatibility locus in the Solanaceae. *Science* 273:1212-1216.
- Rowell, J. B. 1955. Functional role of compatibility factors and an *in vitro* test for sexual compatibility with haploid lines of *Ustilago zaeae*. *Phytopathology* 45:370-374.
- Rowell, J. B., and DeVay, J. E. 1954. Genetics of *Ustilago zaeae* in relation to basic problems of its pathogenicity. *Phytopathology* 44:356-362.

37. Schulz, B., Banuett, F., Dahl, M., Schlesinger, R., Schäfer, W., Martin, T., Herskowitz, I., and Kahmann, R. 1990. The *b* alleles of *U. maydis*, whose combinations program pathogenic development, code for polypeptides containing a homeodomain-related motif. *Cell* 60:295-306.
38. Silva, J. 1972. Alleles at the *b* incompatibility locus in Polish and North American populations of *Ustilago maydis* (DC) Corda. *Physiol. Plant Pathol.* 2:333-337.
39. Slatkin, M. 1987. Gene flow and the geographic structure of natural populations. *Science* 236:787-792.
40. Slatkin, M., and Barton, N. H. 1989. A comparison of three indirect methods for estimating average levels of gene flow. *Evolution* 43:1349-1368.
41. Takahata, N., Satta, Y., and Klein, J. 1992. Polymorphism and balancing selection at major histocompatibility complex loci. *Genetics* 130:925-938.
42. Taylor, M. F. J., Shen, Y., and Kreitman, M. E. 1995. A population genetic test of selection at the molecular level. *Science* 270:1497-1499.
43. Uyenoyama, M. K. 1995. A generalized least-squares estimate for the origin of sporophytic self-incompatibility. *Genetics* 139:975-992.
44. Whitehouse, H. L. K. 1949. Heterothallism and sex in the fungi. *Camb. Phil. Soc. Biol. Rev.* 24:411-447.
45. Wright, S. 1939. The distribution of self-sterility alleles in populations. *Genetics* 24:538-552.
46. Wright, S. 1960. On the number of self-incompatibility alleles maintained in equilibrium by a given mutation rate in a population of a given size: A reexamination. *Biometrics* 16:61-85.
47. Wright, S. 1964. The distribution of self-incompatibility alleles in populations. *Evolution* 18:609-618.
48. Yee, A. R., and Kronstad, J. W. 1993. Construction of chimeric alleles with altered specificity at the *b* incompatibility locus of *Ustilago maydis*. *Proc. Natl. Acad. Sci. U.S.A.* 90:664-668.
49. Yokoyama, S., and Hetherington, M. E. 1982. The expected number of self-incompatibility alleles in finite plant populations. *Heredity* 48:299-303.
50. Zolan, M. E., and Pukkila, P. J. 1986. Inheritance of DNA methylation in *Coprinus cinereus*. *Mol. Cell. Biol.* 6:195-200.