Genetic variation in the aphid *Pemphigus spyrothecae*

Paul C. D. Johnson

Darwin College

A dissertation submitted for the degree of

Doctor of Philosophy at the University of Cambridge

October 2000

PREFACE

I certify that no part of this dissertation has been or is currently being submitted for a degree or diploma or any other qualification at any other university. This dissertation does not exceed the proscribed maximum length of 80,000 words, and is the result of my own work. Any assistance received by other workers is acknowledged specifically in the text.

ACKNOWLEDGEMENTS

A big 'thank you' to:

- My supervisors, Bill Amos and William Foster, for dreaming up this project, and for much-needed enthusiasm, encouragement and guidance
- Everyone who has worked in the Amos lab over the past 4 years
- The Zoology Department, for giving me the opportunity to undertake this research, and NERC for funding it
- Everyone who helped me in the field, particularly Theresa Burg and her family for collecting Canadian samples, and Christian Brändle and his helpers for the samples from Italy, Germany, Austria and Sweden
- All the people in the department who took the time to answer questions and discuss ideas, especially Oscar Gaggiotti, David Stern, Rufus Johnstone, and everyone in the Amos and Foster groups
- Kate Llewellyn, Nick Miller, Nathan Pike and Richard Harrington, for acceding to my regular demands for their data
- Tanita, for collecting aphids, correcting manuscripts, and being patient
- My mother and father, who helped in many ways, and who have also been patient

SUMMARY

Genetic variation in the aphid *Pemphigus spyrothecae*

This project used genetic variation to investigate dispersal, inbreeding and social behaviour in the aphid *Pemphigus spyrothecae*. *P. spyrothecae* is cyclically parthenogenetic, reproducing sexually on the bark of its primary host, *Populus nigra*, and asexually within galls on the leaf petioles. Within the gall, a soldier caste defends and cleans the gall, potentially reducing its own fitness.

P. spyrothecae collected across the UK from 1997 to 1999 and from mainland Europe and America in 1999 were genotyped using seven variable microsatellite markers that I developed in collaboration with William Amos and Kate Llewellyn. Using population genetic analysis, I showed that *P. spyrothecae* populations were temporally stable over three years, and spatially structured. Populations from trees 5 to 1700 km apart were significantly differentiated, and loosely followed an isolation-by-distance model. There was slight evidence of differentiation between neighbouring trees (5 to 500 m apart), but not between samples taken from within trees (less than 5 m apart). By contrast, *P. bursarius*, a closely related species that, unlike *P. spyrothecae*, has a secondary host, showed no differentiation between populations 150 km apart, suggesting that population isolation in *P. spyrothecae* may be a consequence of losing its secondary host. Populations within trees were highly inbred, probably due to selfing between sexuales from the same clone. This finding corresponds with the theory that female-biased sex ratios in *P. spyrothecae* evolved through local mate competition. There was no evidence for a correlation between inbreeding and population density.

Genetic variation was also found within galls. Of 633 aphids in one gall, 619 shared one genotype, while the remaining 14 were immigrants from at least nine other clones. One immigrant was found among 49 aphids from four other galls. Such a low level of clonal mixing probably favoured the evolution of soldiers, and may represent an investment in dispersal by the clone as an insurance against its death.

CONTENTS

Preface ii

Acknowledgements ii

Summary iii

Contents iv

Chapter One: Introduction 1
- Genetic variation in aphids 1
- The aims of this study 2
- Aphid biology 2
  - Aphid parthenogenesis is apomictic 3
  - Biology of the Pemphigidae 4
  - Biology of the study species, *Pemphigus spyrothecae* 4
- Rationale 7
- The source of genetic variation 9
  - Microsatellites 9
  - The polymerase chain reaction 9

Chapter Two: Methods 13
- Sampling 13
- Population density surveys 20
- Laboratory methods 21
  - Development of microsatellite markers 21
    - Preparation and screening a restriction fragment library 21
    - Sequencing of positive clones 22
    - Design and synthesis of PCR primers 22
  - Genotyping of *P. spyrothecae* 22
    - DNA extractions 22
    - PCR amplification of microsatellite markers 23
    - Initial appraisal of variation at *P. spyrothecae* microsatellite markers 24
    - Genotyping of *P. spyrothecae* clones 24
- Statistical methods 25
  - General statistical methods 25
    - The significance level, α 25
    - Difference between proportions 25
    - Difference between means 26
    - Correlation 26
    - Confidence intervals of means 26
    - Outliers 26
  - Statistical methods for analysing genotypic data 26
    - Measure of diversity 26
    - Hardy-Weinberg equilibrium 27
    - Linkage disequilibrium 27
    - Population genetic analysis 27
      - *F*-statistics 27
      - Analysis of variance 28
      - Spatial genetic structure 28
      - Temporal genetic variation 29
      - Inbreeding and selfing 29

Chapter Three: Population Biology **30**
Introduction 30
Identification of the host tree, *Populus nigra* 30
The distribution of *Populus nigra* 31
The distribution of *P. spyrothecae* in Cambridgeshire 35
The distribution of the Lombardy poplar in Modena 36
The distribution of the Lombardy poplar in northwestern America 36
The distribution of *P. spyrothecae* on *Populus nigra* 37
Results 38
Temporal variation in gall density 38
Spatial variation in gall density 40
Variation between sites 40
Variation between trees 41
Association between gall density and latitude 42
Discussion 44
Temporal variation 44
Competition 45
Inbreeding 46
Predation 47
Physical factors 48
Spatial variation 51
Conclusion 52

Chapter Four: Population Genetics **53**
Introduction 53
Studies of aphid genetic variation 53
Allozyme electrophoresis 53
Mitochondrial DNA 56
Multi-locus DNA fingerprinting with probes 56
Multi-locus DNA fingerprinting using PCR 57
Microsatellites 57
Studies of genetic variation in the genus *Pemphigus* 58
P. spyrothecae 58
Results 60
Genotyping of *P. spyrothecae* galls 60
Analysis of PCR reliability 66
Microsatellite variation in the UK 68
Tests of linkage disequilibrium 69
Worldwide variation in *P. spyrothecae* 70
Variation in Europe 70
Variation in America 70
Hardy-Weinberg equilibrium 72
Hardy-Weinberg equilibrium in the UK 72
Hardy-Weinberg equilibrium in Italy 74
Temporal genetic variation in UK populations 75
Spatial population genetic structure 77
Investigating the reasons for heterozygote deficits 77
Genetic isolation between five cities across Europe 77
Genetic differentiation between Europe and America 77
Genetic differentiation between towns in the UK 78
Between sites in Cambridge 79
Spatial genetic structure between trees within sites 80
Analysis of molecular variance in the UK populations 82
Investigation of within-trees homozygote deficits 84
Tests of population genetic structure within trees 85

Inbreeding within trees.......85
Investigation of a link between population density and selfing rate.......87
Spatial genetic structure in the Italian populations.......88
Analysis of molecular variance.......88
Pairwise population comparisons in Italy.......89
Pairwise comparisons between trees.......90
Inbreeding in the Italian sample.......91
Spatial population structure in northwestern America.......93
Discussion.......96
Microsatellite variation in aphids.......96
Spatial population structure.......98
Isolation by distance.......98
Long-range gene flow.......99
Medium-range gene flow.......100
Short-range gene flow.......101
Gene flow between local sites.......101
Gene flow between trees.......101
Gene flow within trees.......101
Inbreeding within trees.......102
Population density and selfing.......103
Long-range versus short-range gene flow.......103
Pemphigus spyrothecae in northwestern America.......106
Temporal genetic stability.......107
Future research.......108

Chapter Five: Clonal mixing.......109
Introduction.......109
The problem of reproductive altruism.......109
Altruism within clones.......110
The rarity of altruism in clonal organisms.......110
Altruism among clonal organisms.......112
Social behaviour in aphids.......112
Ecological incentives.......114
Behavioural pre-adaptations.......114
Genetic integrity.......115
Do galls prevent clonal mixing?.......115
Evidence for inter-gall migration.......116
Field studies.......116
Genetic evidence.......117
Methods.......119
Results.......121
Gall composition.......121
Genotyping success.......122
Genotyping reliability.......123
Genetic variation in Gall 3019.......123
Genetic composition of Galls 3063, 3081, 3291 and 3308.......126
Discussion.......127
Inter-gall migration in *P. spyrothecae*.......127
How prevalent is inter-gall migration in *P. spyrothecae*?.......127
Which aphids migrate?.......129
Immigrant soldiers can leave descendants.......130
Why should *P. spyrothecae* soldiers migrate?.......130
Non-adaptive migration.......130
Adaptive migration.......131
Migration in response to immediate danger.......132

Altruistic migration ... 133
Migration as insurance against future danger ... 133
Altruistic emigration ... 134
Why do only soldiers emigrate? ... 135
Clonal mixing and the evolution of soldiers ... 136
Future work ... 137
Ascertaining the level of clonal mixing in *P. spyrothecae* and other species ... 137
Measuring the potential reproductive success of soldiers ... 138
Do migrants cheat? ... 138

Chapter Six: Conclusion ... 139

References ... 140

Appendix 1: Microsatellite sequences ... 153

Appendix 2: Publication ... 156

CHAPTER ONE: INTRODUCTION

Genetic variation in aphids

Darwin (1859) defined biological evolution as the process of 'descent, with modification'. At the genetic level, evolution is the process by which gene frequencies change in populations. Understanding the principles that determine variation in gene frequencies between and within populations is therefore fundamental to understanding evolution.

Genetic variation has been intensively studied in aphids for a variety of reasons. Many studies have focussed on pest species, which can damage crops directly, by feeding on them, and indirectly, through virus transmission. By investigating gene flow in pest species, it is hoped that more can be understood about their population dynamics, leading to improved forecasting of outbreaks and better understanding of how insecticide resistance spreads among populations (Loxdale et al., 1998).

Genetic variation in aphids has also been studied purely for what it can tell us about their intrinsically fascinating biology, as well as for insights into fundamental biological questions. The appeal of aphids as study animals is based on their diverse and complex life cycles, which usually incorporate both sexual and asexual phases and often involve alternation between two or more host plant species (Dixon, 1973). Accordingly, aphids have attracted researchers interested in, among other topics, the evolution of sex (Simon et al., 1999a), host-specialisation (Vanlerberghe-Masutti and Chavigny, 1998), spatial

genetic structure (De Barro et al., 1995b), genome evolution (Sunnucks et al., 1996) and clonal diversity (Fuller et al., 1999).

The aims of this study

In this study, I have investigated genetic variation in a gall aphid, *Pemphigus spyrothecae*. The apparent confinement of sexually-produced *P. spyrothecae* clones within galls reduced the probability of multiple sampling of clones, allowing genetic variation to be investigated at two levels: between sexually produced clones, and within galls. The aim of studying between-gall variation was to discover the following: the ability of the life cycle of *P. spyrothecae* to maintain genetic variability over time; the potential of *P. spyrothecae* to migrate between hosts over a wide range of distances; and the relative amounts of outbreeding and selfing between and within *P. spyrothecae* clones. Genetic variation was also examined within galls with the aim of investigating the degree of mixing between *P. spyrothecae* clones. Gall-living is thought to have been important in the evolution of social behaviour in aphid species, including *P. spyrothecae*, in part because galls provide a barrier to clonal mixing (Foster and Northcott, 1994).

Aphid biology

Aphids are small (usually less than 7 mm long), soft-bodied insects (Order Hemiptera; Suborder Homoptera) that live and feed on plants, sucking phloem sap through specially adapted mouthparts called stylets (Dixon, 1973). Aphid life cycles are characterised by cyclic parthenogenesis, where several asexual, viviparous generations alternate with one sexual, oviparous generation. The ability to reproduce without fertilisation frees females to develop daughter and granddaughter embryos simultaneously, allowing rapid

population growth. Although the majority of the approximately 4400 known aphid species are cyclic parthenogens, 3% are obligate parthenogens, having lost the annual sexual generation. Aphids are also diverse in the number and species of their host plants. Most aphids are monophagous, spending their entire life cycle on a single species of host plant, but a minority of species host-alternate, dividing their life cycle between a primary host, where sexual reproduction takes place, and a secondary host, where they reproduce parthenogenetically. Usually, the primary host of host-alternators is a single species of woody plant, while one or more herbaceous plants provide the secondary hosts (Dixon, 1973).

Aphid parthenogenesis is apomictic

Aphid parthenogenesis is generally accepted to be apomictic: daughters are genetically identical to their mothers, within the limits of the fidelity of DNA replication. This view was questioned by Cognetti (1961), who reported clonal production of altered genotypes through artificial selection, and observed chromosomal pairing during the development of parthenogenetic eggs. He claimed that recombination occurs during parthenogenesis by a mechanism he termed endomeiosis, and that it generates a significant degree of genetic variation in aphids. Blackman (1979) attempted to replicate Cognetti's results, but failed to show a response to artificial selection for morphological characteristics, and demonstrated that clones heterozygous at an allozyme locus did not produce homozygous offspring. He concluded that the most of the variation observed by Cognetti was generated by environmental variation, contamination and mutation, and that mitotic recombination was probably rare in parthenogenetic lineages. Subsequent genetic studies of laboratory-reared aphid clones also failed to find evidence of endomeiosis (Carvalho et al., 1991; Fukatsu and Ishikawa, 1994).

Biology of the Pemphigidae

The Pemphigidae are a family of 299 species, most of which are host-alternating cyclic parthenogens that form galls on their primary hosts (Blackman and Eastop, 1994). The majority of the primary hosts used by this family are distributed within the genera *Ulmus*, *Zelkova*, *Populus*, *Pistacia* and *Rhus* (elm, poplar, pistachio and sumac, respectively). The choice of primary host is highly conserved. For example, two elm genera (*Ulmus* and *Zelkova*) provide the primary hosts for all species of the subfamily Eriosomatinae, but none of the species of the other two Pemphigid subfamilies, the Fordinae and the Pemphiginae. The Pemphigidae and the closely related Hormaphididae are the only aphid families in which reproductive altruism has been recorded (Stern and Foster, 1996).

Biology of the study species, *Pemphigus spyrothecae*

Pemphigus spyrothecae (Passerini) (Hemiptera: Pemphigidae) is a cyclically parthenogenetic aphid that forms spiral galls on the leaf petioles of black poplar (*Populus nigra*) in the Western palaearctic. Unlike most other species of *Pemphigus*, *P. spyrothecae* does not host-alternate, spending its entire life cycle on black poplar (Lampel, 1969). The life cycle is divided between the bark and the petioles (Figure 1.1). The period on the petioles, which lasts from April until October or November, commences when a foundress hatches from an egg on the bark and walks to a petiole, where she initiates a spiral gall. There follow at least two asexual generations inside the gall, the last of which produces the winged sexuparae, which migrate from the gall to deep cracks on the bark where they give birth to the sexuales (males and females). The sexuales mate, and each female lays a single egg, from which a foundress hatches the following spring.

The aphids born inside the gall are dimorphic, being divided into soldiers and non-soldiers. Soldiers are more sclerotised than non-soldiers, have thicker hind legs (Lampel, 1969) and effectively defend the gall against predators, while non-soldiers show no defensive behaviour (Aoki and Kurosu, 1986; Foster, 1990).

Rhoden (1997) showed that soldiers develop into wingless gall virgins (adult virginoparae), while non-soldiers become alate gall migrants (adult sexuparae). The order of gall generations is shown in Figure 1.1. The foundress bears only soldiers, who moult four times to adult virginoparae. These give birth to equal numbers of soldiers, who become adult virginoparae, and non-soldiers, who develop into adult sexuparae. Adult sexuparae leave the gall to give birth to the sexuals on the bark, while adult virginoparae stay and reproduce in the same way as their mothers. Soldiers remain as first instars for about three times as long as non-soldiers (Rhoden, 1997).

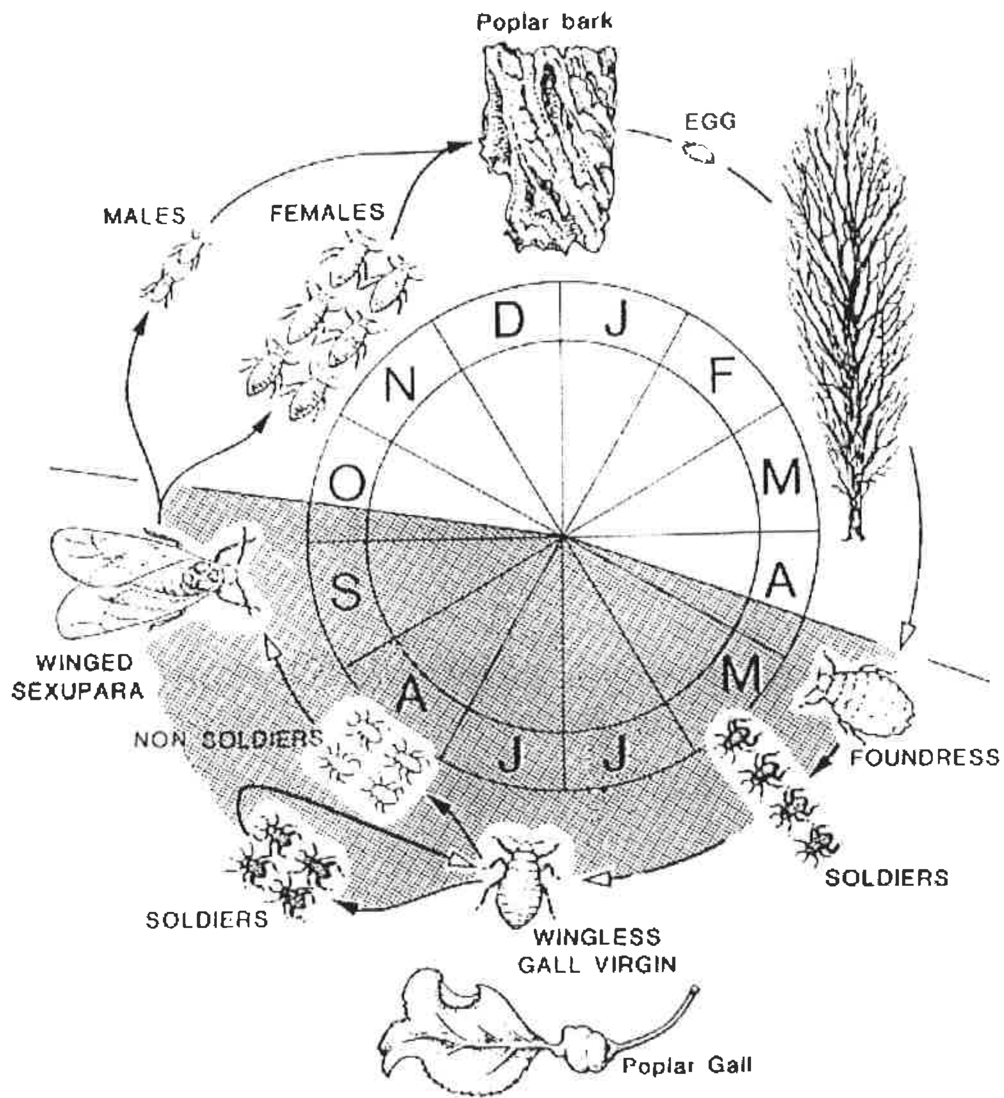


Figure 1.1 Life cycle of *P. spyrothecae*. The shaded portion indicates time spent in the gall, the unshaded part shows time spent on the bark. Dark-headed arrows indicate 'gives birth to', light-headed arrows indicate 'moults to'. Figure adapted from Foster (1994).

Rationale

There are many reasons why *P. spyrothecae* is an attractive species for a study of genetic variation. First, investigation of gene flow between *P. spyrothecae* populations may reveal the genetic consequences of the transition from the ancestral state of host-alternation to monophagy. Whereas most other *Pemphigus* species alternate between a primary host where sexual reproduction occurs, and a secondary host where they undergo several asexual generations, *P. spyrothecae* has lost its secondary host and completes its entire life cycle on *Populus nigra* (Blackman and Eastop, 1994). The loss of either host, though it may bring advantages such as range expansion (Moran and Whitham, 1988), might also invite population isolation and inbreeding. However, the genetic consequences of losing host-alternation have yet to be studied.

Second, assessing the degree of inbreeding in *P. spyrothecae* will allow a prediction of local mate competition (LMC) theory to be tested. *P. spyrothecae* has a strongly female-biased primary sex ratio, which is thought to have evolved in response to LMC (Foster and Benton, 1992; Yamaguchi, 1985). Under the conditions of LMC, brothers compete to mate with their sisters (Hamilton, 1967), so populations should be highly inbred. Foster and Benton (1992) found no evidence for avoidance of brother-sister matings in the laboratory, but the frequency of inbreeding in natural populations is unknown.

Third, investigation of genetic variation in *P. spyrothecae* may shed light on why cyclic parthenogenesis persists in aphid populations, and on the more fundamental question of why sex persists in some species but not others. If genetic diversity within *P. spyrothecae* populations has indeed been severely reduced, either through isolation or inbreeding,

then in this species sex has lost its primary purpose of shuffling genetic variation, and its persistence becomes a mystery.

Fourth, investigation of the level of clonal mixing in *P. spyrothecae* will provide the first evidence for or against the theory that galls have promoted the evolution of social behaviour in aphids by preventing cheating by unrelated immigrants. *P. spyrothecae* is social: a morphologically and behaviourally distinct soldier caste attacks intruding predators and cleans the gall, potentially reducing its own fitness (Aoki and Kurosu, 1986; Benton and Foster, 1992; Foster, 1990; Stern and Foster, 1996). Kin selection theory predicts that such altruistic behaviour is most likely to evolve and persist where mixing between clones is minimal (Hamilton, 1964; Hamilton, 1987). Setzer (1980) found high levels of inter-clone mixing (up to 24%) in two *Pemphigus* species that do not produce obviously morphologically distinct soldiers (although this has not been rigorously tested), but a known soldier-producing species has yet to be investigated.

Finally, knowledge of the dispersal capability of *P. spyrothecae* may have practical benefits. *Populus nigra* is a commercial tree on which *P. spyrothecae* galls can reach extremely high densities. Galls have a markedly negative impact on the health of affected leaves, which suffer reduced growth, paler colour and earlier abscission than unaffected leaves (pers. obs.). Knowing how far and how often *P. spyrothecae* can migrate will allow estimation of the likelihood that an unaffected tree will be parasitised by migrant *P. spyrothecae* alates, and so may inform control strategies.

Are populations genetically isolated, lacking in variation and inbred, as might be inferred from observations of the life cycle and mating behaviour of *P. spyrothecae*? What is the

degree of mixing between clones, and how might the presence or absence of clonal mixing have affected the evolution of social behaviour in *P. spyrothecae*? This study aimed to answer these questions by investigating genetic variation in *P. spyrothecae*.

The source of genetic variation

Microsatellites

Microsatellites are small repetitive DNA elements consisting of up to 300 short (1–4 bp) tandem repeats (Litt and Luty, 1989; Weber and May, 1989). They are scattered throughout eukaryotic genomes, and are often highly variable, having the highest mutation rate of any DNA marker. Microsatellite mutation rates range from 10^{-6} mutations per locus per generation in *Drosophila melanogaster* (Schug et al., 1997), to 10^{-4}–10^{-3} in mammals (Dallas, 1992; Weber and Wong, 1993).

The polymerase chain reaction

Microsatellite variation can be accessed using the polymerase chain reaction (PCR). PCR can amplify a specific DNA sequence from a single cell (Arnheim et al., 1990; Saiki et al., 1985), allowing genetic analysis of very small tissue samples. PCR uses a thermostable polymerase, usually *Taq*, to synthesise multiple copies of a specific region of DNA defined by the sequences of two oligonucleotide primers. The target DNA sequence is amplified to many times its original concentration through about 30 repetitions of three steps: melting of the template at about 90 °C, primer annealing at about 50 °C, and DNA synthesis at 72 °C (Figure 1.2).

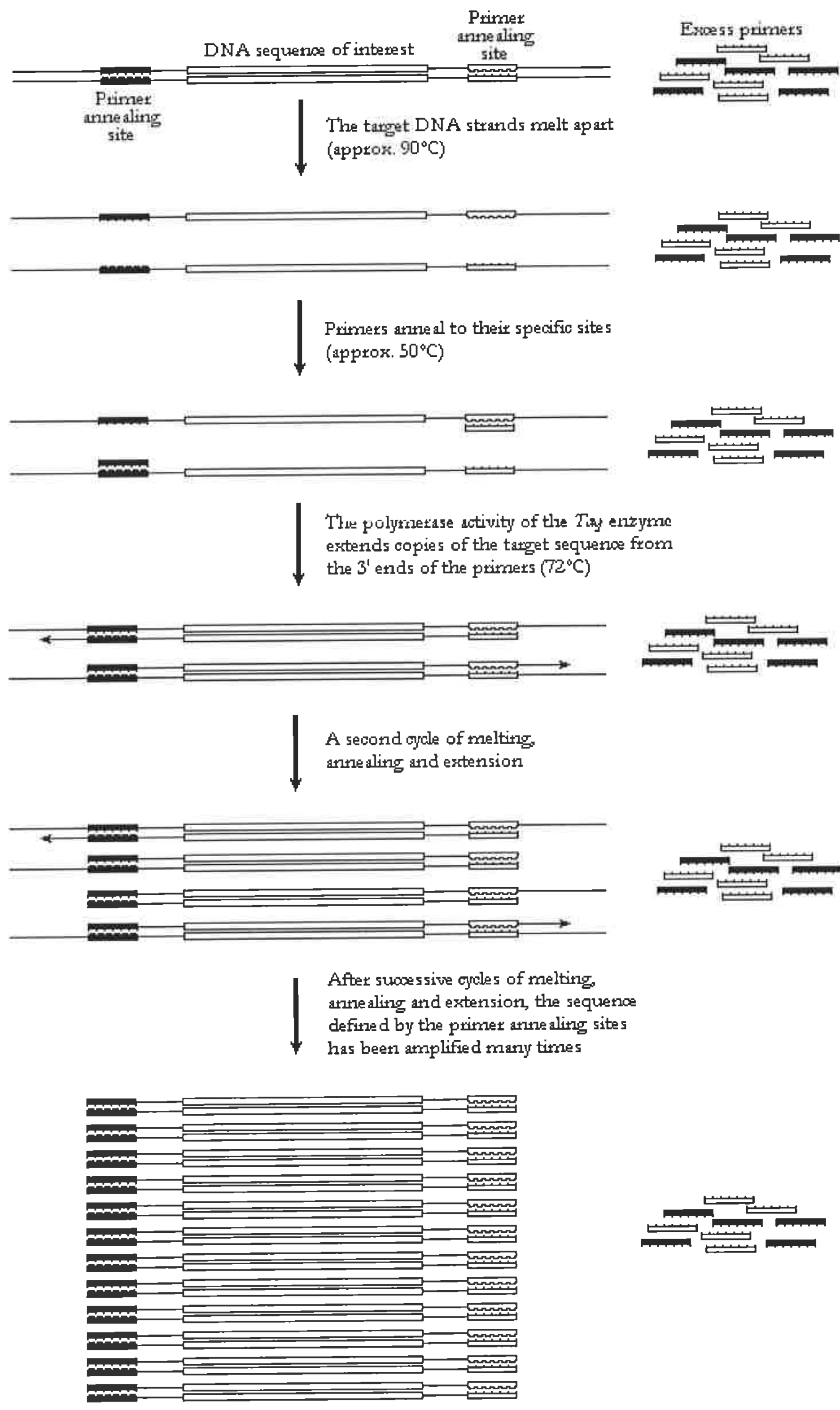


Figure 1.2 The steps of the polymerase chain reaction (PCR). Reagents not shown include *Taq* polymerase and dNTPs.

Following PCR-amplification, microsatellite loci can be visualised by electrophoresis. Scoring of genotypes is usually straightforward, and because microsatellites are co-dominant single locus markers, allele frequencies can be calculated and individuals can be unequivocally designated as heterozygotes or homozygotes, allowing powerful statistical analyses already established for allozymes, which were the previous generation of single-locus markers.

Because of their variability, species-specificity and amenability to statistical analysis, microsatellites were chosen from the many available genetic markers to study genetic variation in *P. spyrothecae* (see chapter 4 for a review of the use of genetic techniques to study aphids).

The only apparent disadvantages of microsatellites, compared with multi-locus markers, are the necessity of cloning for primers and the problem of null alleles. Finding and developing primers for a variable microsatellite locus involves probing a total genomic digest as described for minisatellites, followed by cloning and sequencing of the fragments in which microsatellites are detected. PCR primers are then designed to be complementary to the flanking sequences. This process is expensive and time consuming, but can often be circumvented by the use of primers from closely related species.

A common problem with microsatellites is the presence of null alleles. These are alleles that fail to amplify, usually because of a mutation in the priming site, so that heterozygotes in which one of the alleles is null are mis-scored as homozygotes (Callen et al., 1993). A common null allele can cause a population in Hardy-Weinberg equilibrium to show an artefactual excess of homozygotes. This problem can be combated by

analysing data from many microsatellite loci. As most loci will not have common null alleles, a locus that shows an anomalous homozygote excess should be treated with caution. If the presence of a null allele is suspected, possible remedies include estimating its frequency and treating it as any other allele, or trying to amplify it by redesigning primers.

Chapter Two: Methods

Sampling

P. spyrothecae galls were collected from 16 locations: six in the UK (Figure 2.1), four in mainland Europe (Figure 2.2), and six in North America (Figure 2.3) (Table 2.1). Cambridge and Modena were targeted for detailed sampling (Figures 2.4 and 2.5), where four and three sites, respectively, were sampled within a 20-km radius. The sample set for genetic analysis consisted of 1042 galls collected from 195 Lombardy poplars (*Populus nigra* var. *italica*) in 22 sampling sites over three summers from 1997 to 1999. Table 2.2 gives the numbers of galls collected from each site in each year. Each gall was preserved in 80% ethanol in a sealed plastic 6-ml bijou and stored at 4 °C for up to three years before genotyping. The position of each gall on its host tree was recorded as two measurements: height from the ground, and compass bearing from the centre of the tree.

Table 2.1 Latitude and longitude coordinates of sampling locations

Continent	Location	Latitude	Longitude
Europe	Cambridge, UK	52°13'N	0°08'E
	W. London, UK	51°26'N	0°16'W
	Solihull, UK	52°26'N	1°47'W
	Newcastle-under-Lyme, UK	53°02'N	2°15'W
	Edinburgh, UK	55°57'N	3°13'W
	Hillington, UK	55°50'N	4°23'W
	Modena, Italy	44°40'N	10°54'E
	Vienna, Austria	48°13'N	16°22'E
	Berlin, Germany	52°31'N	13°24'E
	Uppsala, Sweden	59°52'N	17°38'E
N. America	Hope, BC, Canada	49°20'N	121°23'W
	Hell's Gate, BC, Canada	49°48'N	121°23'W
	William's Lake, BC, Canada	52°08'N	122°10'W
	Prince George, BC, Canada	53°50'N	122°50'W
	Radium Hot Springs, BC, Canada	50°48'N	116°12'W
	Highway 95, ID, USA, Canada	48°41'N	116°18'W

Table 2.2 Numbers of galls collected from 22 sites in Europe and North America.

Country	Site	No. of galls (trees)		
		1997	1998	1999
UK	Mill Pit, Cambridge	49 (7)	45 (6)	46 (5)
	Gonville Place, Cambridge	14 (6)	6 (2)	13 (4)
	Hills Road, Cambridge	-	25 (2)	18 (2)
	Coton, Cambridge	38 (9)	56 (7)	43 (7)
	Richmond, London	23 (3)	14 (2)	22 (3)
	Berrylands, London	-	-	12 (1)
	Solihull	20 (2)	-	-
	Newcastle-under-Lyme	59 (5)	47 (8)	7 (4)
	Edinburgh	25 (13)	58 (15)	22 (5)
	Hillington, Glasgow	10 (4)	36 (5)	4 (2)
Italy*	Gaggio, Modena	-	-	34 (3)
	Via N. dell'Abate, Modena	-	-	26 (3)
	Ravarino, Modena	-	-	28 (3)
Austria*	Vienna	-	-	71 (9)
Germany*	Berlin	-	-	29 (6)
Sweden*	Uppsala	-	-	26 (3)
Canada	Hope, B.C.	-	-	41 (16)
	Hell's Gate, B.C.	-	-	9 (1)
	William's Lake, B.C.	-	-	9 (3)
	Prince George, B.C.	-	-	28 (8)
	Radium Hot Springs, B.C.	-	-	11 (3)
U.S.A.	Highway 95, Idaho	-	-	18 (8)

*Collected and kindly donated by Christian Brändle.

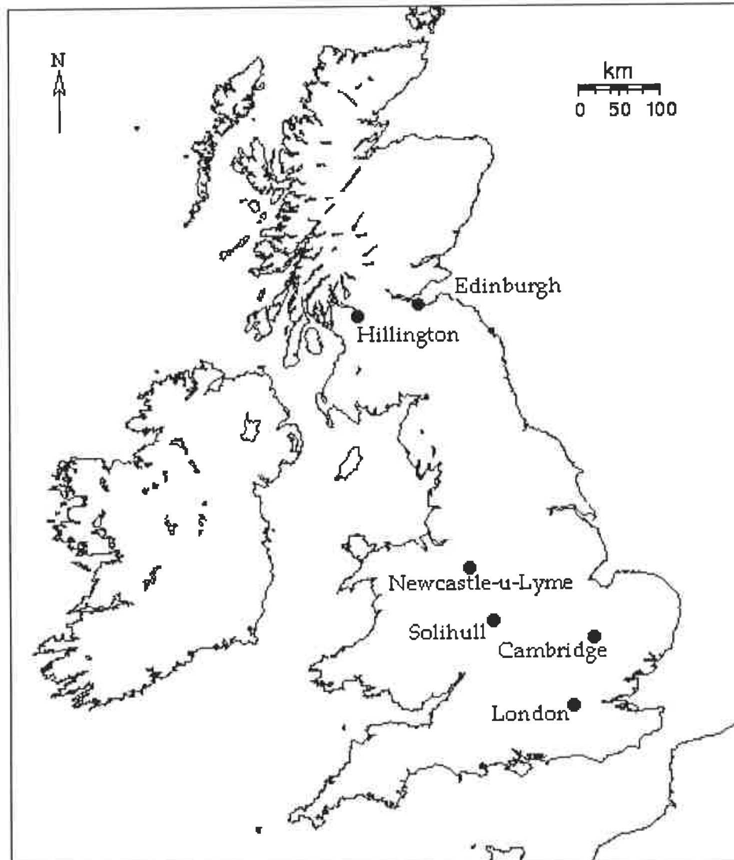


Figure 2.1 Map of the British Isles showing sampling locations.

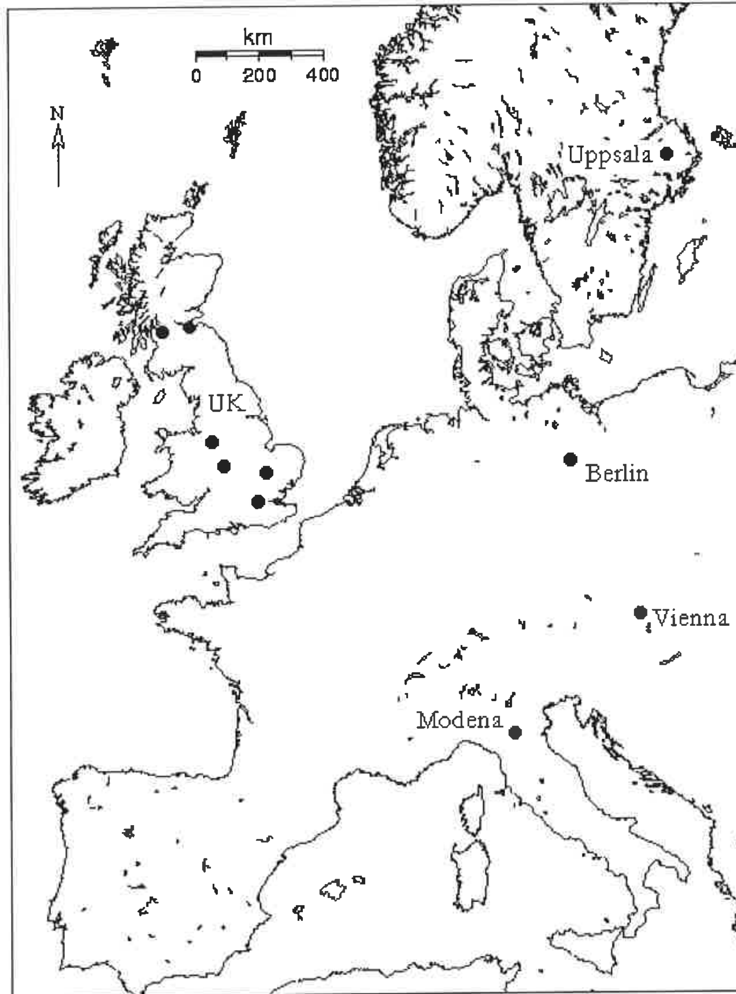


Figure 2.2 Map of western Europe showing sampling locations.

Figure 2.3 Map of northwestern America showing sampling locations.

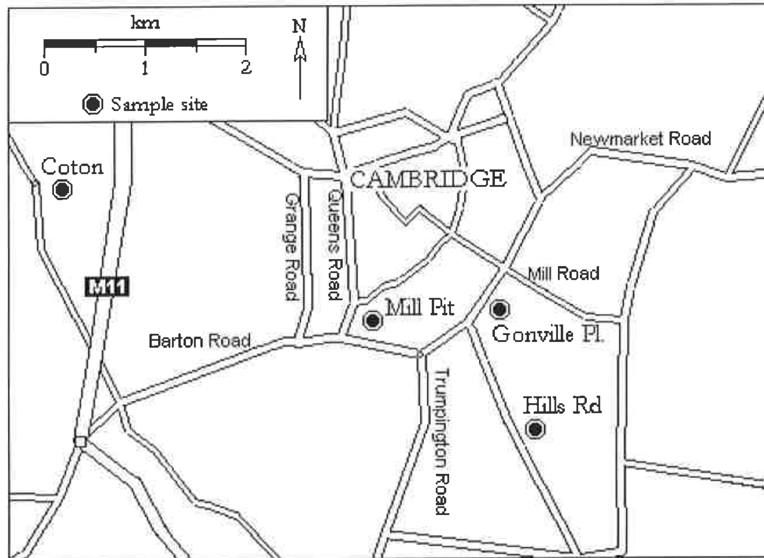


Figure 2.4 Map of the Cambridge area showing sampling locations.

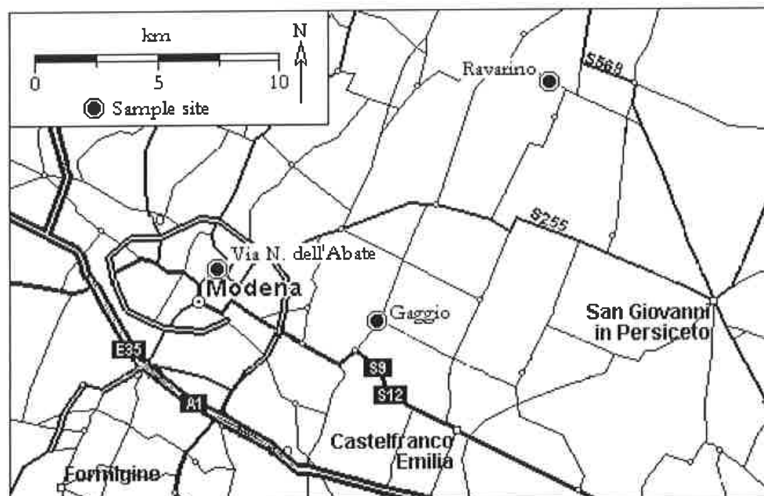


Figure 2.5 Map of the Modena area showing sampling locations.

Population density surveys

Population densities were estimated as the number of galls occupied by *P. spyrothecae* per petiole on a tree. The population densities of clones on trees were estimated by recording the proportion of petioles occupied by a gall. Forty-eight surveys were made on 20 trees over four summers from 1996 to 1999. Twenty branches were selected on each tree, and the twenty petioles nearest to the bark were examined on each branch. The number of petioles occupied by a *P. spyrothecae* gall was recorded. In order to avoid possible bias due to spatial differences in gall density, such as might be caused by the degree of exposure to sunlight or wind, the 20 branches for survey were selected from, as near as was possible, the same positions on each tree. The selection protocol was as follows: (1) four points were defined on each tree at a height of 1 m from the base, and at compass bearings of 0°, 90°, 180° and 270° from the centre of the tree; (2) the five branches nearest to each of the four points, within 45° on either side, were surveyed. The average number of petioles examined per tree was 286 (range 78–400). Gall density does not vary between 1 m and 10 m from the ground (Nathan Pike, pers. comm.), so population densities recorded near to the ground were assumed to be representative of the whole tree.

In order to investigate changes in population density over time, seven of the 20 trees were inspected in each year from 1997 to 1999: Coton 1 and 4, Richmond 9, and Newcastle-under-Lyme 1, 2, 3 and 10 (Table 2.3). The first three of these seven trees were also surveyed in 1996.

Table 2.3 Numbers of petioles inspected in twenty-four surveys of seven trees from 1996 to 1999.

Tree	1996	1997	1998	1999
Coton 1	338	294	389	398
Coton 4	249	277	294	398
Richmond 9	376	281	348	364
Newcastle 1	-	271	377	292
Newcastle 2	-	280	310	251
Newcastle 3	-	313	329	252
Newcastle 10	-	365	317	294

Laboratory methods

Development of microsatellite markers

The development of microsatellites as molecular markers for *P. spyrothecae* required the following four steps: (1) construction and screening of a restriction fragment library of cloned *P. spyrothecae* DNA fragments; (2) sequencing of positively-probed clones; (3) design and synthesis of PCR primers.

Preparation and screening a restriction fragment library

Genomic DNA was phenol-chloroform extracted according to Sambrook et al. (1989) from bulked *P. spyrothecae* taken from a single gall. Two genomic restriction fragment libraries were made, one using *Hae*III alone, the other digested with both *Hae*III and *Alu*I. The restriction digests were separated on an agarose gel, and size-selected fragments in the 300–800 bp range were excised from the gel, purified and ligated into *Sma*I-cut pUC19 using temperature cycle ligation (Lund et al., 1996). DH5α *Escherichia coli* cells were grown from commercial stocks and made competent using $CaCl_2$

(Sambrook et al., 1989). The ligation products were transformed into the competent DH5α cells, and screened with ^{32}P-labelled $(AC)_{15}$ synthetic oligonucleotides according to Sambrook et al. (1989). The plasmid vectors carrying the putative microsatellites were purified and bound to membranes by dot blotting and Southern blotting, then screened with ^{32}P-labelled $(AC)_{15}$.

Sequencing of positive clones

Positive colonies were restreaked and rescreened to confirm the presence of microsatellites prior to BigDye cycle sequencing (PE Applied Biosystems). Sequences were analysed on an ABI PRISM 377 (PE Applied Biosystems), yielding 19 microsatellites (see appendix 1 for sequences).

Design and synthesis of PCR primers

Primers were designed from the 13 microsatellite sequences with sufficient flanking sequence and the longest repeats using the internet-based program Primer 3.0 (Rozen and Skaletsky, 1996). The primers were tested for their ability to reveal genetic variability by PCR amplification from *P. spyrothecae* DNA from a wide geographical range

Genotyping of *P. spyrothecae*

DNA extractions

A genomic DNA extraction protocol for PCR was developed by trial and error. The possibility of PCR inhibitors being present in *P. spyrothecae* or host tissue was tested by comparing PCR results using DNA extracted from different body parts (head, thorax, abdomen, whole aphid) and instars. A suspected inhibitor was tested by adding it to a PCR mix that reliably generated a product in its absence. Experiments were performed to

optimise cell lysing procedures (boiling and freezing) and reagent composition and concentration. The most successful protocol was one adapted from Walsh et al. (1991). DNA was extracted from whole *P. spyrothecae* aphids (2nd- or 3rd-instar sexuparae). Each aphid was washed in 70% ethanol to remove potential host plant-derived PCR inhibitors and dried in air at 55 °C to remove ethanol. In order to break the cell membranes, the aphid was placed in 30 µl H_2O, frozen at -20 °C, then microwaved for 3 min at 650 W. Cells were further lysed by adding 30 µl of 2x extraction buffer (10% Chelex 100™ chelating resin, 0.5 mg/ml proteinase K, 200 mM dithiothreitol) and incubating for 3–8 h at 55 °C. Proteinase K was then inactivated by microwaving for 2 min. DNA extractions were stored at -20 °C. The quality of DNA in the extractions was high, as shown by the presence of a tight band on an agarose gel. DNA concentration, where it was high enough to estimate, ranged up to 5 ng/µl.

PCR amplification of microsatellite markers

PCR conditions were optimised by varying the concentrations of reagents and the profile of the temperature cycle. PCRs were carried out in 10 µl reaction volumes containing 1 µl of 50x dilution of *P. spyrothecae* template DNA, 1x NK PCR buffer (67 mM Tris (pH 8), 160 mM $(NH_4)_2SO_4$, 50 mM KCl, 0.1% Tween 20), 1.5 mM MgCl, 0.1 mM dGTP, 0.1 mM dATP, 0.1 mM dTTP, 0.01 mM dCTP, 400 nM of each primer, 0.25 U *Taq* polymerase (Promega) and 0.1 µCi [α^{32}P]-dCTP. The PCR program was: 2 min denaturing at 94 °C; seven cycles of 40 s at 94 °C, 1 min at 48 °C, 40 s at 72 °C; and 40 cycles of 40 s at 89 °C, 1 min at 52 °C, 40 s at 72 °C. PCR products were resolved on 6% denaturing polyacrylamide gels and visualised by autoradiography. PCR success was strongly dependent on both the instar from which DNA was extracted, and the quantity of DNA extraction used. The highest rate of PCR success (96.3%, n = 218 PCRs) was

achieved when using 0.03% (i.e. 1 μl of a 50 times dilution) of each extraction from a 2nd- or 3rd-instar sexupara (see chapter 5 for a discussion of genotyping reliability).

Initial appraisal of variation at *P. spyrothecae* microsatellite markers

Genotyping of eleven aphids from sites around Europe showed that seven of the 13 loci were polymorphic (Table 2.4). Lengths of PCR products were measured of the eight loci were measured against the M13 standard sequence.

Table 2.4 Characteristics of seven polymorphic microsatellite marker loci in *Pemphigus spyrothecae*: locus name, GenBank accession number, repeat motif, forward and reverse primer sequences, and PCR product size in bp.

Locus	Accession*	Repeat motif	Primer sequence (5′ to 3′)	Size
96PS5	AF246670	$(GT)_{28}$	ACGATACAGATTTAATTTCAATTTTAG	164
			CCAATGTTTGCACGTTAATATC	
96PS20	AF246671	$(GT)_{25}$	TTTACATACATATACGCATGTACAC	190
			TTTTCAGGTCCCCAGTAT	
97PS9	AF246677	$(AC)_{18}$	CCTTATAACCGTCTTGAGAAATCG	189
			ATAAATACGCGCACAAACCG	
97PS12	AF246679	$(AC)_{19}$	AAGCCCGACTTTATACGACG	235
			CACGAGTGTCCGGTTCCTAT	
98PS8	AF246682	$(AT)_{10}N_{62}(GT)_{10}$	CCGTACACCTGTCCCTGC	199
			TTTTTACCGACCAGTATTTTACGAC	
98PS12	AF246683	$(AT)_7(GT)_{16}$	TCTTTCCTCCTAGTGGTC	142
			TTAATGATATTCCTAATACACAC	
98PS26	AF246688	$(GT)_{16}$	GCGCGGATCAATATACACAG	276
			AAGTTCTATGGCGGACAACG	

*Accession numbers can be used to download the microsatellite sequences at the website http://www.ncbi.nlm.nih.gov/.

Genotyping of *P. spyrothecae* clones

One 2nd- or 3rd-instar sexupara was genotyped from each gall listed in Table 2.2. All galls were genotyped at loci 96PS5, 97PS9, 97PS12, 98PS8, and 98PS12. The galls collected in 1999 were also genotyped at loci 96PS20 and 98PS26, including all galls

collected outside the UK. Allele lengths were measured against the M13 standard sequence.

Statistical methods

General statistical methods

The significance level, α

In each statistical test, the null hypothesis was rejected when the probability of type I error, p, was less than the significance level, α. For solitary tests of significance, $\alpha = 0.05$, but where data were analysed in tables of statistical tests the significance level was reduced using the sequential Bonferroni method (Rice, 1989; Sokal and Rohlf, 1995). The smallest p-value in a table of k tests, p_1, was deemed significant only if

$$p_1 < 1 - (1 - \alpha)^{1/k},$$

and the inequality was recalculated for the 2nd-smallest p-value, p_2, replacing k with k-1. This procedure was repeated for successively larger p-values until the inequality was not satisfied.

Difference between proportions

Fisher exact tests were computed by hand (Sokal and Rohlf, 1995), and using the non-parametric statistical analysis program NPS (Aaby, 1999). Chi-square tests were calculated in EXCEL and by hand. Goodness-of-fit was tested by hand using a G-test (Sokal and Rohlf, 1995).

Difference between means

Differences between means were tested non-parametrically, using the Kruskal-Wallis and Wilcoxon two-sample tests (Sokal and Rohlf, 1995).

Correlation

Tests of association were performed by comparing the Spearman's rank correlation coefficient, r_s, with a table of critical values (Clarke, 1980).

Confidence intervals of means

Normal confidence intervals were calculated in EXCEL, and exact binomial confidence intervals were calculated using NPS (Aaby, 1999).

Outliers

Anomalous data were identified using Grubbs test (Sokal and Rohlf, 1995).

Statistical methods for analysing genotypic data

Measure of diversity

Allele frequencies and expected and observed heterozygosities were calculated by locus and population in the program GENEPOP (Levene, 1949; Raymond and Rousset, 1995b). Nei's gene diversity was calculated in the program ARLEQUIN (Schneider et al., 2000). Nei's gene diversity is the probability, based on the observed allele frequencies and the assumption of random mating, that an individual will be heterozygous at one or more loci (Nei, 1987).

Hardy-Weinberg equilibrium

Deviation from Hardy-Weinberg equilibrium (HWE) was tested by locus and population. A locus-specific deviation from HWE indicates that the locus may include a null allele (Callen et al., 1993) or may be linked to a locus under selection, and should be excluded from analysis. Deviation from HWE across many or all loci, on the other hand, is likely to be biologically significant, and suggests that the sample is not drawn from a random-breeding population. HWE was tested within populations using exact tests in ARLEQUIN (Guo and Thompson, 1992).

Linkage disequilibrium

In order to detect physical linkage between loci, which would require the linked loci to be analysed as a single locus, linkage disequilibrium was tested within populations using a likelihood-ratio test in ARLEQUIN (Slatkin and Excoffier, 1996). This test assumes HWE within the populations tested (Excoffier and Slatkin, 1998).

Population genetic analysis

F-statistics

Wright's *F*-statistics (F_{IT}, F_{IS} and F_{ST}) (Wright, 1951) were estimated and tested within and between populations at spatial levels ranging from between continents to within trees, in order to investigate genetic isolation between, and inbreeding within populations. All estimates of *F*-statistics were calculated under the infinite alleles mutation model (θ) (Kimura and Crow, 1964), rather than the stepwise model of microsatellite mutation (ρ) (Slatkin, 1995), because a comparison using computer simulation suggested that θ is a better estimator of F_{ST} when loci are few (< 10) and populations are small (< 100) (Gaggiotti et al., 1999).

Analysis of variance

Deviations from HWE were explored using hierarchical analysis of molecular variance (AMOVA) (Cockerham, 1969, 1973). AMOVA partitions the total genetic variance in a sample into four components, due to differences between populations, between sub-populations, between individuals and within individuals. In a population in HWE, 100% of the genetic variance is found within individuals, and none in the other three categories. Deviations from HWE are reflected by the within-individuals variance component differing significantly from 100%, and by any of the other three components differing significantly from zero. The size of each component reflects its contribution to drawing the population away from HWE. AMOVA was performed in ARLEQUIN (Excoffier et al., 1992; Weir, 1996; Weir and Cockerham, 1984).

Spatial genetic structure

The degree of genetic isolation between pairs of populations was assessed and tested by estimating Wright's F_{ST} in ARLEQUIN (Reynolds et al., 1983; Weir and Cockerham, 1984; Wright, 1951). Confidence intervals for a few pairwise F_{ST} estimates were calculated by bootstrapping over loci in the program FSTAT (Goudet, 1995, 1999; Sokal and Rohlf, 1995). The effective number of migrants per generation between populations, $N_e m$ was estimated according to Reynolds (1983):

$$N_e m = \frac{\frac{1}{F_{ST}} - 1}{4}.$$

Pairwise differences in allele frequencies between samples were also tested, using exact tests in FSTAT (Goudet, 1995; 1999; 1996).

Correlation between genetic isolation (pairwise F_{ST}) and physical distance was tested using Mantel tests in ARLEQUIN (Smouse et al., 1986; Sokal and Rohlf, 1995).

Temporal genetic variation

Variation in allele frequencies over time was tested using exact tests in ARLEQUIN (Goudet et al., 1996; Raymond and Rousset, 1995a).

Inbreeding and selfing

The inbreeding coefficient, F_{IS}, was estimated within populations and tested in ARLEQUIN (Weir and Cockerham, 1984). As sibling *P. spyrothecae* sexuales are produced parthenogenetically, inbreeding can take the form of selfing. The selfing rate, S, defined as the proportion of individuals in a population produced by selfing, was calculated from F_{IS} using the formula

$$S = \frac{2F_{IS}}{1+F_{IS}}$$

(Hartl and Clark, 1997). The selfing rate is meaningful only if it is valid to assume that all excess homozygosity in a population is caused by selfing, to the exclusion of other factors such as population substructure and inbreeding between related clones.

Chapter Three: Population Biology

Introduction

An understanding of the population biology of the study species is an essential element of any population genetics study. Apart from the practical necessity of knowing where to find samples, knowledge of the spatial distribution of the study species aids interpretation of population genetic data. In a species that lives in spatially segregated populations, such as *P. spyrothecae* on *Populus nigra*, genetic variation could be influenced by range of population parameters, including the spatial distribution of patches of available habitat, the spatial distribution of the study species, population size, and by variation in these parameters over space and time.

Identification of the host tree, *Populus nigra*

The major factor determining the distribution of *P. spyrothecae* across the range of this study (Europe and North America) is the distribution of its only host, black poplar *Populus nigra*. The name *Populus nigra* covers a bewildering assortment of wild subspecies, cultivated varieties and hybrids. However, *P. spyrothecae* in Britain is found on two types of tree: the Lombardy poplar *Populus nigra* var. *italica*, and the native wild black poplar, *Populus nigra* subs. *betulifolia*. It appears to be absent from hybrids of *Populus nigra* such as *Populus* x *canadensis* 'Serotina' and *Populus* x *canadensis* 'Eugenei' (Rushforth, 1999; pers. obs.). *Populus nigra* subs. *betulifolia* probably has had little recent impact on *P. spyrothecae* populations in Europe for two reasons. First, it is rare in Europe because of the

destruction of its flood-plain habitat (Mabey, 1996; Rackham, 1986), and second, it often supports very few *P. spyrothecae* galls compared with Lombardy poplars (pers. obs.), possibly because the wild tree is more genetically diverse and may have evolved to become more resistant to aphids than the vegetatively propagated cultivar. While the native population has declined over the last 200 years, the Lombardy poplar has been widely planted across most of Britain (Rackham, 1986). Hence, it is likely that a very large proportion of *P. spyrothecae* lives on Lombardy poplars, and that the distribution of this tree is the major determinant of the distribution of *P. spyrothecae.*

The distribution of *Populus nigra*

Figure 3.1 illustrates the European distribution of three subspecies of *Populus nigra* L., including subs. *betulifolia* but not var. *italica*, which has a wider distribution. Figure 3.2 shows the distribution of *Populus nigra* in Britain, including *Populus nigra* L., which supports *P. spyrothecae*, and *Populus* x *canadensis*, which does not. It is unclear whether this survey includes *Populus nigra* var. *italica*, but the distribution tallies closely with my observations of *italica*. Such surveys are of limited use in showing the distribution of habitat available to *P. spyrothecae* because they often conflate host and non-host varieties (e.g. Perring and Walters, 1982), or mis-identify them (e.g. Burton, 1983). The ability of *P. spyrothecae* to distinguish wild *Populus nigra* subs. *betulifolia* and var. *italica* from hybrids lead Rushforth (1999) to propose using its galls as an identification aid (unfortunately Rushforth deepened the confusion by attributing the galls to *P. bursarius*, but it is clear from his description that he referred to *P. spyrothecae* galls).

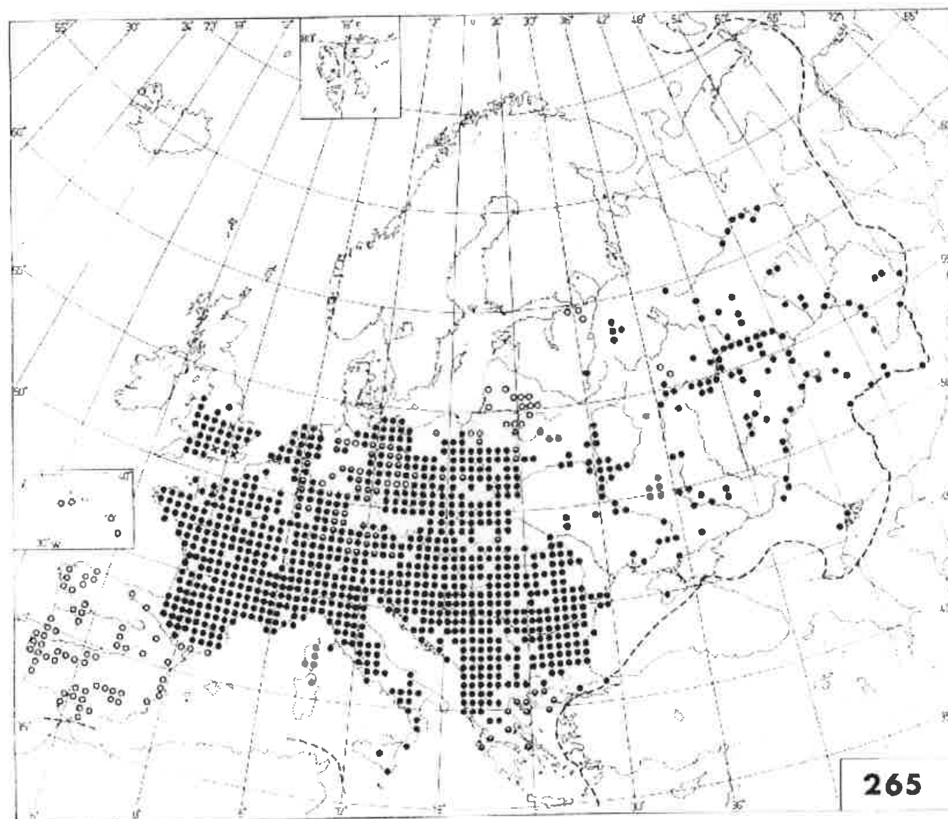


Figure 3.1 Distribution of *Populus nigra* L. in Europe (Jalas and Suominen, 1976).

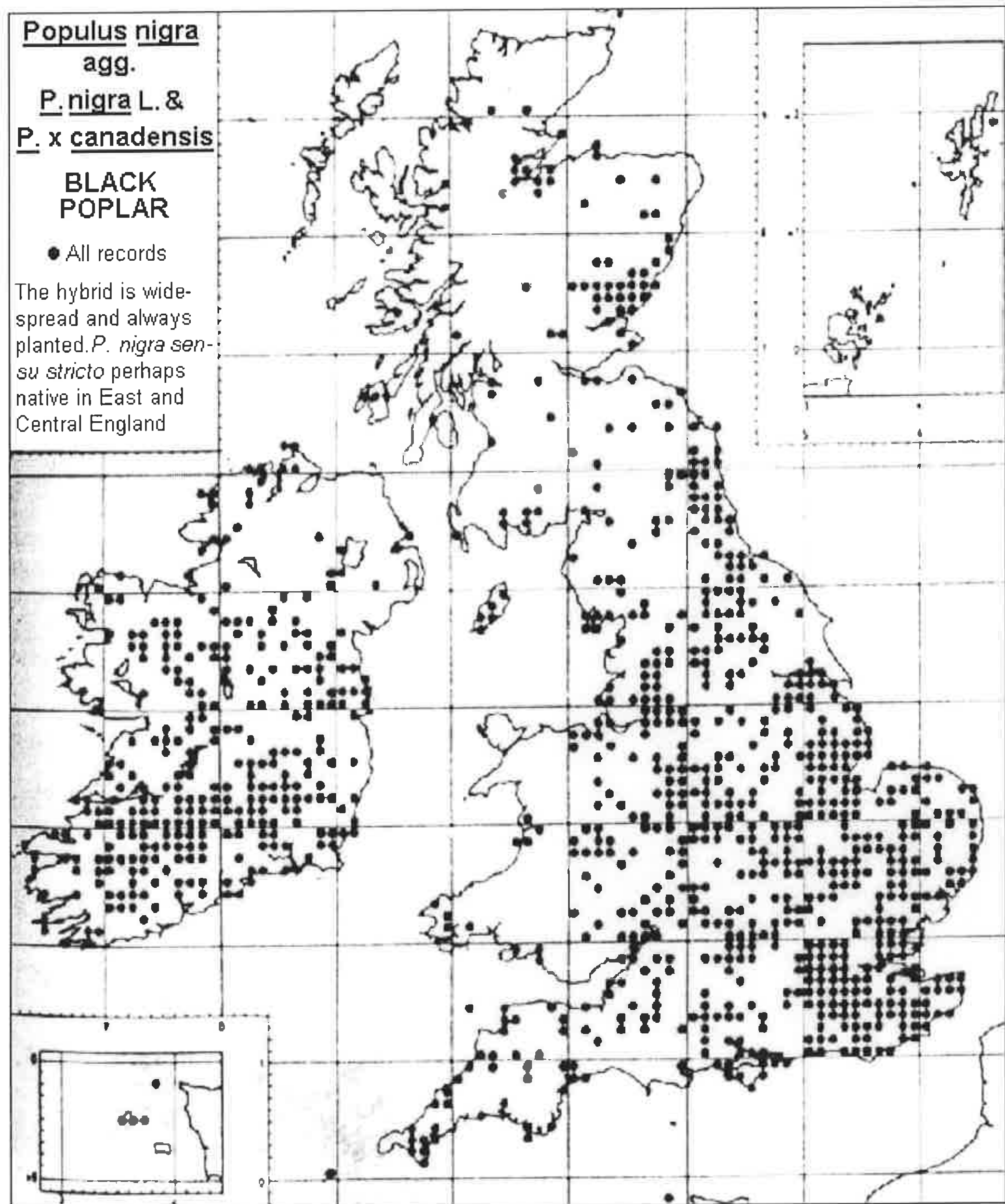


Figure 3.2 Distribution of *Populus nigra* L. and *P.* x *canadensis* in the British Isles (Perring and Walters, 1982).

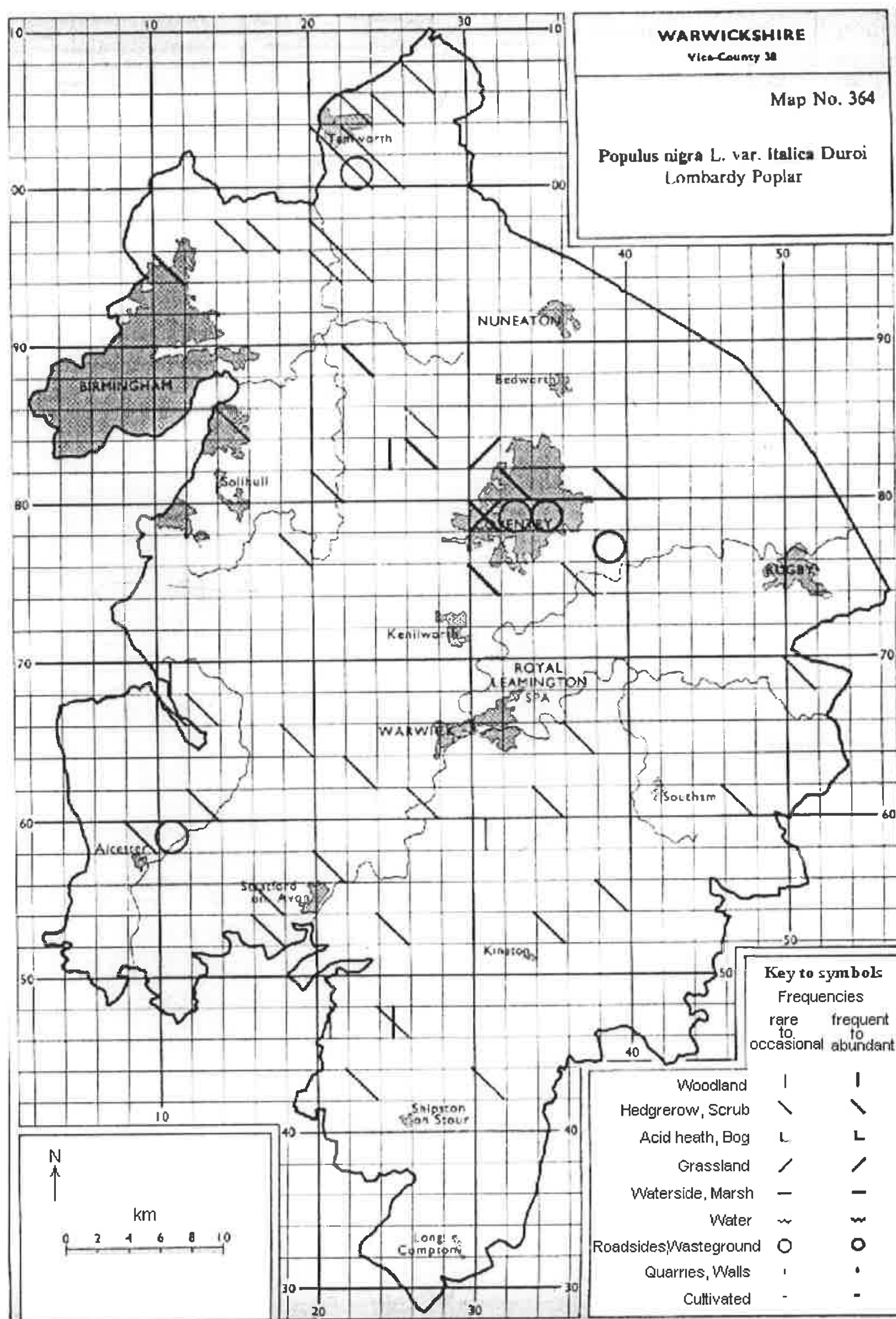


Figure 3.3 Distribution of the Lombardy poplar, *Populus nigra* var. *italica*, in Warwickshire (Cadbury et al., 1971).

Not all surveys have failed to distinguish between hybrids, varieties and native trees. Cadbury et al. (1971) surveyed the county of Warwickshire (area 1981 km^2) for Lombardy poplars, reporting their presence in 53 (11%) of 495 2-km squares (Figure 3.3). No occupied square is more than 10 km distant from its nearest neighbour, and occupied squares in rural areas are typically 2 to 6 km apart. The highest densities of trees were found in urban and suburban areas such as Coventry and Tamworth. Although such detailed data are not available for the rest of the UK, the distribution of Lombardy poplars in Warwickshire is typical of the Midlands and the south of England (pers. obs.), and approximately illustrates the spatial distribution of islands of *P. spyrothecae* habitat over the principal sampling range of this study.

The distribution of *P. spyrothecae* in Cambridgeshire

The location sampled in the finest detail was Cambridge, so knowledge of the distribution of *P. spyrothecae* hosts in the Cambridge area is particularly important to interpreting results of this study. There are thought to be about 200 *Populus nigra* (excluding var. *italica*) in Cambridgeshire (Easy, 1982), equivalent to one tree per 15 km^2. Their distribution is patchy: they are most frequent in the south and west of the county, and rare in the north (Perring, 1964). Only four trees have been recorded in Cambridge (Easy, 1982). There are no detailed records of the distribution of *Populus nigra* var. *italica* in Cambridge or Cambridgeshire. They are common in Cambridge: there are at least seven stands of six or more trees within the city, as well as numerous isolated trees (pers. obs.), and the average number of trees per km^2 is probably at least two.

The distribution of the Lombardy poplar in Modena

The most detailed sampling outside the UK was conducted in and around Modena in Italy. A survey of Lombardy poplars in Modena was made in 1985, showing a high average density of 15 trees per km^2 (Daniele Dallai, Curator of the Botanical Gardens of Modena, pers. comm.).

The distribution of the Lombardy poplar in northwestern America

Both *P. spyrothecae* and its host *Populus nigra* are native European species introduced to America. Although *Populus nigra* var. *italica* had reached North America by 1784 (Spongberg, 1990), the tree and the aphid could not have been introduced to the sampling area (Figure 2.3) before Europeans began to settle there in the early 19th century. Wild *Populus nigra* has also been introduced to North America from Europe, but it is thought to be absent from the northwestern region (B.O.N.A.P., 1998). Therefore, *Populus nigra* var. *italica* planted by hand is the only host of *P. spyrothecae* in this area. Consequently, the *P. spyrothecae* population must be very young, certainly less than 200 years old.

In addition, the distribution of Lombardy poplars in northwestern America is considerably sparser than in Europe (pers. obs.). As in Europe, the density of Lombardy poplars is related to the human population density, and the northwestern American study area is sparsely populated compared with Europe (e.g. England has 373 people per km^2, British Columbia only 3.5). Moreover, much of the terrain is mountainous and rocky and unsuitable for Lombardy poplars. Hence, the patches of available habitat for *P. spyrothecae* tend to be smaller and more isolated than in Europe.

The distribution of *P. spyrothecae* on *Populus nigra*

Very little is known about the distribution of *P. spyrothecae*. Blackman and Eastop (1994) recorded its presence throughout the western palaearctic and on the western seaboard of Canada, but there are no detailed records of the distribution of populations, or sizes, or how these factors vary spatially and temporally. In order to supplement the population genetic data of this study with population biology data, I conducted a survey of *P. spyrothecae* on Lombardy poplars across Britain.

Results

In 48 surveys of 20 trees from 1996 to 1999, the density of *P. spyrothecae* galls on *Populus nigra* varied from zero to 47% of petioles occupied by a gall (Figure 3.4). The median value was 3.8% and the interquartile range was 0.68% to 9.2%.

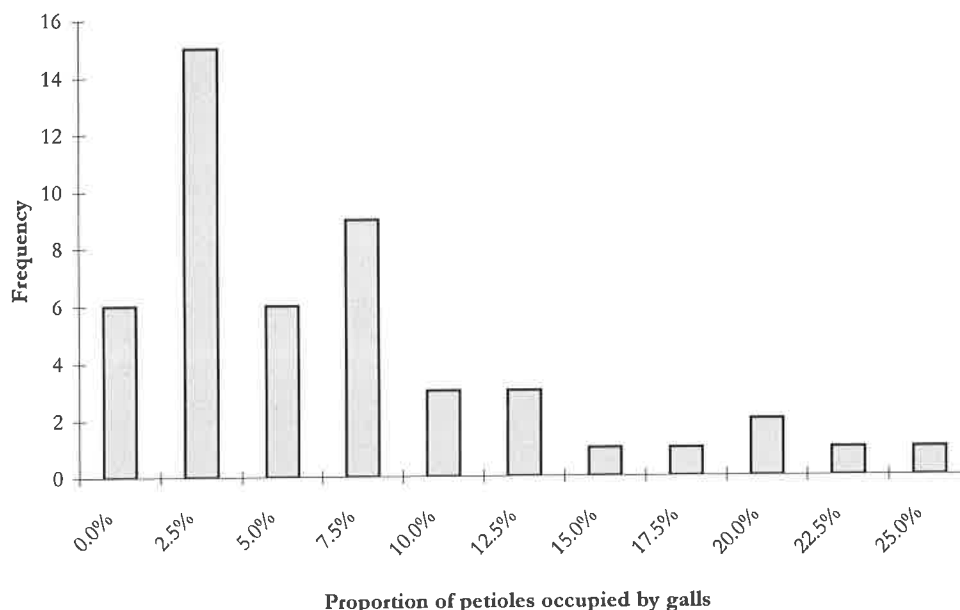


Figure 3.4 Frequency distribution of 47 surveys of *P. spyrothecae* gall density on *Populus nigra* petioles. One outlying survey of 47% is not represented on the histogram to allow detailed presentation of the remaining data.

Temporal variation in gall density

Examination of the gall density data within years showed that they were not normally distributed, as was expected from proportions. The data were arcsine transformed according to Sokal (1995) ($\theta = \arcsin \sqrt{p}$, where p is the untransformed proportion and θ the transformed datum), but the transformation failed to render the data normal. Hence, the data were analysed non-parametrically. Analysis of data from seven trees that were

surveyed in three consecutive years (1997–99) showed that gall density varied significantly over this period (Kruskal-Wallis test: χ^2 = 104.6, d.f. = 2, p < 0.0001) (Figure 3.5). *Post hoc* pairwise Wilcoxon signed ranks tests showed that gall density on the seven trees differed significantly between 1997 and 1998 (W = 0, p < 0.05), and between 1997 and 1999 (W = 0, p < 0.05), but not between 1998 and 1999 (W = 9, p > 0.05).

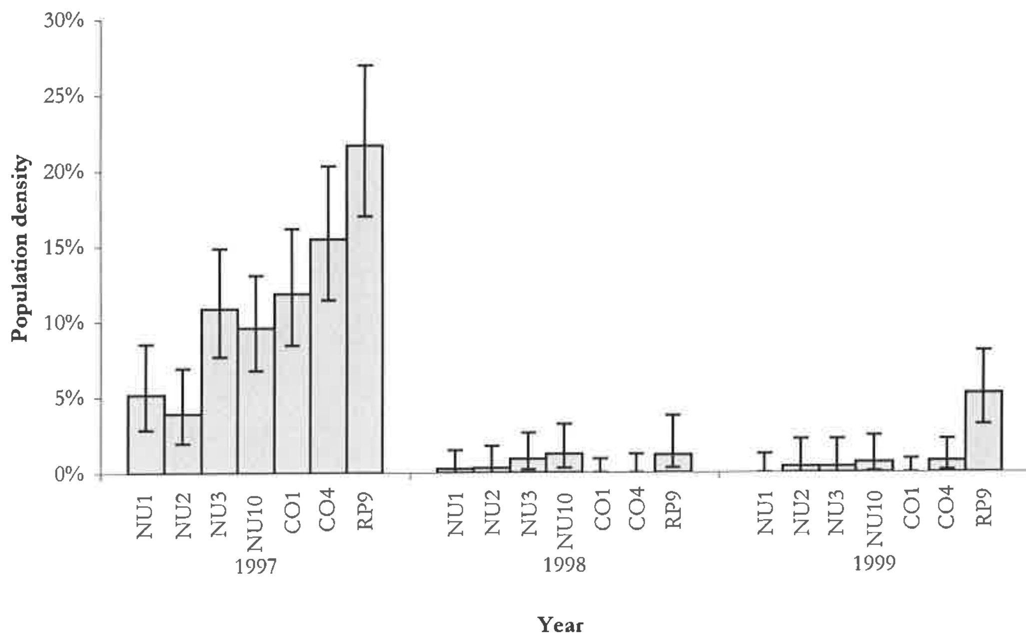


Figure 3.5 Population densities of *P. spyrothecae* on seven trees over three years (1997–99). Error bars show exact 95% confidence intervals for the means (NU: Newcastle-under-Lyme; CO: Coton; RP: Richmond Park)..

Three of the seven trees were also surveyed in 1996 (Figure 3.6). Gall density varied significantly on these trees over the four years from 1996–1999 (Kruskal-Wallis test: χ^2 = 8.56, d.f. = 3, p = 0.04).

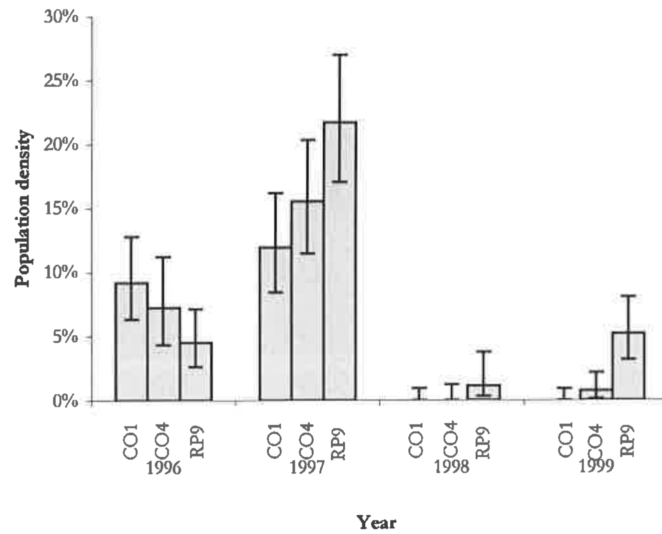


Figure 3.6 Population densities of *P. spyrothecae* on three trees over four years (1996–99). Error bars show exact 95% confidence intervals for the means (CO: Coton; RP: Richmond Park).

Spatial variation in gall density

Variation between sites

Gall densities were compared between Newcastle-under-Lyme and Cambridge using the data presented in Figure 3.7. Gall densities differed significantly between the two sites in 1997 (Mann-Whitney *U*-test, $U = 0, p < 0.05$) but not in 1998 (Mann-Whitney *U*-test, $U = 8, p > 0.05$).

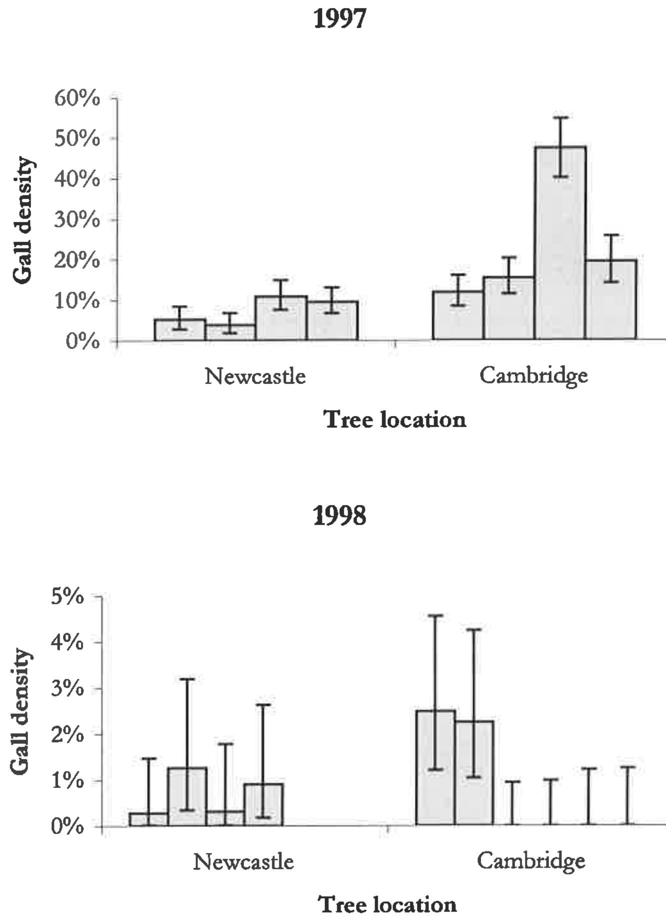


Figure 3.7 Gall densities on all trees surveyed in Newcastle-under-Lyme and Cambridge in 1997 and 1998.

Variation between trees

Differences in gall densities between trees within a poplar stand were tested. There were no significant differences between four trees in Newcastle-under-Lyme sampled in 1997, 1998 and 1999 (Kruskal-Wallis test: $\chi^2 = 2.33$, d.f. = 3, $p = 0.51$) (Figure 3.8).

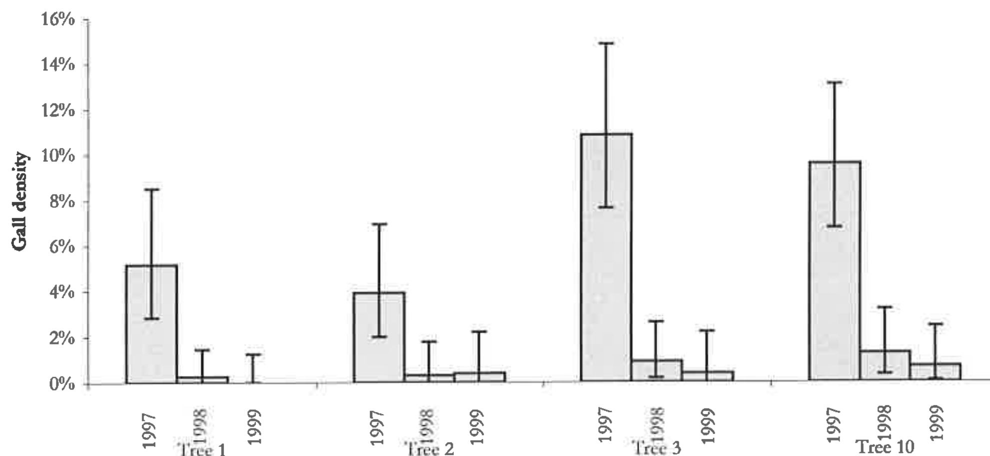


Figure 3.8 Gall densities on four trees in Newcastle-under-Lyme in 1997 to 1999. Error bars show exact 95% confidence intervals for the means.

Association between gall density and latitude

The relationship between gall density and latitude was investigated. In order to remove the effect of annual variation, the distributions of the gall density data within each year were standardised to a mean of zero and variance of one. There was no evidence for a link between gall density and latitude in the UK (Spearman's $r_s = 0.066$, $p = 0.47$) (Figure 3.9).

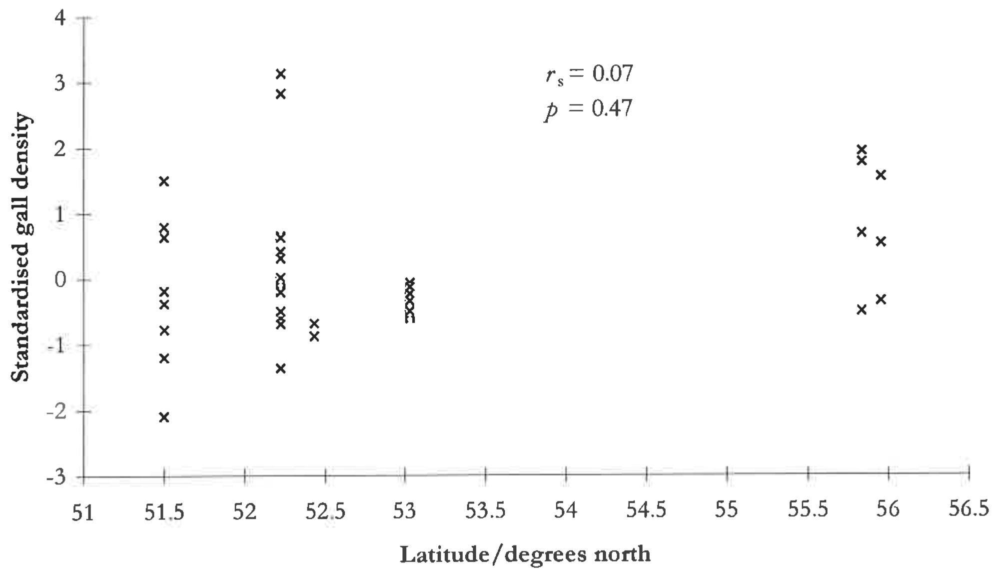


Figure 3.9 A plot of standardised gall density against latitude, showing the degree (Spearman's correlation coefficient, r_s) and significance of association (p).

Discussion

The data presented here show that *P. spyrothecae* population density, measured as the proportion of petioles occupied by galls, varied temporally, between the years from 1996 to 1999. There was also evidence that gall density can vary spatially between sites, but no evidence of variation between trees within sites.

Temporal variation

Population density more than doubled from 1996 to 1997 on the three trees in Coton and Richmond (Figure 3.6), and dropped by a factor of 42 from 1997 to 1998. Increases in aphid populations are easy to explain. Like all aphids, *P. spyrothecae* has a potentially explosive reproductive rate. A gall founded by a single aphid can produce more than 600 alate sexuparae in a year (Rhoden, 1997), each sexupara gives birth on average to five female sexuales (Foster and Benton, 1992), each sexupara can lay one egg, and each egg can hatch one foundress. Hence, an unchecked *P. spyrothecae* population could multiply each year by a factor of 3000. If the population merely doubles in one year (as it did from 1996 to 1997) the question we should ask is not 'What caused the increase?' but 'What are the factors that, on average, reduce the potential annual rate of increase 3000 times, and how does variation in these factors drive variation in numbers?'

Whatever factors are responsible, they probably operate during the half of the year when *P. spyrothecae* is outside the gall. A gall that survives for a whole season produces on average about 133 (± 10 S.E.) alates from one foundress (Nathan Pike, pers. comm.), so, to prevent a drop in numbers *during the gall phase*, as few as 1% galls need survive. The true level of survival appears to be much higher. Brändle (1999) found that at least 40%

of established *P. spyrothecae* colonies survived to the end of the season. The remainder suffered reduction or death by factors including predation and invasion by inquiline aphids (*Chaitophorus* spp.). Thus, the gall phase turns out a large net increase in numbers, which is not surprising in view of the gall's purpose as the colony's shelter. Population crashes must therefore happen at one of the stages outside the gall: as sexuparae, sexuales, eggs or foundresses. Here I consider four factors that might regulate *P. spyrothecae* numbers: competition, inbreeding, predation, and physical factors.

Competition

P. spyrothecae sexuparae and sexuales do not feed outside the gall. However, they might compete for other resources, for example the sexuparae might compete for the best sites for parturition. The sexuparae aggregate in deep cracks on the bark before giving birth to male and female sexuales, possibly to give their daughters access to most sheltered sites for egg-laying (Foster and Benton, 1992). It is possible that the number of sites offering protection is limited, leading to competition, which in turn could drive fluctuation in numbers. As each sexupara gives birth to on average 5 females (Foster and Benton, 1992), competition would have to be very severe to cause any net decrease in population. Hence, intraspecific competition for breeding sites probably could not defeat the reproductive potential of *P. spyrothecae*.

There is also potential for competition between foundresses. *P. betae* foundresses fight on petioles for gall-initiation sites (Whitham, 1979), and this behaviour has also been observed in *P. spyrothecae* (William Foster, pers. comm.). This behaviour is costly, not only to the loser, who can fall to her death, but also for the winner through prolonged exposure to predation and adverse weather conditions. As galls rarely occupy more than

10% of petioles on a tree, it might seem more practical for a foundress to look for a vacant petiole, rather than risk fighting for an occupied one. However, *Populus nigra* buds do not burst synchronously, and *Pemphigus* foundresses initiate galls only on young petioles, so only a fraction of petioles are available to the emerging foundresses (Dunn, 1960a). Hence, intraspecific competition for gall-initiation sites could reduce *P. spyrothecae* numbers between hatching and gall establishment.

P. spyrothecae also competes for gall-initiation sites with other species that form galls on *Populus nigra* petioles, such as *P. bursarius*. However, as *P. spyrothecae* overwhelmingly outnumbers other petiole gall aphids on *Populus nigra* (pers. obs.), the contest is more likely to affect numbers of the less numerous competitors than it is to significantly affect *P. spyrothecae*.

Inbreeding

In small populations, it is possible that sexuparae could be so thinly spread on the bark that their sons and daughters have no opportunity to mate with non-siblings. *P. spyrothecae* sexuales show no inbreeding avoidance, and sibling matings produce no fewer eggs than non-sibling matings (Foster and Benton, 1992). However, inbreeding significantly reduces hatching success in *Sitobion avenae* (Helden and Dixon, 1997), and *P. spyrothecae* eggs produced by sibling parents might also suffer reduced viability, or they might hatch inferior foundresses, reducing the population in the following year. My data indicate that a high proportion of foundresses are the daughters of siblings mating: the average in Cambridge from 1997 to 1999 was 61%. However, there is no evidence for a link between inbreeding and population density in the previous year (see chapter 4). In addition, if inbreeding were both the cause and consequence of population reduction, a

shrinking population could never recover. Inbreeding is therefore probably not important in regulating *P. spyrothecae* populations.

Predation

At least six species of predator have been observed eating *P. spyrothecae* sexuparae and sexuals on the bark of *Populus nigra* (Foster and Benton, 1992). The probability of being eaten on the bark is unknown, but it is likely to be much higher than in the enclosed and defended gall, where predation can affect up to 16% of galls by the end of the season (Brändle, 1999). Predation could also reduce the numbers of eggs, and in spring the foundress is exposed to predators from the moment she hatches until she is protected in her sealed gall. It is possible then that a large majority of sexuparae, sexuales, eggs and foundresses, are eaten by predators before reproducing. Annual fluctuations in *P. spyrothecae* numbers could be driven by variation in predator numbers, which in turn could be dependent on the density of *P. spyrothecae*. Dunn (1960b) proposed such a mechanism of predator-prey density-dependence to explain similar patterns of population density in *P. bursarius* from 1955 to 1958 (two years of abundance followed by two years of scarcity). However, predator-prey dynamics are notoriously difficult to unravel (e.g. Krebs et al., 1998), and data will need to be collected over many more than four years in order to establish the existence of predator-prey cycling in *P. spyrothecae* and *P. bursarius*, before it is possible to investigate its causes. If numbers of predator and prey do cycle together, *P. spyrothecae* and *P. bursarius* could become useful model species for studying predator-prey dynamics, because herbivore food quality, which is suspected of driving cycles in other systems (Krebs et al., 1995), is probably relatively stable in *Pemphigus*.

Physical factors

Competition, inbreeding, and to a lesser extent predation, are factors that vary in a species-specific manner. However, population densities in other aphids in England appear to have followed a similar pattern to that found in *P. spyrothecae* from 1996 to 1999 (Figure 3.10; Table 3.1). *P. bursarius* galls were abundant on *Populus nigra* across the Midlands and the south of England in 1997, but scarce in 1998 (Nick Miller, pers. comm.). The pattern of abundance in 1997 followed by substantially lower numbers in the following two years was also mirrored by counts of aphids of all species caught in suction traps near Bury St. Edmunds in Suffolk (Richard Harrington, IACR Rothamsted, pers. comm.). There is a close correlation between *P. spyrothecae* numbers estimated in this study and the combined numbers of all suction trapped species (Figure 3.11). However, data from more than four years will be required to demonstrate that the correlation is not spurious.

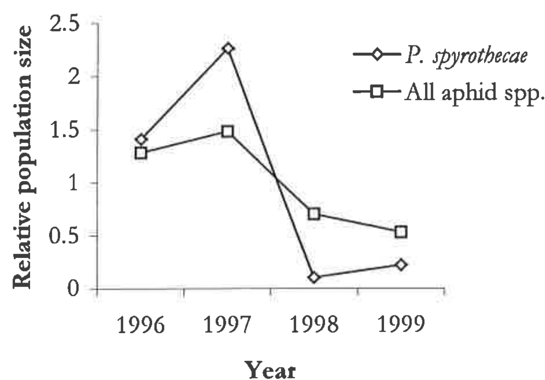


Figure 3.10 Comparison of English *P. spyrothecae* population densities (this study) with numbers of all aphids suction-trapped at Broom's Barn near Bury St. Edmunds. Each data set was standardised for comparison by dividing by the average population density from 1996 to 1997.

Table 3.1 Aphid population densities in England, average April temperatures in central England, and average winter (January to February) temperatures in central England. Data are presented from *P. spyrothecae* (this study), *P. bursarius* (Nick Miller, pers. comm.), and aphids of all species suction-trapped near Bury St. Edmunds (Richard Harrington, pers. comm.). Source of climate data: International Station Meteorological Climate Summary, Version 4.0.

Year	*P. spyrothecae*	*P. bursarius*	All spp.	April *T*/°C	Winter *T*/°C
1996	7.0% (963)	-	35455	8.5	3.4
1997	11.2% (2081)	Abundant	40987	9.0	4.6
1998	0.5% (2364)	Scarce	19349	7.7	6.2
1999	1.1% (2249)	-	14654	9.5	5.4

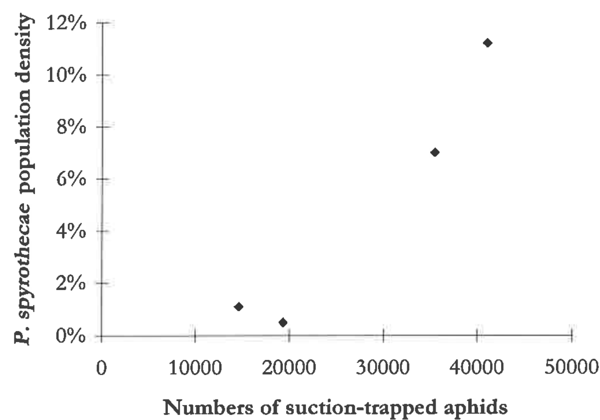


Figure 3.11 Annual population densities in *P. spyrothecae* plotted against numbers of all aphid species suction-trapped at Broom's barn from 1996 to 1999.

If *P. spyrothecae* numbers do co-vary with numbers of many other species over several years, the root of the fluctuations cannot be species-specific, but must be a factor that can affect many species over a large part of England. A candidate for such a factor is climate. For examples, a cold winter could reduce summer populations by killing overwintering eggs. However, although egg mortality over winter can be high (e.g. 84%

in *Rhopalosiphum padi*), there is no evidence that it is aggravated by harsh winters (Dixon, 1985).

Another way in which climate could reduce populations is by killing foundresses before they establish a gall. Many foundresses may fail to find a suitable petiole before they desiccate, starve or are washed or blown from the tree. This stage may account for a significant reduction in numbers, and variation in climate (e.g. temperature, rainfall, wind) might explain the variation in numbers. For example, a cold spell following hatching might dramatically reduce survival. *P. spyrothecae* hatches in April, and April 1998 was the coldest of the 1990s, 0.9 °C below the decade average (Table 3.1). This might explain the fall in numbers from 1997 to 1998. However, numbers of *P. spyrothecae* did not rise significantly following the unusually warm April of 1999 (average 9.5 °C), and numbers of suction-trapped aphids dropped still further, suggesting that variation in April temperature is not the primary cause of population fluctuations.

Temperature is also important for timing hatching in aphids that overwinter as eggs. For example, a 1 °C rise in winter temperature (January to February) correlated with a 16-day advance in the appearance of the peach-potato aphid *Myzus persicae* over the past 35 years (source: IACR, Rothamsted). An aphid that hatches early is more likely to encounter adverse weather conditions, leading to a fall in numbers. Thus, the small populations in 1998 and 1999 could have been reduced as a result of aphids hatching too soon before the arrival of warmer spring weather. The limited data on *P. spyrothecae* population sizes in 1996, when numbers were closer to 1997 levels than to 1998 or 1999, fit this theory.

Winter temperature may also affect population size by influencing egg viability. Hatching of aphid eggs can be triggered by a long period of cold (Hand, 1983), so it is possible that the proportion of *P. spyrothecae* eggs that hatch successfully is lower following warmer winters. Such a trigger mechanism could be an adaptation against hatching before the coldest part of the winter. Low numbers could be explained by the failure of a high proportion of eggs to mature following warm winters. Although global temperatures have been rising steadily for the past 10,000 years, global warming has recently accelerated, and annual minimum temperatures in central England have risen by 1.0 to 1.5 °C over the past 130 years (Parker et al., 1992). This rapid rise in winter temperatures may have defeated the ability of aphid eggs to adapt. Accordingly, the higher numbers of *P. spyrothecae* and other holocyclic aphids in 1996 and 1997 could be explained by the unusually cold winters in those years. This theory would also explain the much higher gall densities of *P. spyrothecae* in the Apennines in Italy and in British Columbia in Canada (pers. obs.), where minimum winter temperatures are lower than in the England. However, this theory is based on limited evidence, and would require more gall density data combined with experiments to show an effect of minimum winter temperatures on egg viability.

Spatial variation

Galls densities varied over a distance of 184 km between Cambridge and Newcastle-under-Lyme in 1997 but not in 1999, suggesting that gall density can vary spatially but that spatial differences are not consistent over time. The 20 trees surveyed in this study were spread over 4°30' latitude (495 km) from London to Edinburgh, covering a difference in average annual temperature of 3 °C, but no relationship between latitude and population density was found, suggesting that the variation was not caused by broad

climatic differences, although smaller scale differences may play a role. At the smallest scale tested, between neighbouring trees (< 200 m), there was no evidence of variation in gall density between trees. There are many possible explanations for the long-distance differences in population density beyond climate variation. Inbreeding was investigated as a factor, and rejected (see chapter 4). Other variables not investigated that might vary over less than 4 km include air quality, host plant quality, local insecticide treatment, disease, parasites, and predator numbers. Another possibility is that aphids move from site to site, and that apparent variation in numbers is due to mass migration. Spatial variation in abundance of other aphids, such as *Aphis sambuci*, has been attributed to large-scale population movement (Dixon, 1985). However, mass migration can be discounted, as the evidence presented in chapter 4 suggests that effective long-distance dispersal between *P. spyrothecae* populations is in the order of about one individual every three years.

Conclusion

The data presented in this chapter show that the frequency of *P. spyrothecae* galls on *Populus nigra* can vary over time and distance. No explanation for these variations was discovered, a number of hypotheses were discussed. Warm winters could reduce numbers either by inducing early hatching or by failing to trigger egg maturation, and spatial variation might be caused by local environmental differences.

CHAPTER FOUR: POPULATION GENETICS

Introduction

Studies of aphid genetic variation

Aphid genetic variation has been studied for more than 20 years, using a variety of genetic markers, and yielding a wealth of information on aphid biology (Table 4.2) (Black, 1990; Loxdale and Lushai, 1998).

Allozyme electrophoresis

Allozyme electrophoresis was the first technique used to study aphid genetic variation (May and Holbrook, 1978; Wool et al., 1978; Furk, 1979). Aphids show little allozyme diversity, and very low heterozygosities (Tomiuk and Wöhrmann, 1980; Hebert, 1987; Tomiuk, 1987;). The reasons for the lack of diversity are unclear, although candidate factors include selection (Sunnucks et al., 1997) and bottlenecks (Loxdale and Brookes, 1988; Blackman, 1990). Despite the paucity of variation, allozymes have been successfully employed in studying many aspects of aphid biology, from distinguishing between species of similar appearance (Blackman and Spence, 1992) to measuring gene flow between populations (Loxdale, 1990).

Table 4.1 Studies of genetic diversity in aphids.

Species	Molecular technique*	Reference	Findings
Acyrthosiphon pisum	Allozymes	(Suomalainen et al., 1980)	Low diversity
Acyrthosiphon pisum	Allozymes	(Simon et al., 1982)	Low diversity
Acyrthosiphon pisum	mtDNA	(Barrette et al., 1994)	Low diversity
Acyrthosiphon pisum	mtDNA	(Boulding, 1998)	No evidence for host races
Acyrthosiphon pisum	RFLP, mtDNA, bacterial DNA	(Birkle and Douglas, 1999)	Low diversity
Acyrthosiphon pisum	Allozymes	(Via, 1999)	Evidence for sympatric host races
Aphis fabae	Allozymes	(Furk, 1979)	Variation detected
Aphis gossypii	RAPD	(Vanlerberghe-Masutti and Chavigny, 1998)	Evidence for host races
Aphis gossypii	Microsatellites	(Fuller et al., 1999)	Clonal diversity
Aphis gossypii	Microsatellites	(Vanlerberghe-Masutti et al., 1999)	Variation detected
Aphis grossulariae	Allozymes	(Tomiuk and Wöhrmann, 1983)	Low diversity
Ceratovacuna nekoashi	AP-PCR	(Fukatsu and Ishikawa, 1994)	No evidence for inter gall migration
Diuraphis noxia	Allozymes, RAPD	(Puterka et al., 1993)	Genetic structure
Drepanosiphum platanoidis	Allozymes	(Wynne et al., 1994)	Genetic structure
Elatobium abietinum	RAPD, rDNA	(Nicol et al., 1998)	Low diversity in introduced species
Hyalopterus pruni	Allozymes	(Mosco et al., 1997)	Reproductive isolation, sympatric speciation
Macrosiphon rosae	Allozymes	(Tomiuk and Wöhrmann, 1981)	Allele frequency variation
Melaphis rhois	Allozymes	(Hebert et al., 1991)	Inter gall migration, genetic structure
Myzus persicae	Allozymes	(Wool et al., 1978)	Low variation
Myzus persicae	Allozymes	(Weber, 1985)	Host adaptation
Myzus persicae	rDNA	(Fenton et al., 1998)	Clonal diversity
Myzus persicae, Macrosiphum euphorbiae	Allozymes	(May and Holbrook, 1978)	Low variation
Myzus persicae, Sitobion avenae	Minisatellites	(Carvalho et al., 1991)	Clonal diversity
Myzus spp.	Allozymes	(Blackman and Spence, 1992)	Species distinction
Myzus spp.	RAPD	(Margaritopoulos et al., 1998)	Two species may be host races
Myzus spp.	Allozymes, microsatellites	(Terradot et al., 1999)	Clone-specific virus transmission ability
Pemphigus bursarius	Microsatellites	(Miller, 2000)	High variation, possible host races
Pemphigus populitransversus, P. populicaulis	Allozymes	(Setzer, 1980)	High levels of clonal mixing
Pemphigus spyrothecae	Microsatellites	(Llewellyn, 1996)	High variation, clonal mixing
Phorodon humuli	Allozymes	(Loxdale et al., 1998)	Marker linked to insecticide resistance
Rhopalosiphum maidis	Allozymes, mtDNA	(Simon et al., 1995)	Low diversity in an introduced species

Rhopalosiphum padi	Allozymes	(Loxdale and Brookes, 1988)	High gene flow
Rhopalosiphum padi	mtDNA	(Martínez-Torres et al., 1992)	High variation, genetic structure
Rhopalosiphum padi	mtDNA	(Martínez-Torres et al., 1996)	High variation, genetic structure
Rhopalosiphum padi	Allozymes	(Simon and Hebert, 1995)	Low diversity in an introduced species
Rhopalosiphum padi	RAPD, mtDNA	(Simon et al., 1996)	Sexual populations more diverse than asexual
Rhopalosiphum padi	RAPD, SCAR	(Simon et al., 1999b)	Markers linked to reproductive mode
Rhopalosiphum padi, Myzus persicae	RAPD	(Cenis et al., 1993)	Clonal diversity
R. padi, M. persicae, A. gossypii	RAPD	(Martínez-Torres et al., 1997a)	Low diversity
Schizaphis graminum	mtDNA	(Powers et al., 1989)	Genetic diversity
Schizaphis graminum	rDNA	(Shufran et al., 1991)	Genetic structure
Schizaphis graminum	Allozymes, rDNA	(Shufran et al., 1992)	High clonal diversity
Schizaphis graminum	rDNA	(Shufran and Wilde, 1994)	High clonal diversity
Schoutedenia lutea	Allozymes	(Tomiuk et al., 1991)	Genetic structure
Sitobion avenae	Allozymes	(Loxdale et al., 1985a)	Variation detected
Sitobion avenae	Allozymes	(Loxdale et al., 1985b)	Spatial and temporal genetic structure
Sitobion avenae	RAPD	(De Barro et al., 1995a)	Spatial and temporal genetic structure
Sitobion avenae	RAPD	(De Barro et al., 1995c)	Host adaptation
Sitobion avenae	Microsatellites	(Sunnucks et al., 1997)	Genetic structure
Sitobion avenae	Microsatellites	(Simon et al., 1999b)	High diversity in sexual and asexual lineages
Sitobion avenae, Metopolophium dirhodum	$(GATA)_4$	(De Barro et al., 1995b)	Microgeographic genetic variation
Sitobion fragariae	Allozymes	(Loxdale and Brookes, 1990)	Low gene flow
Sitobion misacanthi	Microsatellites	(Sunnucks et al., 1996)	Microsatellite and chromosome evolution
Sitobion spp.	mtDNA, RAPD, microsatellites	(Figueroa et al., 1999)	Interspecific differentiation
Sitobion spp.	Microsatellites, SSCP	(Wilson et al., 1999)	Clonal diversity
Various spp.	Allozymes	(Tomiuk and Wöhrmann, 1980)	Low diversity
Various spp.	Allozymes	(Wöhrmann and Hales, 1989)	High impact of selection on diversity

*mtDNA: mitochondrial DNA; RFLP: restriction fragment length polymorphism; RAPD: randomly amplified polymorphic DNA; AP-PCR: arbitrarily primed PCR; rDNA: ribosomal DNA fingerprinting; SCAR: sequence-characterised amplified region; SSCP: single-stranded conformation polymorphism.

Mitochondrial DNA

Following a decade in which allozymes were the only measure of aphid genetic variation, Powers et al. (1989) found a mitochondrial DNA restriction fragment length polymorphism in populations of the greenbug *Schizaphis graminum* using markers. Using the same technique, Barrette et al. (1994) found polymorphisms at only two of 126 restriction sites in 35 clones of *Acyrthosiphon pisum*, and Simon et al. (1995) found no variation among clones of *Rhopalosiphum maidis*, suggesting that aphid mitochondrial DNA might suffer similar lack of variation to allozymes. However, Martínez-Torres et al. (1992; 1996; 1997b) have found extensive mitochondrial DNA variation in populations of *Rhopalosiphum padi*, revealing both spatial and temporal genetic variation.

Multi-locus DNA fingerprinting with probes

Genetic markers can be divided into multi-locus techniques, which assess variation at many sites along the genome simultaneously, and single-locus techniques, which pinpoint one specific site. Application of multi-locus DNA techniques (DNA fingerprinting) to aphid populations allowed the discovery of variation in aphid populations that were monomorphic when analysed using allozyme and mitochondrial DNA markers. Carvalho et al. (1991) used minisatellite DNA fingerprinting to reveal variation within clones of *Sitobion avenae*, and demonstrate the genetic stability of *Myzus persicae* over nine generations, and De Barro et al. (1994; 1995b) used a short probe, $(GATA)_4$, to detect both geographic and host-based genetic differences between aphid clones. In recent years minisatellite and the $(GATA)_4$ probe-based markers have been superceded by techniques employing the polymerase chain reaction (PCR).

Multi-locus DNA fingerprinting using PCR

PCR-based multi-locus techniques, such as randomly amplified polymorphic DNA (RAPD) (Williams et al., 1990) and amplified fragment length polymorphisms (AFLP) (Vos et al., 1995; Mueller and Wolfenbarger, 1999), are widely employed because of their ease of use, short development time and high resolution of genetic differences.

De Barro et al. (1995a) used RAPD-PCR to show spatial and temporal genetic variation in populations of *Sitobion avenae*, Fukatsu and Ishikawa (1994) distinguished between clones of *Ceratovacuna nekoashi* using AP-PCR, and Via (1999) demonstrated host-based genetic isolation between sympatric populations of *Acyrthosiphon pisum*. However, the utility of multi-locus techniques is limited: allele frequencies cannot be measured and heterozygotes and homozygotes generally cannot be distinguished, restricting their analytical power in population genetic analyses. In addition, PCR-based multi-locus techniques can be unreliable due to the extreme sensitivity of PCR. Because multi-locus techniques are not species-specific, contaminant DNA from pathogens or symbionts can produce artefactual differences (e.g. Fenton et al., 1994).

Microsatellites

The development of microsatellites as genetic markers accessed an abundant source of variation in human populations (Litt and Luty, 1989; Tautz, 1989; Weber and May, 1989), and biologists were quick to apply them to the study of other organisms (Amos et al., 1993; Ellegren, 1992). Their first application to aphids demonstrated the considerably greater resolution of microsatellite markers compared with allozymes. Sunnucks et al. (1996) used four microsatellite loci to distinguish between and within chromosome races of *Sitobion misacanthi*, and demonstrated the absence of sex in populations of *Sitobion*

misacanthi and *Sitobion* near *fragariae*. Since this groundbreaking study, microsatellites have been applied to subjects as diverse as population genetics and sex determination. Sunnucks et al. (1997) used the same four loci to reveal high average heterozygosity (50%), low geographical but high between-host differentiation. In a highly imaginative use of microsatellites, Wilson et al. (1997) used an X-linked microsatellite to show that loss of the X during male (XO) sex determination is random in *Sitobion* near *fragariae*. In recent years, the application of microsatellites to problems of aphid biology has boomed, and they havebeen used to demonstrate genetic diversity in the green spruce aphid, *Elatobium abietinum* (Nicol et al., 1998) and the cotton aphid *Aphis gossypii* (Fuller et al., 1999; Vanlerberghe-Masutti et al., 1999). Microsatellites have also been used to distinguish species in the *Myzus persicae* complex (Hales et al., 2000; Terradot et al., 1999).

Studies of genetic variation in the genus *Pemphigus*

Genetic variation has been studied in four *Pemphigus* species. Setzer (1980) found clonal mixing in *P. populitransversus* and *P. populicaulis* using variable allozyme loci, while Miller (2000) developed five microsatellite loci for *P. bursarius*, revealing high levels of variation, and a pattern of genetic structure consistent with a high degree of gene flow between distant populations, but genetic isolation between host-specialised races. Using two of the microsatellite loci employed in this project, 96PS5 and 97PS9, Llewellyn (1996) uncovered high levels of genetic variation in *P. spyrothecae*.

P. spyrothecae

Using seven microsatellite markers, I conducted a wide-ranging study of genetic diversity in *P. spyrothecae*. Samples were collected on a range of scales, from between continents to

within trees, and over a period of three years, allowing investigation of spatial and temporal genetic variation.

Results

Genotyping of *P. spyrothecae* galls

A total of 525 *P. spyrothecae* galls collected in 1999 were genotyped at seven loci, and a further 517 collected in 1997 and 1998 were genotyped at five loci. Ninety-eight galls (9.4%) were excluded from further analysis because PCR failed at more than one locus to generate an unambiguous banding pattern. An example of a clear banding pattern, where alleles can be unequivocally identified, is illustrated in Figure 4.1. The genotypes of the remaining 944 galls were recorded and analysed. Allele frequencies at seven microsatellite loci among the samples collected in 1999 from the UK, Uppsala, Berlin, Vienna, Modena and northwestern America are presented in Table 4.2 and illustrated in Figure 4.2.

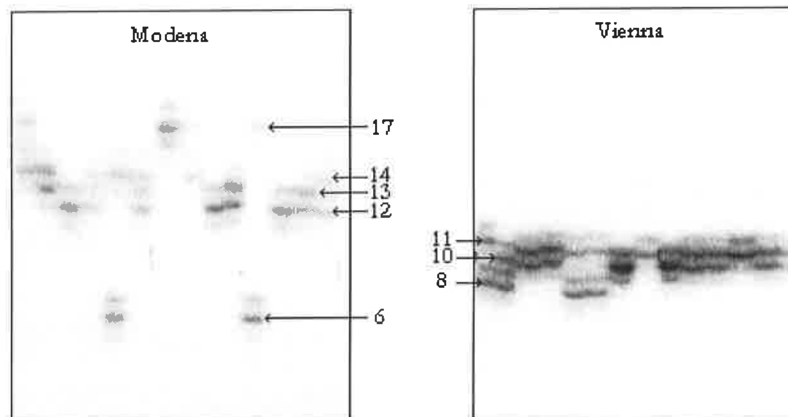


Figure 4.1 Microsatellite variation at locus 96PS20 in one tree in Modena and one tree in Vienna. Allele repeat numbers are indicated. Allele repeat numbers are indicated. Polymerase slippage during PCR amplification of microsatellites characteristically produces a ladder of progressively fainter artifactual bands at steps of one repeat length adjacent to each allele band. Generally, the most intense band can be unequivocally attributed to the allele. A potentially ambiguous pattern can be produced where two alleles of a heterozygote differ in size by one repeat (e.g. lanes 3, 4, 10, 11, 13 and 14 in the Vienna sample) causing the ladders to overlap and produce adjacent bands of similar intensity. Such heterozygotes can be distinguished from homozygotes by comparison with the consistent banding patterns of known single alleles. In the example illustrated, single alleles show one intense band flanked by two fainter bands, while heterozygotes differing by one repeat show two intense bands flanked by two fainter bands.

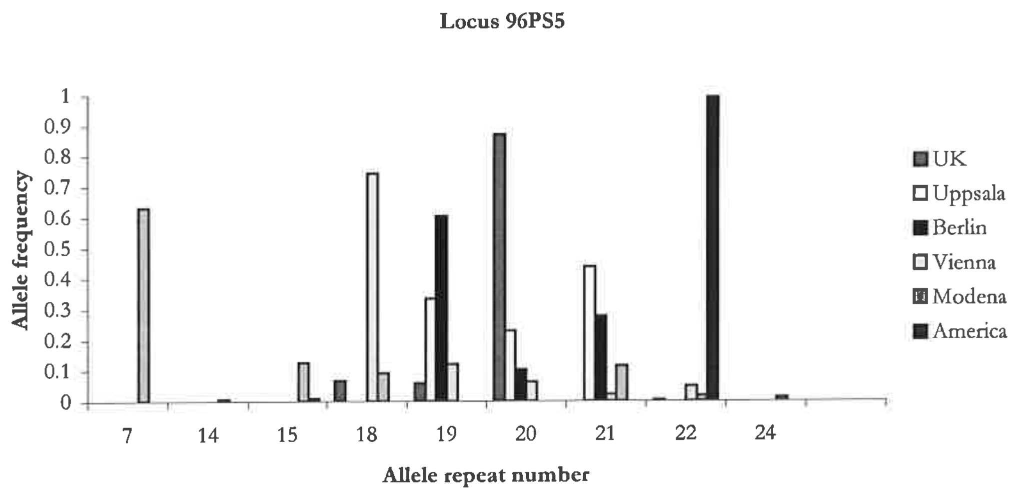
Table 4.2 Allele frequencies by locus and sampling location.

		Sampling location					
Locus	Allele	UK	Uppsala	Berlin	Vienna	Modena	America
96PS5	7	0.000	0.000	0.000	0.000	0.632	0.000
	14	0.000	0.000	0.000	0.000	0.006	0.000
	15	0.000	0.000	0.000	0.000	0.126	0.000
	18	0.068	0.000	0.000	0.743	0.092	0.000
	19	0.059	0.333	0.603	0.121	0.000	0.004
	20	0.87	0.229	0.103	0.064	0.000	0.996
	21	0.000	0.438	0.276	0.021	0.115	0.000
	22	0.003	0.000	0.000	0.050	0.017	0.000
	24	0.000	0.000	0.000	0.000	0.011	0.000
96PS20	1	0.000	0.000	0.173	0.000	0.000	0.000
	6	0.000	0.000	0.000	0.000	0.261	0.000
	7	0.000	0.479	0.365	0.021	0.000	0.000
	8	0.000	0.000	0.000	0.141	0.000	0.000
	10	0.000	0.354	0.346	0.331	0.000	0.000
	11	0.000	0.000	0.000	0.507	0.000	0.000
	12	0.000	0.000	0.000	0.000	0.403	0.000
	13	0.000	0.000	0.000	0.000	0.148	0.000
	14	0.000	0.000	0.000	0.000	0.125	0.000
	15	0.000	0.146	0.000	0.000	0.000	0.000
	16	0.000	0.021	0.115	0.000	0.000	0.000
	17	0.000	0.000	0.000	0.000	0.063	0.000
	22	0.008	0.000	0.000	0.000	0.000	0.000
	24	0.065	0.000	0.000	0.000	0.000	0.000
	25	0.789	0.000	0.000	0.000	0.000	0.991
	26	0.110	0.000	0.000	0.000	0.000	0.000
	27	0.028	0.000	0.000	0.000	0.000	0.009
97PS9	10	0.000	0.960	0.926	0.893	0.794	0.000
	11	0.000	0.000	0.000	0.071	0.000	0.000
	12	0.818	0.040	0.037	0.000	0.000	0.832
	13	0.000	0.000	0.000	0.036	0.000	0.000
	14	0.050	0.000	0.037	0.000	0.206	0.164
	18	0.131	0.000	0.000	0.000	0.000	0.005

Table 4.2 Allele frequencies by locus and sampling location (continued).

Locus	Allele	Sampling location					
		UK	Uppsala	Berlin	Vienna	Modena	America
97PS12	12	0.537	0.154	0.448	0.000	0.000	0.209
	14	0.000	0.000	0.000	0.000	0.256	0.000
	16	0.000	0.000	0.000	0.000	0.180	0.000
	17	0.000	0.000	0.000	0.000	0.233	0.000
	18	0.000	0.000	0.310	0.057	0.297	0.000
	19	0.460	0.846	0.241	0.943	0.035	0.791
	20	0.003	0.000	0.000	0.000	0.000	0.000
98PS8	8	0.000	0.280	0.155	0.448	0.067	0.000
	9	0.000	0.000	0.000	0.067	0.000	0.000
	10	0.568	0.220	0.103	0.127	0.000	0.010
	11	0.000	0.460	0.690	0.313	0.000	0.000
	12	0.017	0.000	0.000	0.045	0.361	0.000
	13	0.000	0.000	0.000	0.000	0.060	0.000
	14	0.008	0.040	0.000	0.000	0.331	0.000
	15	0.282	0.000	0.052	0.000	0.247	0.990
	16	0.011	0.000	0.000	0.000	0.000	0.000
	22	0.113	0.000	0.000	0.000	0.000	0.000
98PS12	16	0.549	0.750	0.370	0.915	0.000	0.027
	17	0.005	0.173	0.370	0.035	0.190	0.000
	18	0.000	0.000	0.000	0.021	0.238	0.000
	19	0.000	0.000	0.000	0.000	0.089	0.000
	20	0.000	0.077	0.259	0.028	0.369	0.000
	23	0.000	0.000	0.000	0.000	0.024	0.000
	24	0.429	0.000	0.000	0.000	0.089	0.973
	26	0.003	0.000	0.000	0.000	0.000	0.000
	43	0.014	0.000	0.000	0.000	0.000	0.000
98PS26	15	0.239	0.596	0.810	0.000	0.557	0.000
	16	0.741	0.154	0.052	0.000	0.443	0.991
	17	0.020	0.000	0.000	0.657	0.000	0.009
	25	0.000	0.250	0.138	0.343	0.000	0.000

Locus 96PS5



Locus 96PS20

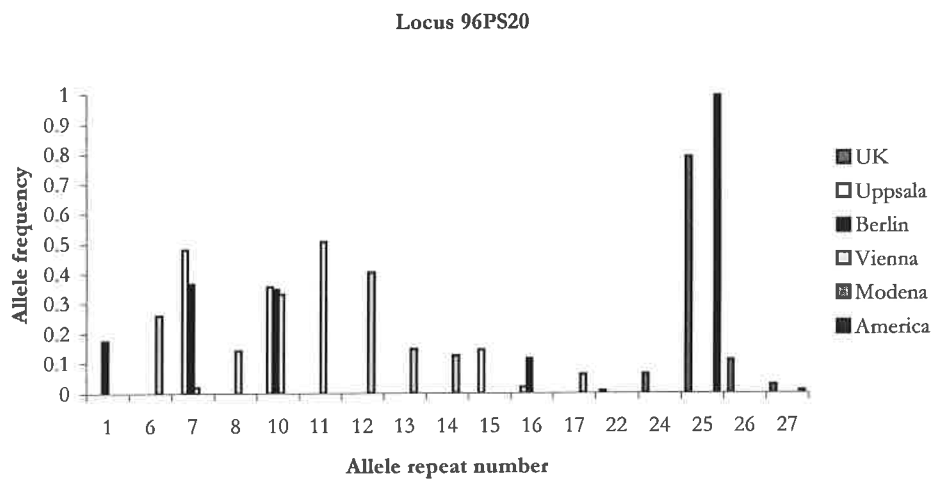
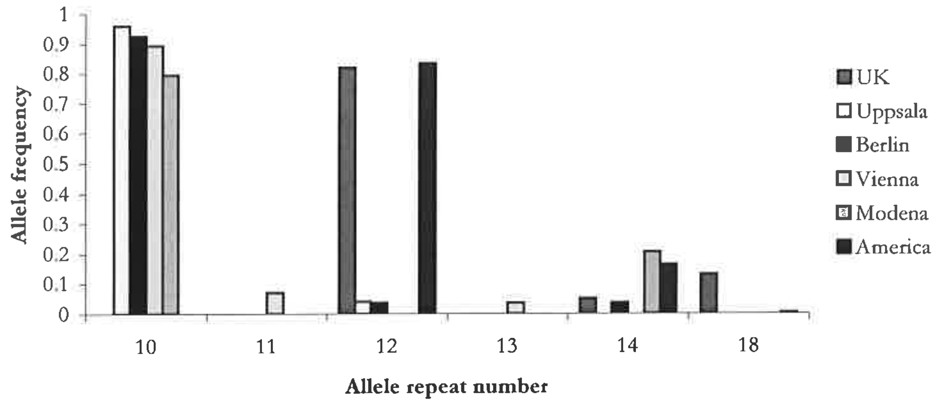
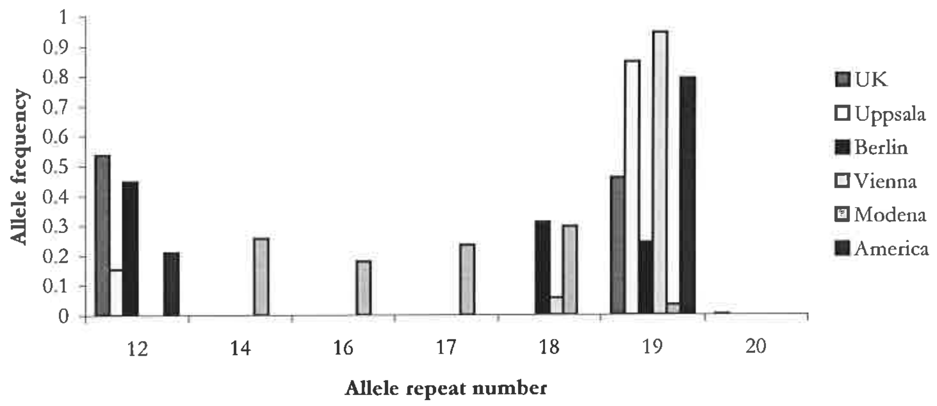
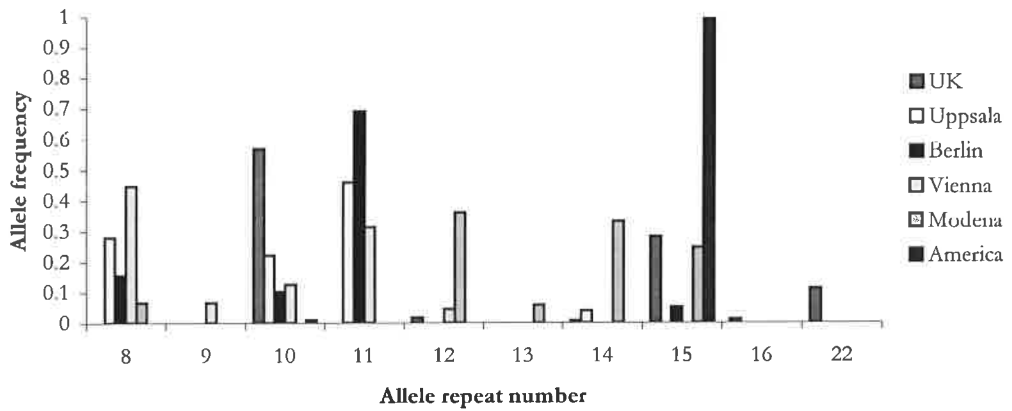
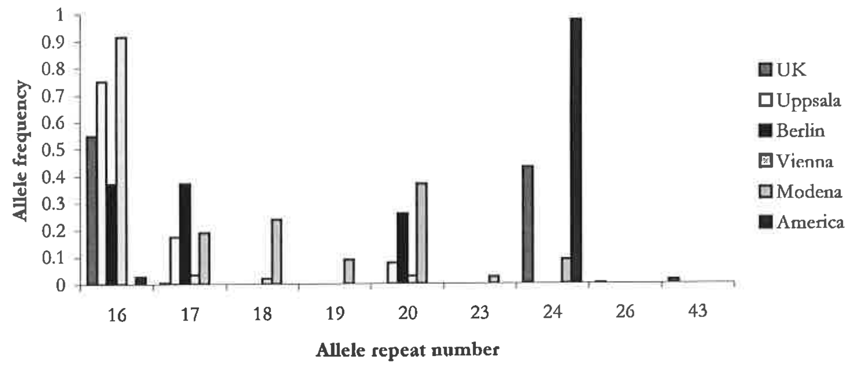


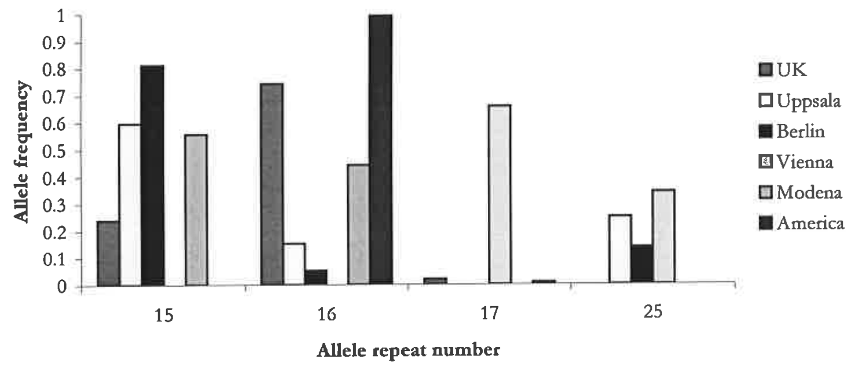
Figure 4.2 Allele frequency distributions at seven microsatellite loci in *P. spyrothecae* samples from six countries. Loci 96PS5 and 96PS20 are illustrated above, and loci 97PS9, 97PS12, 98PS8, 98PS12 and 98PS26 are shown on the following pages.


Locus 97PS9
Allele frequency
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
10
11
12
13
14
18
Allele repeat number
UK
Uppsala
Berlin
Vienna
Modena
America


Locus 97PS12
Allele frequency
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
12
14
16
17
18
19
20
Allele repeat number
UK
Uppsala
Berlin
Vienna
Modena
America


Locus 98PS8
Allele frequency
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
8
9
10
11
12
13
14
15
16
22
Allele repeat number
UK
Uppsala
Berlin
Vienna
Modena
America


Locus 98PS12
Allele frequency
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
16
17
18
19
20
23
24
26
43
Allele repeat number
UK
Uppsala
Berlin
Vienna
Modena
America


Locus 98PS26
Allele frequency
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
15
16
17
25
Allele repeat number
UK
Uppsala
Berlin
Vienna
Modena
America

Analysis of PCR reliability

The PCR success rate in the UK samples was low: only 89% of PCRs yielded an unambiguous banding pattern. In order to find the possible causes of the poor success rate, I examined 95 aphids sampled in Cambridge in 1999 that were genotyped at eight loci (the seven loci in Table 4.2 and one monomorphic locus, 98PS6). This population was chosen because the failure rate was particularly high: only 83% of PCRs yielded a product. The 17% that failed could be attributed either to the presence of null alleles, where an allele fails to amplify due to a mutation in one or both of the annealing sites (Callen et al., 1993; Ishibashi et al., 1996; Van Treuren, 1998; Vanpijlen et al., 1995), or to some non-genetic cause such as degraded DNA or the presence of PCR inhibitors. If null alleles were the sole cause of PCR failure, then the frequency of null alleles must be uniformly high across all seven loci (frequency range: 0.29 to 0.47; average: 0.40), because the failure rate was consistently high over the seven loci (range: 8% to 22%; average: 17%). (The frequency of putative null alleles was estimated as the square root of the failure rate.) This suggests that null alleles are probably not the source of the high PCR failure rate, because they are generally found in a small minority of loci (Ishibashi et al., 1996; Van Treuren, 1998; Vanpijlen et al., 1995). If null alleles were the cause of the high failure rate, then PCR failures should be associated with one or a few loci. A chi-square test provided no evidence for such an association ($\chi^2 = 11.4$, d.f. = 7, $p = 0.12$), so null alleles are unlikely to have been a significant source of PCR failure.

If PCR failure was instead related to degraded template DNA or the presence of PCR inhibitors, then PCR failures should aggregate within specific low quality DNA extractions, instead of being randomly distributed among all individuals. In order to test this possibility I compared the numbers of DNA samples that had failed at each possible

number of loci (from zero to eight) with the number expected if each sample had an equal chance of failure. Under the null hypothesis, the number of loci that failed in each of the 95 samples follows a binomial distribution where $n = 8$ and $\theta = 0.17$. The data deviated significantly from the null expectation (*G*-test: $G = 143$; d.f. = 7; $p < 0.0001$), showing that poor DNA extraction quality is likely to be the major cause of PCR failure. DNA may have degraded due to an aphid's death and decomposition prior to sample collection. Although healthy looking aphids were selected for PCR whenever possible, it is possible that DNA could degrade before visible signs of decomposition appear. Very few aphids that looked unhealthy, for example those covered by fungus, produced a PCR product. The presence of inhibitors, perhaps plant polysaccharides or pathogenic fungi, may also have contributed to PCR failures. However, experiments comparing PCR success between DNA extractions from the thorax and head, the abdomen and the whole aphid provided no evidence that exclusion of the abdomen (and therefore plant-derived gut contents) improves PCR success. However, the minimal success rate among fungus-affected aphids may be a result of fungus-derived inhibitors rather than degraded DNA. PCR trials using DNA extracted from different morphs and instars suggested that 1st-instar virginoparae perform worst and 3rd-instar sexuparae best (see chapter 5). The best evidence that inhibition is important is that the best PCR results were obtained using highly diluted DNA extractions, where each PCR contained 0.03% of an aphid.

This analysis suggests that sample quality, not null alleles, was the significant factor in PCR failure. Sample quality depends on factors that do not bias selectively neutral genotypic data, such as random variation in preservation of individual aphids, or death and decomposition prior to sampling. Thus, it seemed valid to exclude from analysis those samples that failed at two or more loci, without risk of biasing the data.

Having established that the seven loci were unlinked, and that null alleles were not present at high frequencies, the data were deemed suitable for population genetic analysis. However, loci 96PS20 and 98PS26 were excluded from further analysis of the UK populations because they had been used to genotypes samples taken in 1999 only, and would bias comparisons between years. Loci 96PS20 and 98PS26 were used in analysis involving populations outside the UK, all of which were sampled in 1999 and genotyped at all seven loci.

Microsatellite variation in the UK

P. spyrothecae in the main study area, the UK, showed moderate levels of variation at the seven microsatellite loci. The allele frequency distributions illustrate a pattern in which either one or two alleles predominate (Table 4.2; Figure 4.2). Allele numbers ranged from three to nine (average 5.1), and expected heterozygosities from 22% to 63% (Table 4.3).

Table 4.3 Numbers of alleles, sample sizes and observed and expected heterozygosities at seven microsatellite loci among galls collected in the UK from 1997 to 1999. The final column gives the significance level at each locus for the rejection of the null hypothesis of Hardy-Weinberg equilibrium (HWE). Significance levels were adjusted for eight comparisons using the sequential Bonferroni method.

Locus	Alleles	N	H_O	H_E	$p <$
96PS5	5	629	7.7%	21.8%	0.0001
96PS20	5	179	6.2%	37.1%	0.0001
97PS9	5	629	21.7%	33.0%	0.0001
97PS12	3	629	20.4%	51.4%	0.0001
98PS8	9	629	25.9%	63.4%	0.0001
98PS12	5	629	18.1%	50.6%	0.0001
98PS26	4	179	13.2%	39.6%	0.0001
All loci	5.1		16.2%	42.4%	

Tests of linkage disequilibrium

Linkage disequilibrium was tested in 25 samples collected from sites across the UK in 1997 to 1999. Significant linkage disequilibrium was detected between two loci in one sample: 98PS8 with 98PS12 in Coton in 1998 ($p = 0.0001$). No evidence for linkage was found in any of the other 24 samples. Linkage disequilibrium between the two microsatellite loci could have arisen through a variety of factors: close physical linkage on a chromosome; selection for a particular combination of alleles at loci physically linked to the microsatellites; sampling of related individuals; population subdivision; and inbreeding. Wide-ranging factors, such as physical linkage and selection, are unlikely causes because linkage disequilibrium was found in only one population in one year. A more probable explanation is that deviation from random mating in the Coton 1998 sample violated the assumption of HWE implicit in the test (Excoffier and Slatkin,

1998). Exact tests showing heterozygote deficits at both 98PS8 and 98PS12 in Coton in 1998 supported this view ($p < 0.05$).

Worldwide variation in *P. spyrothecae*

Variation in Europe

Levels of genetic variability, measured by Nei's (1987) gene diversity, differed significantly between the five cities sampled in Europe (Cambridge, Uppsala, Berlin, Vienna and Modena) (Figure 4.3). Modena was the most variable site in Europe ($p < 0.0005$), and Vienna was the least variable ($p < 0.0005$). There were no significant differences between Cambridge, Uppsala and Berlin.

Expected heterozygosities, averaged over seven loci, reflected the pattern of the gene diversities. Expected heterozygosities were highest in Modena, and lowest in Vienna (Table 4.4). Figure 4.2 illustrates the great difference in allele frequency distributions between the five European samples, and shows the high incidence of alleles unique to each area. Of the seven loci, 96PS20 shows the most striking divergence. Seventeen alleles were detected at this locus, of which 14 were unique to their sample of origin, including all five alleles in the UK, one out of four in Uppsala, one out of four in Berlin, two out of four in Vienna and all five in Modena.

Variation in America

Genetic variation was very low in the 116 *P. spyrothecae* galls collected from northwestern America. Nei's (1987) gene diversity was significantly lower than in the European samples ($p < 0.0005$) (Figure 4.3). Only three of the seven loci (97PS9, 97PS12 and 98PS12) were polymorphic (that is, where the frequency of the most common allele was

less than 0.99). The average expected heterozygosity was 0.104, much lower than any of the European samples, and only one locus showed more than two alleles (Table 4.4). The allele frequency distributions (Table 4.2; Figure 4.2) show that every one of the 15 alleles in the American sample was also present in the UK sample, while at least five American alleles were absent from each of the other four European samples (Table 4.5). The hypothesis that these data were drawn from a random distribution of shared alleles was rejected using a chi-square test ($\chi^2 = 17.3$, d.f. = 4, $p = 0.002$).

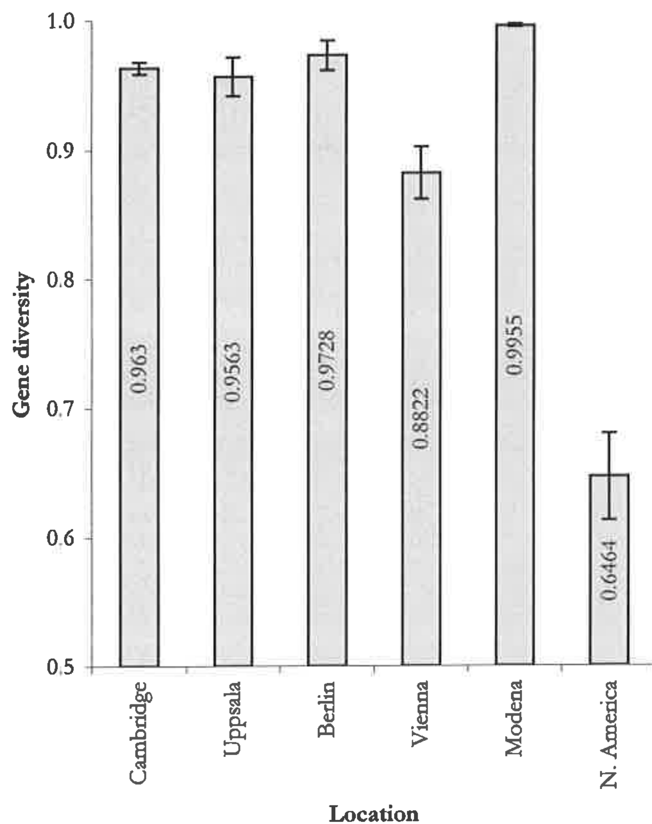


Figure 4.3 Gene diversities in samples from five European cities and northwestern America (Nei, 1987). Error bars show 95% confidence intervals.

Table 4.4 Expected heterozygosities (H_E) at seven microsatellite loci compared between samples of *P. spyrothecae* galls collected 1999 from the UK, Uppsala, Berlin, Vienna and Modena and north-western America. The number of alleles at each locus is given in brackets.

Locus	UK *n* = 179	Uppsala *n* = 26	Berlin *n* = 29	Vienna *n* = 71	Modena *n* = 88	America *n* = 101
96PS5	0.218 (4)	0.659 (3)	0.558 (4)	0.429 (5)	0.565 (7)	0.009 (2)
96PS20	0.371 (5)	0.637 (4)	0.717 (4)	0.617 (4)	0.732 (5)	0.018 (2)
97PS9	0.330 (3)	0.078 (2)	0.143 (3)	0.198 (3)	0.329 (2)	0.283 (3)
97PS12	0.514 (3)	0.265 (2)	0.656 (3)	0.109 (2)	0.763 (5)	0.332 (2)
98PS8	0.634 (6)	0.673 (4)	0.495 (4)	0.684 (5)	0.699 (4)	0.017 (2)
98PS12	0.506 (5)	0.410 (3)	0.671 (3)	0.161 (4)	0.759 (6)	0.052 (2)
98PS26	0.396 (3)	0.569 (3)	0.327 (3)	0.454 (2)	0.498 (2)	0.018 (2)
All loci	0.424 (4.1)	0.470 (3.0)	0.510 (3.4)	0.379 (3.6)	0.621 (4.4)	0.104 (2.1)

Table 4.5 The number and proportion of the 15 alleles found in the northwestern America sample that were also recorded in each European sample.

Sample area	UK	Uppsala	Berlin	Vienna	Modena
Number of alleles shared	15	8	10	6	5
Percentage of alleles shared	100%	53%	67%	40%	33%

Hardy-Weinberg equilibrium

Hardy-Weinberg equilibrium in the UK

P. spyrothecae in the UK does not form a single random-breeding population (Table 4.3). To discover at what level, if any, populations are random-breeding, HWE was tested in the years 1997, 1998 and 1999 within a hierarchy of successively smaller spatial units: within a group of neighbouring sites in Cambridge separated by up to 5 km, within sites consisting of trees separated by up to 200 m, and within single trees. Exact tests of the null hypothesis of HWE versus an alternative hypothesis of heterozygote deficiency were conducted in ARLEQUIN. Significant heterozygote deficits were detected at all spatial levels. *P. spyrothecae* collected from four sites in and around Cambridge that were

separated by up to 5 km did not form a single random breeding population in any of the years 1997 to 1999 ($p < 0.0001$). Deviations from HWE were also detected within sites around the UK (Table 4.6). At the within-trees level, 30 out of 36 trees showed a significant heterozygote deficit ($p < 0.05$) (Table 4.7). Deviation from HWE was tested in 36 trees from around the UK sampled from 1997 to 1999, of which 30 showed a significant heterozygote deficit ($p < 0.05$).

Table 4.6 Exact tests of HWE within five sites around the UK from 1997 to 1999. The maximum linear span of a sample site (S) is given as a measure of spatial scale. The number of trees (T) and galls (G) in the sample set is also given (*$p < 0.05$ after sequential Bonferroni correction).

		1997			1998			1999		
Site	S/m	T	G	*p*	T	G	*p*	T	G	*p*
Mill Pit	200	4	47	<0.001*	2	40	<0.001*	5	35	<0.001*
Coton	230	4	31	<0.001*	4	38	<0.001*	4	40	<0.001*
Hills Rd	5	-	-	-	2	24	0.18	2	18	0.006*
Richmond	70	-	-	-	2	14	<0.001*	3	23	<0.001*
Newcastle	80	2	47	0.003*	2	43	<0.001*	-	-	-

Table 4.7 Exact tests of HWE within trees at five sites from 1997 to 1999.

	In HWE ($p > 0.05$)			Heterozygote deficit ($p < 0.05$)		
Site	1997	1998	1999	1997	1998	1999
Mill Pit	0	0	0	4	2	5
Coton	0	0	1	4	4	3
Hills Rd	-	1	1	-	1	1
Richmond	-	0	1	-	2	2
Newcastle	1	1	-	1	1	-
Total		6			30	

Hardy-Weinberg equilibrium in Italy

The 88 Italian galls deviated significantly from HWE at all spatial levels. Treating the 88 galls as a single population, six of the seven loci deviated significantly from HWE (Table 4.8). Significant deviations from HWE were also detected within the three sites, all of which were out of HWE, and within the nine trees, six of which deviated from HWE (Table 4.9).

Table 4.8 Genetic variability and Exact tests of HWE in 88 *P. spyrothecae* galls collected from three sites in the vicinity of Modena (significance level following sequential Bonferroni correction for seven tests: *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$).

Locus	No of alleles	N	H_O	H_E
96PS5	7	87	0.425**	0.565
96PS20	5	88	0.636	0.732
97PS9	2	80	0.237*	0.329
97PS12	5	86	0.663**	0.763
98PS8	4	83	0.482*	0.699
98PS12	6	84	0.524***	0.759
98PS26	2	79	0.278***	0.497

Table 4.9 Exact tests of HWE in the 88 Italian galls, performed at three hierarchical levels: within the Modena (Mod) area; within the three sites, Modena centre (Cen), Gaggio (Gag) and Ravarino (Rav); and within three trees at each site (significance level following sequential Bonferroni correction for 13 tests: *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$).

Within Modena			Within sites			Within trees		
Area	*n*	Exact test	Site	*n*	Exact test	Tree	*n*	Exact test
Mod	88	p <0.0001***	Cen	26	p < 0.0004***	1	6	0.8078
						2	12	0.0047**
						3	8	0.0223
			Gag	34	p < 0.0001***	1	10	< 0.0001***
						2	14	0.585
						3	10	< 0.0001***
			Rav	28	p < 0.0001***	1	10	0.0004***
						2	9	0.1805
						3	9	0.0064**

Temporal genetic variation in UK populations

Variation in allele frequencies over time was tested between samples collected in 1997, 1998 and 1999 in two sites, Mill Pit and Coton, using exact tests. Because galls were often absent or inaccessible, Coton-4 was the only tree sampled in each year (Table 4.10). When the two sites were treated as single populations, allele frequencies did not vary significantly over a one-year period at either site, but did over the two-year period from 1997 to 1999 at Coton (Table 4.11). However, because different sets of trees were sampled in each year (Table 4.12), the validity of these tests depends on the assumption that allele frequencies did not vary between trees within each site. In order to remove this assumption, variation in allele frequencies was tested between 1998 and 1999 using

pooled genotypes from trees 4, 5 and 6 in Coton, which were sampled in both years. Here allele frequencies did vary significantly from 1998 to 1999 (p =0.04). Annual allele frequency variation was also tested within each of the trees listed in Table 4.10: none of the nine tests showed a significant difference ($p > 0.23$).

Table 4.10 Numbers of galls genotyped in two or more years in Mill Pit and Coton.

Site	Tree	1997	1998	1999
Mill Pit	1	-	20	8
	3	9	-	6
	6	9	-	11
	10	-	9	9
Coton	4	7	6	7
	5	-	13	12
	6	-	6	10

Table 4.11 Exact tests of allele frequency variation between three years in Mill Pit.

Mill Pit pooled data	1997	1998
1998	$p = 0.58$	
1999	$p = 0.58$	$p = 0.87$

Table 4.12 Exact tests of allele frequency variation between three years in Coton (*$p < 0.05$ after sequential Bonferroni correction for three tests).

Coton pooled data	1997	1998
1998	$p = 0.22$	
1999	$p = 0.01$*	$p = 0.05$

Spatial population genetic structure

Investigating the reasons for heterozygote deficits

P. spyrothecae populations deviated from HWE at all spatial scales from 1800 km (across Europe) down to a few metres (within single trees). Two potential sources of heterozygote deficits were explored. The possibility of spatial population genetic structure was investigated because *P. spyrothecae* is distributed in islands of available habitat, and therefore might be prone to genetic isolation. Inbreeding was also considered, because the laboratory observations of *P. spyrothecae* mating behaviour indicate that sib-sib mating could be frequent (Foster and Benton, 1992).

Genetic isolation between five cities across Europe

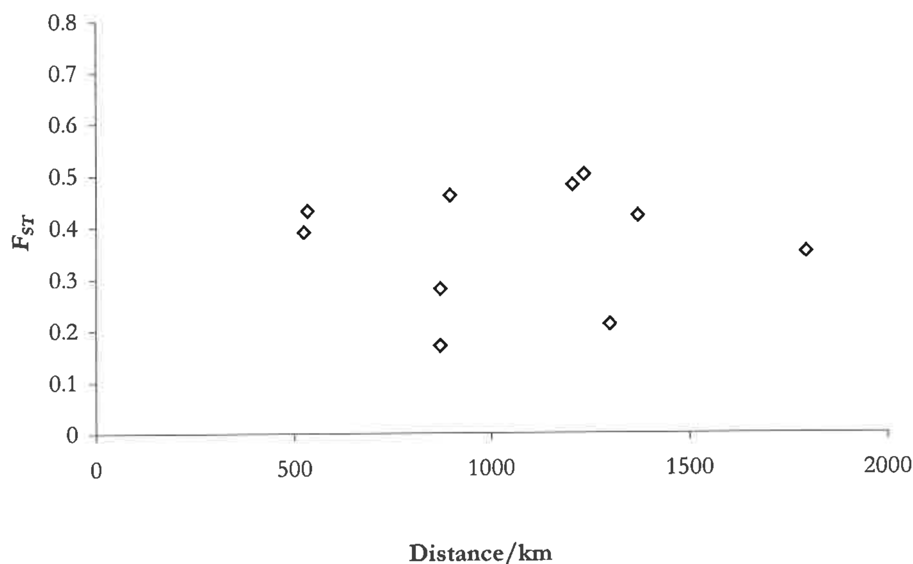
Genetic isolation between *P. spyrothecae* samples from Cambridge, Uppsala, Berlin, Vienna and Modena, estimated using pairwise F_{ST}, was strong and highly significant ($p < 0.0001$) (Table 4.13). The Cambridge population was more genetically isolated than the mainland Europe populations: F_{ST} between Cambridge and mainland Europe ranged from 0.420 to 0.496 (average $F_{ST} = 0.463$), while F_{ST} within mainland Europe ranged from 0.169 to 0.432 (average $F_{ST} = 0.305$). A Mantel test did not support a correlation between physical distance and F_{ST} ($p = 0.53$) (Figure 4.4).

Genetic differentiation between Europe and America

The northwestern America population was significantly differentiated from the five European populations (Table 4.13). Genetic isolation, estimated using F_{ST}, was greater between Cambridge and mainland Europe than between Cambridge and America, and isolation between America and the four mainland Europe populations was considerably greater than between America and Cambridge.

Table 4.13 Beneath the diagonal are pairwise F_{ST} values between *P. spyrothecae* samples from five European cities and northwestern America, all of which were significantly above zero (p < 0.0001). Above the diagonal are physical distances in km between the cities.

	Cambridge	Uppsala	Berlin	Vienna	Modena	America
n	114	26	29	71	88	101
Cambridge		1370	895	1235	1205	7675
Uppsala	0.420		870	1300	1795	7610
Berlin	0.460	0.169		525	870	8140
Vienna	0.496	0.209	0.390		535	8665
Modena	0.475	0.347	0.282	0.432		8840
America	0.400	0.737	0.733	0.747	0.609	



Distance/km

Figure 4.4 The relationship between genetic isolation (F_{ST}) and geographic distance (km) between European populations of *P. spyrothecae*.

Genetic differentiation between towns in the UK

The level of genetic differentiation was examined on a spatial scale ranging from 80 to 500 km between four towns in the UK (Cambridge, London, Newcastle-under-Lyme and Edinburgh) over three years (1997–1999) (Table 4.14). Significant differences were

detected in all six pairwise comparisons in each year, with the exception of Cambridge-Newcastle in 1999. Pairwise F_{ST} between towns varied little from 1997 to 1999 (Table 4.15). A Mantel test showed no evidence for isolation by distance between the four towns ($p = 0.12$). However, a Mantel test has little power to reject the null hypothesis when $n = 4$.

Table 4.15 Tests for the significance of genetic differentiation (F_{ST}) between *P. spyrothecae* samples from four towns in the UK (*$p < 0.05$ [after sequential Bonferroni correction for 18 tests] for the null hypothesis of panmixia between the population pairs).

		F_{ST}		
Pairwise comparison	Distance/km	1997	1998	1999
Cambridge-London	84	0.137*	0.160*	0.131*
Cambridge-Newcastle	184	0.049*	0.052*	0.035
London-Newcastle	215	0.256*	0.304*	0.234*
Newcastle-Edinburgh	330	0.533*	0.636*	0.748*
Cambridge-Edinburgh	460	0.463*	0.509*	0.452*
London-Edinburgh	530	0.639*	0.615*	0.459*

Between sites in Cambridge

To investigate population differentiation on a smaller scale (1 to 5 km), I compared four sites: three in Cambridge, Mill Pit, Hills Road and Gonville Place, and one just outside the city, Coton (Figure 2.4). Hills Rd was sampled in 1998 and 1999, while the other three sites were sampled over three years, 1997–1999. Samples sizes from Mill Pit, Hills Rd and Coton were high, ranging from 18 to 47 galls, but samples from Gonville were small due to scarcity of galls (1997: $n = 14$; 1998: $n = 5$; 1999: $n = 11$). Genetic differentiation was absent between all pairs of populations except the most distant, Hills Rd and Coton, which showed evidence of genetic isolation in both 1998 ($p = 0.002$) and 1999 ($p = 0.005$). Genetic isolation correlated significantly with geographic distance in 1999 (Mantel test; $p < 0.05$) but not in 1998 ($p = 0.12$).

Table 4.14 Genetic isolation (F_{ST}) between *P. spyrothecae* samples from four sites in or near Cambridge. The probability that there was no differentiation between the population pairs is given in brackets. (*$p < 0.05$ after sequential Bonferroni correction for 15 tests).

		F_{ST} (p)		
Pairwise comparison	Distance	1997	1998	1999
Mill Pit-Coton	3.3 km	-0.009 (0.92)	0.010 (0.30)	0.036 (0.022)
Mill Pit-Hills Rd	1.9 km	-	0.026 (0.056)	0.038 (0.082)
Mill Pit-Gonville Pl.	1.0 km	-0.004 (0.71)	-0.032 (0.93)	0.018 (0.41)
Coton-Hills Rd	5.1 km	-	0.059 (0.002)*	0.074 (0.005)
Coton-Gonville Pl.	4.2 km	0.033 (0.20)	0.004 (0.62)	0.063 (0.054)
Hills Rd-Gonville Pl.	1.3 km	-	0.010 (0.38)	0.022 (0.33)

Genetic differentiation was tested between the three sites within Cambridge, Mill Pit, Hills Road and Gonville Place, using exact *G*-tests (not assuming random mating within sites) in FSTAT (Goudet, 1995; 1999; 1996). Population differentiation was detected between Mill Pit and Hills Road in 1999, but not between either of these two sites and Gonville Place (Table 4.15). Mill Pit and Hills Road were also significantly different in 1998 ($p = 0.002$). The Gonville Place sample from 1998 contained only five galls, so it was not included in testing.

Table 4.15 Exact *G*-tests of population differentiation between three sites within Cambridge in 1999. The number of galls in each sample is given in brackets.

	Mill Pit	Hills Road
Hills Road	0.01 (18)	
Gonville Place	0.35 (12)	0.10 (41)

Spatial genetic structure between trees within sites

Population differentiation was tested at an even finer scale: 1 to 200 m, between trees within the same site. Data taken from five sites over three years were analysed (Table

4.16). The number of galls taken from each tree ranged from five to 20. Out of 43 pairwise tests of genetic differentiation between trees, four yielded F_{ST} values greater than zero at the significance level of $\alpha = 0.05$, only two more than the number expected under the null hypothesis of no differentiation. This pattern is exemplified by five trees sampled at Mill Pit in 1999 (Table 4.17; Figure 4.5), where none of the ten pairwise tests of F_{ST} proved significant. Hence, there was no evidence for genetic isolation between trees separated by less than 200 m.

Table 4.16 Numbers of pairwise tests of population differentiation made between trees within five sites sampled from 1997 to 1999. The number of comparisons where F_{ST} is significantly greater than zero is given in brackets ($p < 0.05$).

	No. of comparisons (no. of significant differences)		
Site	1997	1998	1999
Mill Pit	6 (1)	1 (0)	10 (1)
Coton	6 (0)	6 (1)	6 (0)
Hills Rd	-	1 (0)	1 (0)
Richmond Park	-	1 (0)	3 (1)
Newcastle	1 (0)	1 (0)	-

Table 4.17 Tests for pairwise genetic isolation (F_{ST}) between *P. spyrothecae* samples collected in 1999 from five trees at Mill Pit, Cambridge. No F_{ST} values were significantly different from zero after Bonferroni correction for ten tests. Above the diagonal are distances in metres between the sites.

	Tree 1	Tree 3	Tree 6	Tree 9	Tree 10
n	8	6	11	7	9
Tree 1		69	57	58	101
Tree 3	-0.007		58	27	49
Tree 6	0.030	0.027		39	60
Tree 9	0.002	-0.015	0.076		67
Tree 10	0.083	0.111	0.152	0.003	

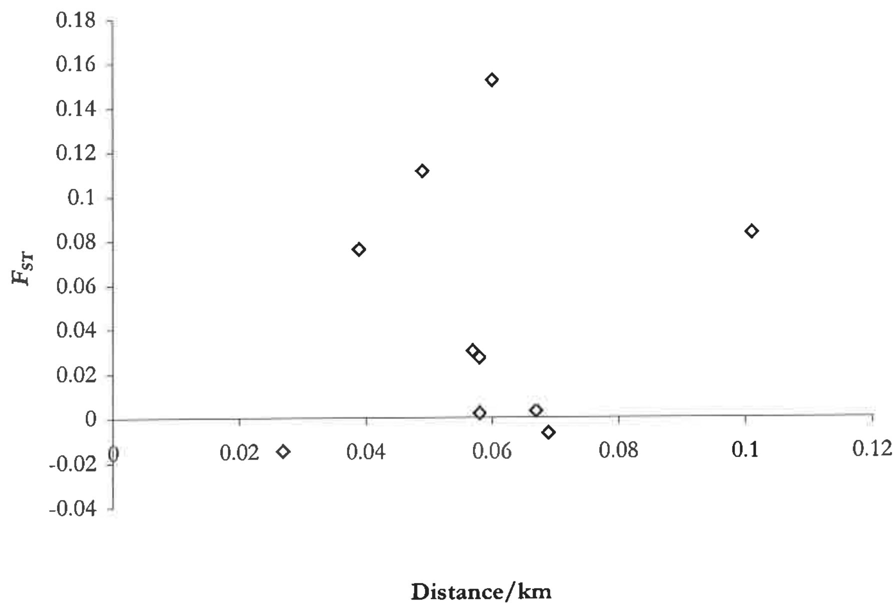


Figure 4.5 Plot of genetic isolation (F_{ST}) against geographic distance (km) between samples of *P. spyrothecae* from five trees at Mill Pit, Cambridge.

Analysis of molecular variance in the UK populations

There are two possible explanations for the failure to detect a pattern of spatial genetic structure between the trees within these five sites. Either there is no pattern, and the populations of *P. spyrothecae* in these sites are panmictic, or the pattern was too subtle to detect due to diminished sample sizes at smaller spatial scales. The deviations from HWE in the UK populations were investigated using ANOVA. The problem of small sample sizes was circumvented by using AMOVA to investigate the degree and significance of genetic variance between trees within these five sites from 1997 to 1999.

The AMOVA table shows the variance components (and associated estimates of *F*-statistics) in the five sites due to differences between sites (F_{CT}), between trees (F_{SC}),

between individuals (F_{IS}, the inbreeding coefficient) and within individuals (F_{IT}) (Table 4.18). (The between-sites variance component has no biological significance here, because it conflates genetic isolation between populations from 2 km to more than 200 km apart.) Genetic isolation between trees within sites, estimated by F_{SC}, showed no consistent pattern over the three years tested: it was low and non-significant in 1997 and 1998, but in 1999 it was significantly greater than zero. The between-individuals variance component, which estimates inbreeding, is responsible for the largest deviation from HWE.

Table 4.18 AMOVA at four hierarchical levels in samples collected from five locations (Mill Pit, Coton, Hills Rd, Richmond and Newcastle-under-Lyme) from 1997 to 1999. Variance components (Var), *F*-statistics, and *p*-values are tabulated by year and spatial level. An estimates of the selfing fraction, *S*, were calculated from the inbreeding coefficient using $S = 2F_{IS}/(1+F_{IS})$ (Hartl and Clark, 1997).

Source of variation			Year		
Between	Within		1997	1998	1999
Sites		Var	3.1%	7.9%	7.2%
		F_{CT}	0.031	0.079	0.072
		p	0.045	< 0.0001	0.0003
Trees	Sites	Var	0.7%	1.7%	4.2%
		F_{SC}	0.007	0.018	0.045
		p	0.52	0.23	0.01
Individuals	Trees	Var	46.4%	32.5%	43.9%
		F_{IS}	0.483	0.359	0.489
		S	0.651	0.528	0.657
		p	< 0.0001	< 0.0001	< 0.0001
	Individuals	Var	49.7%	57.9%	45.3%
		F_{IT}	0.503	0.421	0.547
		p	<0.0001	<0.0001	<0.0001

Investigation of within-trees homozygote deficits

The spatial level responsible for the greatest degree of deviation from HWE in the UK samples was between individuals within trees (Table 4.18). The homozygote deficit within trees could have been caused either by true inbreeding (probably in the form of selfing) or by spatial population structure within trees. These two possibilities are examined below.

Tests of population genetic structure within trees

Spatial genetic structure within trees was tested using samples collected in 1998 from Tree 10 at Newcastle-under-Lyme (n = 18) and Tree 2 at Hills Road (n = 18) (Table 4.19). Tests of pairwise F_{ST} and exact tests revealed no evidence for spatial structure within trees, either between samples on either side of a tree, or at different heights from the ground.

Table 4.19 Pairwise F_{ST} and G-tests between samples of *P. spyrothecae* collected from spatially separate areas within trees.

Site	Tree no.	Spatial division	Sample 1 (n)	Sample 2 (n)	F_{ST} (p)	G-test p-value
Newcastle	10	Bearing	70–90° (11)	190–300° (7)	-0.013 (0.053)	0.29
Hills Rd	2	Bearing	90–170° (9)	180–270° (9)	-0.001 (0.52)	0.60
Hills Rd	2	Height	1.0–1.7 m (9)	1.8–2.4 m (10)	0.040 (0.20)	0.32

Inbreeding within trees

The high degree of inbreeding found within trees using AMOVA was examined by calculating selfing rate estimates (S) in 39 trees sampled in the UK from 1997 to 1999. The average sample size was 9.5 galls. The distribution of S over the 39 trees is illustrated in Figure 4.6.

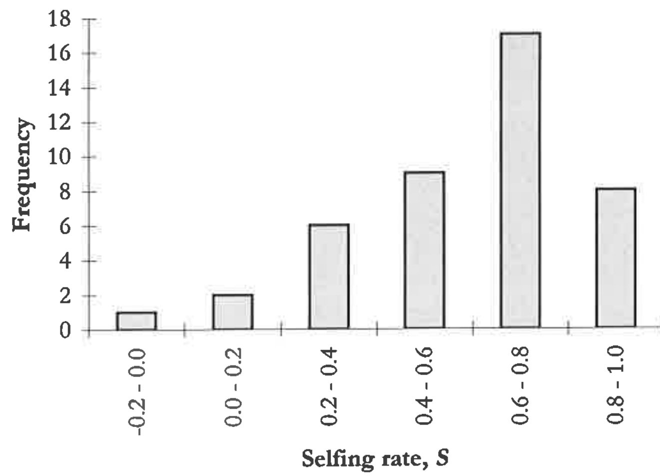


Figure 4.6 Distribution of selfing rate estimates (S) over 39 trees.

Figure 4.7 shows the distribution of S over five loci within the 12 of the 39 trees within which all five loci were polymorphic. In most trees S was consistently high over the five loci, showing that the homozygote deficits are not locus-specific artefacts (e.g. null alleles).

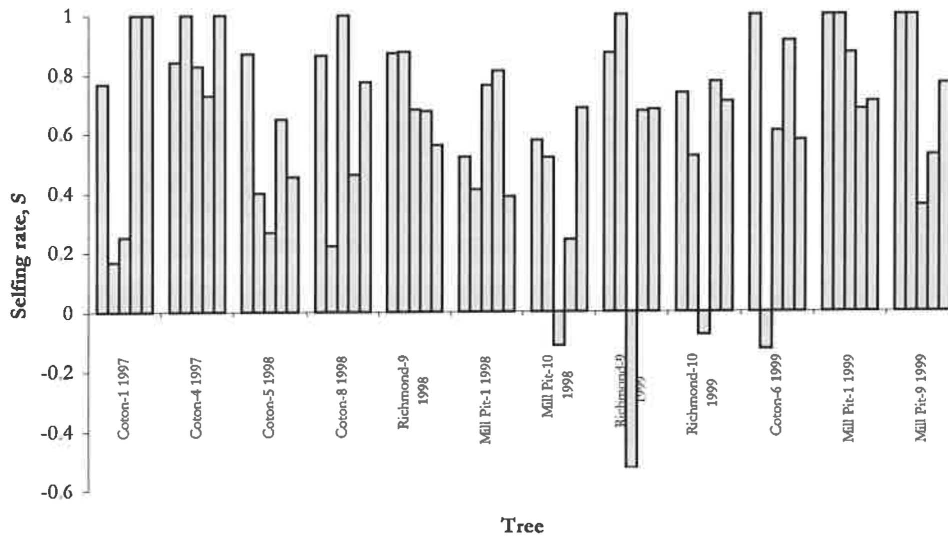


Figure 4.7 Selfing rate estimates (S) at five loci within 12 trees sampled in the UK from 1997 to 1999. Locus-by-locus selfing rate estimates are given within each tree, in the order 96PS5, 97PS9, 97PS12, 98PS8, 98PS12.

Investigation of a link between population density and selfing rate

A potential factor in determining the level of selfing in *P. spyrothecae* populations is the density of sexuales on the bark at the end of the previous summer, which is likely to be closely related to the density of galls on the petioles. However, there was no evidence of correlation between the selfing rate, S, in one year and population density in the preceding year ($r = 0.24, p > 0.05$) (Figure 4.8).

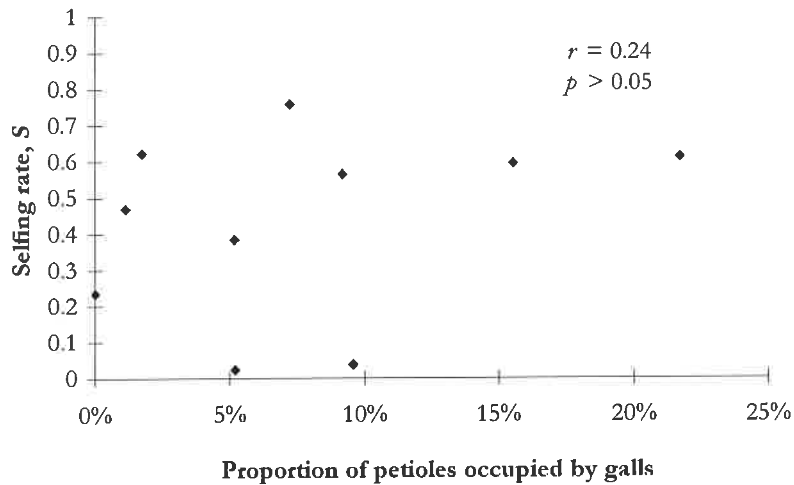


Figure 4.8 A plot of selfing rate, *S*, against population density in the preceding year in ten trees.

Spatial genetic structure in the Italian populations

Analysis of molecular variance

AMOVA showed evidence for hierarchical population genetic structure and deviation from random mating in the Italian populations (Table 4.20). Significant genetic structure was detected both between sites (F_{CT} = 0.058, p = 0.0001) and between trees within sites (F_{SC} = 0.028, p = 0.006). The degree of inbreeding within trees in Italy was high (F_{IS} = 0.194; p < 0.0001), but lower than within trees in the UK, corresponding to a selfing rate of 48%.

Table 4.20 Hierarchical variance components at four levels, associated *F*-statistics and *p*-values for non-differentiation, calculated by averaging locus-by-locus AMOVA.

Level	Variance component	*F*-statistic	*p*
Between sites	5.8%	F_{CT} = 0.058	0.0001
Between trees	2.7%	F_{SC} = 0.028	0.006
Between individuals	17.8%	F_{IS} = 0.194,	< 0.0001
Within individuals	73.8%	F_{IT} = 0.262	< 0.0001

Pairwise population comparisons in Italy

The population structure revealed in the Italian samples using AMOVA was investigated in detail by performing pairwise exact tests for differences in allele frequency between populations. Significant differences in allele frequencies were found between Ravarino and both Modena and Gaggio ($p < 0.0001$), but not between Gaggio and Modena ($p = 0.075$) (Table 4.21). Estimates of F_{ST} confirmed the genetic isolation of the Ravarino site relative to Gaggio and Modena (Table 4.21). Figure 4.9 shows the close but non-significant correlation between genetic and geographical isolation.

Table 4.21 Below the diagonal are estimates of genetic isolation (F_{ST}) between Gaggio, Modena and Ravarino. Above the diagonal are geographical distances (km).

	Gaggio	Modena	Ravarino
n	34	26	28
Gaggio		7.3	12.1
Modena	0.012 (-0.006, 0.029)		16.4
Ravarino	0.092 (0.039, 0.146)	0.109 (0.068, 0.155)	

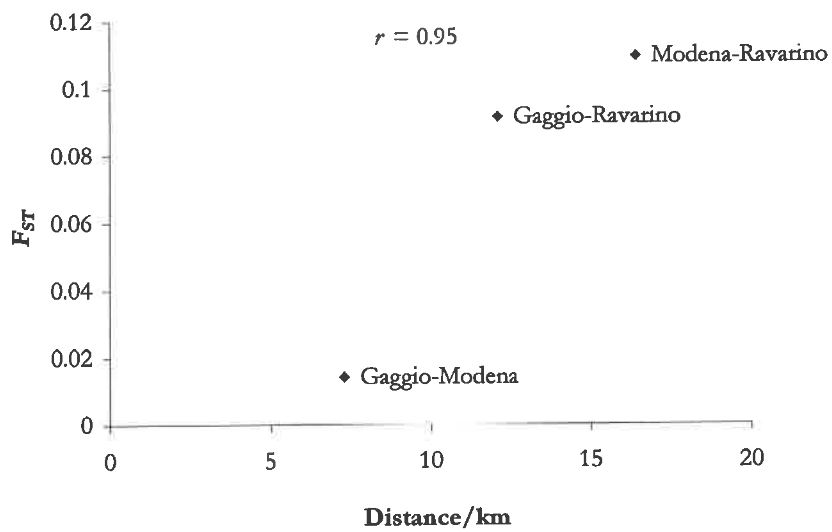


Figure 4.9 A plot of genetic isolation (F_{ST}) against geographical distance (km) between Gaggio, Modena and Ravarino.

Pairwise comparisons between trees

Evidence for spatial genetic structure between trees was found in Gaggio and Modena but not in Ravarino (Table 4.23).

Table 4.23 Pairwise estimates of genetic isolation (F_{ST}) between three trees at each of three sites Gaggio, Modena and Ravarino (*p < 0.05 following Bonferroni adjustment for three tests in each table).

Gaggio	Tree 1 (n = 10)	Tree 2 (n = 14)
Tree 2 (n = 14)	0.052*	
Tree 3 (n = 10)	0.029	0.011

Modena	Tree 1 (n = 6)	Tree 2 (n = 12)
Tree 2 (n = 12)	0.028	
Tree 3 (n = 8)	0.210*	0.085*

Ravarino	Tree 1 (n = 10)	Tree 2 (n = 9)
Tree 2 (n = 9)	-0.007	
Tree 3 (n = 9)	-0.007	-0.012

Inbreeding in the Italian sample

AMOVA identified a high overall level of inbreeding within trees in the Italian sample. In order to investigate spatial variation in levels of inbreeding, inbreeding coefficients (F_{IS}) and selfing rate estimates (S) were calculated within each of the nine trees (Table 4.24). Levels of inbreeding were significantly greater than zero in five trees, and varied widely among the nine trees.

Table 4.24 Inbreeding coefficients (F_{IS}) and associated selfing rate estimates (S) within nine trees in Gaggio, Modena and Ravarino, and the probability that they were derived from random mating populations (p) (*$p < 0.05$ after sequential Bonferroni correction for nine tests).

Site	Tree	n	F_{IS}	S	p
Gaggio	1	10	0.398	0.569	< 0.001*
	2	14	0.032	0.062	0.355
	3	10	0.277	0.434	0.001*
Modena	1	6	-0.058	-0.123	0.653
	2	12	0.196	0.328	0.008*
	3	8	0.155	0.268	0.084
Ravarino	1	10	0.262	0.415	0.005*
	2	9	0.049	0.093	0.311
	3	9	0.270	0.425	0.004*

Isolation by distance

Genetic isolation by distance was detected only once in the European populations, in Cambridge in 1999, although it was investigated at a variety of spatial scales, suggesting that distance is not an important factor in distributing genetic variation within the scales examined. However, when data from all spatial scales are pooled, a clear trend towards isolation by distance emerges (Figure 4.10).

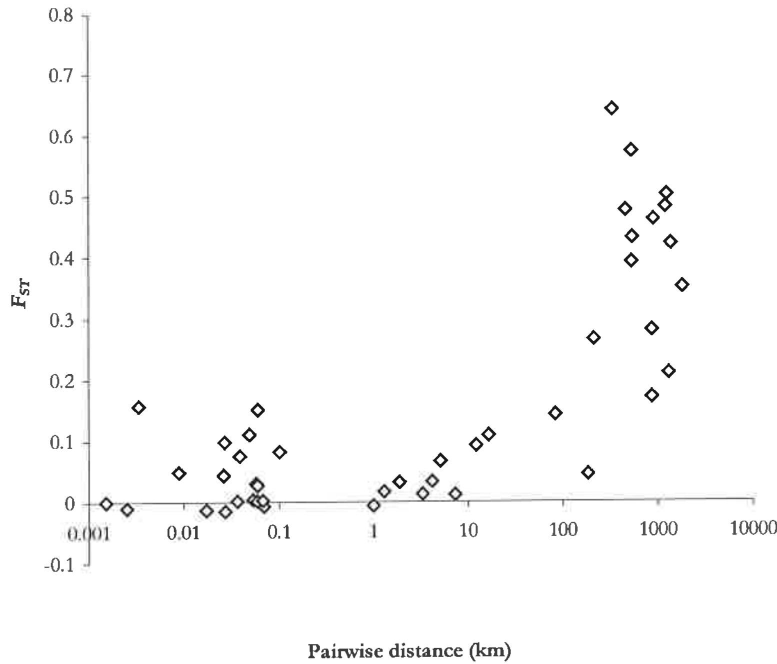


Figure 4.10 The relationship between genetic isolation (F_{ST}) and geographic distance in European populations of *P. spyrothecae.* The plot shows pairwise F_{ST} between European cities (525–1795 km), UK towns (84–530 km), poplar stands in Modena (7–16 km) and Cambridge (1–5 km apart), individual trees in Cambridge sites (3–100 m), and areas within individual trees (1.5–2.5m).

Spatial population structure in northwestern America

Population differentiation between the six northwestern American sampling sites was assessed by calculating and testing pairwise F_{ST}. Only three loci (97PS12, 97PS9 and 98PS12) were sufficiently polymorphic to contribute to this analysis (see Table 4.2). Significant differentiation was found between sites (Table 4.25), but there was no evidence for isolation by distance (Mantel test; $p = 0.97$). This pattern is exemplified by the Hell's Gate site, which was the most genetically isolated but also the nearest site to its five neighbours. The extreme genetic isolation of Hell's Gate was illustrated by comparing the allele frequencies at locus 97PS12 between the six American sites (Figures 4.11 and 4.12). Two alleles were found at this locus in the northwestern America

samples: 12 and 19. Allele 19 was the most frequent in all the samples except Hell's Gate, where it was absent from the seven individuals genotyped.

Table 4.25 Genetic isolation between *P. spyrothecae* samples from six sites in nothwestern America (pairwise F_{ST}, beneath the diagonal) (*$p < 0.05$ after sequential Bonferroni correction). Above the diagonal are physical distances (km) between the sites.

	Hope	Hell's Gate	William's Lk.	Pr. George	Radium	N. Idaho
n	37	7	8	24	9	16
Hope		70	310	510	400	350
Hell's Gate	0.791*		270	470	390	390
William's Lk.	0.279*	0.393*		160	450	520
Pr. George	0.0478	0.613*	0.101		570	690
Radium	0.038	0.878*	0.283*	0.140		180
N. Idaho	0.184*	0.355*	-0.006	0.014	0.214	

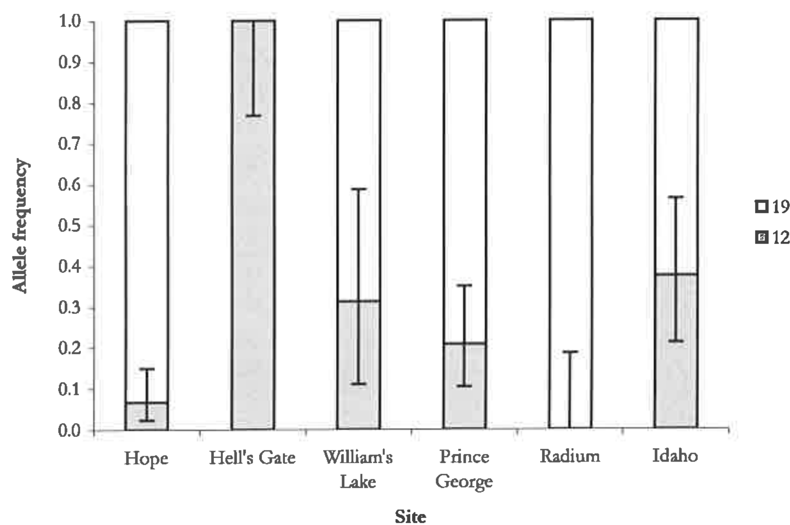


Figure 4.11 Allele frequencies at locus 97PS12 in six sites in northwestern America. Error bars show 95% confidence intervals.

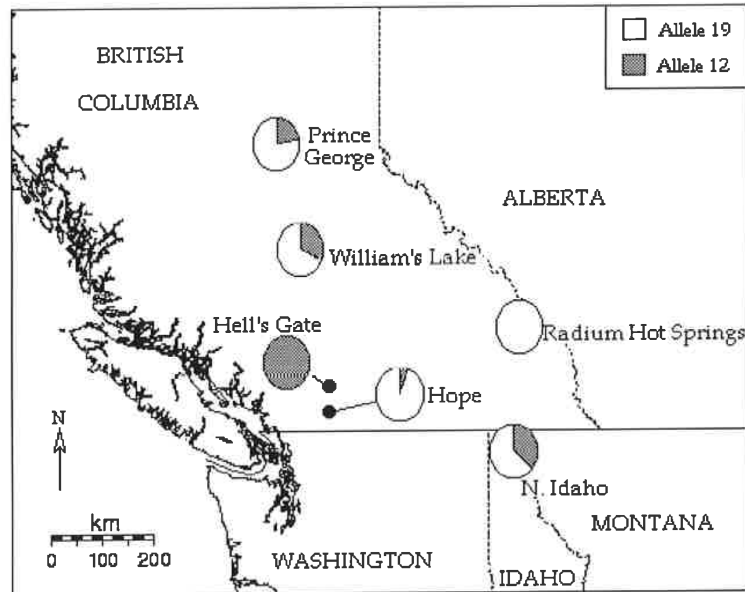


Figure 4.12 Map showing the spatial distribution of alleles 12 and 19 at locus 97PS12 among six sites in northwestern America.

Discussion

Microsatellite variation in aphids

Most studies of genetic variation in aphids have used allozyme analysis (see Table 4.1). The most striking feature of these studies has been the lack of variation discovered: the average expected heterozygosity was only 1.5% compared with 10% in other insects (Simon and Hebert, 1995). By contrast, analyses using microsatellites have found high levels of genetic variation in aphid populations. The results presented here show considerable microsatellite variation in European populations of *P. spyrothecae* (the much less variable North American sample is a special case and is discussed separately). The average expected heterozygosity (H_E) ranged from 38% in Vienna to 62% in Modena, approximately covering the range found in other aphid species that have been studied using microsatellites (Table 4.26). The least variable samples in Europe came from Vienna (average H_E = 38%) and the UK (average H_E = 42%). The relatively low variation in the UK is particularly striking given that the microsatellites were selected for their variability in UK populations, so the amount of variation may even have been exaggerated relative to the other European populations.

Table 4.26 Average expected heterozygosities (H_E) in microsatellite studies of five aphid species, *P. spyrothecae* (Ps), *P. bursarius* (Pb), *Aphis gossypii* (Ag), *Sitobion avenae* (Sa) and *Metopolophium dirhodum* (Md).

Species	Sample source	No. loci	Aver. H_E	Reference
Ps	UK	7	42%	This study
Ps	Uppsala, Sweden	7	47%	This study
Ps	Berlin, Germany	7	51%	This study
Ps	Vienna, Austria	7	38%	This study
Ps	Modena, Italy	7	62%	This study
Ps	NW America	3	10%	This study
Pb	England	5	46%	(Miller, 2000)
Ag	World-wide	8	68%	(Vanlerberghe-Masutti et al., 1999)
Sa	France	5	57%	(Simon et al., 1999a)
Sa	South England	4	50%	(Sunnucks et al., 1997)
Sa	UK	4	60%	Kate Llewellyn, pers. comm.*
Md	UK	1	82%	Kate Llewellyn, pers. comm.*

*Data reproduced by permission of IACR, Rothamsted.

There are a number of factors potentially responsible for the differing levels of microsatellite variation across Europe. *P. spyrothecae* is entirely dependent on its host, *Populus nigra*, and the past and present distribution of the various subspecies and cultivars of the tree must have played a major role in determining the current distribution of genetic variation in the aphid. The high variation in Italy compared with the four populations north of the Alps (UK, Uppsala, Berlin, Vienna) could be explained by post-glacial recolonisation of Northern Europe by *Populus nigra* and *P. spyrothecae*, either from local refugia or by migration from the south. Either scenario should result in a cline of decreasing variation from south to north in the regions once covered by ice, and

considerably greater variation in areas such as Italy that escaped ice cover. However, this theory cannot account for the level of variation in Uppsala being similar to the rest of Northern Europe because Sweden was never recolonised by *Populus nigra* (Figure 3.1), which suggests that the only host of *P. spyrothecae* in Sweden is recently introduced *Populus nigra* var. *italica*. If post-glacial recolonisation were the major factor determining the distribution of microsatellite variation in Europe, then the UK would be genetically closer than Sweden to the mainland European population. Instead, the UK population was the most genetically isolated in Europe, suggesting that events since the spread of *Populus nigra* var. *italica* have had a greater influence on the distribution of genetic variation in Europe.

The significantly higher genetic variation in Italy may also be a result of other factors. For example, populations in Italy may be large compared with northern Europe. Population densities on Italian trees are generally much higher than in the UK (pers. obs.), and Lombardy poplars appear to be much more frequent in Italy (Daniele Dallai pers. comm., and pers. obs.).

Spatial population structure

Isolation by distance

The *P. spyrothecae* populations analysed here follow a pattern of isolation by distance over a broad range of spatial scales. Using F_{ST} as a measure of isolation, UK populations separated by more than 300 km showed extreme and highly significant isolation ($F_{ST} > 0.4$), while populations within 300 km of each other showed less differentiation ($F_{ST} \leq 0.3$), though still significant. Population pairs closer than 4 km were not significantly

differentiated ($F_{ST} \leq 0.05$). However, when examined within spatial scales, the trend of isolation by distance breaks down. For example, in the years 1997 to 1999 the Cambridge population was genetically closer to Newcastle-under-Lyme, 184 km distant, than to Richmond Park in London, 84 km away. Other inversions of isolation by distance were noted among the four populations sampled on mainland Europe. Isolation between Uppsala and Berlin (870 km apart, $F_{ST} = 0.169$) was less than between Berlin and Vienna (525 km apart, $F_{ST} = 0.390$). The degree of genetic isolation between Uppsala and Berlin was comparable with that between London and Newcastle-under-Lyme, less than a quarter of the distance apart (84 km apart, $F_{ST} = 0.131$ in 1999).

Long-range gene flow

As stated above, distant populations pairs tended to yield high values of F_{ST}, suggesting that long-range gene flow in *P. spyrothecae* is rare. An estimate of the effective number of migrants per generation, N_em, can be calculated from F_{ST} (Reynolds et al., 1983). For example, the nine pairwise F_{ST} values between Cambridge, Newcastle-under-Lyme and Edinburgh in 1997 to 1999 ranged from 0.452 to 0.639, which translates to a N_em range of 0.30 to 0.14. In other words, there is one effective migration between these distant populations every three to seven years. However, it is probable that effective migration over such distances by *P. spyrothecae* is rarer even than these tiny estimates suggest, because the model on which N_em is based assumes that there are no stepping stone populations between the two focal populations, an assumption that is violated here. Hence, the finding of rare effective migrations over distances greater than 300 km is likely to be an artefact of gene flow caused by more frequent migrations between a network of local populations.

Medium-range gene flow

Available habitat for *P. spyrothecae*, in the form of black poplars, is distributed over England in patches that are typically about 5 km apart and seldom isolated by more than 10 km (see chapter 3). Hence, the ability to disperse over 5–10 km is likely to have been a major factor influencing genetic isolation between British populations of *P. spyrothecae*.

How much gene flow is there between *P. spyrothecae* populations over distances in this range? Out of the four sites studied in Cambridge, the only two that showed significant differentiation were the furthest apart, Coton and Hills Road (5.1 km apart). The degree of genetic isolation between the populations was consistent from 1998 ($F_{ST} = 0.059$) to 1999 ($F_{ST} = 0.074$), equivalent to three to four effective migrants per year. The only other populations sampled at this scale were the three around Modena, 7.3–16.4 km apart. They showed a similar degree of isolation ($F_{CT} = 0.058$), which equates to four effective migrants per year.

Unfortunately, the usefulness of the Nm statistic is limited by the fact that the only method for estimating confidence intervals gives biased results (Slatkin and Barton, 1989). However, the F_{ST} estimates from which Nm was calculated were consistent between populations separated by 5–15 km, suggesting that the F_{ST} estimates are reliable, and that the average number of annual effective migrants over such distances is close to three to four. However, genes could flow between Coton and Hills Road via stepping-stone populations, so it is possible that the Nm estimates exaggerate the aphid's ability to disperse over 5 km. I do not know of the number or positions of Lombardy poplars between the three Modena populations. However, they are very numerous in Modena, so there are probably more stepping-stone populations between the city-centre population

and Gaggio, which is just outside the city (7 km from the centre, F_{ST} = 0.012), than between Gaggio and Ravarino, which are separated by fields (12 km from the centre, F_{ST} = 0.092).

Short-range gene flow

Gene flow between local sites

The low F_{ST} values in 1997, 1998 and 1999 between the three populations within Cambridge (Mill Pit, Hills Road and Gonville Place, separated by 1.0–4.2 km), indicate that gene flow is high compared with more widely spaced populations. Nevertheless, the Mill Pit and Hills Road samples were genetically distinct, so gene flow can be restricted over distances as low as 2 km.

Gene flow between trees

The trend of finding higher degrees of gene flow over shorter distances extended to the between-trees spatial level. There was no evidence of genetic isolation between trees within Cambridge sites in 1997 and 1998, but there was in 1999. Moreover, significant genetic differences were found between trees within the Gaggio and Modena sites (the distances between these trees are not known). Hence, gene flow can be restricted over distances of less than 200 m, but is sufficiently high to prevent consistent genetic differentiation.

Gene flow within trees

The only spatial level where populations showed no evidence of restricted gene flow was between samples collected from distinct areas within trees. However, the sample sizes at this fine scale were too small (7 to 11 galls) to detect subtle population structure. The

potential for population structure within trees is probably limited by the high mobility of the sexuparae, which walk (or rarely fly) along the branch to the bark, then usually walk more than 1 m down the bark (Foster and Benton, 1992). Sexuparae aggregate on the bark in groups of one to nine (mean 2.5, median 1.6) before parturition (Foster and Benton, 1992). Hence, the potential for mixing with sexuparae from other galls is high.

Inbreeding within trees

Thirty-five of the 45 trees sampled in the UK and Italy showed high levels of inbreeding. If population structure is absent within trees, and aggregates of sexuparae do not tend to consist of relatives, as suggested by this study and by the behavioural observations of Foster and Benton (1992), then inbreeding within trees is probably the result of selfing between sexuales from the same clone. My data suggest that selfing accounts for on average 61% of successful matings in the UK and 32% in the Italian sample. The sedentary nature of the sexuales, their lack of inbreeding avoidance, and the small numbers in groups of sexuparae (Foster and Benton, 1992) probably all contribute to the high rate of selfing. Most of the groups observed by Foster and Benton (1992) contained three or fewer sexuparae (128 out of 191 groups). If we assume that mate choice is random within groups, and there is no mating between groups, the group size frequency distribution observed by Foster and Benton (Foster and Benton, 1992) corresponds to a selfing rate of 44%, which is within the range calculated from my genetic data.

Such a high level of selfing fits the theory that the female-biased sex ratio in aphids evolved through local mate competition (LMC) (Foster and Benton, 1992; Yamaguchi, 1985). Under LMC, brothers compete to mate with their sisters, leading mothers to economise on sons and invest in daughters (Hamilton, 1967).

Population density and selfing

It might be expected that trees on which galls are scarce will give rise to smaller groups of sexuparae on the bark, which in turn should lead to a higher rate of selfing. However, my results yielded no correlation between population density in one year and selfing rate in the following year. This finding is supported by the data of Foster and Benton (1992), which showed no link between the density of sexuparae on the bark and the size of sexupara groups. Population density varied widely in both studies (this study: 0–0.22 galls per petiole; Foster and Benton: 19–168 sexuparae per m^2), so the absence of a link with sexupara group size and selfing rate suggests that the size of sexupara groups might be adaptive.

Long-range versus short-range gene flow

These results support the proposal that long-range dispersal (> 100 km) in many aphid species is infrequent and of little biological significance. Loxdale et al. (1993) argue that dramatic examples of epic journeys by aphids, most famously those found on ice floes 8° south of the North Pole in 1828 by the explorer William Parry, have overshadowed evidence that local migration (< 20 km) is of greater frequency and importance. For example, regular long-distance migration between populations of the rose aphid, *Macrosiphum rosae*, is inconsistent with the discovery of large allele frequency differences between northwestern European countries separated by up to 1600 km (Tomiuk and Wöhrmann, 1984). Other studies have shown that migration can be sufficiently infrequent to restrict gene flow over much shorter distances. Taylor (1979) showed that the median distance travelled by autumn migrants of the damson-hop aphid, *Phorodon humuli*, was 15–20 km and the upper 95% confidence limit was 100–150 km. Moreover,

genetic studies of aphid populations have shown that aphid dispersal can be limited over short distances. Loxdale and Brookes (1990) found differences in allele frequencies at one allozyme locus between populations of the blackberry-grain aphid, *Sitobion fragariae*, only 65 km apart. Hebert et al. (1991) showed that genetic isolation was the primary cause of deviations from HWE in populations of the sumac gall aphid, *Melaphis rhois*, on 52 stands of its primary host, sumac, spread over 452 km, some only a few kilometres apart. Genetic differentiation has also been detected between populations of *Phorodon humuli* less than 30 km apart (Loxdale et al., 1998), and between populations of *Schoutedenia lutea* separated by less than 20 km (Tomiuk et al., 1991).

Not all genetic studies of aphid populations have shown evidence of restricted gene flow. Loxdale et al. (1985b) found only small genetic differences between British and Spanish populations of the grain aphid, *Sitobion avenae*, and Loxdale and Brookes (1988) found evidence of high gene flow between populations of the bird cherry-oat aphid, *Rhopalosiphum padi*, distributed across Britain. Similar absence of population structure has been observed in population genetic studies of the sycamore aphid, *Drepanosiphum platanoidis*, in southern Britain (Wynne et al., 1994), and in *Phorodon humuli* in Southern Germany (Eggers-Schumacher and Sander, 1988). Hence, the evidence of population genetics studies is that the dispersal abilities of aphids differ widely between species, as might be expected in such a diverse group of insects (Loxdale et al., 1993). My data show that *P. spyrothecae* populations are among the most genetically isolated of any aphid species.

Widely divergent dispersal abilities may also be found within the genus *Pemphigus*. A microsatellite study of *P. bursarius* found evidence of high gene flow between three

populations on summer lettuce, approximately 150 km apart (Miller, 2000). By contrast, strong genetic isolation was evident between summer lettuce and winter lettuce populations separated by only 14 km, suggesting that *P. bursarius* may consist of sympatric host races, between which reproductive isolation is strong, but within which dispersal has led to genetic homogeneity.

Miller's (2000) study has the greatest relevance to this study because *P. spyrothecae* and *P. bursarius* are closely related, share a host, *Populus nigra*, and were sampled in the same region, central and southern England. The major difference between *P. spyrothecae* and *P. bursarius* is in their life cycles: *P. bursarius* host-alternates and *P. spyrothecae* does not. While *P. spyrothecae* remains on poplar all year round, *P. bursarius* leaves the primary host from late June until early August to colonise one of its wide range of secondary hosts, which includes cultivated or wild Compositae, especially lettuce (*Lactuca* spp.), chicory (*Cichorium intybus*) and dandelion (*Taraxacum officinale*) (Alleyne and Morrison, 1977). *P. bursarius* must then return to black poplar in the autumn to complete its life cycle. Failure of migrants to find their host might explain why *P. bursarius* population densities are so tiny compared with *P. spyrothecae*. *P. spyrothecae*, on the other hand, has no need to migrate further than the distance between petiole and bark on a single tree. The contrast between the need to migrate in the host-alternator and the lack of such an incentive in the non-host-alternator could explain the apparent difference between their genetic structures. Interestingly, the similar heterozygosities between the two species suggest that the contrasting strategies are equally effective at preserving genetic variation. It is possible that, despite its low population density compared with *P. spyrothecae*, *P. bursarius* maintains genetic diversity by keeping its effective population size high through migration. By contrast, *P. spyrothecae* may have balanced the costs of losing one of its hosts, such as

genetic isolation, with benefits, which might include range expansion (Moran and Whitham, 1988), and avoidance of mortality due to failed migration (reviewed by Moran, 1992).

Pemphigus spyrothecae in northwestern America

The apparent similarity of the American and UK populations relative to the mainland European populations is unlikely to be a dramatic inversion of isolation by distance. The most likely explanation for this anomaly is that the *P. spyrothecae* population in this area was relatively recently founded by a few individuals, either from the British Isles or from a genetically similar population. Both *P. spyrothecae* and its host *Populus nigra* are native European species introduced to America. Although *Populus nigra* var. *italica* reached the east coast of America in 1784, the tree and the aphid could not have been introduced to the northwestern area of North America before Europeans began to settle there in the early 19th century. Wild *Populus nigra* has also been introduced to North America from Europe, but it is thought to be absent from the northwestern region (B.O.N.A.P., 1998), thus *Populus nigra* var. *italica* (Lombardy poplar) planted by hand is the only host of *P. spyrothecae* in this area. Consequently, the *P. spyrothecae* population must be very young, certainly less than 200 years old. The lack of variation suggests that the population has passed through a bottleneck, possibly as a direct consequence of the introduction of only a few individuals, although a subsequent scarcity of hosts may have caused or contributed to the lack of genetic variation.

The source of the introduction of *P. spyrothecae* to the American study area was probably the UK. None of the other European samples contained all the American alleles. Another possible source is a closely related European population that I have not

sampled, although the strong differentiation between *P. spyrothecae* in the UK and the four samples from mainland Europe suggests that the UK population might be unique as a result of isolation.

The extreme differentiation between some of the sites in northwestern America (e.g. between Hope and Hell's Gate, only 70 km apart, F_{ST} = 0.791) supports the view that gene flow is more restricted in northwestern America than in Europe. London and Cambridge, a similar distance apart (84 km), showed much less differentiation (in 1999, F_{ST} = 0.131). The explanation for the relative isolation of northwest American populations is likely to be a sparser network of hosts. *Populus nigra* var. *italica* is planted where people live, generally as a windbreak on farms or for decoration in towns, and the northwestern American study area is very sparsely populated compared with Europe (e.g. British Columbia has 3.5 inhabitants per km^2, compared with 373 in England). Moreover, much of the terrain is mountainous and rocky, and unsuitable for Lombardy poplars, particularly around the isolated Hell's Gate site (pers. obs.).

Temporal genetic stability

Allele frequencies in *P. spyrothecae* populations were generally stable over three years, although they did change significantly in one population between two years. This finding suggests that selection does not have a major effect on *P. spyrothecae* populations, and that populations are large, in agreement with the data presented in chapter 3.

Future research

This research could be continued in many directions. Possibilities that were not examined in this study include:

1. Host races on native and Lombardy black poplars, although this would be difficult to investigate because *P. spyrothecae* galls are so scarce on most native black poplars.
2. Allele frequency variation over the gall season, although this seems unlikely considering the genetic stability of populations between years.

CHAPTER FIVE: CLONAL MIXING

> "A particularly valuable check on kin selection theory will come when it is possible to contrast [aphid] colonies that have soldiers with those that don't." (Hamilton, 1987)

Introduction

The problem of reproductive altruism

If natural selection favours the fittest, how can we explain the evolution of self-sacrifice? The problem posed to evolutionary theory by the existence of altruism was solved by Hamilton's (1964) development of the ideas of Haldane (1923) and Fisher (1930) into kin selection theory (Maynard Smith, 1964). Hamilton realised that an altruistic trait could prosper if the altruist were closely related to the beneficiary; that is, if the beneficiary, as well as the altruist, were likely to carry the altruistic trait and pass it on to its offspring. Hamilton's rule states that an altruistic trait will be selected for if

$$\frac{b}{c} > \frac{r_A}{r_B},$$

where b is the fitness benefit to the recipient, c is the fitness cost to the donor, r_A is the relatedness of the altruist to her offspring, and r_B is the relatedness of the altruist to the beneficiary's offspring: high relatedness, high benefits and low costs will favour the evolution of altruistic behaviour.

Altruism within clones

Hamilton's rule predicts within clones "a complete absence of any form of competition which is not to the overall advantage and also the highest degree of mutual altruism" (Hamilton, 1964). This is because the relatedness between an altruist and her own offspring is equivalent to her relatedness to the offspring of every other member of the clone: $r_A = r_B = 1$. Hence, Hamilton's rule simplifies for clones to the statement that an altruistic trait will spread if $b > c$, that is, if the decrease in the altruist's offspring is less than the increase in her relatives' offspring. This rule can be generalised to cover any behaviour within a clone by taking a 'clone's-eye-view' (Stern and Foster, 1996). If we define w_A as the change in the number of an individual's offspring due to *any* behaviour (for an altruistic behaviour, $c = -w_A$), and w_C as the change in the number offspring of the whole clone including the behaver (for an altruistic behaviour, $b = w_C - w_A$), then $b > c$ reduces to $w_C > 0$. Hence, selection should favour any behaviour that increases the total reproductive output of the clone.

The rarity of altruism in clonal organisms

Considering such apparent theoretical support, it might seem surprising that, although not unknown, altruism is rare in free-living clonal organisms. However, Hamilton's rule predicts only that clones should benefit from altruism, and this is not the same as predicting that it should evolve in species that reproduce clonally (Dawkins, 1979). This is because a clonally reproducing species can consist of many clones. Relatedness is eliminated from the equation only if the altruistic behaviour is directed exclusively towards other members of the same clone, so that $r_A = r_B = 1$ (for the purpose of this argument, a clone is defined as a group of individuals that are genetically identical to the altruist at loci contributing to the altruistic trait). If individuals that do not carry the trait

receive the altruistic behaviour ($r_A = 1$, average $r_B < 1$), then they will accrue the benefits without paying the costs and the 'altruist' clone will eventually be out-competed by the 'cheat' clone. Note that the success of the altruistic clone is not dependent on reciprocation by clone-mates. The clone would prosper even if only one member of the clone behave altruistically, on condition that the altruistic clone benefit more than competing clones. Cheats do not necessarily defeat altruists, but if the altruistic clone is to win then the altruists must somehow direct more of the benefits towards their own clone than towards the cheats. The threshold above which altruists will beat cheats is reached when

$$b_A - c > b_C,$$

where b_A is the benefit to the altruistic clone and b_C is the benefit to the cheats.

Based on these predictions, it is possible to describe a scenario in which altruism should evolve. The crucial element is the delivery of altruism towards clone-mates to the exclusion of other clones. This could be achieved either though the altruistic behaviour evolving in association with existing kin-recognition behaviour, or by the physical separation of other clones from the range of the altruistic behaviour. Altruism should not evolve where populations of clonal organisms within which individuals mix freely persist for many generations. The longer such a population endures, the more time is available for a mutant clone of cheats to arise and displace the altruistic clone. Hence, an altruistic clone that has evolved in physical isolation from non-altruistic clones will persist only if its descendants regularly found new isolated populations. Regular propagation of daughter populations by one or a few individuals ensures frequent assortment of cheats from altruists, allowing clones made up exclusively of altruists to out-compete and displace clones of non-altruists.

Altruism among clonal organisms

Examples of altruistic behaviour in free-living clonal organisms include defender morphs in polyembryonic parasitoid wasps (Cruz, 1981), and soldier aphids (Aoki, 1977; reviewed by Stern and Foster, 1996). Both of these examples conform to the model of social behaviour evolving in short-lived spatially segregated colonies founded by a single individual.

Social behaviour in aphids

The first altruistic behaviour to be recorded in aphids was the defence of the gall by a specialised soldier caste (Aoki, 1977). Subsequently, gall cleaning (the removal of honeydew from the gall) has been recorded (Aoki, 1980; Benton and Foster, 1992). Although less glamorous, cleaning is probably more important than defence to the survival of the gall (Benton and Foster, 1992), and has a wider species distribution (Stern and Foster, 1996).

The behaviour and morphology of soldier aphids vary widely between species (Stern and Foster, 1996). Soldiers can occur on the primary or the secondary host, inside the gall or on the open leaf. They are usually juveniles, in most cases 1st or 2nd instars, although 3rd-instar and adult soldiers have also been recorded. The soldier instar can be dimorphic, where soldiers form a caste morphologically or behaviourally distinct from non-soldiers of the same instar, or monomorphic, where there is no specialised caste. Defenders can also be sterile or reproductive, and some researchers have limited the term 'soldier' to sterile defenders, on the basis that steriles pay greater costs than reproductives (e.g. Aoki, 1987). However, the distinction between sterile and

reproductive defenders is too hazy to justify such a rigid semantic division. For example, in the study species, *P. spyrothecae*, 1st-instar defenders have the potential to leave descendants that will found galls in the following spring, but there are two reasons why only a minority of them are likely to do so. First, the duration 1st instar of defenders is greatly extended compared with that of non-defenders, and second, while 1st-instar non-defenders can grow directly into alate gall migrants, defenders cannot, and only half of their daughters can (Rhoden, 1997). So, are *P. spyrothecae* defenders sterile or reproductive? I follow (Stern and Foster, 1996) in considering the distinction artificial, and classifying all aphid defensive morphs as soldiers.

Although the 50 or so soldier-producing species are distributed in only two families, the Hormaphididae and the Pemphigidae, soldiers have probably evolved at least nine times in this group, and been lost at least once (Stern and Foster, 1996). Such a high number of origins allows independent comparisons of associations between traits (Harvey and Pagel, 1991), making aphids useful subjects for investigating factors that promote or limit the evolution of altruism in clonal organisms.

The trait that associates most dramatically with soldier production in aphids is the habit of forming plant galls (Stern and Foster, 1996). Although only 14% of aphid species form galls, every one of the six or more evolutionary origins of soldiers has occurred within this group (however, not all soldiers operate inside the gall) (Foster and Northcott, 1994). The close association between galls and soldiers suggests that galls were an important factor in the evolution of soldiers (Foster and Northcott, 1994). Stern and Foster (1996) listed six reasons to explain why the evolution of soldier behaviour might have been favoured by living in a gall rather than on an open plant surface, such as a leaf,

stem or root. I divide these reasons into three categories: ecological incentives, behavioural pre-adaptations, and genetic integrity.

Ecological incentives

1. By providing food, shelter and defence, the gall is a valuable resource that is well worth defending. Neighbouring galls are likely to be occupied, so gall aphids have a strong incentive to stay and fight intruding predators. Aphids on leaves, on the other hand, are more likely to have available habitat nearby as a potential refuge.
2. The gall makes defence of the clone against predators feasible by limiting the size and location of predators to the size and location of the gall entrance (unless the predator is so powerful that it can break the gall open).

Behavioural pre-adaptations

3. Soldier behaviour may have evolved from intra-specific fighting. Foundresses in both social and non-social species fight each other for gall initiation sites. The wide phylogenetic distribution of this 'territoriality' trait indicates that foundresses were fighting before specialised soldier castes evolved, so later gall generations may not have had to evolve defensive adaptations from scratch, but instead may have inherited them 'ready-made' from the foundress. Foundress fighting is common in soldier-producing species (e.g. Foster, 1996; Whitham, 1979).
4. Soldiers may also have evolved from specialised migrants (Stern and Foster, 1996). Adaptations for surviving outside the gall, such as muscular legs for walking long distances and a sclerotised exoskeleton to protect against desiccation, may also have pre-adapted soldiers for defence.

5. Soldiers may have evolved from specialised gall-cleaning castes. Similar to fighting foundresses, gall cleaners have a wider phylogenetic distribution than soldiers, and so probably evolved before them. As cleaners labour at the entrance to the gall, they would be the first to encounter invading predators. This circumstance, combined with pre-existing altruistic behaviour, may have provided a route to the evolution of soldier behaviour.

Genetic integrity

6. The gall may guard against cheating by minimising mixing between clones. In aphid clones living on a leaf, clones can mix freely, so that if an aphid dies defending her neighbours she is likely to have given her life for a mixture of relatives (i.e. clone-mates) and non-relatives. The greater the concentration of clone-mates among the beneficiaries, the more likely altruism is to evolve, and gall living provides a potential physical barrier to clonal mixing, as well as an incentive to stay put (see point 1).

Do galls prevent clonal mixing?

The degree to which galls prevent clonal mixing is probably related to gall morphology (Stern and Foster, 1996). Some galls, such as leaf roll galls, are permanently open and would not provide a barrier to migrants. Other galls, including those of *P. spyrothecae*, are sealed in the early part of the season. However, all galls must eventually open to allow alates to emigrate or honeydew and exuviae to be removed, so there is potential for clonal mixing in every species, except those in which the gall empties immediately after opening. The manner of opening varies between species, and probably plays an important role in determining both the value of soldiers and the potential for clonal

mixing. For example, the gall of non-soldier-producing *P. populi* remains sealed until every occupant has developed into a mature winged gall emigrant, whereupon the gall bursts open and all the occupants fly out at once (Furk and Prior, 1975). Thus, *P. populi* minimises the time during which the clone is vulnerable to predation, but also prevents clonal mixing. *P. spyrothecae* takes the opposite approach, making an exit hole, just big enough for sexuparae to pass out, that remains open to predators for the last three months of the season. Clearly, *P. spyrothecae* has both more need of soldiers and more opportunity for inter-gall migration than *P. populi.* Predators that can be repelled by soldier aphids are unlikely to be able to break open a sealed gall, such as the *P. populi* gall, so while completely enclosed galls might promote the evolution of altruism by preventing clonal mixing, at the same time they will have no use for either of the known aphid altruistic behaviours, defence and cleaning. Hence, factors that promote the evolution of altruism in gall aphids, such as predation, go hand in hand with the opportunity for its inhibition through clonal mixing.

Evidence for inter-gall migration

Field studies

There is little direct evidence to support the theory that galls promote segregation between clones. *P. bursarius* 1st instars have been observed to walk out of galls and along petioles and branches (William Foster, pers. comm.). However, Whitfield (1998) found evidence for a very low level of inter-gall migration in the *P. spyrothecae.* Using sticky traps placed on the petioles between the gall and the branch, he caught on average 0.63 emigrants and 0.18 potential immigrants per gall per week (not including alates) during the 10.6-week period when the galls were open. Even if every individual caught on the way to the gall had entered, the immigrant fraction at the end of the season would have

been tiny (95% confidence interval: 0.18%, 0.73%), suggesting that the degree of clonal mixing in *P. spyrothecae* is very unlikely to be high enough to inhibit the evolution of altruism. Intriguingly, every non-alate migrant caught was a soldier (n = 514).

Although Whitfield's (1998) study was significant in establishing that *P. spyrothecae* larvae are not confined to the gall, and that they can walk from petiole to petiole, it is not possible to prove that migrants enter galls by trapping them on the petiole. Nor is direct observation practical, if immigrants enter galls at an average rate of one every six weeks (Whitfield, 1998). Nor is it practical to observe how many galls contribute immigrants to the study gall, or what becomes of an immigrant once it has entered a foreign gall. Although *P. spyrothecae* does not attack conspecific immigrants, and so appears to lack kin discrimination (Foster, 1990), there may be other reasons why an immigrant might not survive or linger in a foreign gall.

Genetic evidence

The limitations of direct observation have prompted suggestions that the clonal composition of *P. spyrothecae* galls would best be determined by genetic analysis (Stern and Foster, 1996; Whitfield, 1998). Aphid clones have previously been differentiated using a variety of techniques. Clones in gall aphids have been distinguished using allozymes (Hebert et al., 1991; Setzer, 1980), AP-PCR (Fukatsu and Ishikawa, 1994) and microsatellites (Llewellyn, 1996). Clonal diversity in species that do not form galls has also been investigated using ribosomal DNA probes (Lupoli et al., 1990), RAPDs (Black et al., 1992), and multilocus DNA fingerprints (Carvalho et al., 1991; De Barro et al., 1994).

Varying levels of clonal diversity have been found within aphid galls using genetic techniques. Setzer (1980) measured levels of clonal mixing of up to 25% in *P. populitransversus* and *P. populicaulis*, while Hebert *et al.* (1991) demonstrated that clonal mixing is virtually absent from galls of *Melaphis rhois*. Soldier behaviour has not been recorded in any of these three species. Fukatsu and Ishikawa (1994) found no evidence of clonal mixing among three aphids from five galls in the soldier-producing species *Ceratovacuna nekoashi*. However, the aim of their study was to demonstrate the effectiveness of the AP-PCR technique rather than to investigate clonal mixing, and the results from this small sample have only sufficient power to reject mixing levels of 19% and higher ($p = 0.05$).

Only one study has found evidence for clonal mixing in a soldier-producing species. Llewellyn (1996) genotyped up to 12 individuals from each of 13 *P. spyrothecae* galls collected in mid-November using two variable microsatellite loci (96PS5 and 97PS9). Five galls contained more than one genotype. However, Llewellyn (1996) pointed out that, as the galls were collected at the end of the migration period of the sexuparae, the mixing observed might have been caused by migrating sexuparae entering alien galls, and therefore might have been restricted to the end of the season. Hence, there is a need to measure accurately levels of clonal mixing earlier in the life of the gall, to assess the potential impact of cheating on the evolution of social behaviour.

The aim of this study was to investigate the clonal composition of *P. spyrothecae* using highly variable genetic markers to genotype each individual within a gall, thereby giving an accurate measure of the number of immigrants and the number of clones in a gall.

Methods

The two most variable of the seven microsatellite markers used in this study, loci 96PS20 and 98PS12, were chosen to genotype all the occupants of one *P. spyrothecae* gall, Gall 3019 from Tree 2, Gaggio, Modena, Italy. PCR was carried out as described in chapter 2. The study gall was selected from the sample set by examining the genotype data presented in chapter 4. It was chosen from the tree that contained the highest degree of variation in order to minimise the probability that two unrelated immigrant aphids could share the same genotype by chance. The gall that showed the rarest genotype among the 14 samples was selected to minimise the chance of an immigrant sharing the gall genotype by chance. In fact, the genotype of Gall 3019 was unique among the 14 galls collected from tree Gaggio 2, and the frequencies of three of its four alleles at the two loci were low (0.11 and 0.04 at 96PS20, and 0.39 and 0.11 at 98PS12). Gall 3019 was heterozygous at both loci, which allowed the incidence of single allele non-amplification to be gauged.

Each aphid within Gall 3019 was categorised according to morph (sexupara or virginopara) and instar (1st, 2nd, 3rd, 4th or adult). To look for evidence of mixing in other galls, 16 aphids were picked from each of four other galls (Galls 3063 and 3081 from Ravarino, 3291 from Uppsala and 3308 from Berlin) and genotyped using the same markers. The number of aphids of each instar was chosen to reflect the proportions within Gall 3019. To prevent errors, PCR was repeated at both loci on 90 1st-instar virginoparae, 90 1st-instar sexuparae, all the aphids from the four galls that were incompletely genotyped, and any individual whose genotype differed from the dominant gall genotype.

Table 5.1 Number of aphids selected for genotyping, origin and collection date of the five galls investigated in this study.

Gall number	No. genotyped	Source population	Date collected
3019	746	Gaggio, Modena	5th September 1999
3063	16	Ravarino, Modena	14th September 1999
3081	16	Ravarino, Modena	14th September 1999
3291	16	Uppsala	17th September 1999
3308	16	Berlin	4th September 1999

Results

Gall composition

The number of aphids of each instar in Gall 3019 is given in Table 5.2 and illustrated in Figure 5.1. There were significantly more 1st-instar virginoparae (soldiers) than 1st-instar sexuparae ($\chi^2 = 19.3$, d.f. = 1, $p < 0.0001$), consistent with observations that the duration of the 1st instar is greatly extended in *P. spyrothecae* (Aoki and Kurosu, 1986; Rhoden, 1997). The total numbers of sexuparae (395) and virginoparae (351) were not significantly different ($\chi^2 = 2.6$, d.f. = 1, $p = 0.11$).

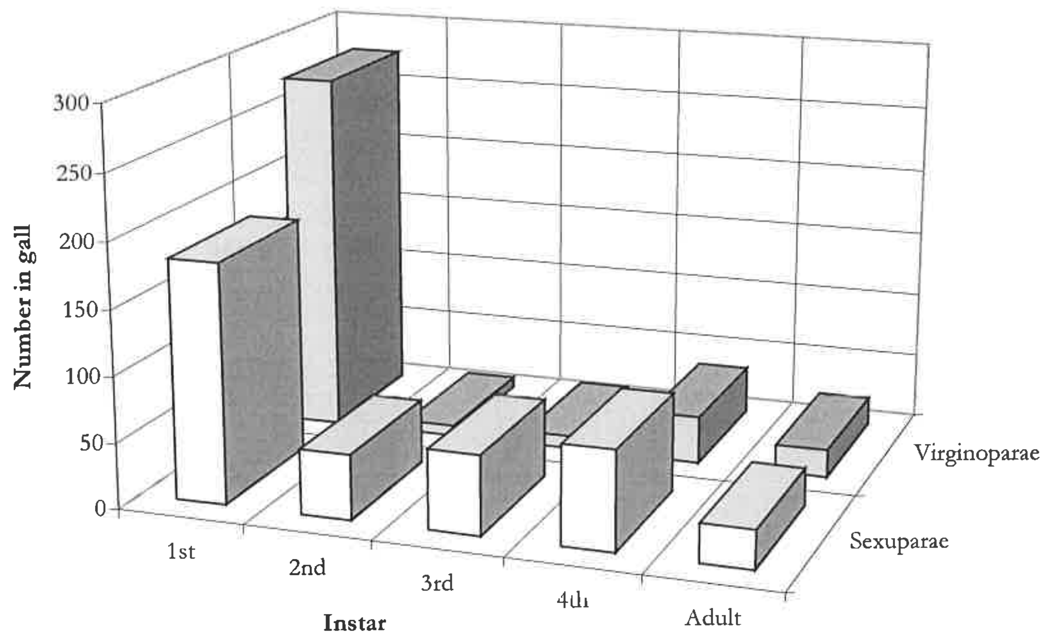


Figure 5.1 The instar and morph composition of Gall 3019.

Table 5.2 Gall composition and associated genotyping success in Gall 3019.

Morph	Instar	Number in gall	No. of loci successfully amplified	
			≥ 1 locus (%)	2 loci (%)
Sexuparae	1st	182	167 (92%)	138 (76%)
	2nd	49	48 (98%)	46 (94%)
	3rd	60	59 (98%)	57 (95%)
	4th	75	69 (92%)	66 (88%)
	Adult	29	23 (79%)	22 (76%)
Virginoparae	1st	276	195 (71%)	143 (52%)
	2nd	8	8 (100%)	6 (75%)
	3rd	10	8 (80%)	6 (60%)
	4th	36	36 (100%)	34 (94%)
	Adult	21	20 (95%)	17 (81%)
	Total	746	633 (85%)	535 (72%)

Genotyping success

Complete genotypes were obtained at two loci (96PS20 and 98PS12) for 577 of the 810 aphids (71.2%). An additional 105 individuals yielded genotypes at only one locus, bringing the total genotyped at one or more loci to 682 (84.2%). The remaining 128 gave no PCR product at either locus. PCR success varied greatly between instars, with 1st-instar virginoparae proving the most troublesome. Compared with their frequency among the 810 aphids analysed (37.0%), 1st-instar virginoparae were significantly over-represented (70.3%) among the 128 aphids that generated no PCR product at either locus ($\chi^2 = 60.8$, d.f. = 1, $p < 0.0001$). The resistance of this instar to PCR may have been caused by its greater sclerotisation and small size compared with other instars.

Genotyping reliability

As the five galls genotyped were heterozygous at both loci, I was able to estimate the incidence of 'false homozygotes', where a heterozygous individual shows a homozygous genotype due to the failure to amplify one allele. False homozygotes can be a problem when, as in this study, amplifying from small quantities and variable qualities of DNA (Taberlet et al., 1996). At locus 98PS12, 27 individuals presented homozygous genotypes after a single PCR. However, repeated PCR showed that 17 of these were heterozygotes, while only two showed the same homozygous genotype and 8 gave no product. Again, 1st-instar virginoparae were the main culprits, contributing 11 of the 17 confirmed false homozygotes. Locus 96PS20 exhibited a lower number of false homozygotes. Repeated PCR of 11 apparent homozygotes showed that nine were in fact heterozygotes while only two returned the same homozygous genotype. Six of the nine false homozygotes were 1st-instar virginoparae. Given the high frequency of false homozygotes in this study, particularly among 1st-instar virginoparae (11.2% at 98PS12, 5.7% at 96PS20), I decided that a homozygous genotype could not be considered reliable until confirmed by at least three independent PCRs.

Genetic variation in Gall 3019

Of the 746 aphids in Gall 3019, 633 yielded genotypes at one or more loci (Table 1). Of these, 515 showed a single two-locus genotype, which is assumed to be the foundress genotype, and a further 104 showed incomplete genotypes that did not differ from the foundress genotype (Table 5.3). The remaining 14 aphids showed genotypes that were distinct from the foundress genotype (Table 5.4), giving an estimated clonal mixing level of 2.2%. An exact 95% confidence interval of 1.2% to 3.7% was calculated for the population clonal mixing level. However, this confidence interval is applicable only

insofar as a uniform level of mixing can be assumed to apply to other galls. Ten of the 14 immigrants (71%) were 1st-instar virginoparae (i.e. soldiers), significantly greater than the proportion of 1st-instar virginoparae within the gall ($\chi^2 = 7.1$, d.f. = 1, $p < 0.01$), and the ratio of 13 virginoparae to one sexupara represents a highly significant skew towards virginoparae ($\chi^2 = 11.8$, d.f. = 1, $p < 0.001$). The 14 immigrants were drawn from at least nine, and possibly as many as 11 clones (as two of the immigrant genotypes were incomplete, the exact number of immigrant clones in the gall could not be determined).

Figure 5.2 Banding patterns at locus 98PS12 of 34 2nd-instar sexuparae from Gall 3019. The repeat number of each allele is indicated. Thirty-three individuals show the gall genotype (1923), while one shows the genotype of an immigrant clone (1718).

Because most of the migrants were 1st-instar virginoparae, and a disproportionate number of this group failed to yield a PCR product, the true proportion of migrants in the gall could have been higher than 2.21%. An unbiased estimate was calculated by scaling up the number of migrants within each instar to the figure expected had the PCR success rate within that group been 100%. Correcting for the skew in PCR success in this way gives an estimated number of migrants of 18 out of 746 aphids, a migrant fraction of 2.45%.

Table 5.3 Genotypes from Gall 3019 at loci 96PS20 and 98PS12, and the number of each genotype. An asterisk indicates a genotype that differs from the gall genotype. Genotypes are presented as two two-digit microsatellite repeat numbers, e.g. a genotype of 1213 at 96PS20 indicates that the repeat sequence was $(GT)_{12}$ in one allele and $(GT)_{13}$ in the other. Dashes indicate that no repeat number could be unambiguously assigned.

Genotype		
96PS20	98PS12	Number
1213	1923	515
1213	----	47
----	1923	50
1213	19--	4
1213	--23	1
--13	--23	1
----	19--	1
1417	1920	1*
0612	1720	1*
0612	1724	2*
0612	1824	2*
--12	1824	1*
0614	1718	2*
1212	1720	1*
1212	17--	1*
1213	1920	1*
1214	1720	1*
1217	1718	1*

Table 5.4 Genetic diversity and instar composition among the 14 immigrants to Gall 3019. The number of independent PCRs that yielded each genotype is given in brackets.

	Genotype		
Aphid ref.	96PS20	98PS12	Instar
10	1417 (3)	1920 (4)	V1
17	0612 (4)	1720 (4)	V1
224	0612 (4)	1724 (4)	V1
464	0612 (4)	1724 (4)	V2
82	0612 (4)	1824 (4)	V1
83	0612 (4)	1824 (4)	V1
85	--12 (2)	1824 (3)	V1
209	0614 (4)	1718 (4)	V1
358	0614 (1)	1718 (3)	V1
338	1212 (4)	1720 (4)	V1
507	--12 (2)	17-- (1)	VA
517	1213 (4)	1920 (1)	VA
275	1214 (4)	1720 (4)	V1
548	1217 (4)	1718 (4)	S2

Genetic composition of Galls 3063, 3081, 3291 and 3308

The 64 aphids from Galls 3063, 3081, 3291 and 3308 yielded 49 genotypes, which included 38 complete two-locus genotypes and 11 incomplete genotypes (Table 5.5). All the incomplete genotypes conformed to the dominant gall genotypes. One of the aphids from Gall 3291, a 2nd-instar sexupara, had a genotype distinct from that of her fellow gall occupants (Table 5.5). This genotype was confirmed by six independent PCRs. Exact 95% confidence intervals for the level of clonal mixing within each of the four galls were calculated (Table 5.5).

Table 5.5 Genotypes at loci 96PS20 and 98PS12 of Galls 3063, 3081, 3291 and 3308. A 95% confidence interval for the clonal mixing level is given in the last column.

		Focal clones		Immigrant clones		
Gall	Source tree	Genotype	No.	Genotype	No.	Mixing level 95% C.I.
3063	Ravarino 1	1214,1724	12	-	0	0%, 26.5%
3081	Ravarino 3	0614,2024	12	-	0	0%, 26.5%
3291	Uppsala 3	1016,1620	9	0707,1617	1	0.25%, 44.5%
3308	Berlin 1	0107,1720	15	-	0	0%, 21.8%

Discussion

Inter-gall migration in *P. spyrothecae*

The data presented here provide direct evidence for inter-gall migration in an aphid that is known to have social behaviour, confirming Llewellyn's (1996) discovery that clonal mixing occurs in *P. spyrothecae*. In addition, this study has measured the level of clonal mixing in a soldier-producing species for the first time. Genotyping of 633 aphids from a single *P. spyrothecae* gall at two variable microsatellite loci identified 14 conspecific immigrants from at least nine alien clones. The level of clonal mixing in this gall is therefore 2.2%. A similar clonal mixing level of 2.0% was found in a survey of 49 aphids from a further four galls, demonstrating that clonal mixing was not unique to Gall 3019.

How prevalent is inter-gall migration in *P. spyrothecae*?

How widely can the results from these five galls be interpreted? That depends on the assumption that all *P. spyrothecae* galls are subject to the same level of clonal mixing, which in turn depends on the variability of the factors that influence inter-gall migration. The incentive to emigrate might be expected to increase as the gall becomes more crowded, but in a study of 60 *P. spyrothecae* galls Whitfield (1998) found no evidence for a

relationship between crowding and emigration rate. Another factor that might be expected to influence migration rate is the density of galls on the tree. Gall density varies widely between trees and between years (see chapter 3), and an emigrant probably has a much greater chance of entering another gall on a tree where 47% of the petioles bear galls, than on one where gall density is less than 1%. There are currently no data to support this supposition, but neither is there evidence that the level of mixing does not vary substantially between galls. Hence, it is not valid to assume that the five galls studied, which were taken from five trees in three countries, all experienced the same level of clonal mixing. However, it is worth noting that the level of mixing found in the four galls presented in Table 5.5 conforms exactly to that expected under the assumption that the level of mixing in Gall 3019 is typical of all *P. spyrothecae* galls.

The proportion of immigrants found in Gall 3019 from Gaggio (2.21%) is three to twelve times greater than that detected in Cambridge by Whitfield (1998) using sticky traps (95% confidence interval: 0.18%, 0.73%). This discrepancy may reflect a true difference in levels of clonal mixing between Cambridge and Gaggio. Although gall densities are unknown from either Whitfield's trees in 1995 or Gaggio-2 in 1999, they tend to be higher on trees in Italy than in Cambridge (pers. obs.), and this might account for the difference in clonal mixing levels. Alternatively, the sticky traps used by Whitfield may have been ineffective, allowing migrants to walk over them. However, Whitfield rejects this possibility on the basis that large numbers of alates were trapped while walking along the petiole to the trunk, as were insects of other species. It is also possible that many aphids fell from gall openings onto leaves below, thus entering a new gall without encountering the sticky traps.

Which aphids migrate?

Thirteen of the 15 immigrants found in the five galls studied were virginoparae, and ten of these were 1st instars and therefore soldiers. This corresponds with Whitfield's (1998) results, which show an even greater bias towards soldiers among migrants: every one of the 514 migrants (excluding alates) caught on his sticky traps was a soldier (95% confidence interval: 99.3%, 100.0%). The absence of non-soldiers in such a large sample is persuasive evidence that only soldiers the leave gall. How then can the presence of five non-soldiers among the 15 immigrants be explained? The most likely explanation is that the three virginoparae (one 2nd instar and two adults) moulted from an immigrant soldier, and that the two 2nd-instar sexuparae were the offspring of immigrant soldiers that had moulted to adulthood.

Immigrant soldiers can leave descendants

The inference that the three non-soldiers found in two foreign galls were moults and offspring of immigrant soldiers is an important, because it provides the first evidence that *P. spyrothecae* migrants not only reach and enter alien conspecific galls, but also can moult and reproduce inside. In particular, the discovery of two alien 2nd-instar sexuparae, one sampled on the 5th September, the other on the 17th September, suggests that gall immigrants have a chance of achieving reproductive. Even if both sexuparae had been sampled immediately after birth, they would have had the potential to moult to alates and emigrate to the bark in as little as 23 days (Rhoden, 1997). Hence, the two alien sexuparae sampled on the 5th and 17th September could have emigrated from their adopted galls as early as the 28th September and the 10th October, respectively. Both dates are well within the observed period of alate emergence, which can continue past the end of October (Brändle, 1999; pers. obs.).

Why should *P. spyrothecae* soldiers migrate?

A migrating *P. spyrothecae* soldier faces many hazards. Once outside the gall, it could quickly dehydrate and starve (Whitfield, 1998; Williams and Whitham, 1986), and it risks predation and failure to find another gall. However, this study has shown that *P. spyrothecae* aphids do move between galls, so either migration happens by accident, or it is an adaptive behaviour.

Non-adaptive migration

P. spyrothecae soldiers cluster near the gall opening (Foster, 1990; Rhoden, 1997), presumably because this is the best place from which to eject honeydew and repel

predators. Soldiers frequently walk in and out of the gall, and sometimes fall out while ejecting honeydew (Benton and Foster, 1992). Migrants might then be soldiers that stray from the gall and fail to find their way back, or fall out with ejected honeydew, or are blown from their natal gall to a different part of the tree. If, as appears likely, *P. spyrothecae* cannot recognise non-kin (Foster, 1990), there would be no resistance to immigration from the native clone. Thus, clonal mixing in *P. spyrothecae* could be an accident, and of no adaptive significance.

Although there is no direct evidence against this view, the existence of non-accidental (i.e. adaptive) inter-gall migration in other aphid species demonstrates that the costs of dispersal are not necessarily prohibitive. The high levels of clonal mixing (up to 25%) found in *P. populitransversus* and *P. populicaulis* are unlikely to have occurred by accident (Setzer, 1980), and a specialised migratory morph has been recorded in *Pachypappa marsupialis* (Aoki, 1979). Some aphids are also known to usurp the galls of other aphids (Akimoto, 1988; Kurosu and Aoki, 1990). *P. spyrothecae* galls are frequently invaded by *Chaitophorus* spp., usually resulting in the death of the original occupiers (Christian Braendle and pers. obs.). Thus, leaving and entering galls are recognised adaptive behaviours in a number of aphid species.

Adaptive migration

There are three main reasons why an aphid should be adapted to inter-gall migration: migration can be provoked by danger; or, in the absence of immediate danger, it can be a 'bet-hedging' strategy, the clone's insurance against the possible extinction of the gall; or it might represent self-sacrifice by the migrant to free resources for those left behind.

Migration in response to immediate danger

The conditions in a gall might change to make dispersal seem a relatively attractive option. An aphid might flee predation, parasitism, disease, or shortage of resources. Despite having less than a 1% chance of reaching another gall, larval *P. betae* are more likely to leave galls that are deteriorating in quality before abscission than healthy galls (Williams and Whitham, 1986). Whitfield (1998) found that *P. spyrothecae* emigrants have a 29% chance of finding another gall, suggesting that the benefits of inter-gall migration might be higher in *P. spyrothecae* than in *P. betae*. *Pemphigus* larvae appear to respond to environmental stress by evacuating the gall, and the 1st instars have the strongest urge to migrate: Williams and Whitham (1986) and Whitfield (1998) observed that *Pemphigus* aphids abandoned their gall after it was removed from the tree, the 1st instars leading the way. However, this evacuation behaviour might be a response to the desiccation of the gall, rather than a direct response to abscission. It is unlikely that migrants from a gall that has fallen to the ground could ever return to a tree and enter another gall, but fallen galls are frequently caught in branches (pers. obs.), particularly in late summer and early autumn, which might allow evacuees to reach a new gall.

Probably no more than a tiny proportion of inter-gall migration can be attributed to emergency evacuation. Aphids of all instars abandon abscinded galls, but only soldiers were caught on Whitfield's (1998) sticky traps. It could be argued that absence of non-soldiers caught leaving galls (0 of 400; 95% confidence interval: 0%, 3.2%) simply indicates sampling bias due to the removal of abscinded galls from Whitfield's sample, while no non-soldiers (0 of 114) were caught approaching galls because they were less able to survive outside the gall than soldiers. However, it is highly unlikely that the

survival advantage of soldiers outside the gall is sufficient to reduce the proportion of non-soldiers from approximately half to less than 3.2%.

Altruistic migration

Altruistic migration, like defence and gall-cleaning, can be viewed from the clone's point of view. A clone should invest in dispersal to the degree that optimises the clone's fitness. Altruistic migration could increase the fitness of the clone in two ways: by colonising new habitat patches, and by relieving competition for scarce resources in the gall.

Migration as insurance against future danger

In organisms living in a patchy environment, such as clones of aphids living in galls, it might pay a clone to invest in individuals that are specialised to colonise other patches (Aoki, 1982; Hamilton, 1987; Hamilton and May, 1977). Aoki (1979) showed that the aphid *Pachypappa marsupialis* produces a morphologically and behaviourally specialised migratory morph that can successfully colonise conspecific galls. Setzer (1980) suggested that the high level of inter-gall migration in *P. populitransversus* and *P. populicaulis* is a bet-hedging strategy in response to high levels of disease, predation and parasitism. However, Hamilton and May (1977) predicted a large investment in dispersal even in stable environments where all other patches are occupied. The proportional investment in migrants should be $1/(2 - p)$, where p is the probability of a migrant surviving dispersal. Whitfield (1998) estimated that $p = 0.29$ in *P. spyrothecae*, which predicts that a clone should invest 58% of its resources in dispersal. This figure clearly does not correspond to the empirical evidence for the low levels of dispersal in *P. spyrothecae* reported here and by

Whitfield (1998). Even if every migrant dies, the model cannot explain an investment in dispersal of less than 50%. The inappropriateness of this model may be partly explained by its assumption that the size and resources of each patch is fixed, allowing intraclonal competition to drive dispersal, while the size of *Pemphigus* galls appears to be determined by the size of the clone (Dunn, 1960a), minimising competition for resources.

Altruistic emigration

All aphid soldiers use the resources of the clone by occupying space. Some, including *P. spyrothecae* soldiers, also feed. As in any society, when resources become scarce, soldiers might no longer pay their way. Under these circumstances, the clone would benefit from the emigration of sufficient soldiers to restore the cost-effectiveness of defence investment. The emigrants would have only a 29% chance of reaching a new gall (Whitfield, 1998), so their individual fitness would be greatly reduced. Hence, clonal mixing might be the result of self-sacrificing soldiers abandoning their deteriorating shelters, like Captain Oates, to increase the share of resources among those left behind. This theory might explain why only soldiers and never non-soldiers leave the gall (Whitfield, 1998). Soldiers take at least 42 days to produce alate sexuparae capable of leaving descendants in the next generation, and only half of their offspring are sexuparae, while a 1st-instar sexupara can become an alate gall emigrant in as little as 22 days (Rhoden, 1997). Hence, a soldier in a starving, crowded gall might be of little value for reproduction.

If, as suggested by this study, 2.2% of gall occupants are immigrants, and if immigrants make up 29% of emigrants (Whitfield, 1998), the average level of emigration in *P. spyrothecae* in the sample of galls is about 8% of individuals. The average proportion of

resources freed by soldiers must be less than this because soldiers are small and grow slowly compared with non-soldiers, and so they are likely to use less space and suck less sap. Such a low level of emigration might appear insufficient to significantly increase the share of resources within a gall. Nevertheless, even a small increase could benefit the clone. In addition, the level of emigration might vary widely between galls, in which case most inter-gall migration would be a consequence of mass exoduses from a few galls. However, Whitfield (1998) found no evidence for an association between resources, measured as gall mass per aphid, and the number of emigrants.

While a few altruistic emigrants might give a small benefit to a starving clone, an aphid carrying an infectious disease or a parasitoid would have a pressing incentive to leave the gall. A similar behaviour has been observed in pea aphids, which parasitised commit 'adaptive suicide' by falling from the host plant when parasitised (McAllister and Roitberg, 1987).

Why do only soldiers emigrate?

As convincingly demonstrated by Whitfield (1998), only soldiers emigrate. If emigration is accidental, its restriction to soldiers is probably a consequence of their proximity to the gall entrance. On the other hand, if emigration is adaptive there are a variety of reasons why it might be restricted to soldiers. If the purpose of migrants is to colonise other galls as insurance against extinction of the clone, then soldiers appear to be better equipped to survive outside the gall than non-soldiers. Soldiers may even have evolved from specialised migrants (Stern and Foster, 1996). If the purpose of emigration is to economise on resources in stressed galls, soldiers should leave first, because defence is dispensable but reproduction is not.

Clonal mixing and the evolution of soldiers

Such a low level of clonal mixing is in accordance with the hypothesis that a background of high clonal segregation should favour the evolution of soldiers (Hamilton, 1987). By contrast, *P. populitransversus* and *P. populicaulis* exhibit relatively high levels of clonal mixing (on average 22% and 24% respectively) (Setzer, 1980). Soldiers have not been observed in either of these species, but they have yet to be rigorously tested.

Stern and Foster (1996) modelled the conditions under which a clone should invest in soldiers. They proposed that, in the absence of clonal mixing, it pays a clone to produce soldiers when

$$\frac{pd_p}{b} > c,$$

where p is the effectiveness of soldiers, d_p is the death rate due to predation, b is the birth rate and c is the cost of producing soldiers. Soldier production is favoured when predation and soldier effectiveness are high and costs are low. Stern and Foster (1996) then investigated the effect of clonal mixing on soldier production. They predicted that in the presence of N other clones not producing soldiers, the benefit to a clone of investing in soldiers, p, is diluted by N, assuming that all clones have equal numbers, so soldier production will be favoured when

$$\frac{pd_p}{Nb} > c.$$

The assumption of equal population sizes may fit species that produce soldiers on open leaves, where several clones may be founded simultaneously, but it is unrealistic in the case of *P. spyrothecae*, where clones are sealed inside galls for the first three to four months of the clonal phase. In this case mixing is most likely to occur through inter-gall

migration by single aphids late in the season, after the gall has opened to allow the exit of honeydew and sexuparae. However, the model is easily made applicable to *P. spyrothecae* by defining a new dilution factor for inter-gall mixing, N_g:

$$N_g = \frac{n_t}{n_t - n_m},$$

where n_t is the total number of aphids in the gall and n_m is the number of immigrants from clones not producing soldiers. Note that N_g is independent of the number of clones that invade the gall. Calculating N_g for Gall 3019, using the estimate of $n_m = 18$ migrants among the $n_t = 746$ inhabitants of the gall, we obtain a dilution factor of $N_g = 1.025$. If the amount of mixing within Gall 3019 is representative of other *P. spyrothecae* galls, then the benefit to *P. spyrothecae* clones of producing soldiers is diminished by a factor of 1.025 due to clonal mixing, insignificant compared with the dilution factor of 1.31 in *P. populicaulis* galls, calculated from Setzer's (1980) data.

Future work

Ascertaining the level of clonal mixing in *P. spyrothecae* and other species

This study has demonstrated that clonal mixing occurs in *P. spyrothecae* galls, and provided evidence that the level of mixing is low. However, many questions concerning the degree, species distribution and adaptive significance of clonal mixing remain unanswered. There is a need to estimate accurately the level of mixing by genotyping a large sample of galls, taken from trees of varying gall densities, at times throughout the life of the gall. In addition, in order to investigate further the relationship between clonal mixing and social behaviour, it will be necessary to measure levels of inter-gall migration in many more aphid species, both social and non-social. Some aphids exhibit social

behaviour outside the gall, and it will be interesting to investigate clonal mixing in these species.

Measuring the potential reproductive success of soldiers

Although my data show that soldier immigrants can reproduce in alien galls, they do not prove that their offspring can ever become alates, and thereby have a chance of leaving progeny in the following year. Proof could be obtained by enclosing galls in muslin bags late in the season, and genotyping all alates that emerge. Alternatively, once a typical pattern of immigration has been discovered it will be possible to simulate mixing by introducing foreign soldiers to a gall at realistic times. The emerging alates could then be trapped and genotyped as above, allowing the proportion whose descendants leave the galls to be measured.

Do migrants cheat?

It would pay a soldier to cheat in a gall populated by non-relatives. Do immigrant soldiers behave differently to native soldiers, perhaps by failing to fight predators? It is currently impossible to observe this in the field, and hardly easier in the laboratory, because soldiers of different clones look identical. It might be possible to separate 'honest' soldiers from cheats by putting a predator in the gall, genotype soldiers from both groups, and discover if immigrants are disproportionately common among the cheats. It would also be interesting to find out if immigrant soldiers cheat in other ways, such as shortening the duration of the 1st instar.

CHAPTER SIX: CONCLUSION

This study has illuminated some important aspects of the biology of a fascinating insect. The strong spatial genetic structure reported here suggests that populations of *P. spyrothecae* are as isolated as those of any aphid that has yet been studied, and that gene flow can be restricted between trees within only 200 metres. The genetic pattern found within trees supports observations of mobile sexuparae and sedentary sexuales (Foster and Benton, 1992), pointing to a mating system that is a balance between two extremes, outbreeding and selfing. Yet, despite population isolation and selfing, *P. spyrothecae* populations maintain high levels of microsatellite variability, possibly due to their high population densities compared with other *Pemphigus* species. Large population sizes, and low or absent selection, may also be responsible for the temporal stability of allele frequencies over the three-year study period.

The low levels of clonal mixing found inside *P. spyrothecae* galls suggest that the gall is an effective barrier to clonal mixing, so that, even if migrant soldiers do cheat in foreign galls, the burden on the native clone is low, and the benefits of sociality are only slightly diluted. The few soldiers that migrate are able to moult and reproduce in foreign galls. This discovery, combined with the restriction of migration to soldiers suggests that migration might be an adaptive specialisation rather than a random event.

REFERENCES

Aaby, C. 1999. NPS: Non-parametric statistical software package designed for MS-DOS. http://www.carsten-aaby.suite.dk/C/NPS/NPS.htm.

Akimoto, S. 1988. The evolution of gall parasitism accompanied by a host shift in the gall aphid, *Eriosoma yangi* (Homoptera: Aphidoidea). *Biological Journal of the Linnean Society*, 35:297-312.

Alleyne, E. H., and F. O. Morrison. 1977. The lettuce root aphid, *Pemphigus bursarius* (L.) (Homoptera: Aphidoidea) in Quebec, Canada. *Annals of the Entomological Society of Quebec*, 22:171-180.

Amos, B., C. Schlötterer, and D. Tautz. 1993. Social structure of pilot whales revealed by analytical DNA profiling. *Science*, 260:670-672.

Aoki, S. 1977. *Colophina clematis* (Homoptera, Pemphigidae), and aphid species with 'soldiers'. *Kontyû*, 45:276-282.

—. 1979. Dimorphic first instar larvae produced by the fundatrix of *Pachypappa marsupialis* (Homoptera: Aphidoidea). *Kontyû*, 47:390-398.

—. 1980. Occurrence of a simple labor in a gall aphid *Pemphigus dorocola* (Homoptera, Pemphigidae). *Kontyû*, 48:71-73.

—. 1982. Soldiers and altruistic dispersal in aphids. In M. Breed, C. Michener and H. Evans (eds.), *The Biology of Social Insects*. Westview Press, Boulder, CO.

—. 1987. Evolution of sterile soldiers in aphids. In Y. Itô, J. L. Brown and J. Kikkawa (eds.), *Animal Societies: Theories and Facts*. Japan Scientific Society Press, Tokyo.

Aoki, S., and U. Kurosu. 1986. Soldiers of a European Gall Aphid, *Pemphigus spyrothecae* (Homoptera, Aphidoidea): Why Do They Molt? *Journal of Ethology*, 4:97-104.

Arnheim, N., H. Li, and X. Cui. 1990. PCR analysis of DNA sequences in single cells: single sperm gene mapping and genetic disease diagnosis. *Genomics*, 8:415-419.

B.O.N.A.P. 1998. Biota of North America Program U.S. Checklist: A Synonymized Checklist of the Vascular Flora of the United States, Puerto Rico, and the Virgin Islands. http://www.csdl.tamu.edu/FLORA/b98/check98.htm.

Barrette, R. J., T. J. Crease, P. D. N. Hebert, and S. Via. 1994. Mitochondrial DNA diversity in the pea aphid *Acyrthosiphon pisum*. *Genome*, 37:858-865.

Benton, T. G., and W. A. Foster. 1992. Altruistic housekeeping in a social aphid. *Proceedings of the Royal Society of London Series B-Biological Sciences*, 247:199-202.

Birkle, L. M., and A. E. Douglas. 1999. Low genetic diversity among pea aphid (*Acyrthosiphon pisum*) biotypes of different plant affiliation. *Heredity*, 82:605-612.

Black, W. C. 1990. Use of molecular probes to characterize aphid populations, *Proceedings, Aphid-Plant Interactions: Populations to Molecules*, pp. 137-151. Oklahoma State University, Stillwater, OK.

Black, W. C., N. M. Duteau, G. J. Puterka, J. R. Nechols, and J. M. Pettorini. 1992. Use of the random amplified polymorphic DNA-polymerase chain reaction (RAPD-PCR) to detect DNA polymorphisms in aphids (Homoptera, Aphididae). *Bulletin of Entomological Research*, 82:151-159.

Blackman, R. L. 1979. Stability and variation in aphid clonal lineages. *Biological Journal of the Linnean Society*, 11:259-277.

—. 1990. Specificity in aphid-plant interactions with particular attention to the role of the alate colonizer. In R. Campbell and R. Eikenbary (eds.), *Aphid-plant genotype interactions*, pp. 251-274. Elsevier, Amsterdam.

Blackman, R. L., and V. F. Eastop. 1994. *Aphids on the world's trees: an identification and information guide*. CAB International in association with the Natural History Museum, Wallingford.

Blackman, R. L., and J. M. Spence. 1992. Electrophoretic distinction between the peach-potato aphid, *Myzus persicae*, and the tobacco aphid, *M. nicotianae* (Homoptera, Aphididae). *Bulletin of Entomological Research*, 82:161-165.

Boulding, E. G. 1998. Molecular evidence against phylogenetically distinct host races of the pea aphid (*Acyrthosiphon pisum*). *Genome*, 41:769-775.

Brändle, C. 1999. The ecology of polyphenism in a soldier-producing aphid. M.Phil. thesis, University of Cambridge.

Burton, R. M. 1983. *Flora of the London area*. London Natural History Society, London.

Cadbury, D. A., J. G. Hawkes, and R. C. Readett. 1971. *A computer-mapped flora. A study of the County of Warwickshire*. Academic Press, London.

Callen, S. F., A. D. Thompson, Y. Shen, H. A. Phillips, R. I. Richards, J. C. Mulley, and G. R. Sutherland. 1993. Incidence and origin of "null" alleles in the $(AC)_n$ microsatellite markers. *American Journal of Human Genetics*, 52:922-927.

Carvalho, G. R., N. Maclean, S. D. Wratten, R. E. Carter, and J. P. Thurston. 1991. Differentiation of aphid clones using DNA fingerprints from individual aphids. *Proceedings of the Royal Society of London Series B-Biological Sciences*, 243:109-114.

Cenis, J. L., P. Perez, and A. Fereres. 1993. Identification of aphid (Homoptera, Aphididae) species and clones by random amplified polymorphic DNA. *Annals of the Entomological Society of America*, 86:545-550.

Clarke, G. M. 1980. *Statistics and Experimental Design*. Edward Arnold, London.

Cockerham, C. C. 1969. Variance of Gene Frequencies. *Evolution*, 23:72-84.

—. 1973. Analyses of Gene Frequencies. *Genetics*, 74:679-700.

Cognetti, G. 1961. Endomeiosis in parthenogenetic lines of aphids. *Experientia*, 17:168-169.

Cruz, Y. P. 1981. A sterile defender morph in a polyembryonic hymenopterous parasite. *Nature*, 294:446-447.

Dallas, J. F. 1992. Estimation of microsatellite mutation rates in recombinant inbred strains of mouse. *Mammalian Genome*, 3:452-456.

Darwin, C. R. 1859. *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life*. John Murray, London.

Dawkins, R. 1979. Twelve misunderstandings of kin selection. *Zeitschrift für Tierpsychologie*, 51:184-200.

De Barro, P. J., T. N. Sherratt, C. P. Brookes, O. David, and N. Maclean. 1995a. Spatial and temporal genetic variation in British field populations of the grain aphid *Sitobion avenae* (F) (Hemiptera: Aphididae) studied using RAPD-PCR. *Proceedings of the Royal Society of London Series B-Biological Sciences*, 262:321-327.

De Barro, P. J., T. N. Sherratt, G. R. Carvalho, D. Nicol, A. Iyengar, and N. Maclean. 1995b. Geographic and microgeographic genetic differentiation in two aphid species over southern England using the multilocus $(GATA)_4$ probe. *Molecular Ecology*, 4:375-382.

De Barro, P. J., T. N. Sherratt, O. David, and N. Maclean. 1995c. An investigation of the differential performance of clones of the aphid *Sitobion avenae* on 2 host species. *Oecologia*, 104:379-385.

De Barro, P. J., T. N. Sherratt, S. D. Wratten, and N. Maclean. 1994. DNA fingerprinting of cereal aphids using $(GATA)_4$. *European Journal of Entomology*, 91:109-114.

Dixon, A. F. G. 1973. *Biology of Aphids*. Edward Arnold, London.

—. 1985. *Aphid Ecology*. Blackie & Son Limited, Glasgow.

Dunn, J. A. 1960a. The formation of galls by some species of *Pemphigus* (Homoptera: Aphididae). *Marcellia*, 30:155-167.

—. 1960b. The natural enemies of the lettuce root aphid, *Pemphigus bursarius* (L.). *Bulletin of Entomological Research*, 51:271-277.

Easy, G. M. S. 1982. Black Poplars in Cambridgeshire. *Nature in Cambridgeshire*, 25:45-50.

Eggers-Schumacher, H. A., and E. Sander. 1988. Spatial and seasonal genetic (allozyme) variation within field populations of *Phorodon humuli* Schrank (Homoptera: Aphididae) on winter and summer hosts in southern Germany (FRG). *Entomologist*, 107:110-121.

Ellegren, H. 1992. Polymerase chain reaction (PCR) analysis of microsatellites - a new approach to studies of genetic relationships in birds. *Auk*, 109:886-895.

Excoffier, L., and M. Slatkin. 1998. Incorporating Genotypes of Relatives into a Test of Linkage Disequilibrium., 62:171-180.

Excoffier, L., P. E. Smouse, and J. Quattro. 1992. Analysis of molecular variance inferred from metric differences among DNA haplotypes: Application to human mitochondrial DNA restriction data. *Genetics*, 131:479-491.

Fenton, B., A. N. E. Birch, G. Malloch, J. A. T. Woodford, and C. Gonzalez. 1994. Molecular analysis of ribosomal DNA from the aphid *Amphorophora idaei* and an associated fungal organism. *Insect Molecular Biology*, 3:183-189.

Fenton, B., J. A. T. Woodford, and G. Malloch. 1998. Analysis of clonal diversity of the peach-potato aphid, *Myzus persicae* (Sulzer), in Scotland, UK and evidence for the existence of a predominant clone. *Molecular Ecology*, 7:1475-1487.

Figueroa, C. C., J.-C. Simon, J.-F. Le Gallic, and H. M. Niemeyer. 1999. Molecular markers to differentiate two morphologically close species of the genus *Sitobion*. *Entomologia Experimentalis et Applicata*, 92:217-225.

Fisher, R. A. 1930. *The Genetical Theory of Natural Selection.* Oxford University Press, Oxford.

Foster, W. A. 1990. Experimental evidence for effective and altruistic colony defense against natural predators by soldiers of the gall-forming aphid *Pemphigus spyrothecae* (Hemiptera, Pemphigidae). *Behavioral Ecology and Sociobiology*, 27:421-430.

—. 1996. Duelling aphids: intraspecific fighting in *Astegopteryx minuta* (Homoptera: Hormaphididae). *Animal Behaviour*, 51:645-655.

Foster, W. A., and T. G. Benton. 1992. Sex ratio, local mate competition and mating behavior in the aphid *Pemphigus spyrothecae*. *Behavioral Ecology and Sociobiology*, 30:297-307.

Foster, W. A., and P. A. Northcott. 1994. Galls and the evolution of social behaviour in aphids. In M. A. J. Williams (ed.), *Plant Galls*, Vol. 49, pp. 161-182. Oxford University Press, Oxford.

Fukatsu, T., and H. Ishikawa. 1994. Differentiation of aphid clones by arbitrarily primed polymerase chain reaction (AP-PCR) DNA fingerprinting. *Molecular Ecology*, 3:187-192.

Fuller, S. J., P. Chavigny, L. Lapchin, and F. Vanlerberghe-Masutti. 1999. Variation in clonal diversity in glasshouse infestations of the aphid, *Aphis gossypii* Glover in southern France. *Molecular Ecology*, 8:1867-1877.

Furk, C. 1979. Field collections of *Aphis fabae* Scopoli s. lat. (Homoptera: Aphididae) studied by starch gel electrophoresis and isoelectric focusing. *Comparative Biochemistry and Physiology*, 62B:225-230.

Furk, C., and R. N. B. Prior. 1975. On the life cycle of *Pemphigus (Pemphiginus) populi*
Courchet with a key to British species of *Pemphigus* Hartig (Homoptera: Aphidoidea). *Journal of Entomology (B)*, 44:265-280.

Gaggiotti, O. E., O. Lange, K. Rassmann, and C. Gliddon. 1999. A comparison of two indirect methods for estimating average levels of gene flow using microsatellite data. *Molecular Ecology*, 8:1513-1520.

Goudet, J. 1995. FSTAT (vers. 1.2): a computer program to calculate F-statistics. *Journal of Heredity*, 86:485-486.

—. 1999. FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.8). Updated from Goudet (1995) http://www.unil.ch/izea/softwares/fstat.html.

Goudet, J., M. Raymond, T. de Meeus, and F. Rousset. 1996. Testing differentiation in diploid populations. *Genetics*, 144:1933-1940.

Green, R. F., and D. L. G. Noakes. 1995. Is a little bit of sex as good as a lot? *Journal of Theoretical Biology*, 174:87-96.

Guo, S. W., and E. A. Thompson. 1992. Performing the exact test of Hardy-Weinberg proportion for multiple alleles. *Biometrics*, 48:361-372.

Haldane, J. B. S. 1923. *The Causes of Evolution*. Princeton University Press, Princeton.

Hales, D., A. C. C. Wilson, J. M. Spence, and R. L. Blackman. 2000. Confirmation that *Myzus antirrhinii* (Macchiati) (Hemiptera. Aphididae) occurs in Australia, using morphometrics, microsatellite typing and analysis of novel karyotypes by fluorescence in situ hybridisation. *Australian Journal of Entomology*, 39:123-129.

Hamilton, W. D. 1964. The genetical evolution of social behaviour, I and II. *Journal of Theoretical Biology*, 7:1-52.

—. 1967. Extraordinary sex ratios. *Science*, 156:477-488.

—. 1975. Gamblers since life began: barnacles, aphids, elms. *Quarterly Review of Biology*, 50:175-180.

—. 1987. Kinship, recognition, disease, and intelligence: Constraints of social evolution. In Y. Itô, J. L. Brown and J. Kikkawa (eds.), *Animal Societies: Theories and Facts*, pp. 81-102. Japan Scientific Society Press, Tokyo.

Hamilton, W. D., and R. M. May. 1977. Dispersal in stable habitats. *Nature*, 269:578-581.

Hand, S. C. 1983. The effect of temperature and humidity on the duration of development and hatching success of eggs of the aphid *Sitobion avenae*. *Entomologia Experimentalis Et Applicata*, 33:220-222.

Hartl, D. L., and A. G. Clark. 1997. *Principles of Population Genetics*. 3rd ed. Sinauer Associates, Sunderland, MA.

Harvey, P. H., and M. D. Pagel. 1991. *The Comparative Method in Evolutionary Biology*. Oxford University Press, Oxford.

Hebert, P. D. N. 1987. Genotypic characteristics of cyclic parthenogens and their obligately asexual derivatives. In S. C. Stearns (ed.), *The Evolution of Sex and its Consequences*, pp. 175-195. Birkhauser Verlag, Basel.

Hebert, P. D. N., T. L. Finston, and R. Foottit. 1991. Patterns of genetic diversity in the sumac gall aphid, *Melaphis rhois*. *Genome*, 34:757-762.

Helden, A. J., and A. F. G. Dixon. 1997. Inbreeding and egg hatching success in *Sitobion avenae*. *Ecological Entomology*, 22:124-126.

Ishibashi, Y., T. Saitoh, S. Abe, and M. C. Yoshida. 1996. Null microsatellite alleles due to nucleotide sequence variation in the grey-sided vole *Clethrionomys rufocanus*. *Molecular Ecology*, 5:589-590.

Jalas, J., and J. Suominen. 1976. Salicaceae to Balanophoraceae, *Atlas Florae Europaeae - Distribution of the Vascular Plants in Europe*, Vol. 3. The Committee for Mapping the Flora of Europe & Societas Biologica Fennica Vanamo, Helsinki.

Kimura, M., and J. F. Crow. 1964. The number of alleles that can be maintained in a finite population. *Genetics*, 49:549-555.

Krebs, C. J., S. Boutin, R. Boonstra, A. R. E. Sinclair, J. N. M. Smith, M. R. T. Dale, K. Martin, and R. Turkington. 1995. Impact of food and predation on the snowshoe hare cycle. *Science*, 269:1112-1115.

Krebs, C. J., A. R. E. Sinclair, R. Boonstra, S. Boutin, K. Martin, and J. N. M. Smith. 1998. Community dynamics of vertebrate herbivores - how can we untangle the web? In H. Olff, V. K. Brown and R. H. Drent (eds.), *Herbivores: Between Plants and Predators*, pp. 447-473. Blackwells, Oxford.

Kurosu, U., and S. Aoki. 1990. Transformation of the galls of *Astegopteryx bambucifoliae* by another aphid, *Ceratoglyphina bambusae*. *Acta Phytopathologica et Entomologica Hungarica*, 25:113-122.

Lampel, G. 1969. Untersuchungen zur morphenfolge von *Pemphigus spirothecae* Pass. *Bulletin Naturforsch Gesellschaft Freiburg*, 58:56-72.

Levene, H. 1949. On a matching problem arising in genetics. *Annals of Mathematical Statistics*, 20:91-94.

Litt, M., and J. A. Luty. 1989. A hypervariable microsatellite revealed by invitro amplification of a dinucleotide repeat within the cardiac-muscle actin gene. *American Journal of Human Genetics*, 44:397-401.

Llewellyn, K. S. 1996. Microsatellite analysis to assess the extent of clonal mixing in the poplar gall aphid *Pemphigus spyrothecae* (Hemiptera: Pemphigidae). MA thesis, University of Cambridge.

Loxdale, H. D. 1990. Estimating levels of gene flow between natural populations of cereal aphids (Homoptera, Aphididae). *Bulletin of Entomological Research*, 80:331-338.

Loxdale, H. D., and C. P. Brookes. 1988. Electrophoretic study of enzymes from cereal aphid populations. V. Spatial and temporal genetic similarity of holocyclic populations of the bird-cherry oat aphid, *Rhopalosiphum padi* (L.) (Hemiptera, Aphididae), in Britain. *Bulletin of Entomological Research*, 78:241-249.

—. 1990. Temporal genetic stability within and restricted migration (gene flow) between local populations of the blackberry-grain aphid *Sitobion fragariae* in south-east England. *Journal of Animal Ecology*, 59:497-514.

Loxdale, H. D., C. P. Brookes, I. R. Wynne, and S. J. Clark. 1998. Genetic variability within and between English populations of the damson-hop aphid, *Phorodon humuli* (Hemiptera: Aphididae), with special reference to esterases associated with insecticide resistance. *Bulletin of Entomological Research*, 88:513-526.

Loxdale, H. D., J. Hardie, S. Halbert, R. Foottit, N. A. C. Kidd, and C. I. Carter. 1993. The relative importance of short-range and long-range movement of flying aphids. *Biological Reviews of the Cambridge Philosophical Society*, 68:291-311.

Loxdale, H. D., and G. Lushai. 1998. Molecular markers in entomology. *Bulletin of Entomological Research*, 88:577-600.

Loxdale, H. D., J. A. Rhodes, and J. S. Fox. 1985a. Electrophoretic study of enzymes from cereal aphid populations. IV. Detection of hidden genetic variation within populations of the grain aphid *Sitobion avenae* (F.) (Hemiptera, Aphididae). *Theoretical and Applied Genetics*, 70:407-412.

Loxdale, H. D., I. J. Tarr, C. P. Weber, C. P. Brookes, P. G. N. Digby, and P. Castañera. 1985b. Electrophoretic study of enzymes from cereal aphid populations. III. Spatial and temporal genetic variation of populations of *Sitobion avenae* (F.) (Hemiptera: Aphididae). *Bulletin of Entomological Research*, 75:121-141.

Lund, A. H., M. Duch, and F. S. Pedersen. 1996. Increased cloning efficiency by temperature-cycle ligation. *Nucleic Acids Research*, 24:800-801.

Lupoli, R., M. E. Irwin, and C. R. Vossbrinck. 1990. A ribosomal DNA probe to distinguish populations of *Rhopalosiphum maidis* (Homoptera: Aphididae). *Annals of Applied Biology*, 117:3-8.

Mabey, R. 1996. The Native Black Poplar: A species in a ghetto?, *British Wildlife*, Vol. 8, pp. 1-6.

Margaritopoulos, J. T., Z. Mamuris, and J. A. Tsitsipis. 1998. Attempted discrimination of *Myzus persicae* and *Myzus nicotianae* (Homoptera:Aphididae) by random amplified polymorphic DNA polymerase chain reaction technique. *Annals of the Entomological Society of America*, 91:602-607.

Martínez-Torres, D., R. Carrio, A. Latorre, J.-C. Simon, A. Hermoso, and A. Moya. 1997a. Assessing the nucleotide diversity of three aphid species by RAPD. *Journal of Evolutionary Biology*, 10:459-477.

Martínez-Torres, D., A. Moya, P. D. N. Hebert, and J.-C. Simon. 1997b. Geographic distribution and seasonal variation of mitochondrial DNA haplotypes in the aphid *Rhopalosiphum padi* (Hemiptera: Aphididae). *Bulletin of Entomological Research*, 87:161-167.

Martínez-Torres, D., A. Moya, A. Latorre, and A. Fereres. 1992. Mitochondrial-DNA variation in *Rhopalosiphum padi* (Homoptera, Aphididae) populations from 4 Spanish localities. *Annals of the Entomological Society of America*, 85:241-246.

Martínez-Torres, D., J.-C. Simon, A. Fereres, and A. Moya. 1996. Genetic variation in natural populations of the aphid *Rhopalosiphum padi* as revealed by maternally inherited markers. *Molecular Ecology*, 5:659-669.

May, B., and F. R. Holbrook. 1978. Absence of genetic variability in the green peach aphid, *Myzus persicae* (Hemiptera: Aphididae). *Annals of the Entomological Society of America*, 71:809-812.

Maynard Smith, J. 1964. Group Selection and Kin Selection. *Nature*, 201:1145-1147.

McAllister, M. K., and B. D. Roitberg. 1987. Adaptive suicidal behavior in pea aphids. *Nature*, 328:797-799.

Miller, N. J. 2000. Population Structure and Gene Flow in a Host Alternating Aphid, *Pemphigus bursarius*. PhD thesis, University of Birmingham.

Moran, N. A. 1992. The evolution of aphid life-cycles. *Annual Review of Entomology*, 37:321-348.

Moran, N. A., and T. G. Whitham. 1988. Evolutionary reduction of complex life-cycles - loss of host-alternation in *Pemphigus* (Homoptera, Aphididae). *Evolution*, 42:717-728.

Mosco, M. C., P. Arduino, L. Bullini, and S. Barbagallo. 1997. Genetic heterogeneity, reproductive isolation and host preferences in mealy aphids of the *Hyalopterus pruni* complex (Homoptera, Aphidoidea). *Molecular Ecology*, 6:667-670.

Mueller, U. G., and L. L. Wolfenbarger. 1999. AFLP genotyping and fingerprinting. *Trends in Ecology & Evolution*, 14:389-394.

Nei, M. 1987. *Molecular Evolutionary Genetics*. Columbia University Press, New York.

Nicol, D., K. F. Armstrong, S. D. Wratten, P. J. Walsh, N. A. Straw, C. M. Cameron, C. Lahmann, and C. M. Frampton. 1998. Genetic diversity of an introduced pest, the green spruce aphid *Elatobium abietinum* (Hemiptera: Aphididae) in New Zealand and the United Kingdom. *Bulletin of Entomological Research*, 88:537-543.

Parker, D. E., T. P. Legg, and C. K. Folland. 1992. A New Central England temperature Series, 1772-1991. *International Journal of Climatology*, 12:317-342.

Perring, F. H., Sell, P. D., Walters, S. M. & Whitehouse, H. L. K. 1964. *A flora of Cambridgeshire*. Cambridge University Press.

Perring, F. H., and S. M. Walters. 1982. Atlas of the British Flora. Botanical Society of the British Isles, London.

Powers, T. O., S. G. Jensen, S. D. Kindler, C. J. Stryker, and L. J. Sandall. 1989. Mitochondrial DNA diversity among greenbug (Homoptera: Aphididae) biotypes. *Annals of the Entomological Society of America*, 82:298-302.

Puterka, G. J., W. C. Black, W. M. Steiner, and R. L. Burton. 1993. Genetic variation and phylogenetic relationships among worldwide collections of the Russian wheat aphid, *Diuraphis noxia* (Mordvilko), inferred from allozyme and RAPD-PCR markers. *Heredity*, 70:604-618.

Rackham, O. 1986. *The History of the Countryside.* Dent, London.

Raymond, M., and F. Rousset. 1995a. An exact test for population differentiation. *Evolution,* 49:1280-1283.

—. 1995b. GENEPOP (version1.2): population genetics software for exact test and ecumenicism. *Journal of Heredity*, 86:248-249.

Reynolds, J., B. S. Weir, and C. C. Cockerham. 1983. Estimation of the co-ancestry coefficient - basis for a short-term genetic distance. *Genetics*, 105:767-779.

Rhoden, P. K. 1997. The evolution of social behaviour in aphids of the genus *Pemphigus.* PhD thesis, University of Cambridge.

Rice, W. R. 1989. Analyzing tables of statistical tests. *Evolution*, 43:223-225.

Rozen, S., and H. J. Skaletsky. 1996. Primer3. Code available at http://www-genome.wi.mit.edu/genome_software/other/primer3.html.

Rushforth, K. D. 1999. *Collins Wildlife Trust guide : a photographic guide to the trees of Britain and Europe.* HarperCollins, London.

Saiki, R. K., S. Scharf, F. Faloona, K. B. Mullis, G. T. Horn, H. A. Erlich, and N. Arnheim. 1985. Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle-cell anemia. *Science*, 230:1350-1354.

Sambrook, J., E. F. Fritsh, and T. Maniatis. 1989. *Molecular Cloning: A Laboratory Manual.* 2nd ed. Cold Spring Harbor Laboratory Press, New York.

Schneider, S., D. Roessli, and L. Excoffier. 2000. Arlequin ver. 2.000: A software for population genetics data analysis. Genetics and Biometry Laboratory, University of Geneva, Switzerland. http://anthro.unige.ch/arlequin.

Schug, M. D., T. F. C. Mackay, and C. F. Aquadro. 1997. Low mutation rates of microsatellite loci in *Drosophila melanogaster. Nature Genetics*, 15:99-102.

Setzer, R. W. 1980. Intergall migration in the aphid genus Pemphigus. *Annals of the Entomological Society of America*, 73:327-331.

Shufran, K. A., W. C. Black, and D. C. Margolies. 1991. DNA fingerprinting to study spatial and temporal distributions of an aphid, *Schizaphis graminum* (Homoptera, Aphididae). *Bulletin of Entomological Research*, 81:303-313.

Shufran, K. A., D. C. Margolies, and W. C. Black. 1992. Variation between biotype-e clones of *Schizaphis graminum* (Homoptera, Aphididae). *Bulletin of Entomological Research*, 82:407-416.

Shufran, K. A., and G. E. Wilde. 1994. Clonal diversity in overwintering populations of *Schizaphis graminum* (Homoptera, Aphididae). *Bulletin of Entomological Research*, 84:105-114.

Simon, J.-C., S. Baumann, P. Sunnucks, P. D. N. Hebert, J.-S. Pierre, J.-F. Le Gallic, and C.-A. Dedryver. 1999a. Reproductive mode and population genetic structure of the cereal aphid *Sitobion avenae* studied using phenotypic and microsatellite markers. *Molecular Ecology*, 8:531-545.

Simon, J.-C., and P. D. N. Hebert. 1995. Patterns of genetic variation among Canadian populations of the bird cherry-oat aphid, *Rhopalosiphum padi* L. (Homoptera, Aphididae). *Heredity*, 74:346-353.

Simon, J.-C., P. D. N. Hebert, C. Carillo, and R. Demelo. 1995. Lack of clonal variation among Canadian populations of the corn leaf aphid, *Rhopalosiphum maidis* Fitch (Homoptera, Aphididae). *Canadian Entomologist*, 127:623-629.

Simon, J.-C., N. Leterme, and A. Latorre. 1999b. Molecular markers linked to breeding system differences in segregating and natural populations of the cereal aphid *Rhopalosiphum padi* L. *Molecular Ecology*, 8:965-973.

Simon, J.-C., D. Martínez-Torres, A. Latorre, A. Moya, and P. D. N. Hebert. 1996. Molecular characterization of cyclic and obligate parthenogens in the aphid *Rhopalosiphum padi* (L.). *Proceedings of the Royal Society of London Series B-Biological Sciences*, 263:481-486.

Simon, J. P., M. A. Parent, and J. L. Auclair. 1982. Isoenzyme analysis of biotypes and field populations of the pea aphid, *Acyrthosiphon pisum*. *Entomologia Experimentalis et Applicata*, 32:186-192.

Slatkin, M. 1995. A measure of population subdivision based on microsatellite allele frequencies. *Genetics*, 139:457-462.

Slatkin, M., and N. H. Barton. 1989. A comparison of three indirect methods for estimating average levels of gene flow. *Evolution*, 43:1349-1368.

Slatkin, M., and L. Excoffier. 1996. Testing for linkage disequilibrium in genotypic data using the EM algorithm. *Heredity*, 76:377-383.

Smouse, P. E., J. C. Long, and R. R. Sokal. 1986. Multiple regression and correlation extensions of the Mantel Test of matrix correspondence. *Systematic Zoology*, 35:627-632.

Sokal, R. R., and F. J. Rohlf. 1995. *Biometry*. 3rd ed. Freeman, New York.

Spongberg, S. A. 1990. *A Reunion of Trees: the discovery of exotic plants and their introduction into North America and European landscapes.* Harvard University Press, Cambridge, MA.

Stern, D. L., and W. A. Foster. 1996. The evolution of soldiers in aphids. *Biological Reviews of the Cambridge Philosophical Society*, 71:27-79.

Sunnucks, P., P. J. De Barro, G. Lushai, N. Maclean, and D. F. Hales. 1997. Genetic structure of an aphid studied using microsatellites: Cyclic parthenogenesis, differentiated lineages and host specialization. *Molecular Ecology*, 6:1059-1073.

Sunnucks, P., P. R. England, A. C. Taylor, and D. F. Hales. 1996. Microsatellite and chromosome evolution of parthenogenetic *Sitobion* aphids in Australia. *Genetics*, 144:747-756.

Suomalainen, E., A. Saura, J. Lokki, and T. Teeri. 1980. Absence of genetic variation within parthenogenetic aphid clones. *Theoretical and Applied Genetics*, 57:129-132.

Taberlet, P., S. Griffin, B. Goossens, S. Questiau, V. Manceau, N. Escaravage, L. P. Waits, and J. Bouvet. 1996. Reliable genotyping of samples with very low DNA quantities using PCR. *Nucleic Acids Research*, 24:3189-3194.

Tautz, D. 1989. Hypervariability of simple sequences as a general source for polymorphic DNA markers. *Nucleic Acids Research*, 17:6463-6471.

Taylor, L. R. 1979. The Rothamsted Insect Survey - an approach to the theory and practice of synoptic pest forecasting in agriculture. In R. Rabb and G. Kennedy (eds.), *Movement of Highly Mobile Insects*, pp. 148-185.

Terradot, L., J.-C. Simon, N. Leterme, D. Bourdin, A. C. C. Wilson, J. P. Gauthier, and Y. Robert. 1999. Molecular characterization of clones of the *Myzus persicae* complex (Hemiptera: Aphididae) differing in their ability to transmit the potato leafroll luteovirus (PLRV). *Bulletin of Entomological Research*, 89:355-363.

Tomiuk, J. 1987. The neutral theory and enzyme polymorphism in populations of aphid species. In J. Holman, J. Pelikan, A. F. G. Dixon and L. Weisman (eds.), *Population structure, genetics and taxonomy opf aphids and Thysanoptera*, pp. 45-62. SPB Academic Publishing, The Hague.

Tomiuk, J., D. F. Hales, K. Wöhrmann, and D. Morris. 1991. Genotypic variation and structure in Australian populations of the aphid *Schoutedenia lutea*. *Hereditas*, 115:17-23.

Tomiuk, J., and K. Wöhrmann. 1980. Enzyme variability in populations of aphids. *Theoretical and Applied Genetics*, 57:125-127.

—. 1981. Changes of the genotype frequencies at the MDH-locus in populations of *Macrosiphum rosae* (L.) (Hemiptera, Aphididae). *Biologisches Zentralblatt*, 100:631-640.

—. 1983. Enzyme polymorphism and taxonomy of aphid species. *Zeitschrift Fur Zoologische Systematik Und Evolutionsforschung*, 21:266-274.

—. 1984. Genotypic variability in natural populations of *Macrosiphum rosae* (L.) in Europe. *Biologisches Zentralblatt*, 103:113-122.

Van Treuren, R. 1998. Estimating null allele frequencies at a microsatellite locus in the oystercatcher (*Haematopus osralegus*). *Molecular Ecology*, 7:1413-1417.

Vanlerberghe-Masutti, F., and P. Chavigny. 1998. Host-based genetic differentiation in the aphid *Aphis gossypii* Glover, evidenced from RAPD fingerprints. *Molecular Ecology*, 7:905-914.

Vanlerberghe-Masutti, F., P. Chavigny, and S. J. Fuller. 1999. Characterization of microsatellite loci in the aphid species *Aphis gossypii* Glover. *Molecular Ecology*, 8:693-695.

Vanpijlen, I. A., B. Amos, and T. Burke. 1995. Patterns of genetic variability at individual minisatellite loci in minke whale *Balaenoptera acutorostrata* populations from 3 different oceans. *Molecular Biology and Evolution*, 12:459-472.

Via, S. 1999. Reproductive isolation between sympatric races of pea aphids. I. Gene flow restriction and habitat choice. *Evolution*, 53:1446-1457.

Vos, P., R. Hogers, M. Bleeker, M. Reijans, T. Vandelee, M. Hornes, A. Frijters, J. Pot, J. Peleman, M. Kuiper, and M. Zabeau. 1995. AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Research*, 23:4407-4414.

Walsh, P. S., D. A. Metzger, and R. Higuchi. 1991. Chelex-100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. *Biotechniques*, 10:506-513.

Weber, G. 1985. Genetic variability in host plant adaptation of the green peach aphid, *Myzus persicae*. *Entomologia Experimentalis et Applicata*, 38:49-56.

Weber, J. L., and P. E. May. 1989. Abundant class of human DNA polymorphisms which can be typed using the polymerase chain-reaction. *American Journal of Human Genetics*, 44:388-396.

Weber, J. L., and C. Wong. 1993. Mutation of human short tandem repeats. *Human Molecular Genetics*, 2:1123-1128.

Weir, B. S. 1996. *Genetic Data Analysis II: Methods for Discrete Population Genetic Data.* Sinauer Associates, Sunderland, MA.

Weir, B. S., and C. C. Cockerham. 1984. Estimating *F*-statistics for the analysis of population structure. *Evolution*, 38:1358-1370.

Whitfield, J. A. 1998. Studies of the Diversity and Evolution of Soldier Aphids. PhD thesis, University of Cambridge.

Whitham, T. G. 1979. Territorial behaviour of *Pemphigus* gall aphids. *Nature*, 279:324-325.

Williams, A. G., and T. G. Whitham. 1986. Premature leaf abscission: an induced plant defence against gall aphids. *Ecology*, 67:1619-1627.

Williams, G. C. 1975. *Sex and Evolution*. Princeton University Press, Princeton, NJ.

Williams, J. G. K., A. R. Kubelik, K. J. Livak, J. A. Rafalski, and S. V. Tingey. 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Research*, 18:6531-6535.

Wilson, A. C. C., P. Sunnucks, and D. F. Hales. 1997. Random loss of X chromosome at male determination in an aphid, *Sitobion* near *fragariae*, detected using an X-linked polymorphic microsatellite marker. *Genetical Research*, 69:233-236.

—. 1999. Microevolution, low clonal diversity and genetic affinities of parthenogenetic *Sitobion* aphids in New Zealand. *Molecular Ecology*, 8:1655-1666.

Wöhrmann, K., and D. Hales. 1989. Life-cycle components and genetic variability in aphids. *Journal of Applied Entomology-Zeitschrift Fur Angewandte Entomologie*, 107:71-77.

Wool, D., H. F. van Emden, and S. Bunting. 1978. Electrophoretic study on genetic variation in British *Myzus persicae* Sulzer (Hemiptera: Aphididae). *Biochemical Genetics*, 16:987-1006.

Wright, S. 1951. The genetical structure of populations. *Annals of Eugenics*, 15:323-354.

Wynne, I. R., J. J. Howard, H. D. Loxdale, and C. P. Brookes. 1994. Population genetic structure during estivation in the sycamore aphid *Drepanosiphum platanoidis* (Hemiptera, Drepanosiphidae). *European Journal of Entomology*, 91:375-383.

Yamaguchi, Y. 1985. Sex ratios of an aphid subject to local mate competition with variable maternal condition. *Nature*, 318:460-462.

Young, J. P. W. 1983. The population structure of cyclic parthenogens. In G. S. Oxford and D. Rollinson (eds.), *Protein Polymorphism: Adaptive and Taxonomic Significance*, pp. 361-378. Academic Press, London.

APPENDIX 1: MICROSATELLITE SEQUENCES

```
DEFINITION  Pemphigus spyrothecae microsatellite 96PS2 sequence.
ACCESSION   AF246669
        1 gggctattgt agtattgtac tataactgtt atttgttgct gcgtcgctat gcataaagcc
       61 cgtgatgatg ataataggca gaagccgccg tgctggtgta ttaactgcca cttgcatgca
      121 cacacacaca cacaggcaca cgcacaaatg catgcaaaac atgttagcgc gcgcgcgcgc
      181 gtgcgcacac acacacacaa ttatgcacgc gcgtgatagg cgccatcaca atggtcacac
      241 acgcacacac gcacacacac acacacacac acacacacac acgcgcgcgc gcgcgcgcgc
      301 agtcacacac gtacacaaac acatgtgttt gttgttttta gcttgc

DEFINITION  Pemphigus spyrothecae microsatellite 96PS5 sequence.
ACCESSION   AF246670
        1 ccggatgnct ctttggtcta aacactttac aacagcgaaa tccaattaat gtgacgacga
       61 aacgttacaa caataatacc ggcaccagcc acttggacgg ttattaacga cctcgggtga
      121 actcgaaatt ttcgattata ccctgtacgc gagtgtttcg tccctttccc agcccccggt
      181 cccttgaaca cacgctcttc tacgttcctc tccggaccgt ctccggggac tttgccgaaa
      241 gccgccgaca catacgcgca ctcccgaaac cgcatcactg acgacgacga tacagattta
      301 atttcaattt tagcctgtcg cgttattaat ttcagcccaa acgtttgtgt gtgtgtgtgt
      361 gtgtgtgtgt gtgtgtgtgt gtgtgtctgt gtgtgtgtgt gtaatgccat tttcgggaga
      421 cactacgcga tattaacgtg caaacattgg cttccggaat atg

DEFINITION  Pemphigus spyrothecae microsatellite 96PS20 sequence.
ACCESSION   AF246671
        1 ccgaaacgaa atgtgtgata tgagacaaac gcaaatcgac acgcgaacga tggagaaaat
       61 aatatatatt atgatcctag accgatactg cgattgattg ttttacatac atatacgcat
      121 gtacacgtac acgtgtgtgt gtgtgtgtgt gtgtgtgtgt gtgtgtgtgt gtgtgtgtgt
      181 gtattttctc cgttgttttc cgatttgaaa aaattgtgat ggtgtttttc gagtcgacgg
      241 ctactggacg gcggcacgac acactgcgtg catatactgg ggacctgaaa aatcgatttg
      301 cgaacaccgc gtcccggggt cgcaccgttt tgttcatgcg catgaatcg

DEFINITION  Pemphigus spyrothecae microsatellite 96SA19 sequence.
ACCESSION   AF246672
        1 ccgtcttata cgtatatatt attacacaca cacacacaca cacacacaca cacacacaca
       61 cacacacaca tttatatata atatacatgg acatatataa tattatatat cattattata
      121 cgtaccagtg atgtgaatga gcactgtgac gcaaatgctc aacgtcgtct tcgaaggatt
      181 gtacaacatt ctgcagtggt tcgctcacgt ctcccggtgc angcttgtcg gtcggtcgat
      241 cggtcggtcg g

DEFINITION  Pemphigus spyrothecae microsatellite 96SA21 sequence.
ACCESSION   AF246674
        1 ccgttcgatt tttgttgttt gtgtacacaa acacatgcac acactgacgc gcgcgcgcgc
       61 gcgcgcacac acacacacac acacacacac acacacgcac acacgcacac actggtaaca
      121 ctaccgcgga tagtgcgcgc acgtattaac acacacacac acgcgtgcgc gcgcgcgcgc
      181 gattgtacaa aacgtacgta aacacgcaca cggacacaca cacacacacg tacgttcacc
      241 gtcaattatg tggtcgtgcc

DEFINITION  Pemphigus spyrothecae microsatellite 97PS2 sequence.
ACCESSION   AF246675
        1 ccgataacaa cataacatga tattcacaat aaacaacgac gcagcgatac gtatttcggg
       61 cactactact attatccgtc ttcggcggca cgaccacata attgacggtg aacgtacgtg
      121 tgtgtgtgtg tgtccgtgtg cgtgtttacg tacgttttgt acaatcgcgc gcgcgcgcgc
      181 acgcgtgtgt gtgtgtgtta atacgtgcgc gcactatccg cggtagtgtt accagtgtgt
      241 gcgtgtgtgc gtgtgtgtgt gtgtgtgtgt gtgtgtgtgt gcgcgcgcgc gcgcgcgcgt
      301 cagtgtgtgc atgtgtttgt gtacacaaac aacaaaaacg aacg

DEFINITION  Pemphigus spyrothecae microsatellite 97PS6 sequence.
ACCESSION   AF246676
        1 cccgtttgga tgtacgccag taaatttata acgcacgcac acgtacgcaa attcacgtac
       61 acacacacac accggcgacg catatacgcg cgcacacatg cacgtagaca agacataatg
      121 aacgctatta aaactgacca caaacttaca ccgtaccgca ccgatcccgt ccaccgtctg
      181 tgtatataac gtacgactgt gtgtgtgtgt gtgtgtgtgt gtgtgcgcat gtgtgtgtgt
      241 gtgtgtgcgt
```

```
DEFINITION  Pemphigus spyrothecae microsatellite 97PS9 sequence.
ACCESSION   AF246677
        1 ccgagggtcg ataacgatga cggtgtatgt gtggatggaa gaaaaaaaaa aaacacctta
       61 taaccgtctt gagaaatcgc acacacacac acacacacac acacacacac acacacgtcg
      121 atgagtgaga gtgtatttta taggatgcgt aaatatagtg gtagtagttg tagtggtata
      181 cgaaaagaaa aacagctggt gcggtctctc ggagttcgtt ctcacggttt gtgcgcgtat
      241 ttattttatg acgaaagaga gaacgagata gatatgggaa aagaaaaaag aaaaagaaaa
      301 aaaaaacgac cgttttgcca gcctcgcgcc tgtgatatta ggggatgggg acggaccacg
      361 tataatatat aaaataatat aataaagag

DEFINITION  Pemphigus spyrothecae microsatellite 97PS11 sequence.
ACCESSION   AF246678
        1 cctatcctaa ataaataaaa taacacacac acacacacac agaaaaacat aatataaaat
       61 aattcacaca cacacacaca cacacaacaa ttaacagatg caaaaataaa tataacgaaa
      121 atggctatta gagcgttaaa agtataaaat ttacaagggc aattcgtgcc gatgtgcaca
      181 atataaaaat gaactgtaat tgatgaaaag gctatttgag cctacaaaat attacggtaa
      241 ttggtgccga aaaaaacaat gtatcaatat tttgcatagc cgtaaaattt tatatagata
      301 cgtacactca gtacaggtct aaatcgcacg gcggatggaa atagacaaaa gaaaaccaac
      361 gcataataat ataatataat ataattcaca aaaaacagat aaagattaaa ttacgtatgt
      421 gtgtgtgatc agtgatgcag tggtggcgaa atgaccgtga aaagtggtat tcttg

DEFINITION  Pemphigus spyrothecae microsatellite 97PS12 sequence.
ACCESSION   AF246679
        1 cgttcaccaa aaaacgtttc tgccgctcgc gcgcgcccgc ccgcctcgtt aaatttctcc
       61 cgtatatacc ggtcgtgcgg ttcgacccac aaccgcggta agcccgactt tatacgacgg
      121 accgtatata tacggacgaa gagcggtggc attccggtcc ggacgaaacg ggtaaactta
      181 tgtatcgcac aggttacaca cacacacaca cacacacaca cacacacaca cacgtattca
      241 catattatac tgaacatgtg cgtgtataac atgcgtatac gagaattata agttcgtcgt
      301 aacgttcacc tcatatagga accggacact cgtgtgcgcg cgcgcacggc atattacgta
      361 tatactttat tattataaat agtccgtaca atttatatat acattatata atacagtaca
      421 tttattatat ctaaaataat atcactgcgt gtaaatatat acatatatat atatatatat
      481 aatatggtaa gctctccggt ggaattccac ccctttaact tttccccgaa ctcggatgtt
      541 ccnggtggaa attaatt

DEFINITION  Pemphigus spyrothecae microsatellite 97PS15 sequence.
ACCESSION   AF246680
        1 gtcttatacg tatatattat tacacacaca cacacacaca cacacacaca cacacacaca
       61 cacacacatt tatatataat tacatggaca tatataatat tatatatcat tattatacgt
      121 accagtgatg tgaatgagca ctgtgacgca aatgctcaac gtcgtcttcg aaggattgta
      181 caacattctg cagtggttcg ctcacgtctc ccggtgcagg cttgtcggtc ggtcgatcgg
      241 tcggtcggtc ggtcggtcgg tacgcagtca attatgtcgg gtacgatata ataatatgtc
      301 cggtgtccac agacgaggaa gaccgacaat gtgatattac ggtcgacgtg gtgcgaattt
      361 gaacgcgcga ctgtcgtgtc tcgcgccgtg ccactccaat gtaccctacc cgcgacccgt
      421 accgtcgcac gc

DEFINITION  Pemphigus spyrothecae microsatellite 98PS2 sequence.
ACCESSION   AF246681
        1 ctgtgtacct aacctgccta cctatataat attataatgt atataagcat cggaccagag
       61 agagagaaag aacatatttt ccgacatgta taatgtataa tgtatatata tatatatata
      121 ttattattat tattattatt attgttgttg ttgtttttgt tgctatacaa aactggcgac
      181 ggtggtcgta ctccgtcccg acggaaagtc atcaaaaata cacacgggat agatagatag
      241 agatatgtgt gtgtgtgtgt gtgtgtgtgt gtgtgtgtaa gatacggata gatataatat
      301 agatagaacg agtttgtgtg cgagacggag acgaccacat atcgctgctt ttataaatgg
      361 aaactatatg accgcagcaa aagcaacagc a

DEFINITION  Pemphigus spyrothecae microsatellite 98PS8 sequence.
ACCESSION   AF246682
        1 ccgactgaat ttatattata taaccgttat aataataatg aagtcatttc gataatgaat
       61 ttatcatcgc gtacaattat aatttaccta tataataata atataatatc ccatatagtt
      121 atattacacg tcattgtgtt tgttttttaa acgagccgta cacctgtccc tgcagtattg
      181 cagttagtag tcgggaattt gtattatata tatatatata tatatcggaa taacataata
      241 taagcaataa ataattgatc gtatgtgaca accgttggtg gcgtattgtg tgtgtgtgtg
      301 tgtgtgtttc atattatttt taatgacatg tcgtaaaata ctggtcggta aaaaaatgg
```

```
DEFINITION  Pemphigus spyrothecae microsatellite 98PS12 sequence.
ACCESSION   AF246683
        1 cctatagtaa taattctgat caaaactgtg taggtagata gagaaaaaat atatcattgt
       61 ataaaatgat ataccaagta tactaataag tcaaataatt gagtgtttat aaagaatcat
      121 tgtctataat aatgatctgt aaggtgccgt acttttttct acattaatga tcaagatatt
      181 gtgatcagta gtgactggtg agagattttt cactatatgt acgtggcaca aatgcacaac
      241 actacaataa tcaatattat actatacgtt atgttataac taataaacag ccgatattta
      301 tatttatttt attattatta tcttttctta taacttataa gatttaatat atctccacaa
      361 aattgcgtag ttaaaacatt tagttagaac ttagtaacag gttggaagca cttttgttcg
      421 ttaatcagat atctataatg aaacacgatg acaatataag gtctgtgatc aatatacaat
      481 tttaatccaa ataataaatt aaaattgtct ttcctcctag tggtcctgtt gtatctatat
      541 atttctcgtt tattagacgt cgctgttctc tttatatcgc caccgtatat atatatatat
      601 gtgtgtgtgt gtgtgtgtgt gtgtgtgtgc ggccnaggaa tanaatttta

DEFINITION  Pemphigus spyrothecae microsatellite 98PS14 sequence.
ACCESSION   AF246684
        1 gctcattttt taaaagcccg atagcagcac cctttcttca atttttccaa taattcgacc
       61 gccgaatttg tcagtaatgg tcgtcggtgc cgaagttttt ttttttttaa gattgataag
      121 aaaaaaaaaa cattataggc atattatatt aaatctcggt tcggatcgtt tcaattataa
      181 cacatgggct tatcgtaata aatttatgtt tctcagtacg gatcattcgt gtcgtaacac
      241 cctgtccgac acgacgaatc gcgggcgggg gaacacatag tactgtncat gattattata
      301 ataatgtgcg cccgttttat ttttcttctt cncnaattta attaataaca tacacacaca
      361 cacacacaca cacacact

DEFINITION  Pemphigus spyrothecae microsatellite 98PS21 sequence.
ACCESSION   AF246685
        1 ttttttatcg aacaaaatat ttgtcagcgg taaagccaca gttgtattgt taattgaatt
       61 tgtacattta tcaatattag gtgccaagtt tttattttca gtgtaatcgg aaaataaccg
      121 ctacataatg ggcacttgtg tttagtgtcc atacaactac acgataaatt ttttgattga
      181 ttttagttca tttttgtaca tgagtatata caatatagaa ataggaaaag ttactgggaa
      241 tatatgtgtg tgtgtgtgtg tgtgtgtgtg tgtgtgtgtg taatatatag agaaattaat
      301 ttcctaaatt gacatttgtt tatcgaaatg gaatctaacg aagtatattt cgtaagaaaa
      361 tcaaatttat attattttat cgaattaata ctgaagtaaa gtaaatattt ctgtataaac
      421 ggtcattggc aaaccaactg tggaatatct tctgaagatc gttcgataca aacgtgttat
      481 taaattacat tgggctattt ttacttaaaa agttatcgta aaacaaatag tttttttgag
      541 taagtatttg

DEFINITION  Pemphigus spyrothecae microsatellite 98PS23 sequence.
ACCESSION   AF246686
        1 ctatgaacat gagtttcaat tcaagtaatg ttatttaatt gtctttctat actaatttct
       61 ggggtcgcaa tttcggtttt tcgttaaaga aaggcatctg tcgactaaat gtagggcaat
      121 atgatattta tatatacatg gatataaaaa ctatatggct aagaaggttt ctagagaata
      181 agtacgttta atcattagcc ccgcaaatcg aatcatttcg ctcgattttc tttttgcgtg
      241 ggcggacgcc tttaaattac acatacacac acacacacac acacacacat acacacacat
      301 atatacttcg gtgtgtatat tatatatatt atttattcga taatataata acgtacacgt
      361 ctaactaagt ataacgcgaa atatacgaaa tggaaacgcc accgtgcact gtatattatg
      421 cactacacta tatacacgag acccgtgtat taatatatat agtgttcggg gaggggggag
      481 ggatgcggag agcgttttgc ggcacgttca cgttcttcat ttagaccatt gtctatatat
      541 tatatgcgcg tcg

DEFINITION  Pemphigus spyrothecae microsatellite 98PS25 sequence.
ACCESSION   AF246687
        1 ctattggttt tatatatttt aatttgtatt gatgataata aatatatgcg gtatttggtt
       61 atgtatttag tccggtattg ggtcgtgttt gtgcggtaat gaaaagcacc gtaactttaa
      121 aacaaattaa aggttaaaaa ataaatattt atgtacaaat attatgaata cgagaaaact
      181 taaaactgta gttacataaa aacgttaaca ccacgagtta ttgttaacgt aacgttatag
      241 aaaatacttg ccatggaggt cgaactcgtg tgtatccctc aagtatatat aataatggtt
      301 ttcaatttca aacattcgaa ttgcgttgaa gcgtcgtaat aatacaaatt ataataatat
      361 tatatactgt attataataa ttttatatac gacgtataaa attaaaaagt ctttaaaaat
      421 cattgggttt ttttttttca tttacgatta tatatataat atatatatat atggaaaatg
      481 gtatttatat aaaacataaa cgaaatccaa tttggaccta ataaatatta tttaacaacc
      541 gattggtgtg tgtgtgtgtg tgtgtgtgtg tgtgtgtgta tatatttatt aatata

DEFINITION  Pemphigus spyrothecae microsatellite 98PS26 sequence.
ACCESSION   AF246688
        1 ccggtattat gcgcggatca atatacacag ggtatatata caccatgtgt cacgccgcgg
       61 cgtgcgtgtg tgtgtgtgtg tgtgtgtgtg tgtgtgtgct tctccgtcgc ctaagtcgaa
      121 tacgaccgtc ctgatgatcc gtccaaagtt atgcgccacc cactaatgtt gttccgattt
      181 ttcctggtgc gtttacaccg gtttatcctt ttattcccga acaattcatg gtgcccgatg
      241 acccttctct tgatcgaatt ccaatacgtt gtccgccata gaactttaac cgagg
```

APPENDIX 2: PUBLICATION

Microsatellite loci for studying clonal mixing, population structure and inbreeding in a social aphid, *Pemphigus spyrothecae* (Hemiptera: Pemphigidae)

PAUL. C. D. JOHNSON,*
KATE. S. LLEWELLYN† and WILLIAM AMOS*

**Department of Zoology, University of Cambridge, Downing Street, Cambridge CB2 3EJ, UK, †Entomology and Nematology Department, IACR-Rothamsted, Harpenden, Hertfordshire AL5 1JP, UK*

Keywords: aphid, clonal mixing, inbreeding, local mate competition, microsatellite markers, *Pemphigus spyrothecae*

Received 3 April 2000; revision accepted 20 April 2000

Correspondence: Paul Johnson. Fax: + 44 1223336676; E-mail: pcdj2@hermes.cam.ac.uk

The aphid *Pemphigus spyrothecae* forms galls on the leaf petioles of the black poplar *Populus nigra*. The colony of aphids within each gall is produced asexually by a single sexually produced foundress (Lampel 1960). *P. spyrothecae* provides an attractive system for studying the relationship between a number of life history traits and the distribution of genetic variation. First, it is social: a morphologically and behaviourally distinct soldier caste attacks intruding predators and cleans the gall, potentially reducing its own fitness (Foster 1990; Benton & Foster 1992; Stern & Foster 1996). Kin selection theory suggests that such altruistic behaviour is most likely to evolve and persist where mixing between aphid clones is minimal (Hamilton 1987). Setzer (1980) found high levels of interclone mixing (up to 24%) in two *Pemphigus* species which do not produce morphologically distinct soldiers, but a soldier-producing species has yet to be investigated. Second, whereas most other *Pemphigus* species alternate between a primary host where sexual reproduction occurs, and a secondary host where they undergo several asexual generations, *P. spyrothecae* has lost its secondary host and completes its entire life cycle on *P. nigra* (Blackman & Eastop 1994). Such a strategy, though it may bring advantages such as range expansion (Moran 1988), might also invite population isolation and inbreeding. Third, *P. spyrothecae* has a strongly female-biased primary sex ratio, which is thought to have evolved in response to local mate competition (LMC) (Yamaguchi 1985; Foster & Benton 1992). Under the conditions of LMC brothers compete to mate with their sisters (Hamilton 1967), so populations should be highly inbred. Foster & Benton (1992) found no evidence for avoidance of brother–sister matings in the laboratory, but the frequency of inbreeding in natural populations is unknown. To investigate these three problems we developed eight variable microsatellite markers in *P. spyrothecae*.

Genomic DNA was phenol–chloroform extracted according to Sambrook *et al.* (1989) from bulked aphids taken from a single gall. Two genomic libraries were made, one using *Hae*III alone, the other digested with both *Hae*III and *Alu*I. Size-selected fragments in the 300–800 bp range were ligated into *Sma*I-cut pUC19 using temperature cycle ligation (Lund *et al.* 1996), transformed into DH5α *Escherichia coli*, and screened with ^{32}P-labelled $(AC)_{15}$ synthetic oligonucleotides according to Sambrook *et al.* (1989). Positive colonies were restreaked and rescreened prior to BigDye cycle sequencing (PE Applied Biosystems). Sequences were analysed on an ABI PRISM 377 (PE Applied Biosystems), yielding 19 microsatellites. Primers were designed from the 13 microsatellite sequences with sufficient flanking sequence and the longest repeats using the program PRIMER 3.0 (Rozen & Skaletsky 1996).

Genomic DNA for polyermase chain reaction (PCR) was extracted from whole aphids preserved in 80% ethanol or 100% isopropanol using a protocol adapted from Walsh *et al.* (1991). Each aphid was washed in 70% ethanol to remove potential plant-derived PCR inhibitors, air-dried at 55 °C, placed in 30 μL H_2O, frozen at −20 °C, then microwaved for 3 min at 650 W. Cells were further lysed by adding 30 μL of 2× extraction buffer (10% Chelex 100™ chelating resin, 0.5 mg/mL proteinase K, 200 mM dithiothreitol) and incubating for 3–8 h at 55 °C. Proteinase K was then inactivated by microwaving for 2 min.

PCRs were carried out in 10 μL reaction volumes containing 1× NK PCR buffer 67 mM Tris (pH 8), 160 mM $(NH_4)_2SO_4$, 50 mM KCl, 0.1% Tween (20), 1.5 mM MgCl, 0.1 mM dGTP, 0.1 mM dATP, 0.1 mM dTTP, 0.01 mM dCTP, 400 nM of each primer, 0.25 U *Taq* polymerase (Promega) and 0.1 μCi [α^{32}P]-dCTP. PCR conditions were: 2 min denaturing at 94 °C; seven cycles of 40 s at 94 °C, 1 min at 48 °C, 40 s at 72 °C; and 40 cycles of 40 s at 89 °C, 1 min at 52 °C, 40 s at 72 °C. PCR

Table 1 Characteristics of eight polymorphic microsatellite marker loci in *Pemphigus spyrothecae*. Exact tests of Hardy–Weinberg equilibrium using GENEPOP (version 3.1d) (Raymond & Rousset 1995) showed significant heterozygote deficits (* $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$)

Locus (GenBank Acc. no.)	Repeat motif	Primer sequence (5′ to 3′)	Size (bp)	No. of alleles	H_O	H_E
96PS5 (AF246670)	$(GT)_{28}$	ACGATACAGATTTAATTTCAATTTTAG CCAATGTTTGCACGTTAATATC	164	7	0.425***	0.565
96PS20 (AF246671)	$(GT)_{25}$	TTTACATACATATACGCATGTACAC TTTTCAGGTCCCCAGTAT	190	5	0.636*	0.732
97PS9 (AF246677)	$(AC)_{18}$	CCTTATAACCGTCTTGAGAAATCG ATAAATACGCGCACAAACCG	189	2	0.237*	0.329
97PS12 (AF246679)	$(AC)_{19}$	AAGCCCGACTTTATACGACG CACGAGTGTCCGGTTCCTAT	235	5	0.663**	0.763
98PS8 (AF246682)	$(AT)_{10}N_{62}(CT)_{10}$	CCCTACACCTCTCCCTGC TTTTTACCGACCAGTATTTTACGAC	199	4	0.482***	0.699
98PS12 (AF246683)	$(AT)_{7}(GT)_{16}$	TCTTTCCTCCTAGTGGTC TTAATGATATTCCTAATACACAC	142	6	0.524***	0.759
98PS23 (AF246686)	$(AC)_{12}$	AAATCGAATCATTTCGCTCG AATATACAGTGCACGGTGGC	213	3	0.128	0.122
98PS26 (AF246688)	$(GT)_{16}$	GCGCGGATCAATATACACAG AAGTTCTATGGCGGACAACG	276	2	0.278***	0.497

products were resolved on 6% denaturing polyacrylamide gels and visualized by autoradiography.

Initial screening of 11 aphids from sites around Europe showed that eight of the 13 loci were polymorphic. To determine levels of variability at the eight loci, one aphid from each of 88 galls was genotyped. The galls were taken from nine trees from three locations in the vicinity of Modena, Italy. The number of alleles at each locus ranged from two to seven, and levels of expected heterozygosity ranged from 0.122 to 0.763 (Table 1). Seven loci showed significant heterozygote deficits, suggesting population subdivision or inbreeding. Screening of *P. spyrothecae* populations from the UK and elsewhere in northern Europe showed levels of variability much lower than in Italy, and many alleles were unique to either area, indicating long-term isolation between these populations. We conclude that the markers described here have revealed ample variation for further genetic analysis of *P. spyrothecae* populations.

Acknowledgements

Natalia Giorgini, Stefano Silvestri and Christian Brändle collected samples. We thank William Foster and Karen Fullard for advice on the manuscript. This study was funded by NERC (Natural Environment Research Council).

References

Benton TG, Foster WA (1992) Altruistic housekeeping in a social aphid. *Proceedings of the Royal Society of London B Series*, **247**, 199–202.

Blackman RL, Eastop VF (1994) *Aphids on the World's Trees: an Identification and Information Guide*. Natural History Museum, London.

Foster WA (1990) Experimental evidence for effective and altruistic colony defence against natural predators by soldiers of the gall-forming aphid *Pemphigus spyrothecae* (Hemiptera: Pemphigidae). *Behavioural Ecology and Sociobiology*, **27**, 421–430.

Foster WA, Benton TG (1992) Sex ratio, local mate competition and mating behaviour in the aphid *Pemphigus spyrothecae*. *Behavioural Ecology and Sociobiology*, **30**, 297–307.

Hamilton WD (1967) Extraordinary sex ratios. *Science*, **156**, 477–488.

Hamilton WD (1987) Kinship, recognition, disease, and intelligence: Constraints of social evolution. In: *Animal Societies: Theories and Facts* (eds Itô Y, Brown JL, Kikkawa J), pp. 81–102. Japan Scientific Society Proceedings, Tokyo.

Lampel G (1960) Die morphologischen und ökologischen Grundlagen des Generationswechsels monözischer und heterözischer Pemphiginen der Schwarz- und Pyramidenpappel. *Zeitschrift Fur Angewandte Entomologie*, **47**, 334–375.

Lund AH, Duch M, Skou Pedersen F (1996) Increased cloning efficiency by temperature-cycle ligation. *Nucleic Acids Research*, **24**, 800–801.

Moran NA (1988) Evolutionary reduction of complex life cycles: loss of host-alternation in *Pemphigus* (Homoptera: aphididae). *Evolution*, **42**, 717–728.

Raymond M, Rousset F (1995) GENEPOP (Version 1.2): population genetics software for exact test and ecumenicism. *Journal of Heredity*, **86**, 248–249.

Rozen S, Skaletsky HJ (1996–98) Primer 3. Code available at http://www-genome.wi.mit.edu/genome_software/other/primer3.html.

Sambrook J, Fritsh EF, Maniatis T (1989) *Molecular Cloning: a Laboratory Manual*. 2nd edn. Cold Spring Harbor Laboratory Press, New York.

Setzer RW (1980) Intergall migration in the aphid genus *Pemphigus*. *Annals of the Entomological Society of America*, **73**, 327–331.

Stern DL, Foster WA (1996) The evolution of soldiers in aphids. *Biological Reviews*, **71**, 27–79.

Walsh P, Metzger DA, Higuchi R (1991) Chelex 100 as a medium for simple extraction of DNA for PCR-based typing from forensic material. *Bio-Techniques*, **10**, 506–513.

Yamaguchi Y (1985) Sex ratios of an aphid subject to local mate competition with variable maternal condition. *Nature*, **318**, 460–462.