Invasion genetics of New World medflies: testing alternative colonization scenarios

Andrew J. Bohonak , Neil Davies , Francis X. Villablanca & George K. Roderick

Abstract

The Mediterranean fruit fly (*Ceratitis capitata*) is an invasive agricultural pest with a wide host range and a nearly global distribution. Efforts to forgo the medfly's spread into the United States are dependent on an understanding of population dynamics in newly established populations elsewhere. To explore the potential influence of demographic and historical parameters in six medfly populations distributed from Mexico to Peru, we created population genetic null models using Monte Carlo simulations. Null expectations for genetic differentiation (F_{ST}) were compared with actual sequence variation from four highly polymorphic nuclear loci. Four colonization scenarios that were modeled led to unique genetic signatures that could be used to interpret empirical data. Unless current gene flow across Latin America was assumed to be very high, we could reject colonizations consisting of multiple introductions, each of low genetic diversity. Further, if simulated populations were small ($N_e = 5 \times 10^2$ individuals per population), small invasions from a single source consistently produced F_{ST} values comparable to those currently observed in Latin America. In contrast, only large invasions from diverse sources were compatible with the observed data for large populations ($N_e \geq 5 \times 10^3$). This study demonstrates that alternative population genetic hypotheses can be tested empirically even when departures from equilibrium are extreme, and that population genetic theory can be used to explore the processes that underlie biological invasions.

Abbreviations: F_{ST} – Wright's F_{ST}, the most common metric for quantifying genetic population differentiation; N_e – effective population size; $N_e m$ – gene flow, represented as the number of successful migrants leaving (and entering) each population per generation

Introduction

Invasions of exotic pests and pathogens constitute a growing threat to agriculture, human health and the natural environment (Metcalf 1995b; Schrag and Wiener 1995). For growers and agricultural managers in the United States, potential threats posed by the invasive Mediterranean fruit fly (*Ceratitis capitata*) have been at the forefront of interdiction and eradication efforts over the past three decades. In California alone, it is estimated that established populations of the Mediterranean fruit fly, or medfly, would result in losses of over one billion dollars annually in lost crop value and jobs (Siebert and Cooper 1995). An understanding of basic

population level processes is necessary for successful interception and eradication of the medfly and other exotic pests. These processes include the source and demography of the invasion, current population structure and the minimum number of individuals required for a successful invasion (e.g. Carey 1991, 1996).

As with other invasive species, population genetic studies can be used to study medfly invasion biology. For example, some of the controversy surrounding the colonization history of *C. capitata* in the United States (Carey 1991, 1992; Saul 1992) can be resolved using genetic markers (Gasparich et al. 1997; Villablanca et al. 1998; Davies et al. 1999; He and Haymer 1999). Similarly, information on demography and dispersal might be inferred from genetic data in populations that are already established. However, simple interpretations of medfly population genetic structure are complicated by severe departures from a drift-gene flow equilibrium. In the last 100 years, the range of this species has been extended globally following repeated invasions, which has impacted the distribution of genetic variation both within and among populations (Malacrida et al. 1992; Gasparich et al. 1997; Gomulski et al. 1998). This suggests that historical and contemporary processes will be difficult to separate using a static picture of genetic composition. Genetic similarities among distant populations might reflect a recent, widespread invasion, current high levels of dispersal or a combination of these factors. After populations become established, the initial genetic signature of range expansions and founder events will be eroded, and genetic population structure will eventually reflect levels of mutation, gene flow and random drift. However, it is often difficult to determine where any species lies on this trajectory between initial conditions and an eventual equilibrium state. Although a number of direct and indirect tests have been proposed to separate these factors (Slatkin 1993; Templeton 1998; Bohonak and Roderick 2001), no single method is applicable to all taxa or all genetic markers.

Parametric bootstrapping provides a way to statistically contrast alternative invasion pathways and mechanisms. Although primarily restricted to applications in systematics (Huelsenbeck et al. 1996), this approach has also been used to test hypotheses, generate confidence intervals and assess sampling bias in ecology (Manly 1997) and population genetics (Tufto et al. 1998). We used parametric bootstrapping to test alternative colonization scenarios for populations of *C. capitata* that have been established in Latin America

for approximately 100 years. Specifically, we used Monte Carlo simulations to generate null models for genetic differentiation during the first 1000 generations following one of four colonization scenarios. These models allowed us to (1) understand how genetic diversity changes in invading medfly populations over time, (2) determine if alternative invasion histories yield statistically distinct patterns of genetic structure and (3) test for concordance between empirical data and each invasion scenario. This approach is flexible enough to permit the future investigation of additional scenarios in light of new ecological or genetic data. Robustness of the conclusions can be assessed by varying simulation parameters within a biologically meaningful context, because the history and biology of *C. capitata* is well known due to its economic importance (e.g. Metcalf 1995a; Roderick 1996b). By extension, our results demonstrate the utility of Monte Carlo simulation studies for investigating demography, gene flow, colonization history and population bottlenecks in other species that are invasive, endangered or comprised of short-lived metapopulations.

Methods

Latin American simulation parameters

We modeled the evolution of gene genealogies using the program ESP (Bohonak and Roderick 2000). ESP utilizes Monte Carlo simulations to follow the evolution of unique alleles (e.g. non-recombining sequences of DNA) across multiple populations, incorporating random genetic drift, mutation and gene flow. Migration, mutation and reproduction are represented as simple Poisson processes, and mating conforms to the standard Wright–Fisher model, with each individual contributing a large number of gametes to the gamete pool (Hartl and Clark 1989). Mutation occurs according to an infinite allele model (Kimura and Crow 1964). Populations can be 'colonized' with user-defined gene frequencies or by sampling from an ancestral gene pool (algorithms described in Bohonak and Roderick 2000). Population differentiation, heterozygosity and allelic diversity are summarized across replicates. The primary advantage of this approach for analysis of relatively young medfly populations is that it facilitates hypothesis testing without the assumption of equilibrium.

To simulate invading populations of *C. capitata*, we first founded 100 Latin American populations using one of the colonization scenarios described below. Following this 'invasion', gene flow within Latin America was assumed to follow an island model (Wright 1931). Whether the island model provides an acceptable approximation to spatially explicit migration patterns will depend on many factors, including the time since colonization and the rate of drift. Under most of the conditions relevant for Latin American medflies (very young and possibly very large populations), founder effects will dominate patterns of differentiation more than ongoing gene flow. As a result, we expected few qualitative differences between island and spatially explicit models of migration.

To follow evolution at four independent genes (loci), mutation rates were set at $\mu = 1.00 \times 10^{-8}/$bp/generation, which approximates mitochondrial and nuclear synonymous substitution rates (Li 1997). Sequence length for each locus was 600 bp. Population differentiation was calculated at 50 generation intervals for 1000 generations, based on a random sample of six populations, with four haplotypes (i.e. sequences) per population for each locus (consistent with the sample sizes of Davies et al. 1999; see Table 1).

Table 1. Allelic diversity of medfly populations at four intron-containing loci.

Population	Actin	Chorion	SOD	Vitellogenin
Latin America				
Brazil	2 (2)	1 (2)	2 (2)	2 (2)
Costa Rica	5 (7)	4 (6)	3 (5)	N/A
Ecuador	3 (5)	2 (3)	3 (3)	4 (4)
Guatemala	3 (8)	2 (8)	2 (3)	5 (7)
Mexico	3 (3)	3 (4)	3 (4)	2 (5)
Peru	3 (6)	2 (2)	3 (4)	2 (4)
Africa				
Malawi	5 (5)	3 (4)	8 (8)	5 (5)
Kenya	9 (12)	7 (9)	7 (8)	8 (9)
Regional				
Latin America (6 populations)	8 (31)	6 (25)	13 (21)	13 (22)
Africa (2 populations)	12 (17)	9 (13)	15 (16)	12 (14)
Other (3 populations)	7 (20)	8 (19)	14 (23)	7 (16)
Global				
(11 populations)	19 (68)	18 (57)	38 (60)	32 (52)

The first number for each entry represents the number of unique alleles, and the total number of haplotypes sequenced follows in parentheses (Villablanca et al. 1998; Davies et al. 1999).

Davies et al. (1999) sequenced only one gene from most medfly individuals at each locus; consequently, our models focused at the level of the haplotype, rather than the individual. Recombination among alleles (defined here as unique haplotypes) was not considered.

For each sample, population differentiation was estimated as Wright's F_{ST} with the statistic θ (Weir and Cockerham 1984). Results using an alternative statistic, Φ_{ST} (Excoffier et al. 1992) were qualitatively similar due to low differentiation among potential source populations (see Davies et al. 1999) and are not reported here.

Sensitivity analysis

To examine the sensitivity of our conclusions to assumptions about effective population size (N_e) or gene flow ($N_e m$, the number of successfully migrating individuals per population per generation), we varied these two parameters across areas of parameter space that seemed biologically plausible. (The use of $N_e m$ to characterize gene flow, rather than the per capita rate m, follows standard population genetic conventions.) In different sets of simulations, populations consisted of $N_e = 2.5 \times 10^2, 5 \times 10^2, 5 \times 10^3$ or 1×10^4 individuals, and migration rates were $N_e m = 5 \times 10^{-5}$, 0.05, 0.5 or 5 individuals per generation. All 16 combinations of these two parameters were simulated 200 times for each of the four colonization scenarios (see below). In total, 12,800 random simulations were conducted. To simplify interpretation, we present here only the 95% confidence intervals for effective population sizes of $2.5 \times 10^2, 5 \times 10^3$ and 1×10^4 individuals, and migration rates of $N_e m = 5 \times 10^{-5}$, 0.5 and 5 individuals per generation.

Single invader colonization scenarios

The demographic and temporal complexities of bioinvasions can be categorized in a variety of ways (e.g. Lovette et al. 1999). In a genetic context, the most relevant factors are the number of colonists per population, diversity within source populations and genetic divergence among source populations. To explore the impact of these factors, we simulated colonizations of Latin America using four alternative scenarios (Table 2). The single source/single invader scenario (SS/SI) represented the rapid proliferation of an invading allele into 100 populations (i.e. all haplotypes were identical at generation 0). This scenario

Scenario		Unique ancestral alleles	Initial θ	Ancestral conditions
	SS/SI: Single source/ Single invader	1	Undefined (0)	One allele
	MS/SI: Multiple sources/ Single invaders	100	1	One unique allele per population, each two mutational steps from all others
	SS/MI: Single source/ Multiple invaders	24–39	-1×10^{-4} to 2×10^{4}	All haplotypes chosen randomly from a single ancestral population, with no bottleneck
	MS/MI: Multiple sources/ Multiple invaders	134–191	0.09–0.10	Each population descends from one randomly chosen ancestral population, with no bottleneck

Unique ancestral alleles is the number of unique alleles present in colonizing individuals over all 100 populations. Arrows in the diagrams represent invasions and open circles denote new populations. Filled circles represent computer-generated, diverse ancestral populations.

represents the invasion of one (a gravid female) or a few individuals from a genetically depauperate source. The multiple source/single invader scenario (MS/SI) began with each population possessing its own unique genotype. MS/SI represents multiple introductions, each of a single, homozygous gravid female or a small group of homozygous individuals. Initial values of F_{ST} were undefined (near 0) and 1 for the SS/SI and MS/SI scenarios, respectively.

Multiple invader colonization scenarios

Conceptually, SS/SI and MS/SI represent extremes along an axis that describes the genetic diversity among source populations. However, diversity within source populations is very low for these two scenarios. We used ESP to create realistically diverse source populations for additional scenarios, using genetic studies of African medflies as a guide. Sub-Saharan Africa is believed to be the ancestral source of all present-day *C. capitata* (White and Elson-Harris 1992), and populations in this region show higher levels of mitochondrial diversity than other areas of the world (e.g. Gasparich et al. 1997). Previous studies of nuclear loci in two populations in Africa have also shown high diversity within and low subdivision among populations (Gasperi et al. 1991; Baruffi et al. 1995; Villablanca et al. 1998). Coalescent analysis (Kuhner et al. 1995) of African populations studied by Davies et al. (1999) suggests an effective population size of

3×10^5 to 3×10^7 medflies in Africa across a range of likely mutation rates.

We randomly generated ancestral source populations in ESP using the following parameters: for each set of ancestral populations, 200 populations of $N_e = 5 \times 10^3$ haplotypes were permitted to evolve for 5×10^5 generations, beginning with a single allele of 600 bp ($\mu = 1.00 \times 10^{-8}$/bp/generation, $N_e m = 2.5$). This simulation was replicated six times, resulting in a total of 1200 ancestral populations. By the end of each replicate, alleles up to nine mutational steps from the original sequence had evolved, and the mean number of alleles per population had increased to 31 (SE = 7). This high diversity and low genetic subdivision ($\theta \approx 0.09$) were consistent with theoretical expectations and with the empirical data from Africa (Table 1, Davies et al. 1999). Final allele frequencies and genealogies were saved for all ancestral populations.

To contrast with the 'single invader' colonizations described above, these genetically diverse ancestral populations were used for 'multiple invader' scenarios that occurred without bottlenecks (Table 2). In the single source/multiple invader scenario (SS/MI), one of the 1200 ancestral populations was randomly chosen as the source for every Latin American haplotype. In contrast, the multiple source/multiple invader scenario (MS/MI) began by randomly selecting one of the six ancestral simulations, and then founding each of the 100 Latin American populations from one of the 200 ancestral populations in that simulation (i.e. within each Latin American population,

all haplotypes descended from only one African ancestral population; see the propagule pool model of Slatkin 1977). The MS/MI scenario represents colonizations from 100 genetically diverse sources at approximately the same time. Unlike MS/SI, MS/MI assumes that no appreciable bottlenecks were associated with the invasion and establishment of *C. capitata* populations.

Analysis of results

Simulation results were compared with DNA sequence data from six populations in Central and South America (Table 1; Villablanca et al. 1998; Davies et al. 1999). For these populations, between 21 and 31 haplotypes have been cloned and sequenced from muscle specific actin intron 1, chorion s36 intron 1, vitellogenin 1 gamma intron 2 and Cu/Zn super oxide dismutase intron 1. Differentiation across these populations is low ($\theta = 0.07$–0.10). In the New World, from 6 to 13 distinct alleles have been found per locus (18–38 globally). These numbers may be conservative estimates of allelic diversity because 'singletons' (base pair positions defined by substitutions seen only once among all haplotypes sequenced) were pooled with more common haplotypes to avoid Taq polymerase artifacts (see Villablanca et al. 1998).

Our primary goal was to reject (in a statistical sense) areas of parameter space that produce levels of genetic differentiation inconsistent with those observed in Latin America over biologically relevant time scales. For each parameter level/scenario combination, we looked for overlap between the theoretical and empirical 95% confidence intervals for F_{ST}. (When generating theoretical confidence intervals, ESP considers F_{ST} as zero when all individuals in the sample are fixed for the same allele.) An empirical CI of (0.068, 0.098) was generated from 10,000 bootstraps (over loci, Weir 1990) of Davies et al.'s (1999) data. The use of Monte Carlo simulations for these types of statistical tests incorporates sampling variation, stochastic variation and variation in the initial state of the colonized populations.

Using a relatively conservative criterion, we failed to reject a scenario-parameter combination if its CI overlapped (0.068, 0.098) at any point between generations 200 and 1000. (*C. capitata* has at least two generations per year, and as many as nine in the tropics; the oldest populations in Latin America are approximately

100 years old; Carey 1989; Enkerlin et al. 1989; Fimiani 1989; Metcalf 1995a; Liedo and Carey 1996.)

Results

When effective population size was large, mutational pressures were sufficient to quickly increase the number of alleles within each group of interacting populations. Even in the SS/SI scenario, fewer than 50 generations were required to increase one allele to more than 40 (when $N_e = 5 \times 10^3$) or 100 ($N_e = 1 \times 10^4$). In contrast, the number of alleles tended to increase only slightly or decrease when $N_e = 2.5 \times 10^2$. For small populations, low allelic diversity led to high variance in population differentiation for the SS/SI scenarios (e.g. compare confidence interval sizes across columns in Figure 1). However, confidence intervals for F_{ST} did not differ dramatically with effective population size in the three scenarios that began with multiple alleles. This suggests that sampling variation and variation in initial state are more important than stochastic variation among replicates.

Equilibrium values for F_{ST} and the rate of approach to equilibrium varied predictably across parameter space, with low migration leading to higher population divergence and slower approaches to equilibrium (compare rows in Figure 1). Within 1000 generations, population differentiation stabilized in the medium and high migration simulations for $N_e = 2.5 \times 10^2$ but none of the simulations for $N_e \geq 5 \times 10^3$.

In eight of nine parameter combinations, MS/SI could be easily distinguished from the other scenarios throughout the first 500 generations after colonization. In six of these cases, the MS/SI confidence interval did not cross the other three scenarios within the first 1000 generations. Confidence intervals for MS/MI and SS/MI overlapped broadly in every case, although the size of this overlap decreased with increasing N_e. For $N_e \geq 5 \times 10^3$, the confidence interval for F_{ST} in the SS/SI scenario rarely increased above zero. In general, the genetic signature of founder events persisted in all scenarios unless migration was high and effective population size was very low.

Because each scenario led to distinct changes in population differentiation over time, conclusions regarding Latin American populations of *C. capitata* were possible. To a large extent, these conclusions depended on effective population size (Table 3). For the smallest populations, observed values of F_{ST} could not be
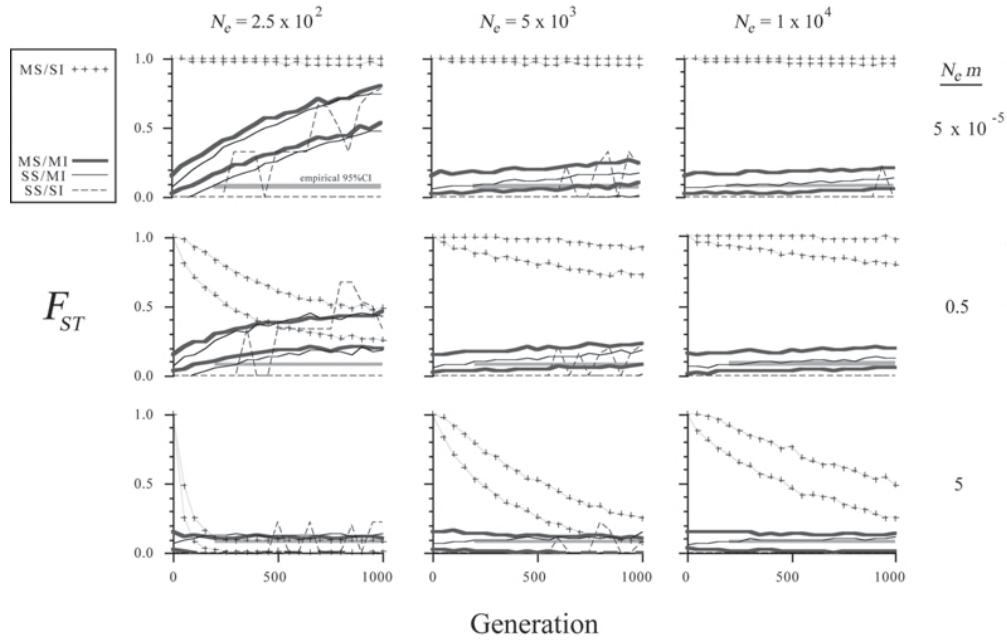
Figure 1. 95% confidence intervals for F_{ST} following one of four colonization scenarios (Table 2). Each plot shows the upper and lower bounds of the four confidence intervals (200 Monte Carlo simulations each) for a particular value of N_e and $N_e m$. Thin gray boxes border a 95% CI calculated from empirical data (vertical axis) and potential ages for Latin American medfly populations (horizontal axis; see text for details).

Table 3. Simulation results.

Population size (N_e)	Gene flow $N_e m$	SS/SI	SS/MI	MS/MI	MS/SI
2.5×10^2	5×10^{-5}			*	*
	0.5			*	*
	5				
5×10^3	5×10^{-5}				*
	0.5				*
	5				
1×10^4	5×10^{-5}				*
	0.5	*			*
	5	*			*

Stars indicate scenario-parameter combinations which can be rejected (see Figure 1).

rejected for either of the single source invasions (SS/SI and SS/MI). Due to drift, population differentiation in the MS/SI and MS/MI simulations quickly increased to levels beyond those observed in Latin America, except when gene flow was very high. Additional simulations (not presented) showed similar results for $N_e = 5 \times 10^2$. However, as gene flow $N_e m$ increased from 0.5 to 5, genetic panmixia led to an inability to distinguish among scenarios, and none of them could be rejected statistically. It is possible that an

increase in sample size and/or the number of loci would decrease the size of the theoretical confidence intervals and increase the statistical power of these tests.

As population size increased to $N_e \geq 5 \times 10^3$, the congruence between SS/SI scenarios and the empirical data began to disappear (Figure 1, Table 3). For these larger populations, only multiple invader colonizations consistently yielded levels of population differentiation comparable to those observed empirically. The approach to equilibrium in the two single invader scenarios slowed enough that theoretical predictions for F_{ST} generally remained either above or below empirical values. In contrast, differentiation for MS/MI (and, to a large extent, SS/MI) was concordant with the empirical values at the population's origins and changed little in 1000 generations.

In summary, multiple invader scenarios (MS/MI and SS/MI) were compatible with the observed data over a wide range of parameter space. However, if medfly introductions can be assumed to occur via extreme bottlenecks (one or a few genotypes per population), a scenario involving multiple introductions (MS/SI) is far less likely than rapid geographic radiation from a single source of low diversity (SS/SI).

Discussion

Despite the potential complexities of interpreting genetic structure in recent populations, we found that different invasion scenarios often produce distinct signals in genetic population structure during the first 1000 generations after colonization. When F_{ST} was used to compare theoretical expectations from parametric bootstrapping with empirical data, the qualitative conclusions were relatively straightforward. The MS/SI model can be rejected unless current gene flow is extremely high across Latin America. If effective population size for Latin American populations of *C. capitata* are small, low-diversity invasions best approximate the empirical data. If population sizes are large, colonizations from moderately diverse sources (such as sub-Saharan Africa) produce levels of differentiation comparable to those observed by Davies et al. (1999).

Some invasive species (such as the Mediterranean fruit fly) have been hypothesized to constitute 'meta-invasions' that are temporally and geographically complex (Davies et al. 1999). Distinguishing simple patterns of colonization from metainvasions may be difficult using the markers and analytic techniques that are currently available, although our analyses here are consistent with a demographic and genetic linkage among invading populations. Using parametric bootstrapping, we have begun to analyze the genetic signatures from colonizations more complicated than the four analyzed here. For example, one might expect two SS/MI invasions 100 generations apart (each contributing to 50% of the populations) to produce levels of differentiation intermediate between SS/MI and MS/MI. However, F_{ST} increases dramatically following the second invasion, due to the temporal lag in establishment. The high estimates of differentiation observed in this and similar metainvasions might resemble simpler scenarios involving widespread introductions with extreme bottlenecks (MS/SI), but there would be no easy way of distinguishing them from the genetic data alone. Inferences from standard phylogeographic analyses might fail, because potential source populations of metainvasive species are often too young for informative mutations to have reached fixation (Gasparich et al. 1997). Consequently, we stress that hypothesis testing via parametric bootstrapping cannot provide definitive proof that a particular invasion scenario has occurred. Rather, this approach provides a flexible framework for (1) *rejecting* specific scenarios, (2) suggesting empirical parameters that require further study (e.g. population size, the extent to which colonization and dispersal are distinct processes) and (3) quantifying the statistical benefits of altered sampling strategies (e.g. increasing sample sizes, number of loci).

Unlike some alternative analyses, our conclusions do not require assumptions such as equilibrium, an infinite number of populations, or the absence of mutation (Roderick 1996a; Bohonak 1999). The flexibility of these simulations permits the plausibility of additional colonization scenarios to be determined as new genetic and ecological evidence is uncovered. More complicated algorithms for migration or mating structure could also be introduced. Assessment of structural patterns within the haplotype networks for each gene can also provide additional insights. For example, if a SS/SI invasion took place, one would predict that the most common New World allele should be present in the ancestral population, and that all other New World alleles should be absent elsewhere. Based on this criterion, a simple SS/SI scenario can be rejected, because Latin American alleles occupy more than one internal node in each of the four haplotype networks (Davies et al. 1999). It remains to be seen whether empirically observed patterns of population differentiation *and* genealogy can be explained by single source colonizations that lie along the continuum between extreme bottleneck (SI invasions) and no bottleneck (MI invasions). In light of the potential complexity of the invasion process and the extreme departures from equilibrium that are involved, the relative advantages of other analytic approaches should also be explored (e.g. Rogers 1995; Bertorelle and Rannala 1998; Templeton 1998).

Detailed knowledge of medfly invasion biology should also guide the types of genetic analyses conducted. For example, it has been hypothesized that one or more *C. capitata* populations may already be established in California (Carey 1991, 1992; but see Voss 1992). If this is true, genetic approaches focused on these established populations may be as important for management as approaches focused on invading individuals. For example, the genotypes of flies caught in successive years could provide the basis for using assignment tests (Rannala and Mountain 1997) to predict the likelihood of a resident population. Alternatively, one could begin with the assumption that one or more hypothetical resident populations does exist, and identify its probable demographic composition based

on theoretical expectations. Because studies of invasive species are used in both predictive and retrospective frameworks, a synthesis of ecological and population genetic data is imperative.

Our results from Latin America medflies emphasize the importance of ecological studies for population genetics. Many simple or complicated colonization scenarios might be rejected *a priori* based on non-genetic data. For *C. capitata*, the ability to test the likelihood of hypothesized colonization scenarios also depends critically on estimating effective population size (Table 3). These estimates could be obtained in several ways. Despite the difficulties of translating census population size (N) to effective population size (N_e) (Gilpin 1991; Nunney and Elam 1994; Frankham 1995), demographic and behavioral studies of *C. capitata* could provide an upper bound for N_e. Alternatively, coalescent analysis of a large number of individuals from populations in the medfly's native range may be informative, if geologic and biogeographic evidence suggests that these populations have been stable for long periods of time. (In reality, the long-term stability of these populations is not fully understood. Although Africa is likely to be the ancestral origin of *C. capitata*, the current cultivation of crops in this region suggests that recent range extensions could have occurred.) Any approach that provides an unbiased estimate of N_e and an associated error rate (e.g. Kuhner et al. 1995) may provide enough information to help distinguish among some colonization scenarios.

As with many invading species, studies of the Mediterranean fruit fly are limited for both technical and theoretical reasons (Roderick et al. 1998). A paucity of variation in traditional genetic markers means that techniques such as SSCP, AFLPs and sequencing of EPIC loci must be employed. Further, the alleles which are present do not tend to sort cleanly into phylogeographic groups, and they are not amenable to simple interpretations of F_{ST} in terms of gene flow and drift. Previous studies (Gasparich et al. 1997; Davies et al. 1999) have made progress in determining one of the most critical management questions: where do invading medflies come from? Here we have addressed more detailed questions concerning invasion demographics, with encouraging results. Alternative colonization scenarios can be distinguished by evaluating empirical data using specific theoretical expectations. This represents a first step towards the application of more comprehensive approaches (e.g. Bowcock et al. 1991; Neigel and Avise 1993; Wakeley and Hey 1997) to invading species.

References

Baruffi L, Damiani G, Guglielmino CR, Bandi C, Malacrida AR and Gasperi G (1995) Polymorphism within and between populations of *Ceratitis capitata*: comparison between RAPD and multilocus enzyme electrophoresis data. Heredity 74: 425–437

Bertorelle G and Rannala B (1998) Using rare mutations to estimate population divergence times: a maximum likelihood approach. Proceedings of the National Academy of Sciences of the United States of America 95: 15452–15457

Bohonak AJ (1999) Dispersal, gene flow and population structure. Quarterly Review of Biology 74: 21–45

Bohonak AJ and Roderick GK (2000) ESP: Evolution in simulated populations. http://www.bio.sdsu.edu/pub/andy/ESP.html

Bohonak AJ and Roderick GK (2001) Dispersal of invertebrates among temporary ponds: are genetic estimates accurate? Israel Journal of Zoology (in press)

Bowcock AM, Kidd JR, Mountain JL, Hebert JM, Carotenuto L, Kidd KK and Cavallisforza LL (1991) Drift, admixture, and selection in human evolution: a study with DNA polymorphisms. Proceedings of the National Academy of Sciences of the United States of America 88: 839–843

Carey JR (1989) Demographic analysis of fruit flies. In: Robinson AS and Hooper G (eds) Fruit Flies: Their Biology, Natural Enemies, and Control, Vol 3B, pp 253–265. Elsevier, New York

Carey JR (1991) Establishment of the Mediterranean fruit fly in California. Science 253: 1369–1373

Carey JR (1992) The medfly in California (Response). Science 255: 515–516

Carey JR (1996) The incipient Mediterranean fruit fly population in California: implications for invasion biology? Ecology 77: 1690–1697

Davies N, Villablanca FX and Roderick GK (1999) Bioinvasions of the medfly, *Ceratitis capitata:* source estimation using DNA sequences at multiple intron loci. Genetics 153: 351–360

Enkerlin D, Garcia L and Lopez F (1989) Mexico, Central and South America. In: Robinson AS and Hooper G (eds) Fruit Flies: Their Biology, Natural Enemies, and Control, Vol 3A, pp 83–90. Elsevier, New York

Excoffier L, Smouse PE and Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. Genetics 131: 479–491

Fimiani P (1989) Mediterranean region. In: Robinson AS and Hooper G (eds) Fruit Flies: Their Biology, Natural Enemies, and Control, Vol 3A, pp 39–50. Elsevier, New York

Frankham R (1995) Effective population size/adult population size ratios in wildlife: a review. Genetical Research 66: 95–107

Gasparich GE, Silva JG, Han HY, McPheron BA, Steck GJ and Sheppard WS (1997) Population genetic structure of Mediterranean fruit fly (Diptera: Tephritidae) and implications for worldwide colonization patterns. Annals of the Entomological Society of America 90: 790–797

Gasperi G, Guglielmino CR, Malacrida AR and Milani R (1991) Genetic variability and gene flow in geographical populations of *Ceratitis capitata* (Wied.) (medfly). Heredity 67: 347–356

Gilpin M (1991) The genetic effective size of a metapopulation. Biological Journal of the Linnean Society 42: 165–176

Gomulski LM, Bourtzis K, Brogna S, Morandi PA, Bonvicini C, Sebastiani F, Torti C, Guglielmino CR, Savakis C, Gasperi G and Malacrida A (1998) Intron size polymorphism of the *Adh*$_1$ gene parallels the worldwide colonization history of the Mediterranean fruit fly *Ceratitis capitata*. Molecular Ecology 7: 1729–1741

Hartl DL and Clark AG (1989) Principles of Population Genetics. Sinauer Associates, Sunderland, Massachusetts

He M and Haymer DS (1999) Genetic relationships of populations and the origins of new infestations of the Mediterranean fruit fly. Molecular Ecology 8: 1247–1257

Huelsenbeck JP, Hillis DM and Jones R (1996) Parametric bootstrapping in molecular phylogenetics: applications and performance. In: Ferraris JD and Palumbi SR (eds) Molecular Zoology: Applications and Performance, pp 19–45. Wiley, New York

Kimura M and Crow JF (1964) The number of alleles that can be maintained in a finite population. Genetics 49: 725–738

Kuhner MK, Yamato J and Felsenstein J (1995) Estimating effective population size and mutation rate from sequence data using Metropolis-Hastings sampling. Genetics 140: 1421–1430

Li W-H (1997) Molecular Evolution. Sinauer Associates, Sunderland, Massachusetts, 487 pp

Liedo P and Carey JR (1996) Demography of fruit flies and implications to action programs. In: McPheron BA and Steck GJ (eds) Fruit Fly Pests: A World Assessment of Their Biology and Management, pp 299–308. St. Lucie Press, Delray Beach, Florida

Lovette IJ, Seutin G, Ricklefs RE and Bermingham E (1999) The assembly of an island fauna by natural invasion: sources and temporal patterns in the avian colonization of Barbados. Biological Invasions 1: 33–41

Malacrida AR, Guglielmino CR, Gasperi G, Baruffi L and Milani R (1992) Spatial and temporal differentiation in colonizing populations of *Ceratitis capitata*. Heredity 69: 101–111

Manly BFJ (1997) A method for the estimation of parameters for natural stage-structured populations. Researches on Population Ecology 39: 101–111

Metcalf RL (1995a) Biography of the medfly. In: Morse JG, Metcalf RL, Carey JR and Dowell RV (eds) The Medfly in California: Defining Critical Research, pp 43–48. University of California, Center for Exotic Pest Research, Riverside, California

Metcalf RL (1995b) The need for research on exotic pests in California. In: Morse JG, Metcalf RL, Carey JR and Dowell RV (eds) The Medfly in California: Defining Critical Research,

pp 5–39. University of California, Center for Exotic Pest Research, Riverside, California

Neigel JE and Avise JC (1993) Application of a random walk model to geographic distributions of animal mitochondrial DNA variation. Genetics 135: 1209–1220

Nunney L and Elam DR (1994) Estimating the effective population size of conserved populations. Conservation Biology 8: 175–184

Rannala B and Mountain JL (1997) Detecting immigration by using multilocus genotypes. Proceedings of the National Academy of Sciences 94: 9197–9201

Roderick GK (1996a) Geographic structure of insect populations: gene flow, phylogeography, and their uses. Annual Review of Entomology 41: 325–352

Roderick GK (1996b) Population genetic studies of tephritid flies of economic importance. In: McPheron BA and Steck GJ (eds) Fruit Fly Pests: A World Assessment of Their Biology and Management, pp 267–271. St. Lucie Press, Delray Beach, Florida

Roderick GK, Davies N, Bohonak AJ and Villablanca FX (1998) The interface of population genetics and systematics: invasion genetics of the Mediterranean fruit fly (*Ceratitis capitata*). In: Zalucki MP, Drew RAI and White GG (eds) Proceedings of the 6th Australasian Applied Entomological Research Conference, University of Queensland, Brisbane, Australia, pp 489–499. University of Queensland Printery, Brisbane, Australia

Rogers AR (1995) Genetic evidence for a pleistocene population explosion. Evolution 49: 608–615

Saul SH (1992) The medfly in California (Letter to Editor). Science 255: 514

Schrag SJ and Wiener P (1995) Emerging infectious disease: what are the relative roles of ecology and evolution? Trends in Ecology and Evolution 10: 319–324

Siebert JB and Cooper T (1995) Embargo on California produce would cause revenue, job loss. California Agriculture 49: 7–12

Slatkin M (1977) Gene flow and genetic drift in a species subject to frequent local extinctions. Theoretical Population Biology 12: 253–262

Slatkin M (1993) Isolation by distance in equilibrium and non-equilibrium populations. Evolution 47: 264–279

Templeton AR (1998) Nested clade analyses of phylogeographic data: testing hypotheses about gene flow and population history. Molecular Ecology 7: 413–418

Tufto J, Raybould AF, Hindar K and Engen S (1998) Analysis of genetic structure and dispersal patterns in a population of sea beet. Genetics 149: 1975–1985

Villablanca FX, Roderick GK and Palumbi SR (1998) Invasion genetics of the Mediterranean fruit fly: variation in multiple nuclear introns. Molecular Ecology 7: 351–365

Voss HJ (1992) The medfly in California (Letter to Editor). Science 255: 514

Wakeley J and Hey J (1997) Estimating ancestral population parameters. Genetics 145: 847–855

Weir BS (1990) Genetic Data Analysis: Methods for Discrete Population analysis. Sinauer Associates, Sunderland, Massachusetts

Weir BS and Cockerham CC (1984) Estimating F statistics for the analysis of population structure. Evolution 38: 1358–1370

White IM and Elson-Harris MM (1992) Fruit flies of economic significance: their identification and bionomics, CAB International, Wallingford, UK, 601 pp

Wright S (1931) Evolution in Mendelian populations. Genetics 16: 97–159