Panglobal distribution of a single clonal lineage of the Irish potato famine fungus

(DNA fingerprinting/*Phytophthora infestans*/migration/bottlenecks/founder effects)

STEPHEN B. GOODWIN, BARAK A. COHEN, AND WILLIAM E. FRY

Department of Plant Pathology, 334 Plant Science, Cornell University, Ithaca, NY 14853

Communicated by R. W. Allard, August 16, 1994

ABSTRACT   More than 300 isolates of the Irish potato famine fungus, *Phytophthora infestans*, collected in 20 countries on five continents, were analyzed for genetic variation at the mating type and two allozyme loci. A subset of more than 200 isolates was also analyzed for DNA "fingerprint" variation. A surprising result was that a single clonal lineage dominated most populations worldwide. All of the variation within this lineage appeared to have arisen by mitotic recombination or by mutation. In addition to the most common clonal lineage, a number of different, but apparently closely related, lineages occurred in the United States and Canada. The low levels of gene diversity in the derived populations compared to the presumed ancestral population in central Mexico indicate that *P. infestans* went through extreme genetic bottlenecks during its dispersal. The genetic data are consistent with the hypothesis that the initial migration of *P. infestans* in the 1840s was from Mexico to the United States and that only a single genetic individual was transported to Europe and subsequently to the rest of the world. If this hypothesis is correct, then the Irish potato famine was caused by a single clonal genotype of *P. infestans*.

Populations that are founded by a few individuals can be very different compared with the populations from which they were derived. Sampling effects, genetic drift, and inbreeding in small founder populations can lead to a severe reduction in the number of alleles and in gene diversity relative to the original source populations (1, 2). Loss of alleles is primarily dependent on bottleneck size, while the reduction in gene diversity depends both on bottleneck size and the rate of population growth after the bottleneck (1). For sexually reproducing organisms, the most extreme bottleneck occurs when a population is founded by a single inseminated female. However, bottlenecks can be even more extreme for organisms that are capable of asexual reproduction; theoretically, a population can be founded by a single clone.

Phytophthora infestans (Montagne) de Bary is an ideal organism for studying the effects of migration-induced bottlenecks on genetic variability of founder populations. This diploid oomycete causes the late blight disease of potato and tomato, and it appears to have evolved on wild potato relatives (*Solanum* species) in a limited area in the highlands of central Mexico (3, 4). At least two migration events out of central Mexico have now distributed *P. infestans* throughout the world. The first documented migration occurred in the 1840s when late blight appeared as a previously unknown disease on potatoes in the United States and Europe (5). This migration caused the Irish potato famine of 1846 and 1847. Evidently, only one mating type (A1) was distributed outside of central Mexico in this migration, so reproduction in these populations was exclusively asexual (3, 6).

The second migration of *P. infestans* probably occurred in the late 1970s (3), when the A2 mating type, and previously undetected alleles at two allozyme loci, began to appear in European populations (6–9). This migration caused profound changes in most of the recipient populations, with new migrant genotypes usually displacing the old genotypes in only a few years (6, 10). Recent analyses also indicate that sexual reproduction may be occurring in some of the new migrant populations in Europe (10). Although the routes of migration are not known with certainty, it may be possible to infer the pattern of past migrations by studying genetic variation within and among extant populations.

The purpose of this study was to quantify the effects of genetic bottlenecks on the genetic structure of asexually reproducing populations of an important plant pathogen. *P. infestans* was chosen for this study because its center of origin is well defined; a large number of well-characterized and easily assayed genetic markers are available; it has a defined life cycle including both sexual and asexual stages; much is already known about its past migrations; and samples are available from culture collections throughout the world. A secondary goal was to determine the genetic relationships among extant populations and thereby attempt to reconstruct the patterns of migration that led to the current panglobal distribution of *P. infestans*.

MATERIALS AND METHODS

Maintenance and Characterization of Fungal Isolates. More than 300 isolates were obtained from 20 countries on five continents (Table 1). Most isolates were received as axenic cultures. Other isolates were obtained from infected potato or tomato tissue by plating the tissue onto selective media containing antibiotics and fungicides (11). Sampling methods varied widely because the isolates were supplied by different collaborators in each location. For most collections, only one isolate was taken from each field, and the isolates were collected from many locations and over multiple years in each country. Following acquisition, all isolates were grown on rye A agar medium (12) at 18°C in the dark and were stored cryogenically at −135°C.

Mating types and genotypes at the two allozyme loci glucose-6-phosphate isomerase (*Gpi*) (EC 5.3.1.9) and peptidase (*Pep*) (EC 3.4.13.18) were determined by using standard techniques (13, 14). Over 200 isolates were included in the DNA fingerprinting analysis. Extractions of total genomic DNA and Southern analyses with DNA "fingerprint" probe RG57 were as described previously (15, 16). Probe RG57 is a highly polymorphic, moderately repetitive nuclear DNA from *P. infestans* that hybridizes to more than 25 different loci in *P. infestans* and related species (13, 15). Each locus segregates for the presence or absence of a band; no bands appeared to be allelic. Two pairs of cosegregating loci

Abbreviation: UPGMA, unweighted pair-group method of averages.

Table 1. Collection sites for *P. infestans* isolates from 20 countries on five continents

Location	Year(s) sampled	Sample size*
Africa		
Rwanda	1984–1987	21 (11)
Australia	1990	5 (2)
Central America		
Costa Rica	1982, 1993	8 (1)
Eurasia		
Belarus	1990	7 (7)
Estonia	1983	2 (2)
Ireland	Before 1984	2 (1)
Israel	1984–1986	9 (8)
Japan	1988	30 (29)
The Netherlands	1980–1985	27 (15)
Philippines	1989, 1991	28 (28)
Poland	1985–1989	62 (34)
Romania	Before 1990	1 (1)
Russia		
Moscow	1979–1990	5 (5)
Ural	1990	2 (2)
Switzerland	1977, 1981	3 (2)
United Kingdom	Before 1984	8 (5)
North America		
Canada	1980, 1982	3 (3)
United States	1946–1987	38 (23)
South America		
Bolivia	1990	1 (1)
Brazil	1984–1987	8 (4)
Peru	1982, 1984–1986	45 (26)

*All isolates were scored for mating type and dilocus allozyme genotype. Numbers of isolates scored for DNA fingerprints are indicated in parentheses.

have been identified, but the remaining loci do not appear to be closely linked (13, 15). Because each + allele is dominant, there are two potential phenotypes at a locus: +, if the locus is heterozygous or homozygous for the RG57 sequence; or −, if the RG57 sequence is absent from both chromosomes at a locus. Homozygotes can be distinguished from heterozygotes by band intensity (this has been verified for most loci by genetic analyses and densitometry) (15), so it is possible to determine the genotype of each isolate at each locus by visual inspection of the autoradiograms. This probe provides a means for identifying clonally related individuals with a high degree of certainty.

Data Analysis. A multilocus genotype was constructed for each isolate by combining data for mating type, dilocus allozyme genotype, and presence or absence of a band at all of the DNA fingerprint loci. Because the genetic basis for some of these traits (e.g., mating type) has not been determined, these are really multilocus phenotypes, but they will be referred to as genotypes here for simplicity. Allele frequency data were calculated for the two allozyme loci and a subset of the DNA fingerprint loci that could be scored genetically (23 loci in total).

Genotypes were classified as "old" or "new" according to the criteria of Spielman *et al.* (6); four allozyme genotypes were classified as new and four as old (see Table 2). Populations were subsequently classified as new or old depending on the genotypes they contained. Within a location, there was usually a temporal separation between old and new genotypes. For example, in Poland all isolates collected before 1988 had a single genotype, and were classified as old, whereas those collected in 1988 or later had additional genotypes and were classified as new. Cluster analysis of populations was based on frequencies of mating type, allozyme alleles, and DNA fingerprint bands in each population, using BIOSYS-1 (17). Trees were constructed by using the

unweighted pair-group method of averages (UPGMA) algorithm from a Rogers' modified genetic distance (17, 18) matrix. Gene diversity analysis of the allele frequency data was according to the methods of Nei and Chesser (19). Diversity of multilocus genotypes in each population was calculated by using the Shannon information statistic, $h = -\Sigma q_i \ln q_i$, where q_i is the frequency of the *i*th genotype. Because sample sizes were uneven and the possible number of genotypes was larger than the sample size in all locations, these values were normalized by ln N (20), the maximum diversity that could be obtained with a sample of size N, as described elsewhere (11, 21).

RESULTS

Mating Type and Allozymes. Genotypes from all locations except Australia, Costa Rica, and Russia could be classified as old or new on the basis of established criteria (Table 2) for mating type and allozyme genotype. Four locations contained old genotypes exclusively, six had new genotypes exclusively, and three had mixtures of old and new genotypes (Table 3). All isolates from Costa Rica had a previously undetected genotype for peptidase (with the *94* allele) and thus could not be classified as old or new. The Australian isolates had an old allozyme genotype but a different fingerprint pattern that was probably the result of a separate migration from Mexico. One Russian isolate had a new allozyme genotype but a fingerprint pattern characteristic of isolates in old populations. The A2 mating type was found in five locations, including Bolivia and Romania where it has not been reported previously. Additional allozyme and mating type data for other locations were reported previously: populations in the United States, Canada, and Peru contained only old genotypes, while those in The Netherlands, Poland, and Japan were mixtures of old and new genotypes (6, 23–25).

DNA Fingerprinting. The most striking result was that most (108/130 = 83%) of the isolates with old allozyme genotypes had the same 15-band fingerprint pattern (Fig. 1). Because these isolates were identical at the mating type (A1), two allozyme (*Gpi 86/100, Pep 92/100*), and 25 DNA fingerprint loci (Fig. 1), they are probably all clonally derived from a common ancestor and thus represent a single genetic individual. This genotype was identified previously in the United States and Canada, and it has been designated US-1 (14). The remaining 22 isolates had genotypes that were identical to US-1 except for a change at one or two allozyme or DNA fingerprint loci. Changes from presence to absence of a band at DNA fingerprint loci 9 and 10 (Fig. 2), or from heterozygosity to homozygosity at one of the allozyme loci, were the most common variations. The US-1 genotype was present in the oldest collections from 13 countries (Rwanda, Japan, Philippines, Ireland, The Netherlands, Poland, Switzerland, United Kingdom, Canada, United States, Brazil, Peru, and

Table 2. Allozymes and mating types characteristic of old and new *P. infestans* genotypes [after Spielman *et al.* (6)]

Classification	Allozyme genotype	
	Gpi	*Pep*
Old genotypes (A1 mating type only)	86/100	92/100
	86/100	100/100
	100/100	92/100
	100/100	92/92
New genotypes (both A1 and A2 mating types)	90/100	83/100
	90/100	100/100
	100/100	83/100
	100/100	100/100

Table 3. Results of allozyme and mating type analyses for previously uncharacterized isolates of *P. infestans*

Location	Genotype*	$N^{\dagger}$	Designation
Australia	A1, *86/100, 100/100*	5	New[‡]
Bolivia	A2, *100/100, 100/100*	1	New
Brazil[§]	A1, *86/100, 92/100*	1	Old
	A1, *86/100, 100/100*	2	Old
	A2, *100/100, 100/100*	2	New
	A2, *100/100, 96/96*	2	New
Belarus	A1, *90/100, 100/100*	7	New
Costa Rica	A1, *100/100, 94/100*	8	—[¶]
Estonia	A1, *90/100, 100/100*	1	New
	SF,[‖] *90/100, 100/100*	1	New
Israel	A2, *100/100, 100/100*	9	New
Peru**	A1, *86/100, 92/100*	11	Old
Philippines	A1, *86/100, 92/100*	25	Old
	A1, *86/100, 100/100*	3	Old
Romania	A2, *100/100, 100/100*	1	New
Russia			
Moscow	A1, *100/100, 100/100*	1	Old[‡]
	A2, *100/100, 100/100*	4	New
Ural	A1, *86/100, 92/100*	2	Old
Rwanda	A1, *86/100, 92/100*	17	Old
	A1, *90/100, 83/100*	2	New
	A1, *90/100, 100/100*	1	New
	A1, *100/100, 100/100*	1	New
Switzerland	A1, *86/100, 92/100*	3	Old
United Kingdom[††]	A1, *86/100, 92/100*	2	Old

*Listed as mating type, *Gpi, Pep*.

[†]Number of each genotype recovered.

[‡]Based on DNA fingerprint pattern.

[§]Brommonschenkel (22) reported two allozyme genotypes for these isolates but did not identify the alleles.

[¶]The peptidase genotype of these isolates is different from both old and new genotypes.

[‖]Self-fertile.

**Allozyme data on additional Peruvian isolates were reported by Tooley *et al.* (23).

[††]Allozyme data for additional United Kingdom isolates were reported by Spielman *et al.* (6).

Russia) on four continents (Africa, Eurasia, and North and South America).

Pairwise comparison of each genotype in each population with US-1 at the mating type, two allozyme, and 25 DNA fingerprint loci revealed the number of genetic differences from US-1. The mean number ($\pm$SD) of genetic differences from US-1 among isolates in old populations was 0.3 ($\pm$0.61). In contrast, the mean number of genetic differences from US-1 among isolates in new populations was 7.1 ($\pm$2.53), and in Mexican populations it was 6.8 ($\pm$1.39).

Cluster Analysis. All of the old populations were closely related and formed a very tight cluster (Fig. 3). New populations formed at least three UPGMA clusters (Fig. 3): western Europe and Rwanda; Eastern Europe and Israel; and Japan. Two of the Mexican populations, Toluca and Los Mochis, also clustered together.

Genetic Diversity in New vs. Old Populations. Mean gene diversity over all loci was similar in old (0.14 $\pm$ 0.028) and new (0.15 $\pm$ 0.081) populations. The new populations generally had more diversity for multilocus genotypes (mean genotypic diversity = 0.51 $\pm$ 0.466) than was present in the old populations (0.18 $\pm$ 0.221). However, this difference was not significant, primarily because only one new genotype each was identified in Australia, Israel, and Japan. There was more genotypic diversity among old isolates in the United States (genotypic diversity = 0.62) than in any of the other old populations, due to additional clonal lineages that appeared to be closely related to US-1 (14). Old populations were not genetically differentiated (G_{ST} among old populations =
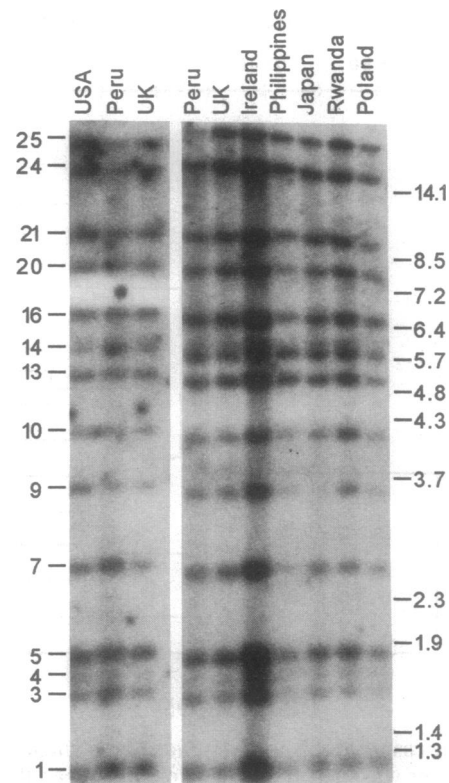

Fig. 1. DNA fingerprints of representative isolates of *P. infestans* with the US-1 genotype from eight countries on four continents. Total genomic DNA (approximately 2 μg per lane) was digested with the restriction enzyme *Eco*RI, and the fragments were separated on a 0.9% agarose gel, transferred to a nylon membrane, and hybridized with ^{32}P-labeled DNA of probe RG57. Bands revealed by probe RG57 are numbered on the left, where 1 is the smallest band and 25 the largest. Approximate locations of size standards (phage λ DNA digested with the restriction enzyme *Bst*EII) are indicated on the right, in kilobases.

0.07), compared with the new populations which were moderately differentiated (G_{ST} among new populations = 0.32).
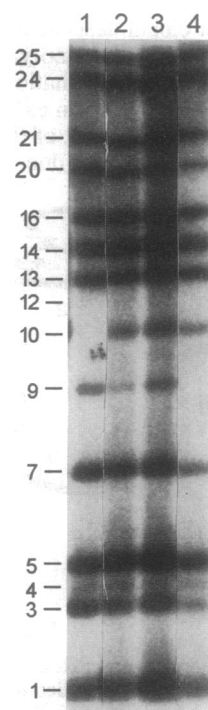

Fig. 2. DNA fingerprints of *P. infestans* isolates from Peru, showing variation within the US-1 clonal lineage. RG57 fingerprint band numbers are indicated on the left. Arrowheads on the right indicate loci that appear to be heterozygous in the US-1 genotype but that have become homozygous for the absence of the RG57 sequence in some isolates, presumably due to mitotic recombination or possibly mutation. Lane 1, isolate 810 has the US-1 genotype except for the loss of a band at locus 10; lanes 2 and 3, isolates 817 and 820 have the typical US-1 genotype; lane 4, isolate 828 has the US-1 genotype except for the loss of a band at locus 9.
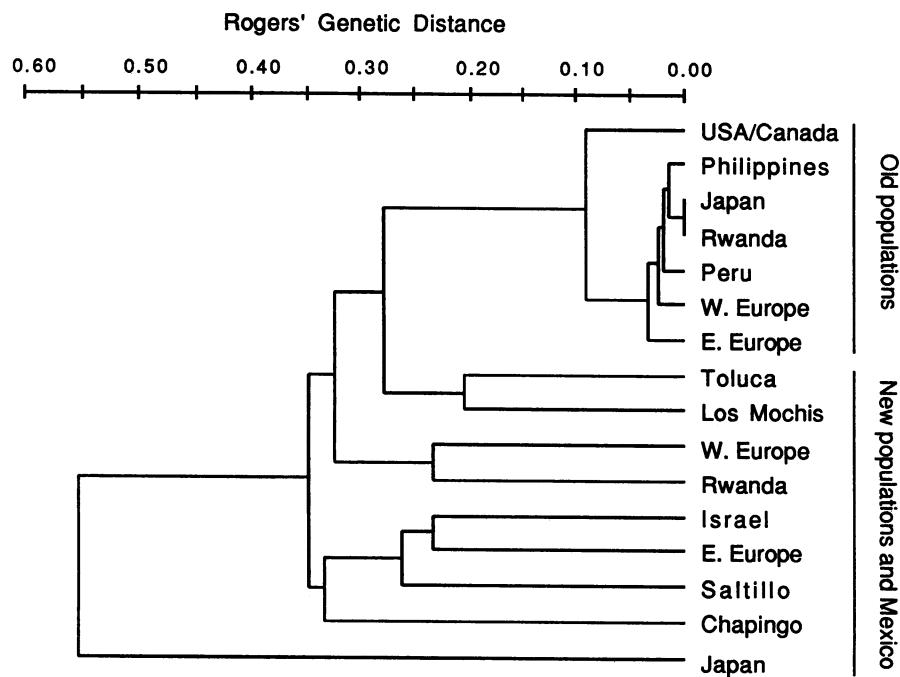
FIG. 3. UPGMA cluster analysis of old populations, new populations, and representative Mexican populations of *P. infestans* based on Rogers' genetic distance. Data for the Mexican populations Toluca, Los Mochis, Saltillo, and Chapingo were from a previous study (11).

DISCUSSION

A surprising result of this analysis was that a single genotype of *P. infestans* was distributed panglobally. When the pattern of genetic variation in extant populations is compared with the historical record (5, 26), it appears that the migration of this genotype occurred in at least three stages (Fig. 4). In the first migration a number of different genotypes, probably from a single source population, were transported from Mexico to the United States, most likely in infected tubers of native Mexican *Solanum* species. By 1845, these genotypes had spread throughout the eastern United States and Canada (5). Probably only one of these genotypes (US-1) was transported to Europe, where late blight disease appeared beginning in 1845 (26). Once into Europe, the US-1 genotype easily could have been spread panglobally, because Europe has been the source of seed potatoes for most of the world. If this scenario is correct, then the Irish potato famine was caused by a single genotype of *P. infestans*, of which US-1 is a direct clonal descendant.

Once out of Mexico, rare variants of the US-1 genotype appear to have arisen by mutation or mitotic recombination. One isolate from Peru, identical to US-1 except for an extra RG57 fingerprint band, was probably a mutant. Most of the other variation within the US-1 clonal lineage was due to

changes from heterozygosity to homozygosity at single allozyme or DNA fingerprint loci. US-1 is heterozygous at both allozyme loci and DNA fingerprint loci 9 and 10 (based on band intensity); mitotic recombination could have generated the observed variants at these heterozygous loci. From the original 92/100 genotype at the *Pep* locus in US-1, for example, mitotic recombination could generate the observed 92/92 and 100/100 homozygotes without affecting any of the other loci (unless they were also heterozygous and distal to the breakpoint on the same chromosome arm). Similarly, a DNA locus that was heterozygous +/− for a particular band could easily become +/+ or −/− through mitotic recombination. The Russian isolate with a new allozyme genotype but the old fingerprint pattern probably arose by mitotic recombination at both allozyme loci. Mutation and mitotic recombination can account for all of the variation in old populations except those in the United States and Canada.

There are at least two potential alternatives to the above interpretation, neither of which seems likely. The first is that the initial migrations contained multiple genotypes, and that selection following the migrations led to the fixation of US-1 in each location. However, it is very unlikely that US-1 would have been selected in each of the 13 countries on four

FIG. 4. Probable paths of migration of *P. infestans* from Mexico throughout the world, beginning in the 1840s. There were at least three waves of migration. The first migration, 1, was probably from Mexico into the northeastern United States in 1842 or 1843. The second migration, 2, was probably from the United States into Europe around 1845. Migrations from Europe continued throughout the world, 3, in subsequent years.

continents in which it has been identified. If the initial migrating population had been genetically diverse, genetic drift and founder effects most likely would have fixed different genotypes in different locations. Furthermore, multiple genotypes could have survived until the present, as seems to have occurred in the United States and Canada (14). The second possibility is that there were multiple migrations out of Mexico. However, this also probably would have resulted in the fixation of different genotypes in different locations. The probability of obtaining US-1 by random recombination in contemporary central Mexican populations is only 2 × 10⁻⁵ [based on data for mating type and two allozyme and 23 polymorphic DNA fingerprinting loci (11)], so it is highly unlikely that US-1 could have arisen independently in more than one location. Therefore, the only likely explanation is that all isolates with the US-1 genotype were derived asexually from a common ancestor by migration, and that this was the only genotype introduced to most locations.

The result of these migrations was a reduction both in the number of alleles and in gene diversity in the derived populations compared with those in central Mexico. Nine alleles were detected at the two allozyme loci (six for *Gpi* and three for *Pep*) in Mexican populations of *P. infestans* (11), compared with only four (two at each locus) in the derived populations. Because a single diploid individual can contain about 50% of the gene diversity in the parent population (1), *P. infestans* populations that were founded by single genetic individuals should retain about half the gene diversity in the source population. Consistent with this prediction, average gene diversity for two central Mexican populations was 0.24 (from table 7 of ref. 11) compared with 0.14 for the same loci in derived populations worldwide.

This appears to be the most severe reduction in genetic diversity associated with colonization yet reported, and is, to our knowledge, the first in which populations were founded by the theoretical minimum effective population size of one genetic individual. Previous studies have reported similar, although less dramatic, decreases in genetic variability in colonizing populations of organisms as diverse as lizards (27), pitcher plants (28), and walnut husk flies (29). However, there was no loss of genetic variation in populations of the face fly that colonized North America from Europe (30). Extreme losses of genetic variability occurred only when the number of founder individuals was very small.

In retrospect, severe genetic bottlenecks associated with the panglobal migration of *P. infestans* are not that surprising, given the biology of the organism and the modes of transport available in the 1840s. *P. infestans* (as mycelium and asexual spores) does not survive without its host; long-distance migration must have been accomplished by the movement of infected potato tubers or living plants in a ship's hold, where survival would be very low. Furthermore, the traditional food of the people of Mexico is maize (4). Because potatoes were not cultivated on a commercial scale in Mexico until the 1950s, and the wild *Solanum* hosts of *P. infestans* in Mexico produce tubers that are small and not commercially marketable, movement of infected tubers would have been very rare. Oospores (the products of sexual reproduction) easily could have survived long-distance transport, but it seems extremely unlikely that oospore-infested soil or tissue from wild *Solanum* species would have been brought from Mexico into the United States or Europe. This may explain why late blight disease did not appear in Europe until more than 250 years after the introduction of cultivated potatoes from South American in the late 1500s.

1. Nei, M., Maruyama, T. & Chakraborty, R. (1975) *Evolution* **29**, 1–10.
2. Maruyama, T. & Fuerst, P. A. (1985) *Genetics* **111**, 675–689.
3. Niederhauser, J. S. (1991) in *Phytophthora*, eds. Lucas, J. A., Shattock, R. C., Shaw, D. S. & Cooke, L. R. (Cambridge Univ. Press, Cambridge, U.K.), pp. 25–45.
4. Reddick, D. (1939) *Chron. Bot.* **5**, 410–412.
5. Stevens, N. E. (1933) *J. Wash. Acad. Sci.* **23**, 435–446.
6. Spielman, L. J., Drenth, A., Davidse, L. C., Sujkowski, L. S., Gu, W.-K., Tooley, P. W. & Fry, W. E. (1991) *Plant Pathol.* **40**, 422–430.
7. Hohl, H. R. & Iselin, K. (1984) *Trans. Br. Mycol. Soc.* **83**, 529–530.
8. Malcolmson, J. F. (1985) *Trans. Br. Mycol. Soc.* **85**, 531.
9. Tantius, P. M., Fyfe, A. M., Shaw, D. S. & Shattock, R. C. (1986) *Plant Pathol.* **35**, 578–581.
10. Sujkowski, L. S., Goodwin, S. B., Dyer, A. T. & Fry, W. E. (1994) *Phytopathology* **84**, 201–207.
11. Goodwin, S. B., Spielman, L. J., Matuszak, J. M., Bergeron, S. N. & Fry, W. E. (1992) *Phytopathology* **82**, 955–961.
12. Caten, C. E. & Jinks, J. L. (1968) *Can. J. Bot.* **46**, 329–348.
13. Goodwin, S. B. & Fry, W. E. (1994) *Exp. Mycol.* **18**, 20–32.
14. Goodwin, S. B., Cohen, B. A., Deahl, K. L. & Fry, W. E. (1994) *Phytopathology* **84**, 553–558.
15. Goodwin, S. B., Drenth, A. & Fry, W. E. (1992) *Curr. Genet.* **22**, 107–115.
16. Drenth, A., Goodwin, S. B., Fry, W. E. & Davidse, L. C. (1993) *Phytopathology* **83**, 1087–1092.
17. Swofford, D. L. & Selander, R. B. (1981) *J. Hered.* **72**, 281–283.
18. Rogers, J. S. (1972) *Univ. Tex. Publ.* **7213**, 145–153.
19. Nei, M. & Chesser, R. K. (1983) *Ann. Hum. Genet.* **47**, 253–259.
20. Sheldon, A. L. (1969) *Ecology* **50**, 466–467.
21. Goodwin, S. B., Saghai Maroof, M. A., Allard, R. W. & Webster, R. K. (1993) *Mycol. Res.* **97**, 49–58.
22. Brommonschenkel, S. H. (1988) Thesis (Universidade Federal de Viçosa, Viçosa, Minas Gerais, Brazil).
23. Tooley, P. W., Therrien, C. D. & Ritch, D. L. (1989) *Phytopathology* **79**, 478–481.
24. Mosa, A. A., Kobayashi, K., Ogoshi, A., Kato, M. & Sato, N. (1993) *Plant Pathol.* **42**, 26–34.
25. Tooley, P. W., Fry, W. E. & Villarreal Gonzalez, M. J. (1985) *J. Hered.* **76**, 431–435.
26. Bourke, P. M. A. (1964) *Nature (London)* **203**, 805–808.
27. Taylor, C. E. & Gorman, G. C. (1975) *Heredity* **35**, 241–247.
28. Schwaegerle, K. E. & Schaal, B. A. (1979) *Evolution* **33**, 1210–1218.
29. Berlocher, S. H. (1984) *Evolution* **38**, 906–918.
30. Bryant, E. H., van Dijk, H. & van Delden, W. (1981) *Evolution* **35**, 872–881.