Genetic expansion of randomly mating founder populations of *Alternaria brassicicola* infecting *Cakile maritima* in Australia

C. C. Linde*, J. A. Liles, and P. H. Thrall

1 Evolution, Ecology and Genetics, Research School of Biology, College of Medicine, Biology and Environment, Bldg. 116, Daley rd, Australian National University, Canberra, ACT 0200, Australia; 2 CSIRO Plant Industry, GPO Box 1600, Canberra, ACT 2601, Australia.

Running title: Genetic expansion of *Alternaria brassicicola* populations

*Corresponding author. Mailing address: 1 Botany and Zoology, Research School of Biology, College of Medicine, Biology and Environment, Bldg. 116, Daley rd, Australian National University, Canberra, ACT 0200, Australia. Phone: +61 2 61257682. Fax: +61 2 612 55573. E-mail: celeste.linde@anu.edu.au
ABSTRACT

Founder populations of fungal plant pathogens are expected to have low levels of genetic diversity, coupled with further genetic drift due to e.g. limited host availability which should result in additional population bottlenecks. This study uses microsatellite markers in the interaction between Cakile maritima and the fungal pathogen Alternaria brassicicola to explore genetic expectations associated with such situations. The host, C. maritima was introduced to Australia approximately 100 years ago but it is unknown whether the pathogen was already present in Australia as it has a wide occurrence, or whether it has been introduced to Australia on brassicaceous hosts. Eleven A. brassicicola populations were studied and all show moderate levels of gene and genotypic diversity. Chi-square tests on frequencies of mating-type alleles, a large number of genotypes and linkage equilibrium among SSR loci, all suggest A. brassicicola reproduces sexually. Significant genetic differentiation was found among populations, but there was no evidence for isolation by distance effects. Bayesian analyses identified eight clusters where the inferred clusters did not represent geographical populations but instead consisted of individuals admixed from all populations. Further analysis indicated that fungal populations are more likely to have experienced a recent population expansion than a population bottleneck. It is suggested that A. brassicicola has been introduced multiple times to Australia, potentially increasing the diversity and size of any A. brassicicola populations already present in Australia; combined with its ability to reproduce sexually, such processes appear to have increased the evolutionary potential of this pathogen through recent population expansions.
INTRODUCTION

Analyses of population genetic diversity provide useful information on the epidemiology and evolutionary history of infectious diseases (7). Infectious agents vary in effective population sizes, transmission abilities and reproduction, all which can be used to infer their population histories (26). Such inferences can in turn be used to determine whether the pathogen has been recently introduced (i.e. a founder population) as well as the number of introductions that have occurred. Introduction of plant pathogens into new geographic areas have occurred commonly in the past for a multitude of organisms, reducing their population diversity in the founder populations but sometimes also causing an epidemic on naïve hosts in the introduced areas. For example, Phytophthora infestans, a pathogen of potato was introduced from Mexico into the US and then only one genotype was introduced into Europe (25), causing the death of more than 1.5 million people due to starvation (13). Cryphonectria parasitica which causes chestnut blight was introduced from Asia into North America (3) where it underwent a host shift from Asian chestnuts to North American chestnuts, threatening the existence of North American chestnut trees.

For both plant and animal pathogens, introduction into new areas can be a major driver of emerging diseases, i.e. host shifts (4, 45, 67). In the case of plant pathogens, range expansion can occur after the introduction of an invasive host species (8, 42), resulting in a pathogen with higher evolutionary potential than represented in the original founder population. Population genetic expansions for plant pathogens could result in the pathogen able to infect a larger number of plant genotypes or even new host species following a host shift (39).

In agricultural systems, many fungal plant pathogens have genetic structures consistent with having experienced a population bottleneck due to founder events (25, 34, 42, 50); others, such as Rhynchosporium secalis on barley (33) and Mycosphaerella graminicola on
wheat (6, 33) eventually underwent a population expansion, possibly due to an increase in host availability. One consequence of this is the accumulation of additional mutations, thereby creating novel genotypes and increasing the effective population size of the pathogen. However, even in well-studied agricultural plant-pathogen systems, tests for population expansions have rarely been conducted. In most cases where there is evidence for the genetic expansion of pathogen populations, it was a result of additional migrants coming into the founder population (e.g. *Phytophthora infestans*, (24), or the occurrence of sexual reproduction as in European populations of *C. parasitica* (42). An increase in genotypic diversity has also been demonstrated in a founder population of Dutch elm disease in Portugal caused by *Ophiostomo novo-ulmi*, which was able to reproduce sexually following the introduction of the other mating type (14). The evolution of founder pathogen populations is likely to be influenced by a range of ecological and life-history parameters including host plant population density, host longevity, and host and pathogen mating systems and dispersal ability (7, 46).

In theory, bottlenecks reduce gene diversity and the number of alleles in populations. Empirically it has been shown that founder events frequently reduce allelic diversity and more rarely gene diversity (16, 22, 29). Thus, we can detect whether a population has experienced a recent population expansion after a demographic bottleneck by using genetic information to determine whether a population has a deficit in gene diversity relative to the number of alleles (17, 36) or experienced a recent population bottleneck in which case an excess in gene diversity relative to the number of alleles under a mutation-drift model will be present (32).

In this study, we investigate the introduced *Cakile maritima-Alternaria brassicicola* host-pathogen interaction in Australia. *Alternaria brassicicola* is a heterothallic haploid fungus that causes black spots on leaves, stems and fruits of a wide range of brassicaceous
hosts. The pathogen has a widespread distribution and is commonly found on both agricultural brassicaceous crops and wild species (54, 62, 66). The pathogen naturally occurs on *Cakile maritima*, a succulent fore-dune annual native to the Mediterranean and Western Europe that was introduced to Australia more than 100 years ago (51). *C. maritima* is a self-incompatible obligate outcrossing species with low levels of self-fertilisation (63), and produces seeds that can survive for long time periods immersed in seawater. Therefore, seed dispersal between populations is possible both in terms of seed survival and ocean current patterns (23), and it is estimated that in Australia, *C. maritima* has spread along the coast at rates of 50-100 km per year (52). Conidia (asexual spores) of *A. brassicicola* are airborne and dispersed through wind and rain-splash, as well as vertically via infected seed transported in ocean currents. In such interactions where there is significant potential for seed-borne pathogen transmission, it is expected that the pathogen population should track the dispersal of the host and, to some extent, display a population structure that reflects the host’s geography.

Whether *A. brassicicola* is able to reproduce sexually is still unknown, although it is presumed to be an asexual fungus as it has no known sexual cycle. However, studies using amplified fragment length polymorphisms (AFLPs) have shown high levels of gene and genotypic diversity in Australian populations of *A. brassicicola* infecting *C. maritima* (10, 11). Although populations were in significant gametic disequilibrium, the authors deemed the levels of gametic disequilibrium to be low and predicted that a sexual stage should be present based on the high genotypic diversity observed (10, 11). Asexual organisms are generally characterized by an over-representation of particular genotypes and therefore genotype diversity is generally low, whereas recombination of alleles creates numerous new genotypes in sexually reproducing populations (38, 41, 64). It is worth noting, in the context of the preceding discussion regarding pathogen dispersal that sexual reproduction in *A. brassicicola*
would result in the formation of sexual spores (ascospores), which are smaller than asexual spores, potentially leading to longer distance wind dispersal.

The aims of this study were to further investigate the population genetic structure of *A. brassicicola* and to determine whether populations are randomly mating. The mating type frequencies of *A. brassicicola* in either natural or agricultural host populations have never been determined. Random mating was investigated by a) analyzing levels of genetic diversity and linkage disequilibrium in *Alternaria* populations characterized with microsatellite markers, and b) quantifying the extent of frequency-dependent-selection of pathogen mating type alleles in these populations. Sexual reproduction can only occur in heterothallic fungi when isolates of opposite mating types are in close physical proximity to each other, therefore the finding that both mating types occur in a population at equal frequencies (i.e. through frequency-dependent-selection on the mating types) would infer that random mating is occurring. Furthermore we investigated the evolution and genetic structure of fungal populations. Specifically we investigated whether; a) local populations belong to one panmictic population representing one population, b) populations have undergone a recent (in the last 100 years) population bottleneck, typical of founder populations, or c) whether populations have expanded genetically as well as geographically, resulting in pathogen populations with an evolutionary potential higher than in the original population (39).

METHODS

**Populations.** Populations of *A. brassicicola* were collected from *C. maritima* growing in three regions along the New South Wales (NSW) south coast of Australia. Three populations were collected previously from Central-Tilba (CT) and four each from Durras (DU) and Moruya-Bodalla (MB) spanning 34°40′S, 150°17′E to 36°19′S, 150°7′E (62). All the populations occur on beaches and are separated by rocky outcrops and headlands. The
sites chosen for this study represent a subset of populations that form part of ongoing work on the metapopulation dynamics of the *Cakile-Alternaria* interaction (Thrall *et al*., 2001, 2005). Collection, isolation and maintenance of isolates were as described in (61). Infected plants were sampled randomly within populations. A total of 401 isolates were analyzed in this study (Table 1).

**Mating type determination.** DNA was either extracted using the protocol of (40) for a previous AFLP study (10), or with a DNeasy® Plant Mini Kit (Qiagen) according to the recommendations of the manufacturer for isolates that were not used in the AFLP study. Initial attempts to determine the mating types of the *A. brassicicola* isolates were made using primers designed by Berbee and co-workers (9). Primers BPHO5 and BPHO4 were used to amplify the *MAT1-1* gene and BPHMG1 and BPHMG2 to amplify the *MAT1-2* gene in a multiplex PCR reaction. Because amplification was poor for the majority of the isolates, new primers were designed from the GenBank accessions of the *MAT1-1* (AY042093) and *MAT1-2* (AY042092) loci. These primers were MAT1-1F (5'-CTCAATGCTTTTGTGATT-3'), MAT1-1R (5'-CCGAGTGTCCAGGAATTT-3'), MAT1-2F (5'-TCTTCAGAGATGCGATGCAC-3') and MAT1-2 R (5'-CTCTTCTTTGCAGACTGTG-3'). The new primers were designed to amplify a 683 bp fragment in *MAT1-1* and a 276 bp fragment in *MAT1-2* isolates to provide unambiguous size distinctions on gels. A multiplex PCR mixture contained: 5-20 ng DNA in 20 µl reactions, 2.5 mM MgCl₂, 2.5 mM of each dNTP, 0.5 U Taq polymerase (Fisher Biotec), 2 µl of 10 x Reaction Buffer (Fisher Biotec), and 2 µM of each of the four primers. PCRs were performed on an Eppendorf Mastercycler with the following conditions: 95 °C for two minutes; 30 cycles of 95 °C for 30 seconds and 55 °C for 1 minute, followed by a final extension step at 72 °C for 45 seconds. PCR products were visualized on 1% agarose gels (0.5 x TBE) and stained with ethidium bromide. Mating type analysis was only conducted on isolates deemed as belonging to different haplotypes
within populations (e.g., clone corrected populations) as determined with microsatellite analyses (see below).

**Microsatellites.** Eleven primer pairs previously designed for *A. brassicicola* (5), were used to characterize the Australian *A. brassicicola* populations. The forward primer of each microsatellite marker was labeled with a M13 (-21) tail (5'-TGTAAACGACGGCCAGT). The reverse primers were retained in their original form, and a third universal M13 (-21) reporter primer was added which was fluorescently labeled with either VIC, FAM or NED (Invitrogen) (55). The following PCR conditions were used for each locus: each 20 µl PCR reaction contained 5-20 ng of DNA template, 1 X PCR reaction buffer, 0.25 mM dNTPs, 2.0 mM MgCl₂, 0.5 units of Taq polymerase, 0.1 µM f-21M13 labeled forward primer, 0.25 µM of reverse primer and 0.2 µM of fluorescently labeled (VIC, FAM or NED) -21M13 universal primer. PCR conditions were an initial 3 min denaturation at 94 °C, followed by 25-35 cycles of 94 °C for 30 seconds, 30 seconds at T_a and 72 °C for 30 seconds to amplify a locus specific fragment. Primers used (5) were Abmic-5 and Abmic-8 in a multiplex reaction with a T_a 60 °C, whereas Abmic-1, -3, -7, -9, -10 and -12 were amplified at a T_a of 55 °C and Abmic-2, -6, and -11 at a T_a of 60 °C. In order to amplify fragments with the universal M13 (-21) reporter primer another eight cycles were carried out at 94 °C for 30 seconds, 50 °C for 30 seconds and 72 °C for 30 seconds. All loci were amplified in single reactions, but only Abmic-5 and Abmic-8 were multiplexed. Only one allele was amplified per locus as expected for a haploid organism. For fragment analyses, 1-5 µl of PCR product was added to 0.35 µl of GeneScan™ -500 LIZ® Size standard (Applied Biosystems) and 9 µl of Hi-Di™ Formamide. The mix was denatured for 3 minutes at 95 °C and then cooled on ice. Samples were loaded onto an ABI 310 Prism Genetic Analyzer. Fragments were analysed using Genescan™ software (Applied Biosystems). Individual alleles at each locus were assigned using fragment lengths.
Data analysis. Isolates with the same alleles at all loci and the same mating type were considered to be clones of the same multilocus haplotype. To determine the population genotypic diversity, the maximum possible genotypic diversity ($\hat{G}/N$) (60) and clone fraction were calculated for each population and region. The clonal fraction was calculated as the occurrence and frequency of clones within a population, $\frac{N-G}{N}$, where $N$ is the sample size and $G$ is the number of haplotypes present. In some cases, multiple pathogen isolates were derived from the same plant individual. To avoid sampling bias where the likelihood of finding the same clone at a small scale increases due to splash dispersal of conidia, an analysis was conducted using only one isolate per genotype per plant. For all other analyses, populations were clone corrected to avoid overrepresentation of alleles present in clones. For genotypic analyses, local populations were also pooled to represent regional populations representing Central Tilba, Durras and Moruya-Bodalla.

Selective neutrality of loci was examined with the Ewens-Watterson test (58, 68) in POPGENE v3.2 (69). Nei’s gene diversity (43) and the number of alleles present in populations were determined in POPGENE v3.2 (69). To measure population relatedness, a modified version of Wright’s $Fst$ for haploids called theta ($\theta$) was used in pairwise estimates of $\theta$ in MULTILOCUS v1.3 (1). Theta was calculated between all pairs of clone corrected populations as well as between the three regions (Central Tilba, Durras and Moruya-Bodalla). To examine whether observations deviated significantly from the hypothesis of no linkage disequilibrium among loci, the observed value was compared to the results of 1000 randomized data sets. Relationships between genetic relatedness ($\theta$) and among-population distances (i.e. isolation by distance) were estimated with Mantel tests in which the significance of correlation was estimated with a 1000 random permutations tests using GENALEX (47). The distribution of genetic variance within and among populations was determined with an analysis of molecular variance (AMOVA) in GENALEX (47). The data
was partitioned into regional populations representing Central Tilba, Durras and Moruya-Bodalla.

Random mating in populations was investigated with a chi-square test ($\chi^2$) on mating type ratios with the null hypothesis that each *A. brassicicola* population did not significantly differ from a 1:1 mating type ratio. Microsatellite data were used to test whether populations were in gametic disequilibrium using the index of association ($I_A$) (15) test in *MULTILOCUS v1.3* (2) running 1000 randomizations. An alternative measure of the index of association ($\tilde{r}_d$) that is less sensitive to the number of loci (2) was also performed in *MULTILOCUS v1.3*.

To examine the levels of admixture of genotypes among populations, populations were analysed with *STRUCTURE v2.2* (20, 21, 49). In this Bayesian approach, multilocus genotypic data are used to define a set of populations with distinct allele frequencies and assign individuals probabilistically to them. The degree of population substructure was investigated using microsatellite data as well as AFLP data from a previous study (11). Only populations from Durras were included in the earlier study, however, the inclusion of additional AFLP analyses conducted on the populations from Central Tilba and Moruya-Bodalla provided AFLP data for a total of 380 isolates. In the program *STRUCTURE v2.2*, an admixture ancestry model-based clustering method (allowing mixed ancestry among individuals from different K populations) with correlated allele frequencies (i.e. allowing allele frequencies among populations to be similar) was used. Four independent runs of one to twenty subpopulations ($K = 1$–20) were performed using 100,000 Markov chain steps after a burn-in period of 50,000 steps. We compared the likelihood estimate of each of the K values assayed to determine the number of K populations present in *A. brassicicola*. The number of genetically discrete populations were estimated based on the maximum log probability of data $lnP(D)$ for different numbers of K, and by using the statistic $\Delta K$ (19), which considers the rate of change in $lnP(D)$ values among successive K runs to account for patterns of dispersal that are not
homogenous among populations.

To determine whether there was an excess (recent population bottleneck) or deficit (recent population expansion) in $H_e$ relative to the number of alleles present in *A. brassicicola* populations, we used BOTTLENECK v 1.2 (17, 48). Rare alleles are lost faster than gene diversity ($H$) and can be further reduced in founder populations or bottlenecked populations that have experienced recent reductions in their effective population sizes. Therefore, $H$ becomes larger than the gene diversity expected under mutation-drift equilibrium ($H_{EQ}$) after going through a bottleneck, because $H_{EQ}$ is based on the observed number of alleles. When population size is restored, the average number of alleles is predicted to increase faster than the gene diversity until reaching mutation-drift equilibrium (44). The Sign and Wilcoxon significance tests were used to determine whether loci displayed a significant excess ($H > H_{EQ}$) or deficit in gene diversity under a mutation-drift equilibrium for loci evolving under the stepwise mutation (SMM) and two phased mutation (TPA70%: SMM and 30% IAM) models (17). A qualitative descriptor of the allele frequency distribution ("mode-shift" indicator) which discriminates bottlenecked populations from stable populations (35) was also investigated.

RESULTS

Gene and genotypic diversity. A total of 401 isolates were analyzed with 11 microsatellite markers of which all but one (Abmic12) were polymorphic. All polymorphic loci were selectively neutral according to the Ewens-Watterson test for neutrality (data not shown). The highest number of alleles per locus within populations was seven in population CT0S (locus: Abmic2). The number of alleles observed among populations as an average across loci ranged from 1.8 to 2.5 (Table 1). Gene diversity ($H$) values were similar across all populations, and ranged from $H = 0.22$ to 0.36 (Table 1).
Across all populations, 164 distinct haplotypes were identified (\( \hat{G}/N\% = 15 \), clonal fraction = 0.59) (Table 1). Genotypic diversities varied among populations, with the highest genotypic diversity observed in DUD (\( \hat{G}/N\% = 61 \), clonal fraction = 0.20) and the lowest in DUI (\( \hat{G}/N\% = 17 \), clonal fraction = 0.70). When populations were clone corrected to represent only one member of a clone per plant to avoid over-representation of clones (e.g. due to splash dispersal of asexual spores), genotypic diversities were always higher due to the smaller sample sizes, and ranged from the lowest in DUI (\( \hat{G}/N\% = 19 \), clonal fraction = 0.68) to the highest in CT18 (\( \hat{G}/N\% = 68 \), clonal fraction = 0.12) (Table 1).

A number of haplotypes were shared among different populations within regions, and even across all three geographic regions examined. Thirty-nine of the haplotypes occurred at least twice in two or more different populations. The most common haplotype was found in six populations and was isolated 29 times, which accounts for 7% of the total sample. This haplotype was found in two adjacent regions, Durras and Moruya-Bodalla, which are approximately 34 km apart. A total of 83 haplotypes were only found once in the entire metapopulation.

**Mating types and linkage disequilibrium.** Haplotypes representing individual populations only were assayed for their mating type allele with the new primers designed from the GenBank accessions of the mating type locus. Of the 211 isolates analysed, 210 produced a single amplicon of the expected size corresponding to a MAT1-1 (683 bp) or MAT1-2 (276 bp) haplotype. One isolate produced a double band. This isolate most likely was contaminated during PCR preparation and was excluded from the analyses.

Both mating types were found in each of the eleven populations of *A. brassicicola*. In seven out of the eleven populations both mating types were found on the same plant. Four of the eleven populations (CT18, DUD, DUI and MBS) differed significantly from a 1:1 mating type ratio based on a \( \chi^2 \) test (Table 2). At the regional scale, only Central-Tilba deviated
significantly from a 1:1 mating type ratio ($P = 0.01$). On average across all populations analysed, mating type frequencies did not differ significantly from a 1:1 ratio (Table 2). Using the $I_A$ and $r_d$ tests, only two populations were not in linkage equilibrium, DUI and DUL (Table 2).

**Population structure.** Estimates of $\theta$ (among-population differentiation) showed low to moderate levels of differentiation in pairwise comparisons between the majority of populations (Table 3). In general, $\theta$ values were higher among populations from different regions, than within the same region. When populations were combined into regional populations, $\theta$ values were all significant but higher between the two regions that are geographically the furthest separated i.e. Central Tilba and Durras (approximately 83 kms apart) ($\theta = 0.10; P = 0.001$) than between Central Tilba and Moruya-Bodalla ($\theta = 0.06; P = 0.001$) or Durras and Moruya-Bodalla ($\theta = 0.06; P = 0.001$). Despite this, there was no significant relationship between genetic distance and geographic distance ($R_{xy} = 0.128, P = 0.157$) or $\Phi_{pt}$ and geographic distance ($R_{xy} = 0.05, P = 0.30$) in a Mantel test, indicating populations have not reached a genetic drift – gene flow equilibrium, typical of recent founder populations.

In a hierarchical analysis of genetic distribution (AMOVA) genetic differentiation was low although significant among regions (df = 2, $\Phi_{pt} = 0.017; P = 0.010$). Significant genetic differentiation ($P = 0.010$) was also observed among populations within regions (df = 8, $\Phi_{pt} = 0.137$); and within populations (df = 200, $\Phi_{pt} = 0.151$). Most of the genetic diversity observed was distributed among individuals within a population (85%). Thirteen percent of the genetic variation was distributed among populations and only 2% among regions.

A cluster based method was used to infer the minimum number of clusters ($K$) required to explain the total sum of genetic variation observed (49) in the SSR and AFLP data sets. Results from STRUCTURE indicated that $K = 8$ or $K = 9$ population groups exist...
among *A. brassicicola* isolates along the NSW coast as estimated from the SSR or AFLP data respectively (Fig. 1). In both data sets, $\Delta K$ gave the strongest indicator of the number of populations estimated (Fig. 1). In the AFLP data set, populations mostly corresponded to the broad geographic regions from which isolates were collected. Clusters 1, 2, 5, 6 and 7 consisted exclusively of isolates collected from Durras, although the clusters did not strictly represent individual subpopulations (i.e. DUD, DUI, DUG, DUL), as these were admixed among clusters. Clusters 3 and 4 consisted of a mixture of Moruya and Durras isolates, whereas clusters 8 and 9 consisted of a mixture of Central Tilba and Moruya isolates. Patterns were less clear in the SSR data set, with all clusters having representatives of most subpopulations analysed (Fig. 2). Grouping of individuals by geographical location (populations) clearly shows admixture among populations as most populations consist of individuals from all eight inferred clusters (Fig. 3).

**Evidence for population expansion.** All SSR loci followed a stepwise mutation method (SMM) of evolution (Table 4). When all local populations were combined into a single population representing the southern coast of NSW, the sign tests in BOTTLENECK showed a significant $H$ deficit in 10 of the 11 loci under both a SMM ($P = 0.0015$) and TPM ($P = 0.0018$) model of evolution, indicating recent population expansion. Only one locus (locus 9) under both models of evolution showed a significant gene diversity excess, indicating a recent population bottleneck (Table 4). With the Wilcoxon test, the probability that all loci had a $H$ deficit was $P = 0.0017$ for TPM and $P = 0.0005$ for SMM. In regional populations, the Wilcoxon’s test also showed a significant $H$ deficit in 10 of the 11 loci under a SMM in Durras ($P = 0.0017$) and Moruya-Bodalla ($P = 0.0005$), and in 8 of the 11 loci in Central Tilba ($P = 0.0105$), indicating recent population expansion in regional populations. Sign values gave similar probability values for regional populations. Allele frequencies
followed a normal L-shaped distribution, indicating populations are at mutation-drift equilibrium and did not experience a recent population bottleneck in all comparisons.

**DISCUSSION**

In an earlier study, Bock and colleagues (10, 11) examined the genetic diversity of *A. brassicicola* populations along the southeastern coast of Australia, using AFLPs as molecular markers. The current study substantially expands on that earlier analysis by additionally including pathogen populations from two other regions, Central Tilba and Moruya-Bodalla thus allowing evaluation of genetic patterns across multiple spatial scales. These data were complemented with additional analyses using microsatellites, which are expected to be more sensitive to recent phylogeographic and demographic events (e.g. bottlenecks or population expansions) due to their faster rate of evolution. The high variation and co-dominant nature of microsatellites make them particularly sensitive for detecting population changes (59). For example, microsatellite mutation rates were estimated to be several orders of magnitude greater than regular non-repetitive DNA for the ascomycetous fungus *Neurospora* (18).

The results from the analyses of microsatellite data indicated moderate levels of gene and genotypic diversity among all populations. Of the 401 isolates included in this analysis, 211 haplotypes (52%) could be identified. Given the high variability of SSR markers, it was somewhat surprising to find that AFLPs were able to distinguish more haplotypes (91%) in a previous study (202 from 222) (11). However, only 12 microsatellite loci were used in this study, of which one was monomorphic. Compared to three primer combinations used with AFLPs resulting in 47 to 69 loci per primer combination, in this case AFLPs clearly provide a locus rich marker with higher resolution.

**Population structure and mating.** Significant pairwise population differentiation among most populations indicates significant spatial structure. Local beach populations of
Cakile maritima are separated by natural barriers (rocky outcrops), which most likely provide enough isolation for population structure to develop in the pathogen, despite effective dispersal of seed in ocean currents. At the same time, the absence of isolation by distance (no significant association between population differentiation and geographic distance) \( (R_{xy} = 0.05; P = 0.30) \) indicates that populations have not reached a gene flow/genetic drift equilibrium (57). This might be because these populations represent recent founder populations (30) where a sufficient number of generations may not have passed since colonization of C. maritima 100 years ago. Alternatively, dispersal in ocean currents is not restricted to adjacent populations and therefore does not represent a stepping stone dispersal gradient among populations.

Further evidence for population structuring in the NSW population, complementing analyses by the population-based method (θ estimates), is found in the number of clusters produced by STRUCTURE. Although STRUCTURE identified fewer clusters \( (K = 8 \text{ with SSR data, } K = 9 \text{ with AFLP data}) \) than the 11 populations analysed, significant structure does exist. The structure is made up of clusters that are admixed from individuals representing different populations within a region, or even individuals from different regions, indicating a common source population/s for the southern NSW populations, or gene flow/migration among populations. A common founder source is also suggested by the AMOVA analyses where most (85%) of genetic diversity was found within populations and only 2% among regions.

Linkage equilibrium, equal mating type allele frequencies and moderate levels of genotypic diversity in most populations indicate that A. brassicicola is reproducing sexually, and that mating is not hampered by an absence of a particular mating type allele. This result is in contrast to an earlier AFLP analysis of a more limited sample, where the results suggested that populations were all in linkage disequilibrium (11), although the authors
deemed the populations to be randomly mating due to the high genotypic diversities observed. Only two populations from Durras (DUI and DUL) showed significant linkage disequilibrium among SSR loci. In the case of DUI, the small sample size would have reduced the power of the analyses. Both mating types were detected in all populations and most populations also had equal proportions of mating types alleles, indicating frequency-dependent selection (37), and providing further support for the occurrence of sexual reproduction in A. brassicicola.

For populations with unequal proportions of mating types (CT18, DUD, DUI and MBS), other than small sample sizes (as in DUI), this may be due to less frequent sexual reproduction, or the possibility that selection may have favored one mating type over the other by chance. Note that either MAT1-1 or MAT1-2 isolates could be in excess, indicating the lack of a selection bias favouring a particular mating type. In either case, it is possible that populations are reproducing sexually but have not yet reached equilibrium among alleles at loci. For example, if the frequency of recombination is low (e.g. recombination rate $r = 0.05$), it would take more than 50 generations for linkage disequilibrium to disappear (27).

**Population expansion.** The levels of genetic diversity observed in A. brassicicola are much higher than would be generally expected for an organism that has been recently founded, as genetic diversity in such situations is usually purged through bottlenecks and genetic drift. Founder pathogen populations are also subject to genetic drift as a result of e.g. low host availability or unfavourable climatic conditions that limits their evolutionary potential. However, many fungal pathogens that are founders in new geographic areas exhibit relatively high levels of genetic diversity compared to their centres of origin e.g. R. secalis (33) and M. graminicola (34). A similar phenomenon was observed for several aquatic invasive species (53), as well as for invasive plants (65) and animals (12). In these organisms, genetic diversity of invasive species has been shown to be equal or higher than
that of native populations; in the case of invasive animals, approximately 80% of genetic 
diversity present in native populations is maintained in introduced populations (12). This 
phenomenon is referred to as ‘the genetic paradox’, i.e. how do newly founded populations 
overcome low genetic diversity and expected low evolutionary potential to become 
established outside their natural range?

Three scenarios might explain the high levels of gene and genotypic diversity observed in 
*A. brassicicola*: a) multiple founder populations resulting in the admixture of populations, or 
b) population genetic expansion where rare alleles have accumulated in populations due to 
mutations etc., or 3) *A. brassicicola* has a worldwide distribution and has been present in 
Australia prior to the introduction of *C. maritima*. All three scenarios are plausible as 
evidenced by high levels of population admixture identified in *STRUCTURE*. Also, many 
haplotypes were shared among populations within as well as among regions, indicating 
multiple introductions of the same haplotype. This is possible for fungal pathogens with a 
pronounced asexual reproductive phase. Furthermore, our analysis showed that all individual 
populations were admixed and consisted of haplotypes from multiple populations within a 
region, or even from other regions (especially SSRs), indicating multiple founder events.

Evidence for recent population expansion was detected using *BOTTLENECK* (48), where all 
regional populations showed a significant deficit in gene diversity under the SSM, indicating 
that the number of alleles increase faster than the gene diversity (17, 35). Because 
*BOTTLENECK* detects only recent bottlenecks or expansions occurring within the last ± 100 
years, i.e. 0.2–4.0 *N*\(_e\) generations ago (17, 35), the population expansions observed here are 
likely to be associated with migration events from neighbouring populations. Whether *A. 
brassicicola* has a natural occurrence in Australia is unknown but is suspected given its 
ubiquitous presence; however, it is worth noting that in the coastal zone where *Cakile* occurs, 
there is a paucity of other potential hosts. Moreover, it is also probable that the pathogen has
been re-introduced to Australia with the importation of brassicaceous crops and other exotic hosts such as *C. maritima*. Naturally, *A. brassicicola* disperses either as airborne spores, or with seed in ocean currents (as can be the case for *C. maritima*), therefore population expansion due to admixture by dispersal is likely.

Predicting the ecological and evolutionary dynamics of novel plant-pathogen interactions (or existing interactions occurring in new environments as a result of introductions) is of considerable practical interest (45). For example, genetic population expansion of introduced pathogens could pose a significant threat to native plant communities or agricultural crops as a result of an increase in the pathogen’s evolutionary potential. A host shift or host range expansion has to be considered as possible for introduced pathogens that are expanding genetically. *Alternaria* produces a suite of toxins (31) that once introduced into Australia, has the ability to introgress into other *Alternaria* populations through gene flow and sexual recombination, thereby increasing the evolutionary potential and pathogenicity of the pathogen. While the majority of native brassicaceous species in Australia are found in the more arid inland zones (28), there are also a range of introduced weedy species, as well as economically significant agricultural and vegetable crops that could potentially act as alternative hosts for *A. brassicicola* (e.g. *Brassica napus* [canola] and *B. oleracea* [cabbages]) in Australia (56).

In contrast to previous studies on *A. brassicicola* using AFLPs, microsatellites were able to uncover recent population demographic events that were not detected with the previous markers. Populations were identified that are likely to have experienced significant genetic expansion or admixture as a result of multiple introductions in the last ±100 years. This conclusion is supported by high observed levels of gene and genotypic diversity, as well as the STRUCTURE analyses. SSR markers also indicated that most populations were in linkage equilibrium and therefore randomly mating. This study highlights the ability of an invasive
fungal plant pathogen to invade, establish and then genetically expand to represent a pathogen population with a high evolutionary potential.

ACKNOWLEDGEMENTS

We thank Mark Kinnear, Luke Barrett and Caritta Eliasson for technical assistance. Jeremy Burdon and Clive Bock contributed to the collection of many of the *Alternaria* isolates used in this study.

REFERENCES


TABLE 1. Sample size (N), number of unique haplotypes, genotypic diversity, clonal fraction, mean number of alleles and gene diversity, as determined with eleven microsatellite loci on eleven populations of *Alternaria brassicicola* representing three regions in New South Wales, Australia; $\hat{G}/N\% = \text{genotypic diversity (60)}, \text{Na} = \text{mean number of alleles across eleven loci}, H = \text{Nei's gene diversity (43)}.$

<table>
<thead>
<tr>
<th>Region</th>
<th>Population</th>
<th>N</th>
<th>Number of haplotypes</th>
<th>$\hat{G}/N%$</th>
<th>Clonal fraction</th>
<th>Clone-corrected for identical genotypes on the same leaf</th>
<th>Na</th>
<th>H</th>
</tr>
</thead>
<tbody>
<tr>
<td>Central Tilba</td>
<td>CT17B</td>
<td>27</td>
<td>18</td>
<td>49</td>
<td>0.33</td>
<td>24 18 63</td>
<td>0.25</td>
<td>2.1</td>
</tr>
<tr>
<td></td>
<td>CT18</td>
<td>18</td>
<td>14</td>
<td>56</td>
<td>0.22</td>
<td>16 14 68</td>
<td>0.12</td>
<td>2.4</td>
</tr>
<tr>
<td></td>
<td>CT0S</td>
<td>51</td>
<td>24</td>
<td>25</td>
<td>0.53</td>
<td>41 24 29</td>
<td>0.41</td>
<td>2.5</td>
</tr>
<tr>
<td>Durras</td>
<td>DUD</td>
<td>25</td>
<td>20</td>
<td>61</td>
<td>0.20</td>
<td>25 20 61</td>
<td>0.20</td>
<td>2.2</td>
</tr>
<tr>
<td></td>
<td>DUG</td>
<td>59</td>
<td>25</td>
<td>19</td>
<td>0.58</td>
<td>41 22 24</td>
<td>0.39</td>
<td>2.4</td>
</tr>
<tr>
<td></td>
<td>DUI</td>
<td>37</td>
<td>11</td>
<td>17</td>
<td>0.70</td>
<td>34 11 19</td>
<td>0.68</td>
<td>2.0</td>
</tr>
<tr>
<td></td>
<td>DUL</td>
<td>35</td>
<td>26</td>
<td>49</td>
<td>0.26</td>
<td>33 26 60</td>
<td>0.21</td>
<td>2.5</td>
</tr>
<tr>
<td>Moruya-Bodalla</td>
<td>MBH</td>
<td>35</td>
<td>17</td>
<td>27</td>
<td>0.51</td>
<td>34 17 28</td>
<td>0.50</td>
<td>1.8</td>
</tr>
<tr>
<td></td>
<td>MBJ</td>
<td>15</td>
<td>42</td>
<td>0.40</td>
<td>15</td>
<td>49</td>
<td>0.35</td>
<td>2.1</td>
</tr>
<tr>
<td>--------</td>
<td>------</td>
<td>----</td>
<td>----</td>
<td>------</td>
<td>----</td>
<td>----</td>
<td>------</td>
<td>-----</td>
</tr>
<tr>
<td>MBON</td>
<td>57</td>
<td>24</td>
<td>20</td>
<td>0.58</td>
<td>50</td>
<td>24</td>
<td>25</td>
<td>0.52</td>
</tr>
<tr>
<td>MBS</td>
<td>32</td>
<td>17</td>
<td>21</td>
<td>0.47</td>
<td>25</td>
<td>17</td>
<td>35</td>
<td>0.32</td>
</tr>
<tr>
<td>Total</td>
<td>401</td>
<td>211</td>
<td>15</td>
<td>0.59</td>
<td>346</td>
<td>208</td>
<td>20</td>
<td>0.53</td>
</tr>
</tbody>
</table>
TABLE 2. Mating type frequencies and multilocus association tests for eleven *Alternaria brassicicola* populations representing three regions of the New South Wales south coast, Australia. $\chi^2$, based on a 1:1 ratio mating type with one degree of freedom, * ($P < 0.05$) and (** $P < 0.01$) indicates mating type frequencies that deviate significantly from a 1:1 ratio; $I_A$, index of association as calculated in Multilocus v 3.1 (2); $\overline{r_d}$, multilocus association as calculated in Multilocus v 3.1 (2); *, significant multilocus association at $P < 0.001$. The number of randomizations was 1000 in all multilocus association tests.

<table>
<thead>
<tr>
<th>Population</th>
<th>MAT1-1</th>
<th>MAT1-2</th>
<th>Total</th>
<th>$\chi^2$</th>
<th>$I_A$</th>
<th>$\overline{r_d}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>CT0S</td>
<td>16</td>
<td>8</td>
<td>24</td>
<td>2.667</td>
<td>0.088</td>
<td>0.012</td>
</tr>
<tr>
<td>C17B</td>
<td>11</td>
<td>7</td>
<td>18</td>
<td>0.889</td>
<td>0.004</td>
<td>0.001</td>
</tr>
<tr>
<td>CT18</td>
<td>11</td>
<td>3</td>
<td>14</td>
<td>4.571*</td>
<td>-0.037</td>
<td>-0.005</td>
</tr>
<tr>
<td>DUD</td>
<td>4</td>
<td>16</td>
<td>20</td>
<td>7.200**</td>
<td>0.166</td>
<td>0.021</td>
</tr>
<tr>
<td>DUG</td>
<td>12</td>
<td>13</td>
<td>25</td>
<td>0.040</td>
<td>0.078</td>
<td>0.013</td>
</tr>
<tr>
<td>DUI</td>
<td>9</td>
<td>2</td>
<td>11</td>
<td>4.455*</td>
<td>0.559*</td>
<td>0.094*</td>
</tr>
<tr>
<td>DUL</td>
<td>17</td>
<td>9</td>
<td>26</td>
<td>2.462</td>
<td>0.191*</td>
<td>0.024*</td>
</tr>
<tr>
<td>MBH</td>
<td>9</td>
<td>8</td>
<td>17</td>
<td>0.059</td>
<td>-0.131</td>
<td>-0.023</td>
</tr>
<tr>
<td>MBON</td>
<td>10</td>
<td>14</td>
<td>24</td>
<td>0.667</td>
<td>-0.006</td>
<td>-0.001</td>
</tr>
<tr>
<td>MBJ</td>
<td>7</td>
<td>7</td>
<td>14</td>
<td>0.000</td>
<td>0.214</td>
<td>0.036</td>
</tr>
<tr>
<td>MBS</td>
<td>2</td>
<td>15</td>
<td>17</td>
<td>9.941**</td>
<td>-0.028</td>
<td>-0.028</td>
</tr>
</tbody>
</table>

**Regions:**

<table>
<thead>
<tr>
<th>Region</th>
<th>MAT1-1</th>
<th>MAT1-2</th>
<th>Total</th>
<th>$\chi^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Central Tilba</td>
<td>38</td>
<td>18</td>
<td>56</td>
<td>7.143**</td>
</tr>
<tr>
<td>Durras</td>
<td>42</td>
<td>40</td>
<td>82</td>
<td>0.049</td>
</tr>
<tr>
<td>Moruya-Bodalla</td>
<td>28</td>
<td>44</td>
<td>72</td>
<td>3.556</td>
</tr>
<tr>
<td>---------------</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>-------</td>
</tr>
<tr>
<td>Total</td>
<td>108</td>
<td>102</td>
<td>210</td>
<td>0.171</td>
</tr>
</tbody>
</table>
TABLE 3. Pairwise comparisons of population differentiation (below the diagonal) among clone corrected populations of *Alternaria brassicicola*.

<table>
<thead>
<tr>
<th></th>
<th>CT17B</th>
<th>CT18</th>
<th>CT0S</th>
<th>DUD</th>
<th>DUG</th>
<th>DUI</th>
<th>DUL</th>
<th>MBH</th>
<th>MBJ</th>
<th>MBON</th>
</tr>
</thead>
<tbody>
<tr>
<td>CT18</td>
<td>0.19*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CT0S</td>
<td>0.20*</td>
<td>0.15*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DUD</td>
<td>0.04</td>
<td>0.19*</td>
<td>0.19*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DUG</td>
<td>0.12*</td>
<td>0.15*</td>
<td>0.15*</td>
<td>0.10*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DUI</td>
<td>0.03</td>
<td>0.29*</td>
<td>0.22*</td>
<td>0.05*</td>
<td>0.14*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DUL</td>
<td>0.16*</td>
<td>0.29*</td>
<td>0.30*</td>
<td>0.13*</td>
<td>0.15*</td>
<td>0.08</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MBH</td>
<td>0.17*</td>
<td>0.30*</td>
<td>0.18*</td>
<td>0.17*</td>
<td>0.09*</td>
<td>0.10*</td>
<td>0.14*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MBJ</td>
<td>0.07*</td>
<td>0.28*</td>
<td>0.21*</td>
<td>0.07</td>
<td>0.11*</td>
<td>0.04</td>
<td>0.07*</td>
<td>0.02*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MBON</td>
<td>0.18*</td>
<td>0.17*</td>
<td>0.14*</td>
<td>0.26*</td>
<td>0.18*</td>
<td>0.26*</td>
<td>0.34*</td>
<td>0.27*</td>
<td>0.28*</td>
<td></td>
</tr>
<tr>
<td>MBS</td>
<td>0.17*</td>
<td>0.19*</td>
<td>0.20*</td>
<td>0.13*</td>
<td>0.04</td>
<td>0.18*</td>
<td>0.15*</td>
<td>0.06</td>
<td>0.09*</td>
<td>0.23*</td>
</tr>
</tbody>
</table>

*Indicates theta values significant at $P < 0.001$. 
TABLE 4. Comparison of observed gene diversity ($H_E$) with expected gene diversity at mutation-drift equilibrium ($H_{EQ}$) calculated from the observed number of alleles under the stepwise mutation model (SMM) and two phase model (TPM) (17). D/E is the number of loci showing a deficit/excess of gene diversity. Significance estimates of gene diversity excess ($H_E > H_{EQ}$) or deficiency, on average across loci were obtained using the one-tailed Wilcoxon’s test and the sign test (17, 48); ** = $P \leq 0.01$.

<table>
<thead>
<tr>
<th>Group</th>
<th>Mutation model ($P = 0.0000$)</th>
<th>D/E</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Central Tilba</td>
<td>SMM</td>
<td>8/3**</td>
<td>7/4</td>
</tr>
<tr>
<td>Durras</td>
<td>SMM</td>
<td>10/1**</td>
<td>10/1**</td>
</tr>
<tr>
<td>Moruya-Bodalla</td>
<td>SMM</td>
<td>10/1**</td>
<td>9/2**</td>
</tr>
<tr>
<td>Total</td>
<td>SMM</td>
<td>10/1**</td>
<td>10/1**</td>
</tr>
</tbody>
</table>
FIG. 1. Comparison of ln (K) and delta K values for (A) the SSR dataset and (B) the AFLP data set calculated from the STRUCTURE v2.2 output where the hypothesized number of populations ranged from 1 to 20.
FIG. 2. Cluster analyses of *Alternaria brassicicola* populations from the NSW coast of Australia. Results from STRUCTURE v2.2. Each individual is represented as a bar, divided into K colours where K is the number of clusters assumed. Individuals are sorted according to Q, the inferred clusters. (A) Microsatellite data, K = 8, (B) AFLP data, K = 9.
FIG. 3. Cluster analyses of Alternaria brassicicola populations from the NSW coast of Australia. Results of microsatellite data from STRUCTURE v2.2. Each individual is represented as a bar, divided into K colours where K is the number of clusters assumed i.e. K = 8. Individuals are sorted according to geographic region and original population 1-12. Pop1 = CT05, pop2 = CT17B, pop3 = CT18, pop4 = DUD, pop5 = DUG, pop6 = DUI, pop7 = DUL, pop8 = MBC, pop9 = MBH, pop10 = MBJ, pop11 = MBON, pop12 = MBS.