Know-how for Horticulture™

Use of microsatellite tracking to determine the source of Qfly outbreaks in the Fruit Fly Exclusion Zone

Dr. Stuart Gilchrist
Fruit Fly Research Centre

Project Number: AH01013

AH01013

This report is published by Horticulture Australia Ltd to pass on information concerning horticultural research and development undertaken for Australian Horticulture.

The research contained in this report was funded by Horticulture Australia Ltd with the financial support of Australian Horticulture though across industry programs.

ISBN 0 7341 0986 5

Published and distributed by:
Horticultural Australia Ltd
Level 1
50 Carrington Street
Sydney NSW 2000
Telephone: (02) 8295 2300
Fax: (02) 8295 2399
E-Mail: horticulture@horticulture.com.au

HAL

Know-how for Horticulture™

HORTICULTURE AUSTRALIA LTD

FINAL REPORT

PROJECT AH01013 (01/09/2001-30/06/2004)

Use of microsatellite tracking to determine the source of Queensland fruit fly outbreaks in the Fruit Fly Exclusion Zone

Dr. A. Stuart Gilchrist

Fruit Fly Research Centre, University of Sydney

FRUIT FLY RESEARCH CENTRE

University of Sydney

Horticulture Australia

Horticulture Australia Project Number AH01013

Program/Project Leader
Dr A. Stuart Gilchrist

Principle Investigator
Dr. A. Stuart Gilchrist, Fruit Fly Research Centre, University of Sydney
Address: Fruit Fly Research Centre, A12, University of Sydney, NSW 2006, Australia
E-mail: stuartg@bio.usyd.edu.au
Telephone: +61 2 9351 2298
Fax: +61 2 9351 4771

Collaborative Investigators
Dr. Alan Meats, Fruit Fly Research Centre, University of Sydney
Dr. John Sved, Fruit Fly Research Centre, University of Sydney
Mr B C Dominiak, NSW DPI, Orange.

Purpose of Report

In order to determine the origins of pest populations of Queensland fruit fly (Q-fly) in southeastern Australia, a comprehensive database of DNA "profiles" of possible source populations needed to be established. This report is an account of the analysis of collections of Queensland fruit fly samples collected across the subtropical and temperate range in eastern Australia. Samples were profiled using DNA microsatellites. In combination with DNA profiles from individual towns bordering and inside the Fruit Fly Exclusion Zone, these results were used to determine the origins of outbreak populations in the FFEZ and Adelaide between 2002 and 2004.

Funding:
Horticulture Australia Ltd. Industry Levy
Australian Research Council
NSW Department of Primary Industries
Primary Industries and Resources South Australia
Department of Primary Industries, Victoria
Riverina Citrus

Collaborating Institutions:
NSW Department of Primary Industries
Primary Industries and Resources South Australia
Department of Primary Industries, Victoria

Date of Report: September 2004

Contents

Media summary

In the 1999-2000 season, there were widespread outbreaks of Queensland fruit fly (Q-fly) in the Fruit Fly Exclusion Zone (FFEZ), affecting especially the Riverina region. In subsequent seasons, outbreaks have continued to plague the FFEZ. In order to combat these outbreaks, we need to know where these outbreak flies are coming from.

This project was initiated in order to use DNA fingerprinting technology to identify the sources of the outbreak flies. The main part of the project involved building an extensive database of DNA profiles from all possible source populations. For this, we collected samples from Brisbane in the north to Wodonga in the south, west to Alice Springs and within the FFEZ itself. A large number of samples were also collected along the Western Slopes of NSW. At least 8 distinct potential source populations were identified. Significantly, at least 4 of these were established in the FFEZ.

We then analyzed the DNA fingerprints of outbreak flies from the FFEZ. Our results showed that between 2002 and 2004:

 most outbreaks could be tracked to nearby sources. There was little indication that outbreak flies were arriving directly from Sydney, Brisbane or other coastal sites.

 the major source of outbreak flies was the population from the Wagga-Albury region;

 the population in Deniliquin was also a significant source of outbreak flies; and

 there appeared to be considerable movement of flies along the Murray Valley.

In previous years, Q-fly control in NSW has concentrated on the region north of Wagga. The relatively few flies entering the north-eastern corner of the FFEZ may reflect the effectiveness of those control measures. Nevertheless, the Wagga-Albury region and Deniliquin are currently major sources of outbreaks in the FFEZ and should attract equal control efforts. Our results also suggest that local traffic (as opposed to long-distance travelers from the east coast) may be the major route by which Q-fly disperse and enter the FFEZ.

Technical Summary

The quarantined fruit growing areas of the Riverina region of NSW (part of the Fruit Fly Exclusion Zone; FFEZ) suffered a region wide outbreak of *Bactrocera tryoni* (Queensland fruit fly; Q-fly) in the season 1999-2000. Subsequent seasons saw little overall reduction in the numbers of Q-fly trapped in the FFEZ. To improve current control efforts, we need to know the origin of outbreak flies and whether and where persistent populations of Q-fly have established within the FFEZ.

The most appropriate methodology to address this question is to use DNA microsatellite markers to profile all possible source populations and then to use these profiles to identify the source population of outbreak flies (i.e. assignment testing). This project enabled annual collections and microsatellite profiling of endemic populations of Q-fly from Queensland, NSW and Alice Springs. It complemented an ARC SPIRT (Strategic Partnerships – Industry with Research & Technology) grant that concentrated on profiling Q-fly populations nearer the FFEZ. The results of both these profiling efforts were combined and analyzed jointly to construct a complete picture of Q-fly population structuring in eastern Australia at a resolution that could be used to identify migrants from DNA microsatellite profiles.

The profiling of populations outside the FFEZ found that, for the purposes of assignment, there were two major and a number of minor source populations. The largest population (the "North" group) extended from Sydney northward in a broad coastal strip extending inland to include the Western Slopes of the Great Dividing Range. This population appeared unchanged between years. A second population grouping (the "South" group) emerged on the Western Slopes, south of Wagga, extending at least to the Victorian border. The change of genetic profile between the two groups occurred gradually in the region between Parkes and Wagga. This pattern is consistent with continuing stepwise migration of flies southward along the Western Slopes. In contrast, all towns sampled on the Western Plains supported populations that were more genetically distinct, presumably originating from small founder propagules, that are unaffected by subsequent migration.

Within the FFEZ, we were also able to identify overwintering populations of Q-fly. Deniliquin and Hay in particular supported persistent populations in all three seasons sampled (2001/2 to 2003/4). In 2002/3, Leeton and Barooga also supported distinct populations of Q-fly. Other towns in the southeastern corner of the FFEZ (Narrandera, Tocumwal, Corowa and Wahgunyah) each supported sizeable populations that appeared to be part of the South group flies to the east.

Assignment testing showed that the pattern of migration into the FFEZ consisted mainly of dispersal events from neighbouring populations, particularly from the South group. There was little evidence of large numbers of migrants arriving from the major east coast populations (i.e. Brisbane or Sydney). Within the FFEZ, Deniliquin appeared to be a major source of migrants to other FFEZ towns.

Introduction

Bactrocera tryoni (Queensland fruit fly; Q-fly) is the major horticultural pest in eastern Australia due to its ability to infest 82 host fruits (Anon. 1996) including almost all commercially grown horticultural crops. The species is originally thought to have been largely confined of the rainforests of tropical northeastern Australia, where population densities are very high (>10^3 individuals/ha). A significant range expansion has occurred during the last 200 years since European settlement, following the introduction of exotic horticultural crops into temperate Australia (Lewontin &Birch, 1966). Unrestricted movement of produce prior to 1900 may also have been important (Froggatt, 1909). However, as late as 1908, Q-fly was not established in Sydney (which is well within the present range) yet was common in areas only 100km to the north (Froggatt, 1909), suggesting a stepwise rather than instantaneous range expansion. The current distribution of Q-fly extends from northern Queensland south to Victoria in a broad coastal zone. That zone extends inland at least to the western slopes of the Great Dividing Range. The southern distribution of Q-fly appears to be limited by climate (cooler winter and drier summer conditions in the south and west respectively) rather than the availability of host fruits (Meats, 1981; Sutherst *et al*., 2000). In southern inland areas, away from the relatively humid coastal strip, Q-fly is thought to survive summer moisture stress mainly in urban centres where lawn, garden and orchard irrigation provide necessary moisture (Mavi *et al*., in prep). Temperate Q-fly are only able to overwinter as adults and the extent of overwintering populations in inland towns is unclear (since spring populations could also be due to immigration).

Economic interest in Q-fly is based on the fact that the marginal habitat in the southwest coincides with some of Australia's major horticultural production areas. The marginal habitat (and correspondingly smaller, often transient populations) makes it feasible to attempt to eliminate the fly entirely from that region. To this end, NSW DPI, the Victorian DPI and PIRSA have declared a "Fruit Fly Exclusion Zone" (FFEZ) in the southwest of the species range, which includes major horticultural cropping areas (Figure 1). Quarantine roadblocks surround the FFEZ and a permanent trapping grid of over 3000 traps within the Zone provides continuous monitoring of the presence or absence of fruit fly. Where outbreaks of fruit fly are detected, responses are determined by the Code of Practice for Management of Q-fly (Anon., 1996). Responses include the setting of additional traps and eradication measures by. Successful control of the fly allows produce from the FFEZ to be certified for domestic and international export as coming from a fruit fly free area. Produce certified as originating in a Q-fly free-area enjoys greater market access than non-certified produce.

Despite quarantine and control measures, outbreaks of Q-fly occur each year within the FFEZ. Between 1992 and 1998, an average of only 39 flies were caught per year within the FFEZ. However, in 1999, 251 were caught, while in the first six months of the 2000 season, 1418 were flies were trapped (Meats & Clift, 2003). In the years 2001-2004, similarly high numbers of flies have been trapped each season in the FFEZ. A major question concerns the origin of these outbreak populations. Agricultural authorities have been uncertain of the origin of outbreak flies and, as a consequence, lack the information to formulate an efficient strategy to prevent reinfestations.

Population structure of Q-fly

At present, Q-fly are present in large numbers to the east of the FFEZ but largely absent from the north, west and south. To the west of this zone heat and moisture stress are too high for Q-fly to survive (Meats, 1981; Yonow &Sutherst, 1998) except in urban areas where they benefit from artificial irrigation. Some, but not all larger towns support populations: Wilcannia and Alice

Springs have Q-fly populations, while Broken Hill has been fly free for at least two decades. Q-fly are certainly present to the south of the FFEZ in Victoria, but the cooler usually climate limits their ability to build populations with serious pest status (Meats, 1981; Yonow & Sutherst, 1998). As a result of this distribution, the most likely source of outbreaks in the FFEZ were the populations to the east and northeast of the FFEZ. To determine the exact origin of the immigrants into the FFEZ, some defining character or characters are needed that allow the geographic origin of immigrants to be determined.

Prior to the advent of molecular techniques, (Bateman, 1967) investigated population structure in Q-fly by looking for differentiation in heat and cold tolerance of flies sampled along the eastern coast of Australia. No cline was found and no simple conclusions about population structure could be drawn. Using allozymes, McKechnie (1974) also found no evidence for population structuring.

DNA microsatellites provide a much more sensitive tool for the investigation of population structuring. Typically, numerous individuals from each of a series of populations are typed for a reasonable numbers of markers (~5-30 microsatellites). Each microsatellite can have numerous alleles, the presence and frequency of which will usually vary between different populations. Various statistical procedures can be used to infer the characteristics of each population (e.g. population size and level of inbreeding) and the relation between different populations (e.g. the degree of differentiation or the geographic pattern of differentiation).

Using six microsatellite markers, Yu *et al.* (2001) investigated population structure in the main endemic regions of Q-fly. They showed that the major endemic populations form three distinct subpopulations that were surprisingly stable over a five-year period. These populations are numbered 1-3 in Figure 1. Using the same six microsatellites, Sved *et al.* (2003) surveyed outbreak flies from within and around the FFEZ and, using a shared allele test, concluded that FFEZ flies were most likely to have originated from populations neighbouring the north-eastern edge of the FFEZ, rather than from the more distant but vastly more populous core populations. It had previously been assumed by regulatory authorities that the large endemic populations on the NSW east coast or in Queensland were the source of infestations (e.g. Dominiak *et al.*, 2000). This view assumed that long-range passive dispersal (as larvae in fruit) from distant but large populations lead to more FFEZ outbreaks than shorter range dispersal from areas closer to the FFEZ. Unassisted dispersal of Q-fly is probably limited to tens of kilometres (Fletcher, 1974; MacFarlane *et al.*, 1987; Meats, 1998).

Prompted by the likely importance of the marginal populations of Q-fly (Sved *et al.*, 2003) and the uncertainty surrounding passive migration rates, we used a larger set of microsatellites developed at the FFRC to study population structuring across the entire range of Q-fly in eastern Australia. Specifically, from a set of 29 microsatellites we found 21 that could be confidently used to establish DNA profiles of samples of Q-fly from across the entire subtropical and temperate range in eastern Australia. From that information, we were able to identify the distinct populations present in 2002-4. We then tested the likelihood that outbreak flies trapped in the FFEZ and Adelaide came from each of those distinct populations.

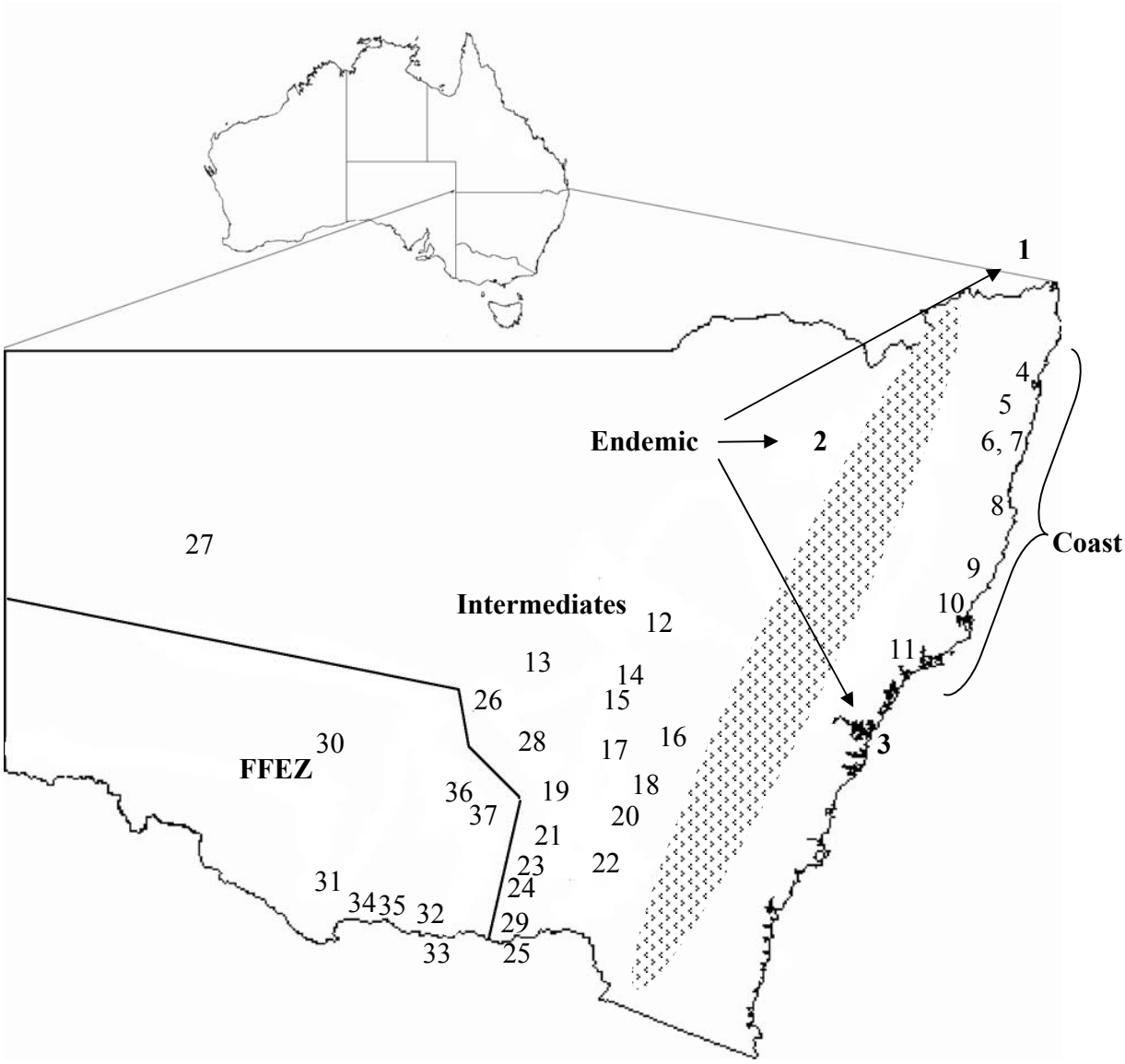


Figure 1. Map showing each sample site, numbered according to Table 1. The bold numbers (1-3) indicate the Endemic populations, 4-11 are the Coast sites, 12 -29 are Intermediate populations and 30-37 represent FFEZ populations. The hatched region indicates the Great Dividing Range mountain chain.

Materials and Methods

1. Fruit Fly Samples

Note: in this document, seasons are referred to by the second year e.g. 2002 refers to the 2001-2002 season.

(a) Source populations

Sampling of Q-fly for this project covered four distinct areas, as indicated in Figure 1 and listed in Table 1. The first area consists of three locations where flies are endemic and have already been shown to have a stable population structure between years (Yu *et al.*, 2001). These samples (the "Endemic" samples) were from the subtropical and warm temperate areas of eastern Australia (i.e. Queensland, Northern NSW and Sydney, sites 1-3). Samples were collected in late summer 2001, 2002 and 2003.

The second group of flies was collected from sites along the NSW coast between Maclean and Newcastle (approximately 400km). These "Coast" samples were collected to determine whether the fruit fly populations in the smaller coastal towns differed from those in the two flanking cities, Brisbane and Sydney.

The third group of samples consisted of the "Intermediate" populations found in inland NSW, to the southwest of the endemic populations but outside the FFEZ. This sampling included the Risk Reduction Zone (RRZ), an 80km wide strip abutting the north-east of the FFEZ. The genetic structuring of the Intermediate populations has not been examined previously and the extent of overwintering among these populations is not clear. A major sampling effort was undertaken in 2002 with the aim of sampling all major towns in the RRZ at the same point in time (February-March 2002). This was necessary to avoid confounding variation between towns with variation between years. This 2002 "snapshot" of the genetic structure of the region was used as the basis of analysis of samples from 2003 and 2004.

The fourth group of samples shown in Table 1 was the FFEZ group collected on the permanent trapping grid maintained within the FFEZ. These samples include only those FFEZ sites where sufficiently large numbers were trapped to enable accurate DNA profiling of flies. These populations, although geographically close to the Intermediate populations, differ in that most are subjected to extensive control programs. Trapping records show that before 1998 fruit fly was essentially absent from the FFEZ. Therefore we can be certain that the FFEZ samples either originate from very recently founded populations or are actual immigrants (or their offspring) into the FFEZ.

A fifth group of flies, not shown in Table 1, were the FFEZ flies that did not appear to be part of established populations. These were the actual "outbreak" flies and their selection from among all the flies trapped in the FFEZ is discussed the Results section 5(e) below. Similarly, we analyzed two small collections of flies from Adelaide from 2002 and 2004. Although Adelaide is outside the FFEZ, it is not thought to support a permanent population of Q-fly (Maelzer *et al.*, 2004). Accordingly, both the Adelaide samples were also analyzed as isolated outbreaks.

Samples of Q-fly from Alice Springs were also collected. Analysis showed that this population is extremely differentiated from all eastern populations and that it could not be a source population for any of the FFEZ or Adelaide flies. Therefore, it was not included in the analysis resent below.

Region/Year	Site (Fig.1)	Population	*N*	*He*	*Ho*	*f*	HWE	Allelic Richness
Endemic								
2001	1	Queensland	52.0	0.601	0.563	0.064		
	2	Northern NSW	38.4	0.595	0.561	0.057		-0.74
	3	Sydney	66.3	0.590	0.568	0.037	*	-0.31
2002		Queensland	39.3	0.580	0.541	0.069	*	
		Northern NSW	46.5	0.603	0.570	0.055		-1.04
		Sydney	71.1	0.578	0.522	0.099	*	-0.69
2003		Queensland	39.1	0.584	0.571	0.022		
		Northern NSW	40.1	0.608	0.577	0.053		-0.21
		Sydney	40.1	0.585	0.573	0.020		
Coast								
2002	4	Maclean	31.4	0.594	0.560	0.058	*	-0.45
	5	Grafton	26.6	0.621	0.572	0.080	*	-0.52
	6	Sawtell	33.0	0.603	0.573	0.049		0.41
	7	Coffs Harbour	17.4	0.551	0.540	0.020		-0.77
	8	South West Rocks	28.1	0.600	0.580	0.033		-0.40
	9	Taree	28.7	0.583	0.532	0.088		-0.73
	10	Foster	31.5	0.608	0.560	0.080	*	0.25
	11	Newcastle	32.3	0.576	0.539	0.066		-0.72
Intermediate								
2002	12	Dubbo	24.7	0.593	0.573	0.035		-0.71
	13	Condobolin	39.8	0.580	0.528	0.090		-1.67
	14	Parkes	19 .9	0.576	0.562	0.024		-0.38
	15	Forbes	32.2	0.566	0.553	0.024		-0.79
	16	Cowra	19.3	0.559	0.551	0.014		-0.60
	17	Grenfell	29.3	0.597	0.585	0.019		-1.29
	18	Young	46.3	0.622	0.627	-0.008		-0.99
	19	Temora	19.4	0.611	0.586	0.042		-1.41
	20	Cootamundra	41.6	0.603	0.597	0.010		-0.97
	21	Wagga	54.9	0.615	0.591	0.040		-1.76
	22	Tumut	19.8	0.608	0.558	0.084		-0.76
	23	The Rock	19.8	0.618	0.621	-0.005		-0.95
	24	Henty	18.9	0.582	0.562	0.036		-0.85
	25	Wodonga	39.7	0.593	0.544	0.082	*	-1.28
	26	Lake Cargelligo	68.8	0.593	0.547	0.079	*	-2.85*
	27	Wilcannia	36.5	0.520	0.473	0.091	*	-2.55*
2003		Condobolin	14.5	0.576	0.575	0.002		-0.90
		Parkes	39.0	0.606	0.553	0.089	*	-1.25
	28	West Wyalong	22.4	0.526	0.511	0.029	*	-1.07
		Wagga	150.9	0.623	0.598	0.041	*	-1.95
	29	Albury	25.2	0.600	0.598	0.004		-0.77
FFEZ								
2002	30	Hay	24.2	0.519	0.485	0.065	*	-2.82*
	31	Deniliquin	33.8	0.517	0.496	0.042		-3.34*
	32	Corowa	20.1	0.589	0.560	0.049		-1.46
	33	Wahgunyah	21.9	0.561	0.510	0.094		-0.85

Region/Year	Site (Fig.1)	Population	*N*	*He*	*Ho*	*f*	HWE	Allelic Richness
	34	Tocumwal	12.2	0.568	0.515	0.103		-1.15
2003	35	Barooga	31.3	0.509	0.561	-0.104	*	-2.52*
		Hay	25.7	0.510	0.480	0.060		-2.46*
		Deniliquin	48.4	0.517	0.482	0.069	*	-2.59*
	36	Leeton	75.7	0.594	0.589	0.008	*	-1.92*
	37	Narrandera	12.2	0.559	0.477	0.152	*	-0.87
2004		Deniliquin	68.3	0.541	0.505	0.067	*	-4.22*
		Narrandera	23.0	0.589	0.547	0.074	*	-1.34

Table 1. Summary statistics of the samples of Q-fly analysed. Sites correspond to those shown in Figure 1. *N* indicates the number of individuals, H_{exp} and H_{obs} are expected and observed heterozygosities and *f* is the inbreeding coefficient. In the HWE column, an asterisk identifies samples not in Hardy Weinberg equilibrium. Allelic Richness indicates for each sample the average difference (mostly a reduction) between observed and expected numbers of different alleles per locus. Expected numbers of alleles were calculated in relation to the combined Queensland samples (see text for details). An asterisk indicates significant reduction in allelic richness after multiple comparisons.

2. Microsatellite typing

Individuals were typed using the standard fluorescent PCR methods detailed in Yu *et al.* (2001). Samples from all sites were typed for up to 29 microsatellites (Kinnear *et al.*, 1998; Wang *et al.*, 2003). In total, 1877 flies were included in the analysis and their basic descriptive statistics (expected and observed heterozygosity and coefficient of inbreeding) are shown in Table 1. The microsatellite typing data was 91% complete.

3. Statistical Analysis

The likely presence of null alleles was detected using tests for homozygote excess implemented in the MICRO-CHECKER software (van Oosterhout *et al.*). We tested for departure from Hardy-Weinberg equilibrium (HWE) proportions within populations across loci using the GDA program (Lewis & Zaykin, 2002). Probabilities of departure from HW equilibrium were combined across loci using Fisher's technique (Sokal &Rohlf, 1995).

Pairwise genetic differences between populations were tested using the exact test procedure of FSTAT version 2.9.3 (Goudet, 2002) with 5000 permutations. Mantel tests were also performed using FSTAT. Population structuring was further investigated using STRUCTURE software (Pritchard *et al.*, 2000; Falush *et al.*, 2003). The program was used, initially, to partition individuals into clusters, which it does on the basis of forming the most likely groupings that are in Hardy-Weinberg and linkage equilibrium. Subsequently, the assignment testing function of the STRUCTURE software was also used to determine the likely geographic origins of individual flies. In all cases 50,000 repetitions were used after a burn-in of 50,000 repetitions. Rather than rely on a single assignment method (Cegelski *et al.*, 2003), we also used a frequency based assignment method (Paetkau *et al.*, 1995) as implemented in the GENECLASS 2.0 software (Piry *et al.*, in press) as an alternative to the Baysian analysis of the STRUCTURE software. A hierarchical analysis of molecular variance (AMOVA) was performed using ARLEQUIN Version 2.0 (Schneider *et al.*, 2000). Estimates of population size from temporal data were made using the MNE 1.0 software (Wang & Whitlock, 2003). In all simultaneous statistical tests, critical significance levels were corrected using the sequential Bonferroni test to enable overall significance to be examined.

Allelic richness, or the number of different alleles present, of each of the Intermediate and FFEZ samples was calculated as it is more sensitive than heterozygosity measures to the effects of short, severe bottlenecks, such as occur during the founding of new populations by a limited number of individuals (Leberg, 2002 and references therein). Allelic richness was calculated by rarefaction (Hurlbert, 1971) and significance tested using a resampling process. Rarefaction calculates the number of different alleles expected in a sample given the allele frequencies in a larger reference sample. In our case, we used the combined Queensland samples as a reference population. To estimate the significance of allelic richness variation, we drew 5000 random resamples from the reference population, matching the size of the particular Intermediate or FFEZ sample. The number of resamples containing the same or fewer alleles than the observed sample was used to calculate the probability of the hypothesis of equal allelic richness between the reference populations and the particular Intermediate or FFEZ samples. Probabilities were combined across loci and the results adjusted for multiple comparisons.

To visualize geographic structuring of the populations, many studies use phylogenetic trees. However, such trees rely on assumptions such as constant population size and isolation of populations. Q-fly populations, particularly in inland areas, violate these assumptions as they are recently founded, variable in population size and are likely to have significant migration rates. As an alternative, we also used the ordination method correspondence analysis (CA, or reciprocal averaging). CA produces one or more axes that summarize major sources of variation and along which the sites can be arranged (e.g. She *et al.*, 1987; Canon *et al.*, 2001; Cruciani *et al.*, 2002). CA differs from the closely related principle components analysis because PCA preserves metric distances calculated from allele frequency data while CA preserves a chi-square distance (Gauch, 1982). The ordination axes represent a simple gradient of the underlying variables along which the sites are distributed. It is notable that ordination has been found to be useful in genetic analysis of some animal breeding programs, where the applicability of hierarchical analysis is limited by the high rates of population mixing (Canon *et al.*, 2001; Rosenberg *et al.*, 2001). Ordination was performed using JMP statistical software (SAS Institute, 1994).

Results

1. Null alleles

We tested for the presence of null alleles using the 2001 and 2002 data from the Endemic and Intermediate regions. We surveyed 29 microsatellites in 16 populations, giving 464 tests for null alleles. Assuming a Type I error rate of 0.05, we would expect at least 22 tests to be significant across the 16 populations. Therefore, we assumed that a particular locus was likely to harbour null alleles where more than 2 of the 16 populations showed a significant excess of homozygotes. This process identified 8 microsatellites that were likely to harbour nulls and these were excluded from further analysis. Prior to the removal of these 8 microsatellites, all the 17 populations all showed significant departures from HWE (after sequential Bonferroni correction). After removal of the 8 microsatellites likely to be harbouring nulls, only 3 populations still showed a deviation from HWE, indicating that null alleles are likely to be the man reason for departure from HWE in these populations. As a result, all following analysis was performed using only the 21 microsatellites that showed no significant levels of null alleles.

2. Microsatellite variation

Microsatellite variation within each sampling site, measured as observed heterozygosity, showed little variation between sites. Expected heterozygosity varied between 0.47 and 0.60, while

observed heterozygosity varied between 0.47 and 0.62. The relative consistency of the heterozygosity values across all sites indicated that only small amounts of microsatellite variation has been lost during the range expansion of Q-fly into temperate regions.

Changes in allelic richness provide a more sensitive measure of the loss of microsatellite variation resulting from population bottlenecks. When the combined Queensland samples were used as the source population, almost all other sites showed a reduction in allelic richness (Table 1). This is consistent with our view of all temperate populations being ultimately derived from the Queensland populations. However, the only populations that showed a statistically significant reduction in allelic richness were those from the western plains (as opposed to the western slopes) of NSW. Two of the significant sites were outside the FFEZ: Lake Cargelligo and Wilcannia. The remaining significant sites (Deniliquin, Hay, Barooga and Leeton) were all in the FFEZ but away from its eastern edge. This strongly suggests a reduction in effective migration rates of fruit flies on the western plains in comparison to the western slopes.

The results of tests for Hardy Weinberg equilibrium (HWE) are also shown in Table 1. Despite the removal of loci likely to harbour null alleles, nine of the non-FFEZ sites still showed a deviation from HWE. This is not a surprising result given the large number of microsatellite loci used. Further possible reasons for the deviations are presented in the Discussion.

3. Removal of Sterile Release flies

Attempts to control Q-fly in the FFEZ using Sterile Insect Technique (SIT) began in 1999. The SIT method relies on the release of vast numbers of sterile flies, which are intended to mate with wild females, thereby reducing the number of fertile eggs laid. Previous analysis of the Q-fly SIT strain has shown that it is highly genetically distinct from wild flies and has a relatively stable microsatellite profile across years (Gilchrist *et al.*, 2004). The same work showed that the STRUCTURE software could be used to identify SIT flies from among a mixture of wild and SIT flies with virtually 100% accuracy.

All SIT flies are marked with a pink fluorescent dust to allow them to be distinguished from wild flies. Despite all flies trapped within the FFEZ having been visually checked for pink fluorescent dust (positive samples are removed), there exists a possibility that some flies may have lost the identifying dust. Consequently, all FFEZ flies analyzed in the current study were tested against 3 reference samples of SIT flies (one each from 2001, 2002 and 2003). Also included in the analysis were representative wild samples from Endemic regions, Dubbo and Wagga (all 2002). These samples could be safely assumed not to contain SIT flies since SIT trials outside the FFEZ finished in April 2001.

To be conservative, any fly showing a probability > 0.5 of being an SIT fly was excluded from the analysis. Of over 450 FFEZ flies tested, 19 showed a probability > 0.5 of being SIT flies. These were excluded from all subsequent analysis. The Adelaide samples (32 flies) were also tested but found to contain no SIT flies.

Table 1 (following page). Results of exact tests for genetic differentiation of the samples shown in Table 1. Values below the diagonal are pairwise F_{ST} values. Above the diagonal are the results of the exact tests: NS = not significant (adjusted p = 0.0001). Tocumwal was not included due to its small sample size. Newc. = Newcastle, Coota = Cootamundra, Wahgun = Wahgunyah, Deni = Deniliquin.

	QLD	NSW	Sydney	Maclean	Grafton	Sawtell	Coffs	SW Rocks	Taree	Foster	Newc.	Dubbo	Condobolin	Parkes	Forbes	Cowra	Grenfell	Young	Temora	Coota.	Wagga	Tumut	The Rock	Henty	Wodonga	Lake Carg	Wilcannia	Hay	Deni.	Corowa	Wahgun.
QLD		NS	NS	NS	*	NS	NS	NS	NS	NS	NS	NS	*	NS	*	NS	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
NSW	-.002		NS	NS	NS	NS	NS	NS	NS	NS	NS	NS	*	NS	*	NS	*	*	*	*	*	*	*	*	*	*	*	*	*	*	NS
Sydney	.005	.001		NS	NS	NS	NS	NS	NS	NS	NS	NS	*	NS	NS	NS	*	*	*	*	*	*	*	*	*	*	*	*	*	*	NS
Maclean	.000	-.001	.004		NS	NS	NS	NS	NS	NS	NS	NS	*	NS	NS	NS	*	*	*	NS	*	NS	*	*	*	*	*	*	*	*	NS
Grafton	.017	.015	.023	.022		NS	NS	NS	NS	NS	NS	*	*	*	*	NS	*	NS	*	*	*	*	*	*	*	*	*	*	NS	*	NS
Sawtell	.002	-.002	.000	-.001	.025		NS	NS	NS	NS	NS	NS	*	NS	NS	NS	*	NS	*	NS	*	*	*	*	*	*	*	*	*	*	NS
Coffs	.004	.006	.010	.003	.020	.005		NS	NS	NS	NS	NS	NS	NS	NS	NS	*	NS	*	NS	NS	NS	*	*	NS	*	*	*	*	*	NS
SWR	.003	.001	.005	-.001	.022	.002	.007		NS	NS	NS	*	*	NS	NS	NS	*	NS	*	*	*	NS	*	*	*	*	*	*	*	*	NS
Taree	.009	.008	.006	.014	.022	.012	.016	.016		NS	NS	NS	*	NS	NS	NS	*	*	*	NS	*	*	*	*	*	*	*	*	*	*	NS
Foster	.005	.001	.000	.003	.012	-.003	.007	.008	.005		NS	NS	*	NS	NS	NS	*	NS	*	NS	NS	NS	*	NS	*	*	*	*	*	NS	NS
Newc.	.006	.004	.005	.005	.023	.001	.014	.011	.007	.001		NS	*	NS	*	NS	*	*	*	*	*	*	*	*	*	*	*	*	*	*	NS
Dubbo	.016	.013	.016	.021	.042	.008	.024	.022	.021	.014	.010		*	NS	NS	NS	*	NS	*	*	*	NS	*	*	*	*	*	*	*	*	NS
Condo.	.032	.034	.032	.033	.051	.037	.022	.040	.033	.032	.034	.042		*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	NS
Parkes	.007	.012	.010	.011	.029	.010	.000	.012	.011	.004	.008	.018	.022		NS	NS	*	*	*	NS	*	NS	*	NS	*	*	*	*	*	*	NS
Forbes	.016	.016	.013	.016	.037	.010	.012	.019	.022	.016	.016	.020	.045	.021		NS	*	NS	*	*	*	*	*	*	*	*	*	*	*	*	NS
Cowra	.012	.006	.007	.014	.038	.015	.019	.014	.024	.020	.016	.021	.032	.018	.021		NS	*	*	NS	*	*	*	*	*	*	NS	*	*	*	NS
Grenfell	.013	.013	.015	.017	.030	.014	.020	.017	.016	.007	.012	.022	.030	.021	.027	.022		*	*	*	*	*	*	*	*	*	*	*	*	*	NS
Young	.014	.011	.015	.007	.020	.010	.005	.013	.011	.006	.013	.019	.030	.012	.017	.026	.021		*	*	*	*	*	*	*	*	*	*	*	*	NS
Temora	.033	.030	.034	.037	.041	.030	.035	.033	.043	.027	.030	.022	.065	.036	.047	.037	.024	.031		*	*	*	*	*	*	*	*	*	NS	*	NS
Coota.	.013	.013	.017	.012	.033	.016	.015	.017	.016	.018	.021	.020	.040	.014	.025	.019	.024	.016	.029		NS	*	*	*	*	*	*	*	*	*	NS
Wagga	.021	.014	.023	.017	.029	.020	.021	.026	.021	.013	.021	.024	.036	.024	.031	.031	.016	.016	.024	.011		NS	NS	NS	*	*	*	*	*	*	NS
Tumut	.019	.013	.024	.017	.033	.020	.017	.021	.019	.013	.024	.019	.038	.015	.040	.027	.022	.015	.020	.016	.003		*	*	*	*	*	*	*	*	NS
The Rock	.029	.027	.029	.023	.041	.026	.030	.028	.023	.015	.027	.040	.040	.033	.046	.055	.026	.021	.047	.030	.009	.019		NS	*	*	*	*	*	*	*
Henty	.045	.038	.044	.039	.046	.045	.037	.039	.052	.036	.045	.051	.057	.044	.059	.055	.046	.028	.042	.038	.018	.016	.026		*	*	*	*	*	*	NS
Wodonga	.041	.041	.042	.042	.040	.035	.038	.046	.035	.026	.031	.037	.053	.035	.033	.058	.023	.031	.033	.034	.023	.033	.035	.044		*	*	*	*	*	NS
Lake Carg	.047	.041	.052	.042	.044	.052	.039	.044	.055	.049	.053	.052	.059	.055	.061	.053	.051	.041	.073	.052	.053	.042	.064	.061	.075		*	*	*	*	*
Wilcannia	.039	.041	.046	.056	.054	.048	.055	.039	.041	.052	.049	.041	.081	.048	.059	.052	.062	.057	.071	.056	.064	.043	.070	.078	.086	.082		*	*	*	NS
Hay	.091	.089	.087	.085	.114	.081	.098	.099	.103	.078	.089	.101	.090	.107	.118	.102	.081	.094	.115	.104	.081	.097	.077	.112	.117	.116	.134		*	*	NS
Deni.	.077	.079	.080	.084	.091	.080	.101	.093	.085	.066	.083	.088	.110	.108	.097	.110	.061	.082	.097	.099	.088	.109	.096	.126	.093	.123	.125	.147		*	NS
Corowa	.036	.024	.024	.030	.056	.019	.043	.031	.040	.018	.029	.021	.059	.033	.052	.031	.037	.031	.035	.030	.023	.020	.028	.039	.054	.068	.065	.083	.103		NS
Wahgun..	.018	.020	.022	.020	.030	.022	.019	.022	.010	.014	.014	.026	.046	.019	.031	.035	.027	.020	.039	.028	.019	.012	.017	.029	.027	.062	.034	.105	.099	.031	

4. Spatial structuring in 2002

By analyzing only samples collected simultaneously across all regions (in early 2002), we were able to delineate the population structuring of Q-fly without the confounding effects of temporal variation. Samples used in this analysis came from all regions (Endemic, Coast, Intermediate and FFEZ) and consisted of at least 20 flies from each site. Many additional sites were sampled in 2002 but small numbers of individuals from a site can give misleading estimates of the genetic composition of the resident population. Therefore smaller samples were excluded from the 2002 spatial structuring analysis. In Table 1, some sites show a sample size slightly less than 20, but this sample size is calculated from the sample size averaged across all microsatellite loci typed. Therefore, where typing failed for some individuals at some loci, the calculated sample size fell below 20. For the FFEZ, two sites were included that had sample sizes of only 12 (Tocumwal 2002 and Narrandera 2003) as trapping records (Table 8) indicated that those two towns might also have supported sizeable populations at that time.

(a) Genetic differentiation among 2002 samples

Table 2 shows the results of the exact tests for pairwise genetic differentiation among the 31 samples after correction for multiple comparisons. The Table shows the samples grouped according to region. Of the 465 pairwise comparisons, 305 showed significant genetic differentiation after correction for multiple comparisons (adjusted $p = 0.0001$), including populations separated by as little as 48km (Young and Cootamundra). Most significant differences involved sites in the Intermediate region and FFEZ, indicating the greater genetic differentiation in those areas than in the Endemic or Coast regions.

Of the 160 non-significant pairwise comparisons, most involved pairings between sites that formed an obvious geographic grouping: all the Endemic, Coast and the four most northeasterly of Intermediate sites (Dubbo, Parkes, Forbes and Cowra). Thus the genetic differentiation is concentrated on the Western Slopes and Western Plains of NSW, suggesting that homogenizing forces such as migration are significantly lower in those regions.

(b) Distribution of genetic variance

Hierarchical analysis of molecular variance was performed on the basis of a division of the 2002 sites into four regions:

1. The Endemics and the Coastal sites;
2. The Intermediate sites;
3. Lake Cargelligo and Wilcannia and;
4. FFEZ sites.

Lake Cargelligo and Wilcannia were separated from the other Intermediate sites since the genetic distances between these and the remaining Intermediate sites were noticeably high, suggesting highly differentiated populations in these two towns. Genetic distances between the FFEZ sites and all other sites were also large. Table 2 shows that 96.7% of microsatellite variation occurred between individuals within sites. Only small, but significant, amounts of variation were found between regions (1.0%) and among sites within regions (2.3%).

Source of variation	d.f.	Variance Components	F-statistics	P	Percentage variation
Among Regions	3	0.048 (V_a)	F_{RT} = 0.010	< 0.0001	1.0
Among Sites within Regions	27	0.111 (V_b)	F_{SR} = 0.023	< 0.0001	2.3
Within Sites	2315	4.731 (V_c)	F_{ST} = 0.033	< 0.0001	96.7
Total	2345	4.890(V_t)			

Table 2. Analysis of molecular variation (AMOVA) of all the 2002 populations.

(c) Correspondence analysis of 2002 samples

Figure 2 shows the results of the correspondence analysis of the 2002 data. The Endemic sites form a central group around is which is clustered the Coast sites, indicating that all the Coast sites are closely related to the Endemics. The Intermediate sites show a range of distances from the Endemics, with the most southerly sites (The Rock, Henty and Wodonga) also being the most distant from the "core" Endemic group. It is notable that the Intermediates are clustered on one side of the Endemic sites, rather than spread evenly on either side, suggesting that geographic structure correlates with the genetic structure. This suggests a pattern of relatedness between the Intermediates, following an approximate north-south axis. The two Intermediate sites that did not follow this trend were Wilcannia and Lake Cargelligo, which are the two Intermediate sites isolated on the western plains. It is also noteworthy that of the FFEZ sites, Hay and Deniliquin (also on the western plains) show the most distant relation to the Endemics, while Corowa, Wahgunyah and Tocumwal (all along the Murray River) are more closely related.

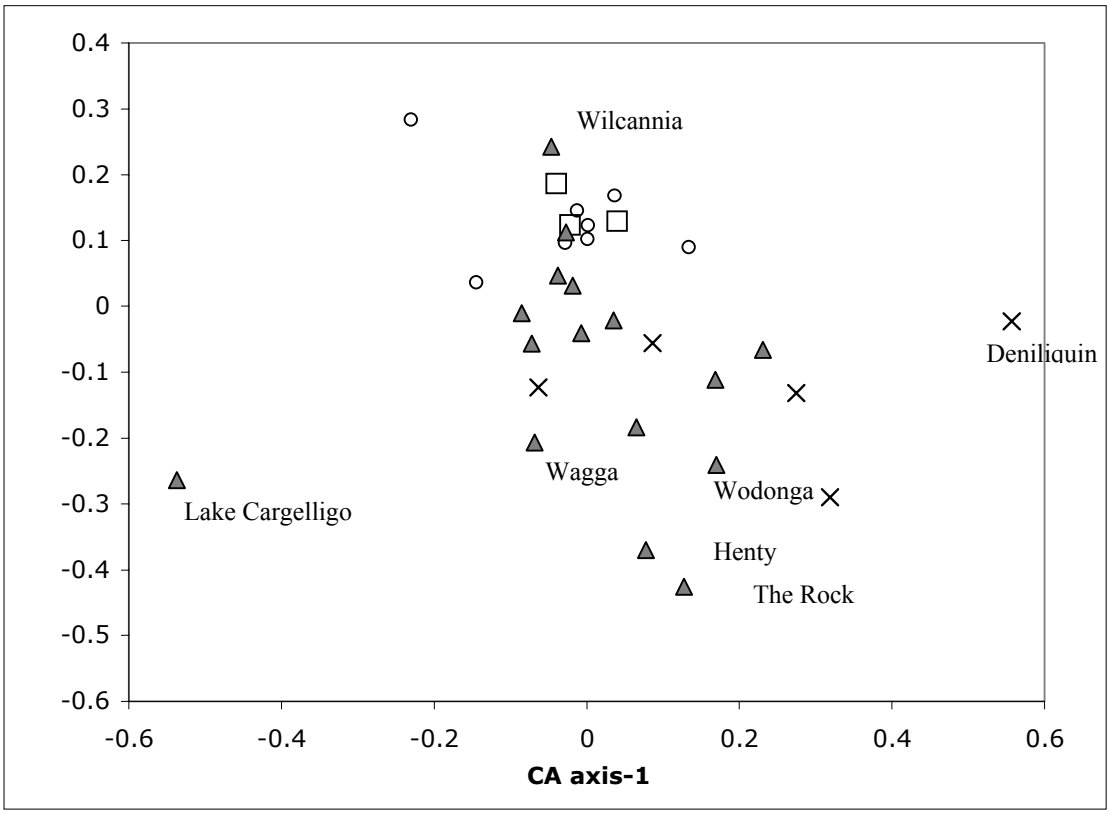


Figure 2. Correspondence analysis of the 2002 samples

(d) STRUCTURE analysis of 2002 samples

The patterns of genetic distances among sites (Section (a) above) showed few significant differences between the Endemic and Coast sites. This suggests relatively little if any significant population structuring among the Endemic and Coast sites. Analysis of these populations using the STRUCTURE software reflected this view: STRUCTURE could not partition the Endemic and Coast sites into any distinct clusters.

However, analysis of the Intermediate and FFEZ samples using the STRUCTURE software showed a high degree of population genetic structuring. To begin this analysis, the Northern NSW sample was used as a "learner "sample. Since it is immediately adjacent to the Intermediate area (and hence a likely source of migrants) and is a relatively stable population (Yu *et al.*, 2001), it provided a user-defined cluster to assist clustering of the remaining samples.

STRUCTURE attempts to group individuals into a number of clusters (K; defined by the user) without reference to their geographic origin. The program analyzes the data repeatedly, each time using a different value of K, and for each analysis, the likelihood of the resulting K clusters is calculated. Typically, the likelihoods plateau as K increases and the point at which the plateau begins indicates the most likely value of K. This value of K was then adopted as the most likely number of distinct clusters in the data. For the 2002 samples, values of K ranging from 1 to 10 were tested. The posterior probabilities of K began to plateau at K = 6. Thus we concluded there were six distinct populations in the 2002 samples.

For each fly, STRUCTURE calculates the probability that the fly belongs to one of the K clusters (with the probabilities summing to 1). When a fly belonged to a cluster with a probability > 0.9, we classified that fly as belonging to that cluster. We used this criterion to decide which flies were "members" of which clusters. The results for K = 6 showed four distinct clusters and 2 less obvious clusters. Each of the four distinct clusters consisted almost exclusively of flies from one each of the 4 most westerly sites: Lake Cargelligo, Wilcannia, Hay and Deniliquin. For example, the "Deniliquin" cluster consisted entirely of flies from Deniliquin and no flies from other towns. In each case, a least 78% of the individuals sampled in one of the four towns showed average membership > 0.9 of the corresponding cluster (Table 4). Only 5 individuals from these four sites showed a membership > 0.9 of any other cluster. Thus each of the four clusters had a specific geographic basis.

The remaining two clusters were far less distinct. Following the recommendations accompanying the software, the 4 distinct clusters were removed from the data and the remaining 466 individuals were tested for clustering separately. Those 466 remaining flies all came from the sites either spread along the western slopes (>200m ASL) or on the Murray River (Figure 1). For that reduced dataset, values of K ranging from 1 to 6 were tested. The plateau of the posterior probabilities of K began at K = 2, although the plateau was not sharply defined. Results for K= 2 showed that of the 466 individuals, only 195 showed membership > 0.9 of either cluster. The pattern of cluster membership among the samples showed a cline, with more northerly sites having more members of one cluster and more southerly sites more members of the alternate cluster (Table 5). These results remained unchanged whether or not the Northern NSW sample was used as a learner sample. The significance of the clinal pattern was tested using the Mantel test, which revealed a significant relation between geographic separation and genetic distance for the populations in the two clusters ($r = 0.47$, $P < 0.01$).

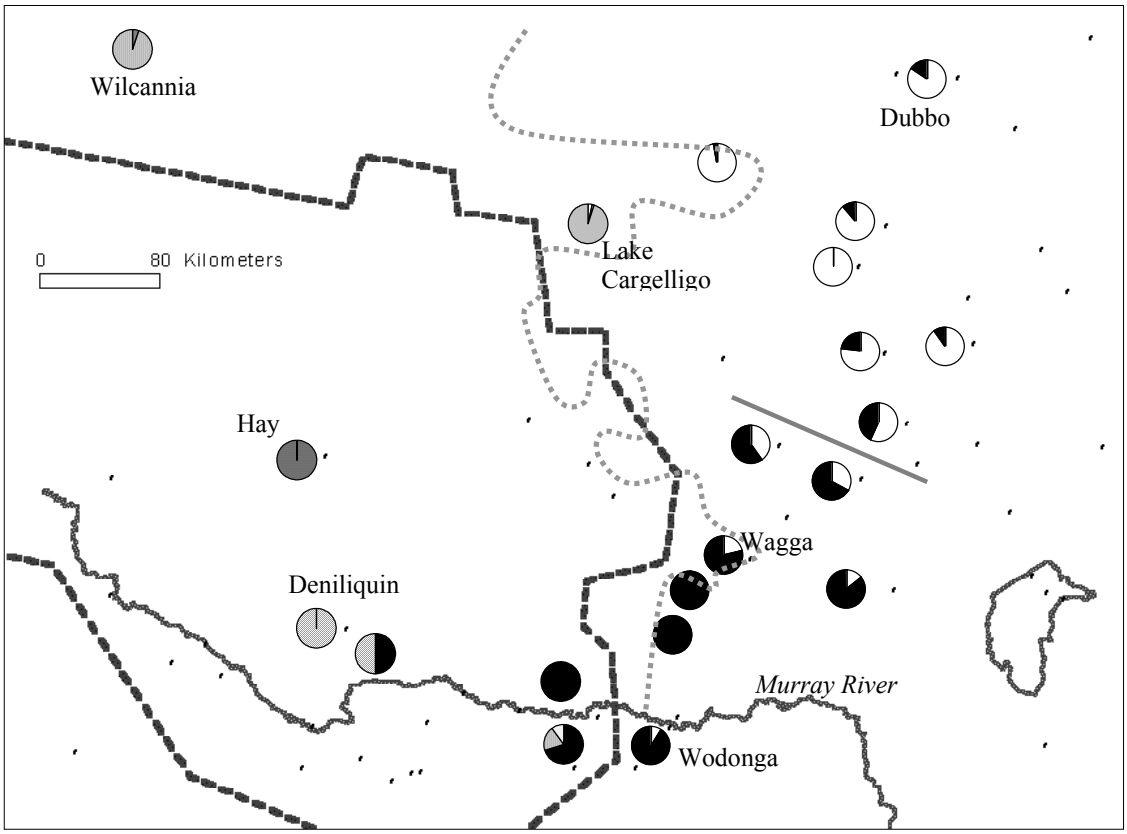


Figure 3. Map of the FFEZ and Intermediate region showing the pattern of genetic differentiation among the 2002 sites and the 6 source populations identified in that year. Dots on the map indicate towns with human population ~1500 or more. Each pie graph represents the sample from that town, and is divided according to the proportion of individuals showing membership of different source populations. White denotes the North group; black the South group, while Lake Cargelligo, Wilcannia, Hay and Deniliquin all show distinct genetic clusters (shown by the different fill patterns). The heavy dashed line is the border of the FFEZ, while the grey dashed line approximates the 200m contour, dividing the Western Slopes to the east from the Western Plains to the west. The short grey line shows the demarcation used in the present study between the North and South regions.

Cluster	2002		2003		2004	
	Local membership (average membership)	Membership of other clusters	Local membership (average membership)	Membership of other clusters	Local membership (average membership)	Membership of other clusters
Lake Cargelligo	58/74 (0.91)	3				
Wilcannia	36/43 (0.90)	2				
Hay	24/26 (0.96)	0	23/28 (0.90)	0		
Deniliquin	37/40 (0.95)	0	48/53 (0.92)	0	69/73 (0.95)	3
Barooga 03			28/33 (0.91)	1		
Leeton03			55/82 (0.80)	5		

Table 4. Genetic clustering results for the Western Plains populations. For each cluster, the number of flies that were members ($p > 0.9$) of that town's cluster are shown as a fraction of the total number of flies trapped at that site in that year. The figure in brackets is the mean probability of membership of all flies trapped at that site in that year. The second column for each year shown then numbers of flies that showed a membership of another cluster, $p > 0.9$).

Based on this cline, we simply divided the samples from the eastern side of the FFEZ into two groups for the remaining STRUCTURE analysis: a "North" cluster and a "South" cluster. The South cluster includes Temora and Cootamundra and sites to the south (shown on Figure 3). Further subdivision of these groups could have been made e.g. at K = 3, the Condobolin site formed another geographically based cluster. However, the authors of the STRUCTURE software strongly suggest adopting only the minimum value of K consistent with prior knowledge of the organism (Pritchard *et al.*, 2000). Also, by using only two groups, we hoped to minimize the Type 1 error rate in detecting relationships between samples inside and outside the FFEZ.

This cline is illustrated in Figure 3, which shows the final six 2002 groupings identified by the STRUCTURE analysis. For each town, a pie graph shows the proportion of the fruit flies trapped in that town belonging to one of the six groups. The populations along the western slopes (elevation > 200m ASL) show a change from the North to the South genotype with increasing latitude. All the sites on the western plains (elevation < 200m ASL) form genetically distinct groups with no obvious relation to the cline along the western slopes. The only exceptions to this are the three towns along the Murray River (Corowa, Waygunyah and Tocumwal), which are both approximately 130m ASL.

The genetic distances between the 8 source populations are shown in Table 6. All source populations were significantly genetically differentiated. The smallest F_{ST} value was for the North-South pair (0.011). This ranking would have been predicted on the basis of the STRUCTURE results, as these two groups were the last to separate (and intergrade to some extent).

Site	N	Membership of North cluster > 0.9	Membership of South cluster > 0.9
Dubbo	26	12	2
Condobolin	42	19	0
Parkes	20	6	0
Forbes	34	21	0
Cowra	20	8	1
Grenfell	31	10	0
Young	49	10	6
Temora	20	2	3
Cootamundra	43	2	7
Wagga	57	6	23
Tumut	20	0	6
The Rock	20	0	15
Henty	20	0	14
Wodonga	42	0	12
Corowa	22	0	10
Total	466	96	99

Table 5. The number of flies showing membership ($p > 0.9$) of the North or South groups. The towns are arranged in approximate north-south order.

	North	South	Lake Carg.	Wilcannia	Hay	Deniliquin	Leeton	Barooga
North		*	*	*	*	*	*	*
South	0.011		*	*	*	*	*	*
Lake Carg.	0.042	0.047		*	*	*	*	*
Wilcannia	0.046	0.051	0.081		*	*	*	*
Hay	0.083	0.085	0.115	0.133		*	*	*
Deniliquin	0.075	0.080	0.124	0.118	0.148		*	*
Leeton	0.041	0.035	0.068	0.083	0.110	0.103		*
Barooga	0.102	0.093	0.140	0.165	0.150	0.175	0.116	

Table 6. Numbers below the diagonal are the pairwise genetic distances (F_{ST}) between the 8 source populations identified in the present study. Exact tests indicated that all were pairs were significantly genetically differentiated.

5. Spatial structuring between years

(a) Analysis of Endemic samples 2001-2003

The analysis of the 2002 samples indicated that the Endemic samples were all closely related. The close relationship between the Endemic samples was also found to persist across years, as shown by the exact tests for pairwise genetic differentiation (Table 7). In general, the Endemic samples did not significantly differ between years. Within the Queensland sites and the Sydney sites there were no significant differences between years. Only one intra-Northern NSW comparison showed a significant difference (2002 vs. 2003). Therefore, we concluded that these sites represent continuous populations persist across a number of years.

	QLD01	QLD02	QLD03	NSW01	NSW02	NSW03	Sydney01	Sydney02	Sydney03
QLD01		NS	NS	NS	NS	NS	NS	NS	*
QLD02	-.001		NS	NS	NS	NS	NS	NS	*
QLD03	.002	.004		NS	*	NS	*	NS	NS
NSW01	.002	.005	.004		NS	NS	NS	NS	NS
NSW02	-.002	-.000	.000	.002		*	*	NS	*
NSW03	.001	-.000	.008	-.000	.003		NS	NS	NS
Sydney01	.005	.004	.005	-.002	.005	.004		NS	NS
Sydney02	.007	.005	.004	.001	.001	.009	.002		NS
Sydney03	.007	.010	.005	.002	.004	.006	-.000	.001	

Table 7: Results of exact tests for genetic differentiation of the Endemic samples over 3 years. Values below the diagonal are pairwise F_{ST} values. Above the diagonal are the results of the exact tests: NS = not significant (adjusted p = 0.0014).

To test whether flies over-winter in the Intermediate region, we examined the relation between fly samples from the same town over a number of years. Data for more than one year was collected from three Intermediate sites (Condobolin, Parkes and Wagga) and three FFEZ sites (Deniliquin, Hay and Narrandera). Figure 4 shows the results of a correspondence analysis of all the populations from all years (i.e. all the sites listed in Table 1). It shows essentially the same population structuring as Figure 2, but the additional data from 2001, 2003 and 2004 has changed the numerical details. It is clear that all pairs of populations are closely related across years, suggesting that at least in those six sites, Q-fly appears able to successfully overwinter in numbers sufficient to maintain the genetic identity of the population.

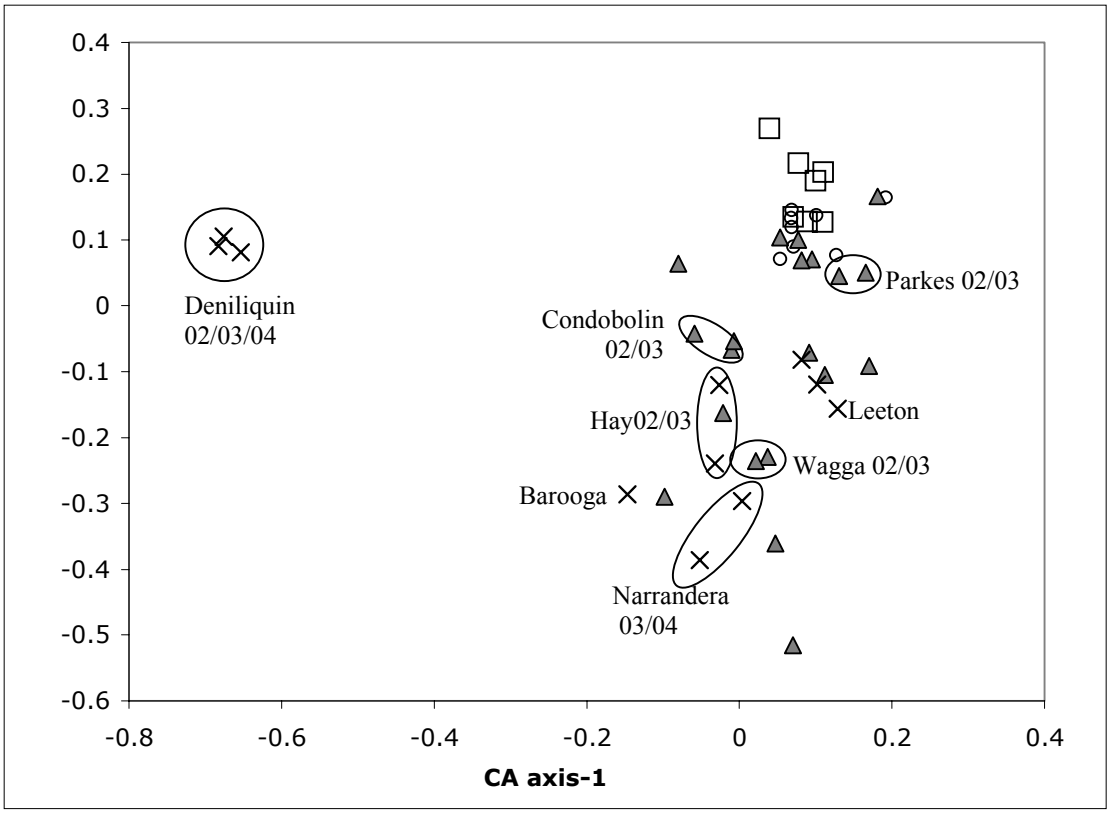


Figure 4. Correspondence analysis of the 2002-2004 samples. The circles enclose samples from the same town in different years.

(c) STRUCTURE analysis of 2003 and 2004 samples

The six population groups established for the 2002 season provided the basis for an investigation of population structuring in the following 2 seasons (i.e. 2003 and 2004). Flies from the two later years were tested to determine if they clustered with any of the six 2002 groups. This provided another test of whether flies were overwintering in any of the towns and, in particular, whether the cline along the western slopes persisted between seasons.

For the 2003 samples, STRUCTURE was run with K = 6 to 15. Posterior probabilities began to plateau at K=8, indicating that 8 clusters best explained the 2003 data. The 8 clusters consisted of the six 2002 groups as well as two new groups: one based on the Barooga sample and one based on the Leeton sample. Table 4 shows that these two new clusters were specific to each town, with most flies from that town only clustering in that town's cluster. The remaining 2003 samples from both the Intermediate region and the FFEZ all followed the pattern of genetic relations found in 2002. Specifically, the five 2003 Intermediate sites consisted of a mixture of the North and South genotypes, with the proportions being consistent with the continued existence of the north-south cline. The 2003 Narrandera sample (within the FFEZ) appeared to be a part of that same cline. The samples from Deniliquin and Hay clustered tightly with the 2002 samples from the same two towns, indicating the continuity of those populations between years.

A similar analysis was carried out for the two 2004 samples from Deniliquin and Narrandera, using the 8 clusters identified in 2002 and 2003. No additional clusters were identified and the 2004 samples clustered tightly with the samples from the same two towns from previous years (Table 4).

Figure 5 shows the geographic structuring of the 2003 and 2004 samples. It is clear that the composition of these samples is consistent with the population structuring observed in 2002 (Figures 2 and 3). Thus it appears that FFEZ sites yielding samples of 20 or more flies are likely to harbour populations of sufficient size to overwinter.

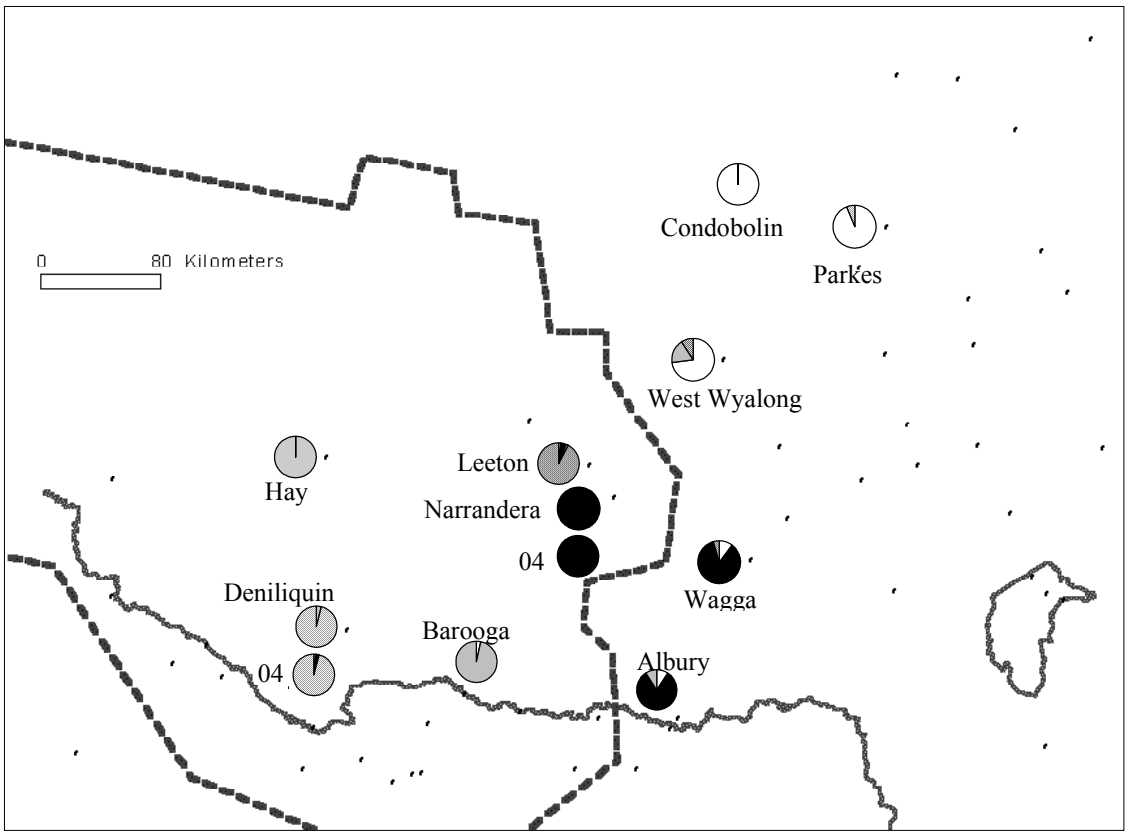


Figure 5. Genetic composition of the 2003 and 2004 samples, showing the additional two genetic clusters identified in 2003 (Leeton and Barooga). All samples were from 2003 with the exception of Narrandera and Deniliquin, for which the 2004 samples is indicated by "04". The fill patterns of each pie graph represents different genetic profiles and correspond to those used in Figure 2.

(d) STRUCTURE analysis of Outbreak samples (assignment testing)

In the preceding sections, the STRUCTURE software was used to group individuals into clusters that could be examined for correspondences between the clusters and their geographic distribution. In this section, we utilized the second function of STRUCTURE. That function measures the probability that additional individuals (of unknown origin) originate from each of the previously identified clusters (assignment testing). This function was used to address the ultimate question posed in the present study: what is the origin of immigrant flies trapped in the FFEZ? Before this can be answered, we have to distinguish between FFEZ flies originating from:

- populations already established within the FFEZ (local recruits) and;
- flies originating from populations outside the FFEZ (migrants).

To reveal the patterns of migration into the FFEZ, we need to remove from the analysis all those flies produced from populations already established within the FFEZ. The preceding analysis showed that (depending on the year) there were at least four extant populations in the FFEZ (Deniliquin, Hay, Barooga and Leeton). Therefore, these four towns were treated as sources of migrant flies, and not as possible examples of migrant flies. It is also possible that there could have been additional undetected populations in the FFEZ during this time. These were most likely to have existed in towns where numerous flies were trapped, but not analyzed in our study. Table 8 shows the trapping records for the entire FFEZ for 2002 until early 2004. That data includes many flies not analyzed in this study. Table 8 shows that, with two exceptions, all trappings of over 10 flies in any six-month period were accounted for in our results: they formed either distinct populations themselves (Deniliquin, Hay, Barooga and Leeton) or were made up of flies clustering with some combination of the 8 source populations (Corowa, Narrandera and Tocumwal; see Figures 3 and 6). These towns are shown in bold in Table 8. The exceptions were Berrigan and Finley, for which only six and two flies respectively were analyzed.

Therefore, our final sample of likely immigrants *excluded* all flies from the resident populations in Deniliquin, Hay, Barooga, Leeton, Corowa, Narrandera and Tocumwal. The only flies from these towns not excluded were two small samples from the 2004 season that were available from Hay and Leeton (4 and 11 flies respectively). Subsequent trapping results for the 2004 season (Table 8) suggest that populations in Hay and Leeton did overwinter, but the two samples analyzed were too small draw definite conclusions. Instead they were included as possible migrants and were expected to show a high probability of originating from their respective prior populations (analogous to a positive control). Such a result would be further evidence of overwintering within the FFEZ. The six flies from Berrigan and Finley were also included. In addition, samples from three small non-FFEZ towns bordering the southeast corner of the FFEZ were included (Coolamon, Ganmain and Lockhart). Accurate trapping records are not available for these three towns, but none are in horticultural production areas. In total, 94 individual flies were used in the final assignment tests, most of which were likely to be migrants into the FFEZ since they were trapped in towns least likely to have resident populations.

For each unknown individual, the assignment testing function of STRUCTURE calculates the probability that the individual originates from each of the possible source populations. Across all possible source populations, these probabilities sum to one. It was necessary to select a level of probability at which an individual would be classed as having originating from the most likely source. For our results, we chose $p = 0.9$ for a single source, which means that the probability of the individual originating from all other sources combined < 0.1. Results using a far less stringent

criterion of p = 0.5 were also calculated but were not qualitatively different from the p = 0.9 results (results not shown).

The final results are shown in Figure 6, where the distribution of flies coming from each possible source population is shown on a separate map. There is an evident geographic correspondence between each source and the flies originating from that source. For example, flies from South grouping were found mainly in the southeastern corner of the FFEZ. Similarly flies from Deniliquin were found only in towns to the west of the Riverina. Wilcannia was not expected to be a major source of migrants into the FFEZ and (as expected) produced no identifiable migrants.

(e) Frequency-based analysis of FFEZ samples

For the frequency-based analysis of the FFEZ outbreak samples, we used the eight population clusters identified using the STRUCTURE software as the reference populations. The limitations of this method can be seen in Table 9, which shows the results of the assignment of all the individuals in the reference populations back to those reference populations (self-assignment). A common method of evaluating statistical support for this type of data is to calculate LOD scores: i.e. the $\log_{10}$ of the ratio of largest and second largest probabilities that the individual originates from each source. A score of 1 indicates that the first assigned source is 10 times more likely than the second most likely source.

The results shown in Table 9 were calculated using LOD > 1 as the criterion for assignment. Success of the method was lowest for the North and South groups due to their genetic similarity, which in effect renders them almost equally likely as source populations. However, success rates approaching 90% were evident for the more genetically diverged populations. This suggests that frequency-based testing would only be useful for testing migrants from the more diverged populations and not the North and South groups.

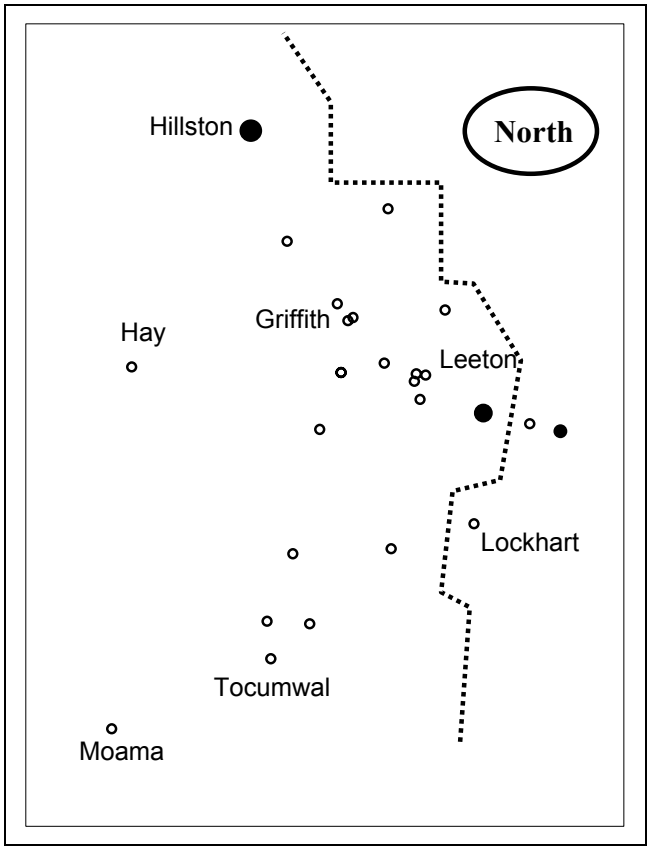
In Figure 7, the right-hand graphs show the distribution of migrants from each of the source populations, excluding the North and South cluster. These graphs show similar distributions of migrants to the STRUCTURE results.

		Assignment result									
		North	South	Lake Carg	Wilcan.	Deni.	Hay	Leeton	Barooga	Unassigned	Total
Origin site	North	38.3	5.4		1.4	0.5		0.4		54.1	222
	South	2.1	34.9	0.5		0.8	0.8	1.6		59.4	384
	Lake Cargelligo	4.1		82.4		1.4			1.4	10.8	74
	Wilcannia	2.3			86.0					11.6	43
	Deniliquin	0.6				89.2				10.2	166
	Hay						85.2			14.8	54
	Leeton	1.2				1.2		70.7		26.8	82
	Barooga								90.9	9.1	33

Table 9. Percentage of flies from each of the 8 source populations that were assigned to each source population using a frequency-based assignment method. Numbers on the diagonal indicate assignment to the "correct" source population. Correct assignments for the North and South groupings were very low.

Season	2002	2002-3		2003-4				
Town	Jan-Jun	July-Dec	Jan-Jun	July-Dec	Jan-Jun	Traps in Use	Max. flies/trap	Flies analyzed
Barellan	2	3				2	1.5	3
Barooga	1	2	45	1	6	11	4.1	
Berrigan	4		42	25	73	13	5.6	6
Cobram	1					1	1.0	
Coleambally			1		1	7	0.1	1
Conargo				1		2	0.5	
Corowa	991	26	1		12	8	123.9	
Darlington Point	1	2	6	3	4	16	0.4	11
Deniliquin	286	126	53	637	4971	43	115.6	
Finley		2	16	5	8	16	1.0	2
Goolgowi	5		1			4	1.3	8
Griffith	1	3	1	1	4	77	0.1	1
Grong Grong	1	3	1		1	2	1.5	3
Hanwood	1	1			6	40	0.2	1
Hay	169	49	12	1	98	21	8.0	4
Hillston	1	1	3	6	1	36	0.2	6
Jerilderie		1	8	3	12	11	1.1	2
Kamarah					1	1	1.0	
Lake Wyangan	1				4	17	0.2	
Leeton Combined	8	43	30	23	27	103	0.4	11
Leeton	4	39	21	14	16	40	1.0	6
Corbie Hill	2	1				10	0.2	1
Cudgell	1				1	5	0.2	1
Merungle Hill		1	1	2	1	6	0.3	
Stanbridge	1				2	9	0.2	1
Stoney Point			2			4	0.5	
Wamoon		1	2	1	2	11	0.2	
Whitton					1	5	0.2	
Yanco		1	4	6	4	13	0.5	2
Mathoura					54	9	6.0	
Moama			2	5	3	18	0.3	1
Narrandera	1	10	9	39	72	33	2.2	
Nericon					2	2	1.0	
Oaklands	1					2	0.5	
Paynter's Siding		1	1			1	1.0	
Rand	4	1	1			1	4.0	
Rankin Springs			1			1		1
Tabbita			2	1		3	0.7	
Tharbogang	2		1		5	27	0.2	2
Tocumwal	44	8		1	46	18	2.6	
Warburn			1			1	1.0	
Yenda	3		3		3	35	0.1	

Table 8. Six-monthly trap catches for all sites in the FFEZ from January 2002 until March 2004. The column "Max. flies/trap" indicate the ratio of the largest six-month trapping to the number of traps in place. The final column indicates how many flies from that site were analysed as FFEZ outbreak flies (samples from outside the FFEZ from Lockhart, Coolamon and Ganmain are not shown).



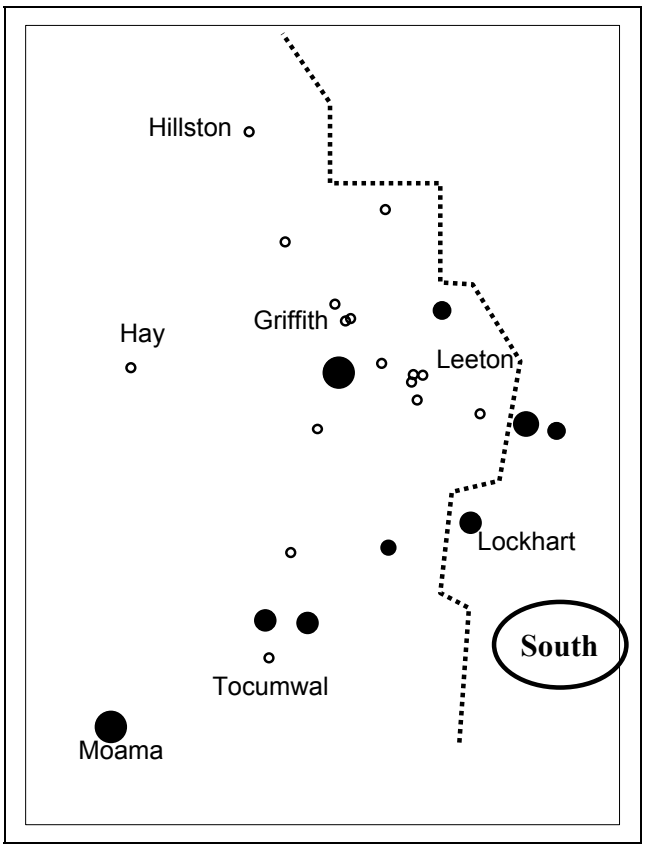
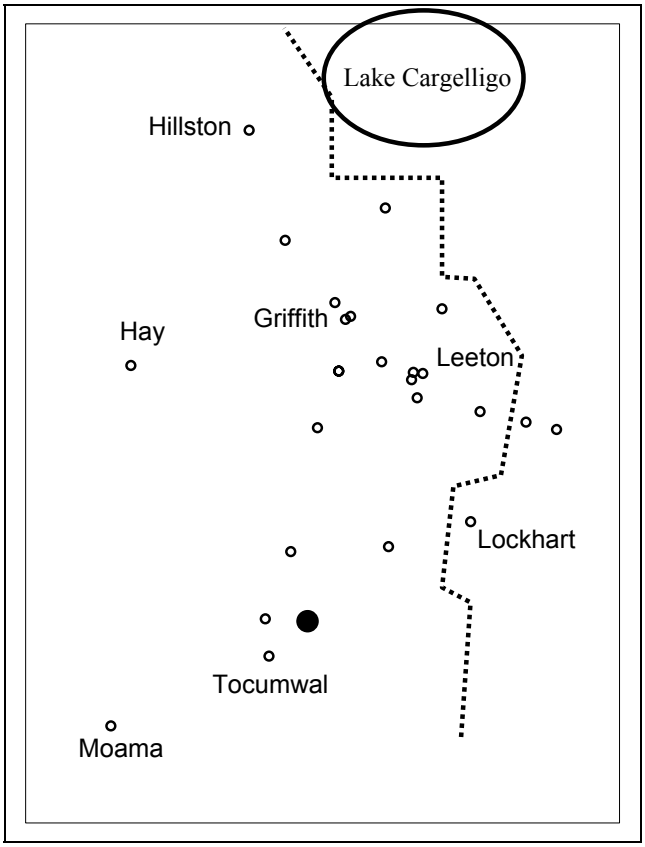
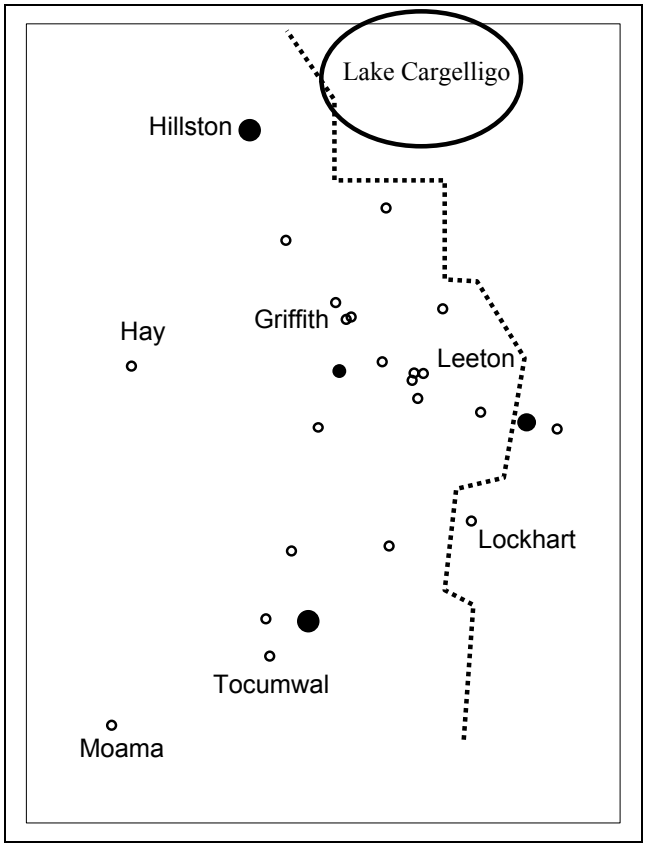


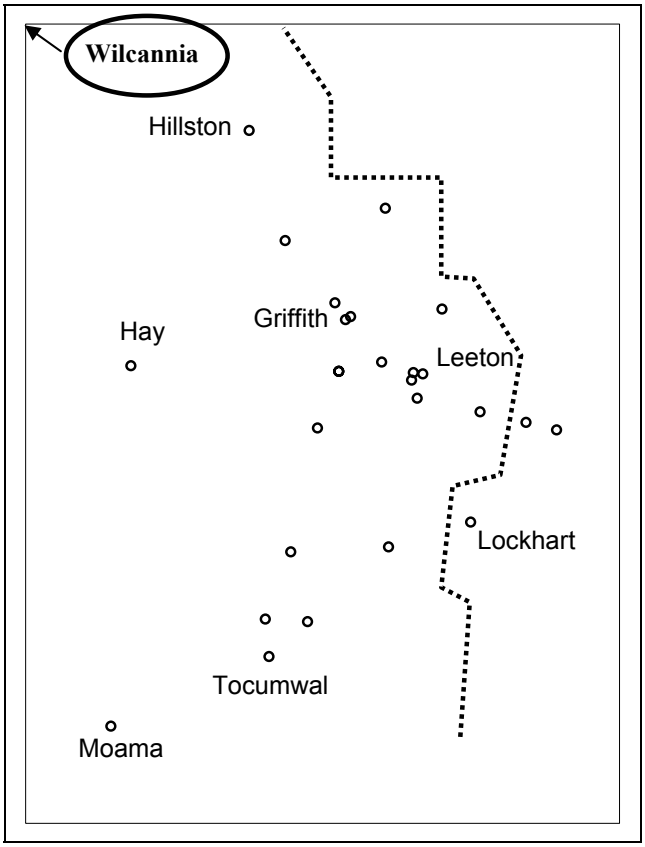
Figure 6. In the following 12 maps of the FFEZ, the black dots show where migrant flies from each source population (circled) were trapped in or near the FFEZ. Empty circles indicate a trapping site at which flies were trapped but none showed a high probability of originating from the particular source. The maps on the left-hand side of each page show the STRUCTURE results, while the right-hand maps show the results of the frequency-based calculations for the same source (not shown for the North or South groups; see text for discussion).

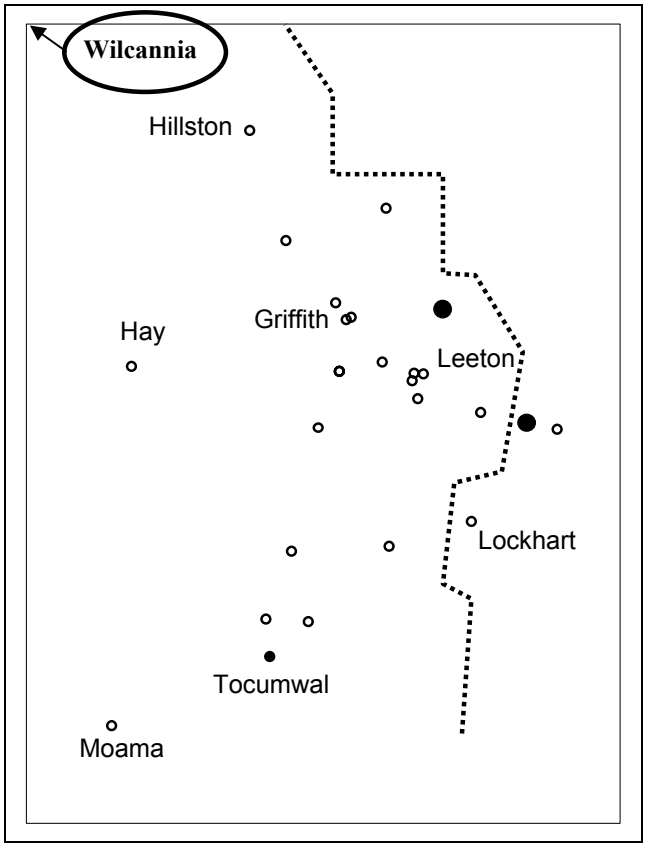
The size of each black circle is reflects the proportion of flies trapped at that site that came from a particular source: i.e. a larger circle means that a greater proportion of flies trapped came from a given source.

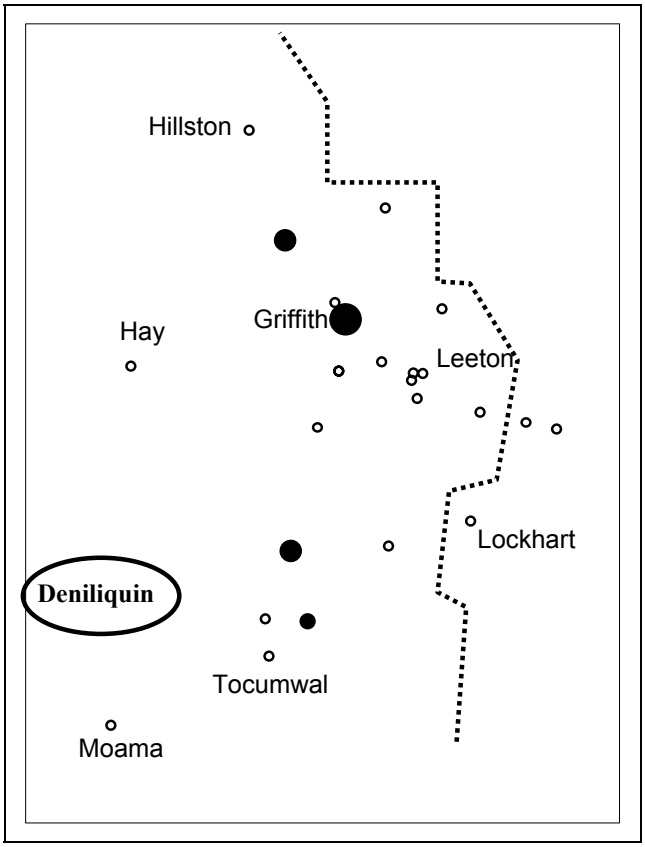
The dotted line indicates the approximate FFEZ border

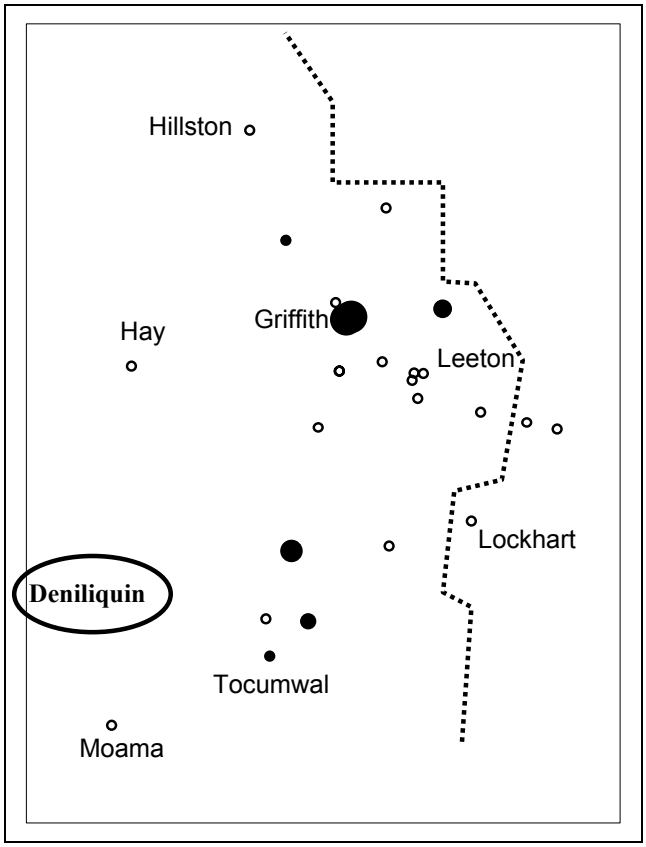

Lake Cargelligo
Hillston
Griffith
Hay
Leeton
Lockhart
Tocumwal
Moama

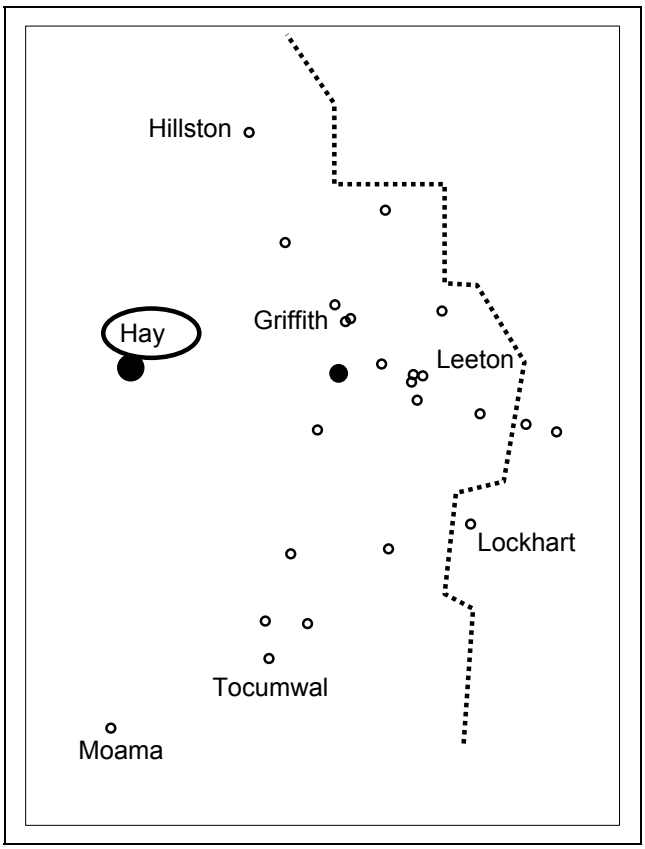

Lake Cargelligo
Hillston
Griffith
Hay
Leeton
Lockhart
Tocumwal
Moama

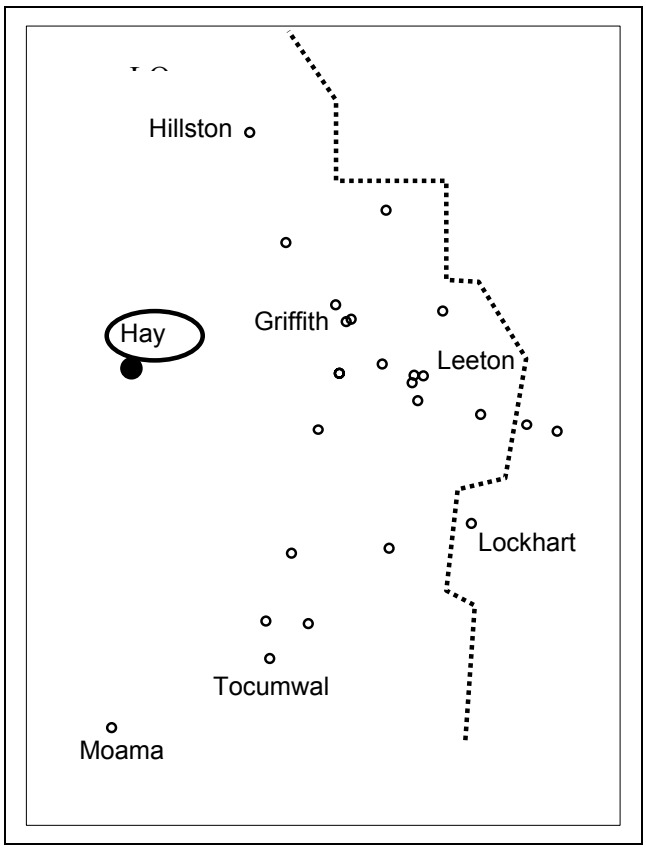

Wilcannia
Hillston
Griffith
Hay
Leeton
Lockhart
Tocumwal
Moama

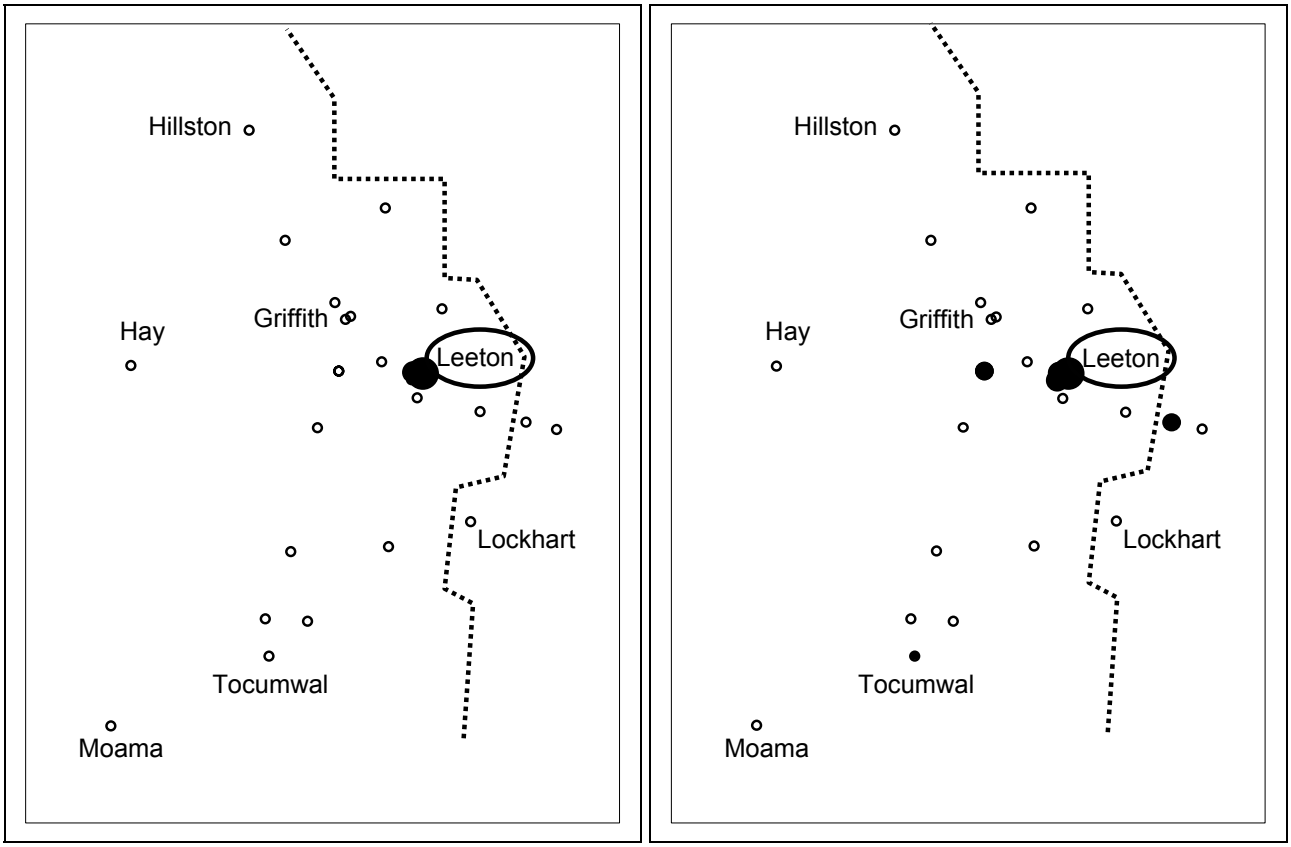

Wilcannia
Hillston
Griffith
Hay
Leeton
Lockhart
Tocumwal
Moama


Hillston
Griffith
Hay
Leeton
Lockhart
Deniliquin
Tocumwal
Moama


Hillston
Griffith
Hay
Leeton
Lockhart
Deniliquin
Tocumwal
Moama


Hillston
Griffith
Hay
Leeton
Lockhart
Tocumwal
Moama


Hillston
Griffith
Hay
Leeton
Lockhart
Tocumwal
Moama


Hillston
Griffith
Hay
Leeton
Lockhart
Tocumwal
Moama
Hillston
Griffith
Hay
Leeton
Lockhart
Tocumwal
Moama

(f) Analysis of 2002 and 2004 Adelaide samples

In 2002 and 2004 samples of Q-fly (22 and 10 flies respectively) were trapped in Adelaide (Appendix 1). None of the flies tested showed any similarity to the Factory strain. Therefore, it can be confidently concluded that none were factory flies.

Similarly to the analysis of FFEZ flies, the assignment testing function of Structure was used to determine the origins of the flies (including a reference sample from Alice Springs). In 15 of 32 cases, flies showed high probability (> 0.9) of membership of the North cluster. 25 of the 32 showed a combined probability > 0.9 of coming from either the North or South groupings. Only one fly of the 32 showed a high probability of being associated with any other cluster. Fly number 12 (Appendix 1) trapped in April 2002 gave a probability > 0.9 of being a member of the Leeton 2003 group. This result must be coincidental since there were relatively few flies trapped in Leeton in 2002 and they were unlikely to be a source of flies for Adelaide outbreaks. On that basis, the results indicate that the Adelaide outbreak flies originate either from the Endemic east coast populations or from an unknown source population.

6. Estimates of effective population sizes

Effective population size (N_e) estimates were calculated for the populations shown in Table 10. It should be stressed that N_e estimates are usually much lower than census population estimates. N_e estimates reflect the average minimum breeding population size (most likely reached at the end of winter). The estimates were limited to these populations for which data from at least two time-points was available, since N_e was calculated from temporal variation in microsatellite frequency. Traditional estimates of N_e ignore migration. However, since migration appears to be significant in the Intermediate regions, N_e was also estimated simultaneously with migration from a source population. For the Intermediates, the source used was the Dubbo sample. For two of the FFEZ populations (Deniliquin and Hay), no migration estimates were made since these appeared to be distinct populations. For the Narrandera sample, Wagga was used as the source since these two populations consisted largely of South group flies.

The results show that N_e estimates in the Intermediate region are much lower than those in the Endemic region, as the relative degree genetic differentiation in those two regions would suggest. Migration from likely source populations lowers these N_e estimates since migration from another population may account for some of the temporal variation in microsatellite frequencies.

Region/Town	N_e (no migration)	Ne (migration)	Migration rate
Qld	**8110** (1102; >9000)		
Northern NSW	**8942** (1130; >9000)		
Sydney	**966** (470; 8031)		
Parkes	**195** (95; 1220)	**47** (29; 87)	**0.099** (0.041; 0.178)
Condobolin	**389** (92; > 9000)	**39** (23; 94)	**0.074** (0.017; 0.155)
Wagga	**1220** (473; > 9000)	**285** (158; 589)	**0.026** (0.011; 0.050)
Narrandera	**124** (59; 1248)	**18** (12; 28)	**0.281** (0.161; 0.524)
Deniliquin	**217** (146; 359)		
Hay	**96** (54; 230)		

Table 10. Estimates of effective population sizes (N_e; shown in bold) for the towns indicated. Where N_e was jointly estimated with the migration rate, both Ne and migration rate are shown. Numbers in parentheses are 95% confidence limits.

Discussion

1. Genetic structuring Q-fly populations

The first aim of this study was to delineate populations of Q-fly in eastern Australia at a resolution that was appropriate for assignment testing. Our results, using the current set of microsatellites, can be summarized as follows.

Firstly, the area between Sydney and Brisbane and inland to Dubbo, Forbes and Cowra forms one indistinguishable source population (the North group). Such homogeneity, at a scale far exceeding the natural dispersal capability of Q-fly, must be due to the effects of both large population size and significant rates of assisted migration. Our results appear to contradict the result of Yu *et al.* (2001), who found that the Queensland, Northern NSW and Sydney samples formed genetically distinct groups. However, in that study population differences were tested using chi-squared contingency table analysis. That method is the most straightforward and sensitive way of detecting genetic differences between a given set of populations. However, the aim of this project was to map distinguishable source populations for the FFEZ outbreak flies. For this purpose, the level of "detail" among populations uncovered by chi-square analysis is too fine. Assignment methods cannot reliably distinguish between the different Endemic populations. For present purposes, chi-square results are not strictly relevant and to avoid confusion, we did not report chi-square results, although we note that they were entirely consistent with the results of Yu *et al.* (2001).

Secondly, moving southwards along the Western Slopes of NSW, we observed the gradual emergence population structure in Q-fly, with the second cluster (the South group) becoming more common in the south. This structuring is most likely due to smaller population sizes in towns along the Western Slopes (Table 10) and some reduction in effective migration rates in comparison to the North group. All populations along the Western Slopes south of Dubbo consist of some mixture of the North and South groups, with the proportion of South flies increasing as latitude increases (i.e. a clinal pattern). Clinal patterns are usually formed where short-range migration predominates over long-range dispersal (here short-range means between neighbouring towns), since a preponderance of long-distance dispersal would obliterate the clinal pattern.

Thirdly, for the populations on the Western Plains, there was a clear distinction between the plains towns and the Murray River towns. The plains towns (Deniliquin, Hay, Leeton, Wilcannia and Lake Cargelligo) showed high genetic distances from the eastern populations (in the range F_{ST} = 0.05 – 0.1). Since the FFEZ populations are all relatively young, that genetic differentiation must result mainly from population bottlenecks rather than gradual divergence of a larger population. The significant reduction in allelic richness of the Western Plains populations was strong evidence that they had suffered bottlenecks events. These bottlenecks have two effects. First, they make the genetic profile of the western plains towns relatively distinct. This means that assignment tests are quite successful in identifying individuals originating from these towns. Secondly, it obscures the origins of the original colonizers of those towns. However, since we have numerous other sites (with transient outbreaks) to assess migration into the FFEZ, the obscurity of the origins of Deniliquin, Hay and Leeton was not of practical importance. For the towns on the Murray River (Corowa, Wahgunyah and Tocumwal) genetic distances from the eastern populations was lower (in the range F_{ST} = 0.02 – 0.04). These towns all contained individuals likely to have originated from the South group (Figure 3 and 5). Since that identifiable genetic profile survived in these populations, it suggests that either large founder populations or sufficiently high continuing

migration along the Murray River. Once established, it appears that these populations all overwinter successfully in the FFEZ (Figure 4).

2. Were there other unknown source populations?

Any test of the origins of individual fruit fly samples depends on having reference samples of all possible source populations. This project was undertaken in order to establish as complete a sample as possible. As discussed in the Introduction, it seems unlikely that there are other major source populations for flies entering the FFEZ that we failed to sample. Climate to the north, west and south dramatically limits population growth compared to the east coast.

However, it is entirely possible that small populations could contribute some migrants to the FFEZ. In the absence of a sample of the source population, outbreak samples would most likely be assigned, almost by default, to the Endemic regions. This result arises from the fact that the Endemic regions contain much larger permanent populations of Q-fly, which must be the original source of all Q-fly that have spread to the south and west. Therefore, the Endemic populations contain all the variation present in the derived inland populations, including any unknown source populations. This means that if an FFEZ outbreak fly cannot be definitely matched to one of the known inland source populations, the Endemic populations are the next most likely source. The most likely sample for which this could have occurred were the Adelaide flies. While our results suggest that the Adelaide flies all came from the Endemic region, it is worth noting that Adelaide is the furthest sample from any of the identified source. Accordingly, the possibility that there is some unidentified source (e.g. along the Murray River, west of Tocumwal) cannot be excluded. Nevertheless, since it seems fly populations in inland towns are related to human population size (Maelzer *et al.*, 2004), any such unidentified source is likely to be small and therefore contribute few migrants to the FFEZ.

3. Would assignment be successful?

The ability to assign individuals to their correct source populations has been extensively simulated in various studies e.g. Cornuet *et al*. (1999) and Paetkau *et al*. (2004). Success depends on the following factors:

- the numbers of microsatellites;
- the number of individuals typed from each of the source populations; and
- how genetically distinct the source populations are from one another.

100% correct assignment can usually be expected when at least 10 microsatellites are used, 30-50 individuals are typed from each source population and F_{ST} (the measure of genetic differentiation) is near 0.1. For our data, the limiting factor was the degree of genetic differentiation between the source populations (Table 6). Only Hay, Deniliquin and Barooga show pairwise F_{ST} values around 0.1. This meant that migrants from these populations could be identified with confidence. For the North and South groups however, F_{ST} is only 0.011, suggesting that would be harder to correctly distinguish between the North or South group when a migrant was most likely to have come from one of those two groups. This is the likely reason that the STRUCTURE software could only assign about 60% of the FFEZ migrants to a source population with $p > 0.9$. Our successes were probably due to our very large sample sizes and 21 microsatellite loci, which to some extent overcame the lack of genetic differentiation. The frequency-based methods are usually less sensitive than Baysian methods (i.e. as in STRUCTURE; Cornuet *et al*. 1999) and this was borne out in our results: the frequency-based method of assignment could not reliably differentiate the North and South groups.

4. Origins of FFEZ Outbreak flies

To gain an accurate reflection of the patterns of movement of Q-fly into the FFEZ, it was necessary to first distinguish migrants from resident flies and then secondly, to deduce the actual patterns of movement of those migrants. These processes are discussed in the following sections.

(a) Distinguishing migrants from residents

Firstly, we had to distinguish actual migrants (or flies moving between towns in the FFEZ) from flies that are the offspring of established resident populations. If this were not done, the results would be highly biased by the larger groups of residents. As an extreme example, in early 2004 the vast majority of flies trapped in the FFEZ originated from the Deniliquin population. Were all these flies assumed to be immigrants, we would have to hypothesize a huge rate of migration into the FFEZ, and Deniliquin in particular, from an unknown Deniliquin-like population outside the FFEZ. This example is extreme, but more similar subtle biases could result from small, undetected resident populations. Were all the possible source populations external to the FFEZ, this problem would not exist.

Using clustering software, we were able to detect several established populations within the FFEZ. We have assumed that where the majority of the flies from the same vicinity were genetically very similar, it is extremely unlikely that they could all be immigrants. Instead, it is more likely that they were offspring of a single population already in that town. The genetic continuity between years of Deniliquin, Hay and Leeton strongly supported that assumption. As a rule of thumb, we found that where average trap catches in larger towns exceeded 1 fly per trap over a 6-month period (Table 8), then that sample of flies formed a single population grouping (Berrigan and Finley may have represented exceptions to this rule, but we had insufficient samples to draw any firm conclusions about those two towns). These groupings were either linked to neighbouring populations in the Risk Reduction zone through continuing migration or had a more distinct DNA profile (e.g. Deniliquin, Hay and Barooga). The significance of this is that it suggests that uncontrolled outbreaks can rapidly develop into on-going populations. Coupled with the ability of Q-fly populations to overwinter throughout the FFEZ between 2002 and 2004, we conclude that any town in which large numbers of flies are trapped has the potential to develop into a source of outbreak flies.

It is possible that smaller trappings might also represent resident populations; e.g. in Goolgowi. This is especially so because in the clustering methods used here, small groups of flies (~5) will add little to the posterior probabilities of clustering solutions and will thus be overlooked. However, from trapping records we know that (apart from Berrigan and to a lesser extent, Finley) we have accounted for all the large populations in the FFEZ. If any of the remaining trappings shown in Table 8 represented additional source populations they could only contribute small numbers of migrants, especially in comparison to the large populations outside the FFEZ.

(b) Patterns of immigration into the FFEZ

Among the flies likely to be migrants rather than residents, actual migrants would be expected to show a high similarity to one or other of the source populations since both parents would have been from that same source population. However, similarity of individuals to any source could be reduced by chance variation or their being first generation offspring of migrants. In both cases, only a reduced similarity to a source would be apparent. We adopted a conservative approach of only identifying as migrants those individuals with a probability > 0.9 of originating from a particular source population (Figure 6).

The major conclusion of these results is that short-range migration appears to be the most common mode of spread of Q-fly in the FFEZ (as it is on the Western Slopes). Figures 2, 5 and 6 shows that for each of the source populations, migrants were found only in the closest towns. Here, "short-range" refers to typical distances between one town and its neighbours (i.e. ~100-200km) and is used in the context of human-assisted rather than natural dispersal of Q-fly. The underlying reason for this is likely to be that the number of short distance vehicle journeys between towns will greatly outnumber the number of longer vehicle journeys from the Endemic regions. It is possible that some movement could be due to natural dispersal but its contribution is likely to be minimal (e.g. Fletcher, 1974; MacFarlane *et al.*, 1987).

The main features of the pattern of migration into FFEZ were as follows.

- The South group is a major source of migrants into the eastern FFEZ. These migrants were found not only in horticultural production areas (e.g. the Riverina), but also in most towns in the southeastern corner of the FFEZ. In numerous towns the only identifiable migrants came from the South group, e.g. Narrandera and the small towns immediately bordering the FFEZ (Coolamon, Ganmain and Lockhart). Previously, Q-fly control has been centred on towns to the north of Wagga (Edge, 2001), which has presumably been responsible for the relatively low rates of infestation in the northeast of the FFEZ. However, our results suggest that Wagga, Albury and Deniliquin should receive equal treatment.
- In contrast, the North group contributed relatively few flies to the FFEZ. This could be due to control efforts or to greater road distances from the North group to major towns in the FFEZ.
- Deniliquin appears to be a significant source of migrants within the FFEZ. Among the larger samples, Deniliquin flies were identified only in Tocumwal ~70km southeast (Figure 3). Figure 6 shows that among the outbreaks, Deniliquin flies in only the western Riverina and nowhere outside the FFEZ.
- The South group was present in most towns along the Murray River, as far west as Moama, presumably due to high migration rates along the river. It is possible that this reflects high volumes of vehicle traffic along the Murray River, rather than any particular effect of the river itself.

Technology Transfer

There was no technology transfer involved in this project

Recommendations

The recommendations from this study are as follows:

1. Q-fly outbreaks in the FFEZ between 2002 and early 2004 originated mainly in the Wagga-Albury region and Deniliquin. Therefore, control in these centres, and any other centres with large Q-fly populations, appears vital in reducing the incidence of outbreaks.
2. Deniliquin is an example of an outbreak that has developed into a source population (Leeton in 2003 is another example). Any large simultaneous trappings of Q-fly within a town or locality should be assumed to represent an established local population.
3. Our general finding was that movement of Q-fly appears most likely to occur between neighbouring towns. Future research should quantitatively investigate the relation between local traffic levels and Q-fly spread.

Acknowledgements

We would like to thank our collaborative investigators, their support staff and the numerous individual fruit fly collectors who made this work possible. Sasha Curthoys was especially efficient in classifying and cataloguing the samples. XuiMei Liang, Van Pam Bich and Alison Ling provided invaluable technical assistance.

References

Anon. (1996) Code of Practice for Management of Queensland fruit fly. Standing Committee on Agriculture and Resource Management, Department of Primary Industries, Canberra.

Bateman MA (1967) Adaptations to temperature in geographic races of the Queensland fruit fly, Dacus (Strumeta) tryoni. *Australian Journal of Zoology* **15**, 1141-1161.

Canon J, Alexandrino P, Bessa I*, et al.* (2001) Genetic diversity measures of local European beef cattle breeds for conservation purposes. *Genetics Selection Evolution* **33**, 311-332.

Cegelski CC, Waits LP, Anderson NJ (2003) Assessing population structure and gene flow in Montana wolverines (Gulo gulo) using assignment-based approaches. *Mol Ecol* **12**, 2907-2918.

Cornuet JM, Piry S, Luikart G, Estoup A, Solignac M (1999) New methods employing multilocus genotypes to select or exclude populations as origins of individuals. *Genetics* **153**, 1989-2000.

Cruciani F, Santolamazza P, Shen PD*, et al.* (2002) A back migration from Asia to sub-Saharan Africa is supported by high-resolution analysis of human Y-chromosome haplotypes. *American Journal of Human Genetics* **70**, 1197-1214.

Dominiak BC, Rafferty TD, Barchia IM (2000) A survey of travellers carrying host fruit of Queensland Fruit fly, Bactrocera tryoni (Froggatt), into a fruit fly free area in 1996/97. *General and applied Entomology* **29**, 39-44.

Edge V (2001) Technical Review of the Tri-Satate Stratergy for Queensland Fruit Fly. Standing committee on Agriculture and Resource Management.

Falush D, Stephens M, Pritchard JK (2003) Inference of population strucutre using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics* **164**.

Fletcher BS (1974) The ecology of a natural population of the Queensland fruit fly, Dacus tryoni. V The dispersal of adults. *Australian Journal of Zoology* **22**, 189-202.

Froggatt WW (1909) Report on parasitic and injurious insects 1907-1908. NSW Department of Agriculture, Sydney.

Gauch HG (1982) *Multivariate analysis in community ecology* Cambridge University Press, Cambridge.

Gilchrist AS, Sved JA, Meats A (2004) Genetic relations between outbreaks of the Queensland fruit fly, *Bactrocera tryoni* (Froggatt) (Diptera: Tephritidae), in Adelaide in 2000 and 2002. *Aust J Entomol* **43**, 157-163.

Goudet J (2002) FSTAT: a program to estimate and test gene diversities and fixation indices.

Hurlbert S (1971) The non-concept of species diversity: A critique and alternative parameters. *Ecology* **52**, 577-586.

SAS Institute (1994) JMP statistics made visual. SAS Institute Inc., Cary, N.C.

Kinnear MW, Bariana HS, Sved JA, Frommer M (1998) Polymorphic microsatellite markers for population analysis of a tephritid pest species, *Bactrocera tryoni*. *Molecular Ecology* **7**, 1489-1495.

Leberg PL (2002) Estimating allelic richness: Effects of sample size and bottlenecks. *Molecular Ecology* **11**, 2445-2449.

Lewis PO, Zaykin D (2002) GDA: Software for the Analysis of Discrete Genetic Data, Version 1.1.

Lewontin RC, Birch LC (1966) Hybridization as a source of variation for adaptation to new environments. *Evolution* **20**, 315-336.

MacFarlane JR, East RW, Drew RAI, Betlinski GA (1987) Dispersal of irradiated Queensland fruit flies, *Dacus tryoni* (Froggatt) (Diptera, Tephritidae) in southeastern Australia. *Australian Journal of Zoology* **35**, 275-281.

Maelzer DA, Bailey PT, Perepelicia N (2004) Factors supporting the non-persistences of fruit fly populations in South Australia. *Australian Journal of Experimental Agriculture* **44**, 109-126.

Mavi HS, Dominiak BC, Nicol HI (in prep) Modification of the growth potential of Queensland fruit fly *Bactrocera tryoni* (Froggatt) (Diptera: Tephritidae) by urban landscape irrigation in inland New South Wales, Australia.

McKechnie SW (1974) Allozyme variation in the fruit flies *Dacus tryoni* and *D. neohumeralis*. *Biochemical Genetics* **11**, 337-346.

Meats A (1981) The bioclimatic potential of the Queensland fruit fly, *Dacus* (*Bactrocera) tryoni*, in Australia. *Proceedings of the Ecological Society of Australia* **11**, 151-161.

Meats A (1998) Predicting or intrepreting trap catches resulting from natural propagules or releases of sterile fruit flies. An actuarial and dispersal model tested with data on *Bactrocera tryoni*. *General and applied entomology* **28**, 29-38.

Meats A, Clift AD (2003) Analysis of movement of Q-fly, particularly related to the use of sterile insect technique (SIT) in eradication programs. Horticulture Australia Ltd.

Paetkau D, Calvert L, Stirling I, Strobeck C (1995) Microsatellite analysis of population structure in Canadian polar bears. *Molecular Ecology* **4**, 347-354.

Paetkau D, Slade R, Burden M, Estoup A (2004) Genetic assignment methods for the direct, real-time estimation of migration rate: a simulation-based exploration of accuracy and power. *Molecular Ecology* **13**, 55-65.

Piry S, Alapetite A, Cornuet JM*, et al.* (in press) GeneClass2: a software for gentic assignment and first generation migrants detection. *Journal of Heredity*.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of Population structure using multilocus genotype data. *Genetics* **155**, 945-959.

Rosenberg NA, Burke T, Elo K*, et al.* (2001) Empirical evaluation of genetic clustering methods using multilocus genotypes from 20 chicken breeds. *Genetics* **159**, 699-713.

Schneider S, Roessli D, Excoffier L (2000) Arlequin version 2.000: A software for population genetics data analysis. Genetics and Biometry Laboratory, University of Geneva, Switzerland.

She JX, Autem M, Kotulas G, Pasteur N, Bonhomme F (1987) Multivariate analysis of genetic exchanges between *Solea aegyptiaca* and *Solea senegalensis* (Teleosts, Soleidae). *Biological Journal of the Linnean Society* **32**, 357-371.

Sokal RR, Rohlf FJ (1995) *Biometry*, 3rd edn. W. H. Freeman and Company, New York.

Sutherst RW, Collyer BS, Yonow T (2000) The vulnerability of Australian horticulture to the Queensland fruit fly, *Bactrocera* (*Dacus*) *tryoni*, under climate change. *Australian Journal of Agricultural Research* **51**, 467-480.

Sved J, Yu H, Gilchrist AS, Dominiak BC (2003) Inferring modes of colonization for pest species using heterozygosity comparisons and a shared allele test. *Genetics* **163**, 823-831.

van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) Micro-checker software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes*.

Wang JL, Whitlock MC (2003) Estimating effective population size and migration rates from genetic samples over space and time. *Genetics* **163**, 429-446.

Wang Y, Yu H, Raphael K, Gilchrist AS (2003) Genetic delineation of sibling species of the pest fruit fly *Bactocera* (Diptera: Tephritidae) in Australia using microsatellites. *Bulletin of Entomological Research* **93**, 351-360.

Yonow T, Sutherst RW (1998) The geographical distribution of the Queensland fruit fly, Bactrocera (Dacus) tryoni, in relation to climate. *Australian Journal of Agricultural Research* **49**, 935-953.

Yu H, Frommer M, Robson M*, et al.* (2001) Microsatellite analysis of the Queensland fruit fly *Bactrocera tryoni* (Diptera: Tephritidae) indicates spatial structuring: implications for population control. *Bulletin of Entomological Research* **91**, 139-147.

Appendix 1

2002 and 2004 Adelaide flies.

Fly	Trap Site	Trap Date
1	Hawthorn	7/3/02
2	Thebarton	13/3/02
3	Magill	15/3/02
4	Magill	16/3/02
5	Thebarton	18/3/02
6	Oaklands Park	19/3/02
7	Warradale	19/3/02
8	Warradale	19/3/02
9	Thebarton	25/3/02
10	Magill	28/3/02
11	Thebarton	28/3/02
12	Thebarton	1/4/02
13	Rosslyn Park	16/4/02
14	Magill	2/5/02
15	Lockleys	3/5/02
16	Magill	6/5/02
17	Warradale	10/5/02
18	Warradale	10/5/02
19	Lockleys	10/5/02
20	Lockleys	14/5/02
21	Warradale	14/6/02
22	Warradale	14/6/02
2004 season		
23	McLaren Flat	28/11/03
24	McLaren Flat	“
25	McLaren Flat	“
26	Hazelwood Park	2/12/03
27	Tusmore	16/12/03
28	Norton Summit	23/12/03
29	Hindmarsh	24/12/03
30	Kensington Pk.	30/12/03
31	Clapham	22/1/04
32	Pooraka	28/1/04