Molecular Ecology (2001) **10**, 761–778

Differential success in northwards range expansion between ecotypes of the marble gallwasp *Andricus kollari*: a tale of two lifecycles

GRAHAM STONE,* RACHEL ATKINSON,* ANTONIS ROKAS,* GYÖRGY CSÓKA† and JOSÉ-LUIS NIEVES-ALDREY‡

**Edinburgh University, Institute of Cell, Animal and Population Biology, The Ashworth Laboratories, The King's Buildings, West Mains Road, Edinburgh EH9 3JT, UK, †Forest Research Institute, Department of Forest Protection, 3232 Mátrafüred, P.O. Box 2, Hungary, ‡Museo Nacional de Ciencias Naturales, José Gutiérrez Abascal 228006 Madrid, Spain*

Abstract

The Marble gallwasp *Andricus kollari* has a native range divided into two geographically separated lifecycles. In Eastern Europe and Turkey, the lifecycle involves a sexual generation on Turkey oak, *Quercus cerris*, while in Iberia and North Africa the sexual generation host is cork oak, *Q. suber*. Over the last 500 years, *A. kollari* has expanded its range into northern Europe, following human planting of *Q. cerris* from Italy and the Balkans. We ask: (i) what is the genetic relationship between eastern and western distributions of *Andricus kollari*? Can we determine which lifecycle is ancestral, and how long ago they diverged? (ii) To what extent have eastern and western native ranges contributed to northwards range expansion? (iii) Is there any evidence for hybridization between the two life cycle types? We present analyses of allozyme data for 13 polymorphic loci and of sequence variation for a 433 bp fragment of the mitochondrial cytochrome *b* gene. These show: (i) that four haplotype lineages (one in Spain, two in Hungary/Italy and one in Turkey) diverged more or less simultaneously between 1 and 2 million years ago, suggesting the existence of at least four refuges through recent ice age cycles. Our data cannot resolve which lifecycle type is ancestral. (ii) Populations north of putative refuges are divided into two sets. Populations in south-west France are allied to Spain, while all remaining populations in northern Europe have been colonized from Italy and the Balkans. (iii) The transition from one race to another in south-west France is marked by abrupt transitions in the frequency of refuge-specific private alleles and corresponds closely to the northern limit of the distribution of cork oak. Although hybrids were detected in north-west France, none were detected where the two lifecycles meet in south-western France. The biology of oak gallwasps predicts that any hybrid zone will be narrow, and limited to regions where *Q. cerris* and *Q. suber* meet. Our data suggest that eastern and western *A. kollari* are effectively separate species.

Keywords: *Andricus*, gallwasp, glacial refugia, host race, invasion, range expansion

Received 4 June 2000; revision received 26 August 2000; accepted 26 August 2000

Introduction

Patterns of genetic diversity in many European plants and animals are dominated by histories of postglacial range expansion. During the last glacial, from 115 000 to ≈ 15 000 years ago, many terrestrial organisms were confined to refuges in southern Europe (Huntley 1990; Hewitt 1996).

Correspondence: Dr Graham N. Stone. Fax: 0131 667 3210; E-mail: graham.stone@ed.ac.uk

For many animals and plants, ice age refuges in southern Iberia, Italy, the Balkans and Turkey remain centres of genetic diversity (reviewed by Hewitt 1996, 1999). Refugial populations of many organisms were separated for long enough for refuge-specific alleles and haplotypes to evolve (reviewed by Hewitt 1996, 1999). Screening of such polymorphism in postglacial populations allows assessment of the genetic contribution of particular refuges to northward migration following retreat of the ice sheets. Northern European populations of some species show genetic input

from several refuges, while in others, large areas are dominated genetically by migrants from a single refuge (reviewed by Taberlet *et al*. 1998; Hewitt 1999). Differential genetic contributions from alternative refuges can sometimes be explained by physical features that restrict gene flow, such as the orientation of watersheds for aquatic organisms (Bernatchez & Wilson 1998). In other cases, biotic interactions may be important. For parasites or predators, escape from refuges can only follow range expansion by specific hosts or prey. Where hosts and parasites exist together in subdivided refuge populations, genetic divergence in host populations may lead to associated genetic divergence in parasite populations, the evolution of host-specific ecotypes, and ultimately speciation. If alternative parasite gene pools are available to contribute to northwards range expansion, which one succeeds may then be determined by which refuge or refuges contribute genotypes to the expanding host population.

Most analyses of postglacial range expansion concern events that occurred thousands of years ago, and it is rarely possible to track differential contributions from available refuges as they happen. In these cases, patterns generated by the range expansion process may have been modified by subsequent mutation, drift and selection. This study minimizes the impact of such modification by analysing an example of recent range expansion over a far shorter timescale. The Marble Gallwasp *Andricus kollari* (Hartig, 1843) (Hymenoptera: Cynipidae, Cynipini) is an obligate parasite of oak trees, and native to southern Europe, where it has two discrete lifecycles (described below). This species has invaded northern Europe over the last 500 years following human introduction of an oak host, Turkey oak (*Quercus cerris*). We examine the relationship between the two lifecycle types, and ask which has contributed more colonists to the invasion of northern Europe, and why.

Geographic variation in the lifecycle of A. kollari

As for many oak gallwasps, the lifecycle of *A. kollari* involves obligate cyclical parthenogenesis between a sexual generation in the spring, and a parthenogenetic generation in the summer/autumn (Askew 1984). In most oak cynipids, the galls of the two generations develop on the same oak host. *A. kollari* belongs to a group of gallwasps, found only in the genera *Andricus* and *Callirhytis*, whose lifecycles involve *host alternation* (or heteroecy; Folliot 1964), in which the sexual and parthenogenetic generations develop on different oak taxa (Nieves-Aldrey 1982, 1992; Stone & Sunnucks 1993; Cook *et al*. 1998). The natural distributions of such gallwasps are restricted to regions in which both host oak taxa occur.

A. kollari has a natural distribution divided geographically into two lifecycle types. From Italy east through the Balkans and Turkey the sexual generation develops in tiny bud galls (originally known as *A. circulans* Mayr, 1870) on Turkey oak (*Q. cerris*). The parthenogenetic generation develops in larger 'marble galls' on oaks in the section Quercus (*Q. robur, Q. petraea, Q. pubescens*, and others) (Beijerinck 1902; Marsden-Jones 1953; Folliot 1964). The natural distribution of this lifecycle in Europe is restricted to the area south and east of the Alps where *Q. cerris* and suitable oaks in the section Quercus naturally occur together (Fig. 1).

A. kollari is also found in Spain, Portugal and north-western Africa (Houard 1912; Nieves-Aldrey 1987; Ros-Farré & Pujade 1998). As in the eastern part of its range, the parthenogenetic generation develops in marble galls on oaks in the section Quercus. There is no *Q. cerris* in Iberia, and the sexual generation host is cork oak *Q. suber*. This oak is closely related to *Q. cerris* (Manos *et al*. 1999) and occurs throughout the western distribution of *A. kollari* (Fig. 1). The sexual generation of *A. kollari* in Spain has only been recently identified by careful rearing experiments (Pujade-Villar 1991, 1992; J. Pujade-Villar & R. Folliot, personal communication), and here we use analytical methods applied to other cyclical parthenogens (Sunnucks *et al*. 1997; Simon *et al*. 1999; Gómez & Carvalho 2000) to confirm the presence of sex across our sampled western distribution sites.

Human planting of Q. cerris *and gallwasp range expansion*

Over the last 500 years, *Q. cerris* has been planted extensively north of its native range (Fig. 1), precipitating rapid colonization of northern Europe by four host-alternating *Andricus* species (Docters van Leuwen 1959; Hails & Crawley 1991; Stone & Sunnucks 1993; Stone *et al*. 1995; Schönrogge *et al*. 1998). Three of these species — *A. ambiguus* (formerly identified as *A. corruptrix*), *A. lignicola* and *A. quercuscalicis* — have natural distributions restricted to the native range of *Q. cerris* (Fig. 1), and had a single eastern origin for range expansion into Europe. The fourth species is *A. kollari*, which differs from the other three in that either or both of eastern and western lifecycles could have contributed to northwards range expansion.

The ability of eastern and western lifecycles of *A. kollari* to contribute migrants to range expansion will depend on the ability of each generation in the lifecycle to locate and to induce galls on available oak hosts. The sexual generation must lay eggs, and the asexual generation induce galls, on available oaks in the section Quercus (*Q. petraea, Q. pubescens, Q. pyrenaica* and *Q. robur*). Extensive surveys of chloroplast DNA (cpDNA) (Ferris *et al*. 1993, 1995; Dumolin-Lapègue *et al*. 1997; Petit *et al*. 1997) show section Quercus species across northern continental Europe to contain haplotypes derived from both Iberian and central European refuges, suggesting that both eastern and western lifecycles of *A. kollari* should encounter suitable host genotypes. Evidence from other range expanding *Andricus* also suggests that the refugial origin of section Quercus oaks has little impact on their susceptibility as hosts; gallwasps

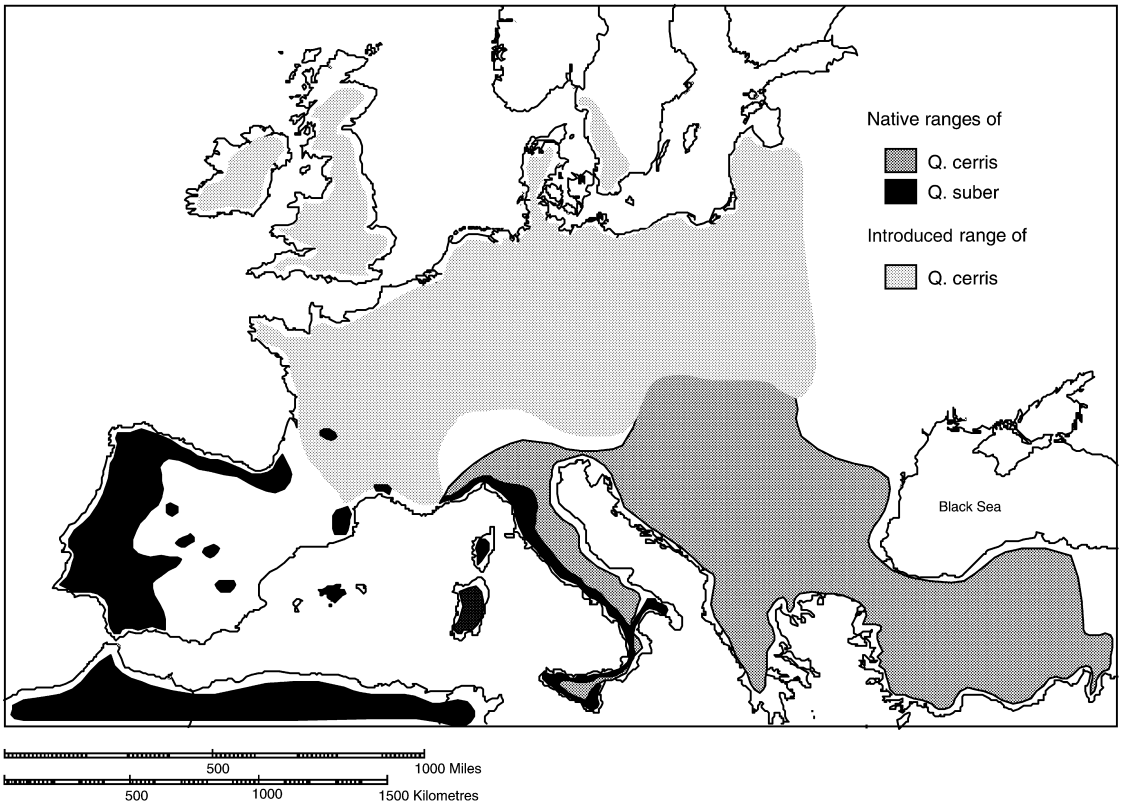

Fig. 1 The natural distributions of *Quercus cerris* and *Q. suber*, and the introduced distribution of *Q. cerris*. Suitable section Quercus oaks are found throughout the distributions of both these oak species. The natural distribution of the eastern lifecycle of *Andricus kollari* corresponds to the native range of *Q. cerris*. The lifecycle involving *Q. suber* is known to occur in Spain, but it is not known if this host is also used in Italy. Oak distributions in Europe are based on maps in Jalas & Suominen (1987) and Toumi & Lumaret (1998), and in Turkey on maps in Yaltirik (1984). *Q. suber* is probably planted in some of the unshaded areas in Iberia, but we know of no published study of its introduced distribution.

originating in the Balkans are able to exploit native British oaks, even though these are derived from the Iberian refuge (Stone & Sunnucks 1993; Dumolin-Lapègue *et al*. 1997; Csóka *et al*. 1998).

The asexual generation must lay eggs, and the sexual generation induce galls, on *Q. cerris* or *Q. suber*. Because cork oak is not found naturally north of south-western France (Fig. 1), and is extremely rare as an introduced tree outside its native range, *Q. cerris* is effectively the only available sexual generation host in northern Europe. Colonization from the eastern native distribution would be by immigrants for which the introduced Turkey oak is the natural sexual generation host. In contrast, range expansion from Iberia would necessitate a host shift from *Q. suber* to *Q. cerris*.

We address the following questions arising from the divided distribution of *A. kollari* and its colonization of northern Europe.

1 What is the genetic relationship between eastern and western lifecycles? Do they represent lineages independently derived from a larger ancient distribution, or has one been founded recently from the other? If so, can we infer which distribution (and which lifecycle) is ancestral, and how long the parts of the native range of *A. kollari* have been divided? If the two populations represent long-separated refugia, we expect refuge-specific genetic differences to have evolved. If one refuge has been recently colonized from the other, genetic variation in the recently colonized region is predicted to be a subset of that found in the source (Nei *et al*. 1975; Nichols & Hewitt 1994). This pattern is seen clearly in comparisons between native and invaded ranges for another invading oak gallwasp (Stone & Sunnucks 1993; Sunnucks & Stone 1996).

2 To what extent have eastern and western native ranges contributed migrants to northwards range expansion following human dispersal of *Q. cerris*?

3 Is there any evidence for gene flow between the two lifecycle types? Is there a hybrid zone where the two lifecycles meet?

We answer these questions through analyses of allele frequency data for 13 polymorphic allozyme loci for 46 populations (1457 individuals) distributed through Spain, France, Holland, Germany, Italy, Hungary and Turkey, and

of sequence data from a subset of 27 individuals for a 433 bp fragment of the mitochondrial cytochrome *b* gene.

Materials and methods

Sampling sites and sample preparation

Sampling sites are shown in Fig. 2, and the number of individuals screened for allozymes for each population is shown in Appendix I (because of its large size, the dataset is available from the Molecular Ecology website, URL http://www.blackwell-science.com/products/journals/suppmat/mec/mec1211/mec1211sm.htm). The western distribution is represented by 284 individuals from eight sites in Spain, while eastern populations are represented by 447 individuals from 12 sites in Hungary, 80 individuals from two sites in Italy and 42 individuals from one site in Turkey. Populations north of putative refuges are represented by 539 individuals from 18 sites in France, and 68 individuals from four sites in Germany and Holland. Sequence data were generated for 27 of the individuals screened for allozymes. Twelve were selected from the eastern and western native distributions — three from Hungary (Gödöllõ), three from Italy (two from Ruffeno, one from Greve), two from Turkey (Antalya) and four from Spain (two from Potes, one from Salamanca and one from Quintanilla). The remaining 15 individuals were from French populations, selected to traverse the apparent genetic divide in France (see below). Ten individuals were drawn from sites north of Bordeaux (two from St. Malo, two from Mortain, five from Forêt de Crécy, one from Nantes) and five individuals from south of Bordeaux (two from Amoux, two from Mugron, one from Bayonne).

All populations were sampled through rearing of parthenogenetic females (body mass 10–15 mg) from marble galls. Individual wasps were stored at −80 °C until required, homogenized in extraction buffer (Peakall & Beattie 1991), and screened immediately as described below. Samples were subsequently stored at −80 °C and used for DNA extraction.

Allozyme screening

An initial screening of commonly used allozyme systems was carried out using cellulose acetate electrophoresis (Zip-zone, Helena Laboratories) on three buffer systems across a range of pH, as described by Stone & Sunnucks (1993). The buffers were 40 mM sodium phosphate (pH 6.3; Stone & Sunnucks 1993), 0.1 M Tris-EDTA-maleate-$MgCl_2$ pH 7.6 (Richardson *et al*. 1986; buffer F), and 25 mM

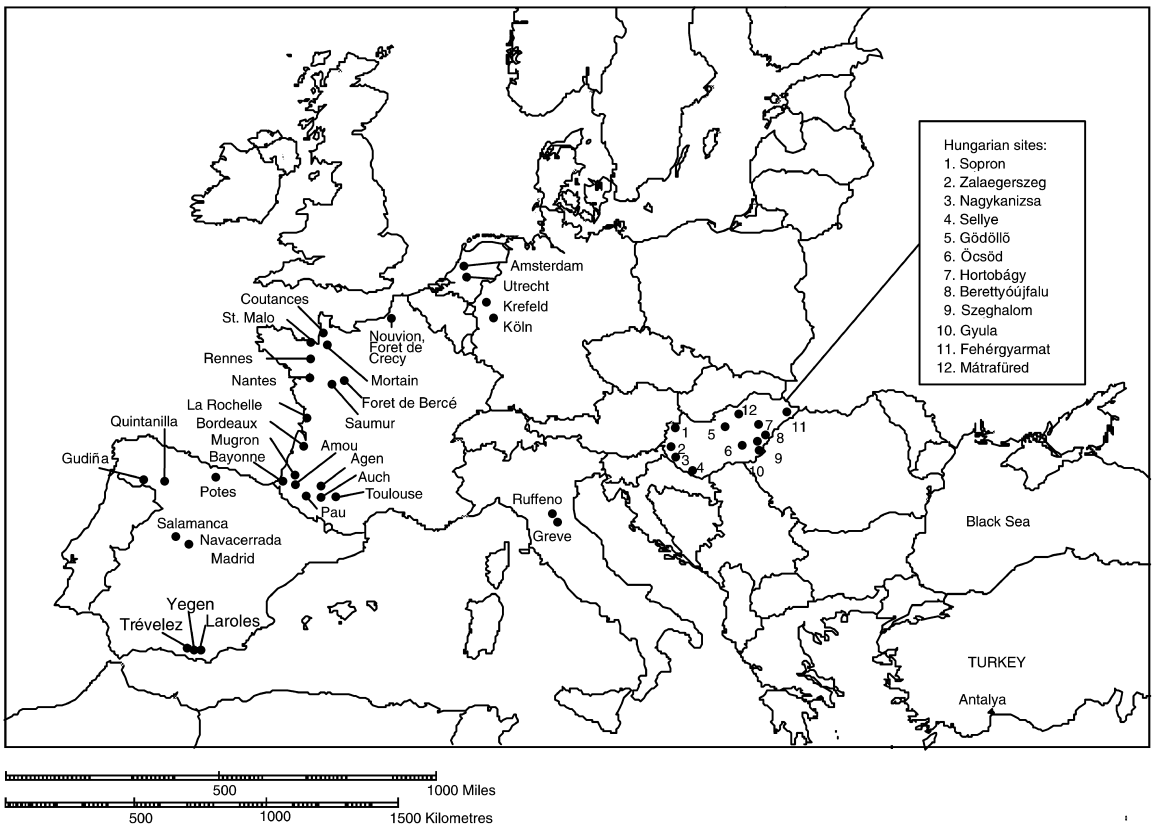

Fig. 2 Sample sites for *Andricus kollari* populations used in this study.

Tris-Glycine pH 8.5 (Richardson *et al*. 1986; buffer I). Substrate staining was carried out using protocols described by Richardson *et al*. 1986. The following 13 polymorphic loci were screened for all individuals, grouped by the running buffer used: AK, αGPD1 and αGPD2, PEP-b (sodium phosphate pH 6.3); GOT-s, GOT-m, GPI, MDH-s, MDH-m, ME, 6PGD (Tris-EDTA-maleate-$MgCl_2$ pH 7.6); HK and PGM (Tris-Glycine pH 8.5).

Analyses of allele frequency data

Summary analyses of allele frequency data were carried out using GENEPOP version 3.1 (Raymond & Rousset 1995) and GENETIX version 4.0 (Belkhir 1999). The generality of a sexual generation in the lifecycle of spanish *Andricus kollari* was checked by analysing the extent of departure from Hardy–Weinberg (HW) and linkage equilibrium. Persistent asexual reproduction, whether apomictic or automictic, is predicted to result in substantial departure from HW equilibrium across all loci. Furthermore, apomixis results in complete linkage across all loci, and automixis commonly results in complete homozygosity (Suomalainen *et al*. 1987; Sunnucks *et al*. 1997; Simon *et al*. 1999; Gómez & Carvalho 2000). Departures from HW equilibrium were tested using exact tests (Weir 1991) incorporated in GENEPOP, with *P* values estimated using a Markov chain method (Guo & Thompson 1992). Significance levels were adjusted for multiple tests using a Bonferroni correction (corrected threshold *P* value = $1 - (1 - \alpha)^{1/k}$ where k is the number of tests and α is the desired threshold value of 5%). Linkage disequilibrium was tested using the linkdos algorithm (Garnier-Géré & Dillman 1992) incorporated into GENEPOP and GENETIX. Significance in the first case is assessed by Fisher's exact tests, and in the second by permutation. We apply these two approaches to minimize the impact on our conclusions of assuming a given underlying distribution of a test statistic. Again, significance levels for multiple tests were adjusted by applying a Bonferroni correction. FSTAT (Goudet 1995) was used to compute *F*-statistics using the method of Weir & Cockerham (1984).

To assess the relationships between populations, and the dependence of such relationships on methods used, dendrograms linking populations were generated for two commonly used genetic distances — Nei's standard genetic distance and Cavalli-Sforza and Edward's chord distance. In the absence of a priori models, the two genetic distances were selected to represent different sets of assumptions concerning the underlying causes of genetic differentiation in *A. kollari*. Nei's unbiased genetic distance (Nei 1978) is formulated for an infinite isoalleles model of mutation, in which there is a rate of neutral mutation and each mutant is to a completely new allele. It is assumed that all loci have the same rate of neutral mutation, and that the genetic variability initially in the population is at mutation-drift equilibrium, with the effective population size of each population remaining constant. Cavalli-Sforza and Edward's chord distance assumes that there is no mutation, and that all gene frequency changes are by genetic drift alone (Cavalli-Sforza & Edwards 1967). This measure does not assume that population sizes have remained constant and equal in all populations. We used three tree building algorithms — maximum parsimony (MP), neighbour-joining (NJ, Saitou & Nei 1987) and maximum likelihood (ML). MP analysis was carried out using the Wagner routine of BIOSYS-1 version 1.2 (Swofford & Selander 1981). NJ was carried out using the NJBP program (Jean-Marie Cornuet, INRA Laboratoire de Modélization et Biologie Évolutive, Montpellier), and ML estimation was carried out using PHYLIP (Felsenstein 1993). For NJ trees, bootstrapping was carried out over both populations and loci, with 1000 bootstrap replicates. Because of the dramatic increase in computational complexity associated with ML estimation for a large number of populations, we used 100 bootstrap replicates for this method.

Cytochrome b *sequencing and phylogenetic analysis*

DNA was extracted from single adults using Proteinase K/SDS digestion followed by salting out of protein and precipitation of DNA in ethanol. Each dried DNA pellet was then resuspended in 50 μL of pure water. A 433 bp sequence of the cytochrome *b* gene was amplified using the primers CB1 (5′–3′ TATGTACTACCATGAGGACAAATATC) and CB2 (5′–3′ ATTACACCTCCTAATTTATTAGGAAT) (Jermiin & Crozier 1994; Stone & Cook 1998) in polymerase chain reaction (PCR) (35 cycles of denaturation at 92 °C for 60 s, annealing at 50 °C for 60 s, and extension at 72 °C for 120 s). PCR reactions had a total volume of 25 μL, with 1.0 μL DNA extract, 2.5 μL of PARR buffer (BIORAD), 1 μL 50 mM $MgCl_2$ (final concentration 2 mM), 0.5 μL 10 mM dNTP's, 0.35 μL of the primers CB1 and CB2 (each at 20 mM), 0.25 μL of *Taq* (Promega), and the balance made up with pure autoclaved water. All PCR and sequencing reactions were carried out on a PTC-200 DNA engine (MJ Research). The PCR product was gel-purified by electrophoresis through a 1% agarose gel stained with ethidium bromide. Gel fragments were excised from lanes showing a band at the correct size, and DNA extracted using the QIAquick gel extraction kit (QIAgen cat. 28704). Sequencing was direct from the purified PCR product. Mitochondrial genes are known to have nuclear pseudogene copies in insects (Sunnucks & Hales 1996; Zhang & Hewitt 1996; Bensasson *et al*. 2000), and if sequencing from an apparent single band suggested multiple copies of the cyt *b* gene, specific DNA sequences from particular individuals were amplified by cloning. When necessary, 2–4 μL of a PCR were used in standard ligation and transformation reactions using the TOPO TA cloning kit (Invitrogen cat. K4500-01). Plasmid DNA was

purified using the QIAprep spin miniprep kit (QIAgen cat. 27104). Sequencing was carried out using Perkin-Elmer Big Dye Terminator chemistry and an ABI 377 sequencer.

Sequences were all 433 bp long, and were aligned using CLUSTAL-W (Thompson *et al*. 1994). All sequences were found to possess the same reading frame, and contained no stop codons. Sequences for 27 individual *A. kollari* are deposited in GenBank, with acquisition numbers AF242739–AF242762 and AF242764–AF242766. Phylogenies were generated by MP and ML using test version 4.0b3 of PAUP* (Swofford 1998). We do not present NJ analyses of the sequence data for two reasons. First, MP and ML have been demonstrated to recover more closely a single true topology in simulation studies (Huelsenbeck & Hillis 1993). Second, ML methods allow the testing of specific models of substitution, and the identification of a model most appropriate to a given dataset (for a discussion of the impact of substitution models on phylogeny construction see Swofford *et al*. 1996 and Yang 1996). We used the program Modeltest 3.0 (Posada & Crandall 1998) to identify the most appropriate substitution model for our data. The model supported both by ML ratio tests and on the basis of Akaike information criterion was the Hasegawa–Kishino–Yano (HKY) model, which allows unequal transition and transversion ratios and unequal base frequencies (Hasegawa *et al*. 1985) with corrections for among-site variation (HKY + G + I). The following parameters returned by Modeltest were used in PAUP*: Transition and transversion (Ti/Tv) ratio = 18.23. Base frequencies A = 0.3492, C = 0.1183, G = 0.1102, T = 0.4223. Proportion of invariable sites = 0.733. Changes are gamma distributed, with a shape parameter 0.5123. Trees were rooted by using sequence for two other members of the *A. kollari* species group — *A. polycerus* (GenBank accession number AJ228457) and *A. conglomeratus* (GenBank accession number AJ228568) — known to be close relatives in a sister-clade to the *A. kollari* species group based on an earlier sequence-based phylogenetic analysis of the genus (Stone & Cook 1998). All trees were generated using 100 random additions in a heuristic search, using the tree bisection-reconnection (TBR) algorithm of PAUP. Ten trees were held at each step, with positions weighted equally. For both tree-building methods, we generated 100% consensus trees and subjected them to bootstrap analysis using full heuristic searches for 100 replicates. With the exception of minor rearrangements of terminal taxa, the tree topologies recovered by MP and ML were identical, with significant bootstrap values for the same nodes. Genetic distances given are for the Kimura 2-parameter model, which allows comparison with a large number of existing studies.

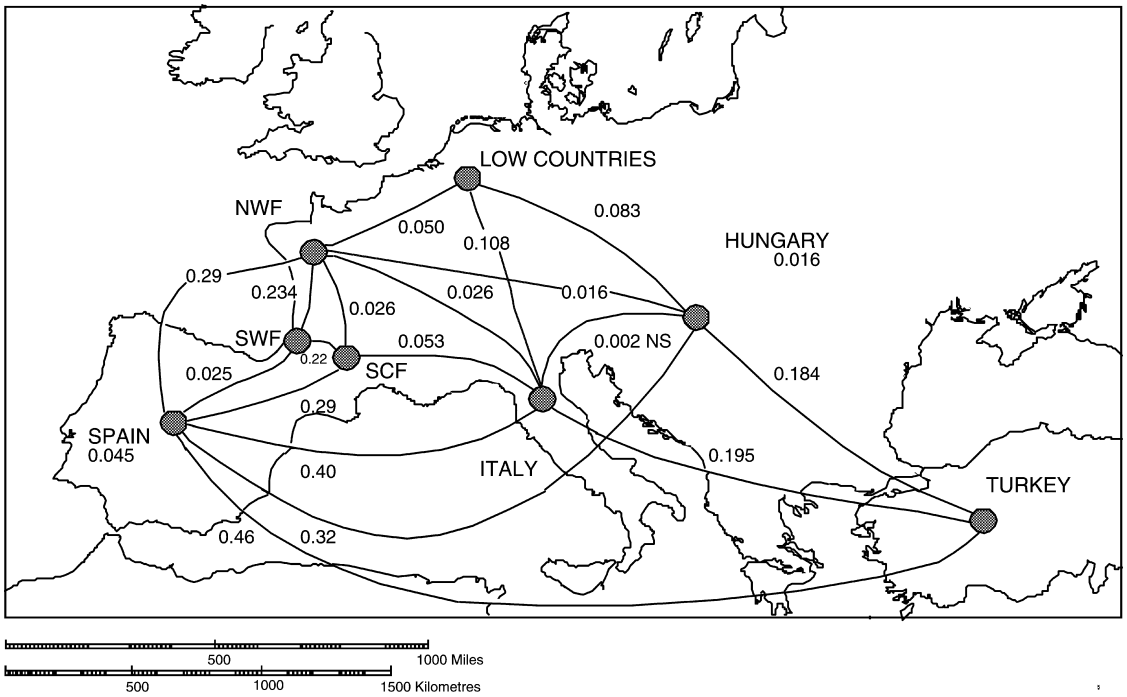

Fig. 3 Genetic subdivision between subsets of *Andricus kollari* populations, as indicated by pairwise F_{ST} calculations. All values except that between Italy and Hungary are highly significant by the permutation testing procedure in GENETIX. Each shaded circle represents a set of populations pooled together for pairwise F_{ST} between regions. Abbreviations for summed populations are as follows: NWF = north-west France (Coutances, Forêt de Bercé, Forêt de Crécy, La Rochelle, Nantes, Nouvion, Rennes, Saumur, St. Malo); SCF = south central France (Agen, Auch, Toulouse); SWF = south-west France (Amoux, Bayonne, Mugron, Pau); Low Countries = Utrecht, Amsterdam, Krefeld, Köln and data for 18 individuals from several sites in Belgium. Mean pairwise F_{ST} values are also included among populations within each of Hungary and Spain.

Results

Analyses based on allozyme allele frequency data

Relationships between the eastern and western native distributions. The 13 polymorphic loci possessed a total of 62 alleles over all 46 populations, with frequencies summarized for each population in Appendix I (URL http://www.blackwell-science.com/mec/). Hungary, Italy and Spain all possess regionally private alleles, and neither the eastern nor the western distributions is a simple genetic subset of the other. The only area not to have private alleles is Turkey, whose single sampled population contains a subset of the alleles found in Hungary. Four alleles were found only in Hungary (GOT-s allele 5, GPI allele 5, MDH-s allele 4 and PGM allele 5), one allele was found only in Italy (AK allele 5), and two alleles were found only in Spain (MDH-m allele 3 and AK allele 4). A further three alleles were found almost entirely in Spain, with only a single copy of each allele found in central and eastern European populations; αGPD1 allele 1 had a single copy in Sopron (north-west Hungary), αGPD1 allele 2 had a single copy in Gyula (south-east Hungary) and αGPD1 allele 3 had a single copy in Ruffeno (Appenines, Italy).

Several analytical methods suggest that Italian and Hungarian populations are far more similar to each other than either are to Spain or to Turkey. First, Hungarian and Italian sites share 17 alleles absent from Spain, while Spain and Italy share only one allele, αGPD1 allele 3, absent from Hungary (furthermore, as noted above only one copy of this allele is found in Italy). Second, F_{ST} between pooled Italian and pooled Hungarian sites (0.002, nonsignificant by permutation testing) is far lower than that between either and Spain or Turkey (Fig. 3). Third, a close relationship between Italy and Hungary, and their genetic divergence from Spain, is supported by analyses based on allele frequencies (Fig. 4) for all three tree-building methods and genetic distances used. Bootstrap support for the divide is more than 95% by ML or NJ, and for the latter the same result is obtained whether bootstrapped over loci or individuals.

Population genetic support for a sexual generation in western distribution Andricus kollari. All Spanish populations were in HW equilibrium, confirming the generality of a sexual generation in the *A. kollari* lifecycle. The small number of significant departures from HW equilibrium occurred elsewhere in the sampled range (Bordeaux for aGPD1, Bayonne for GOT-s, Fehergyarmat for GOT-m, and Mátrafüred, Agen, Forêt de Bercé and Mortain for PEP-b).

Significant linkage disequilibria, both across sets of populations (global tests) and for individual populations, are shown in Table 1. Both eastern and western populations show some significant linkage, although for no pair of loci is linkage supported by both permutation and exact tests across all populations in a set. Some linkage relationships are supported for both Spanish and Hungarian sets of populations (those linking GOT-m with PEP-b, PGM and 6PGD, and PGM with 6PGD), while others are only supported for one set of populations (Table 1).

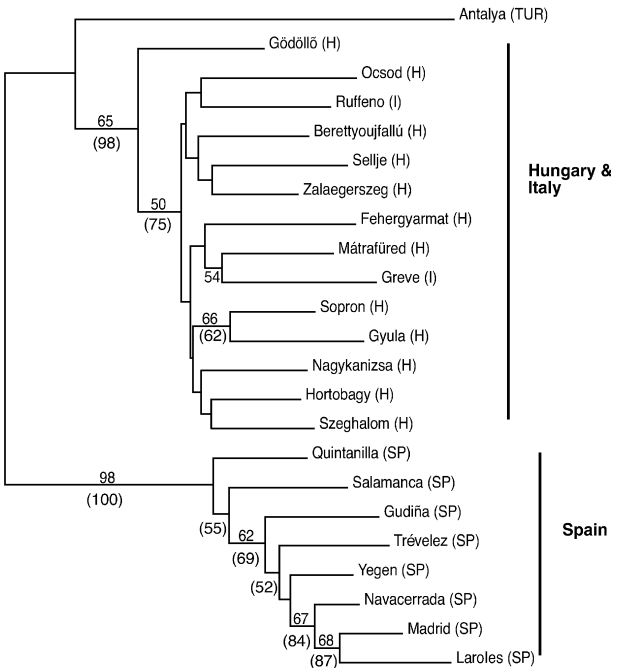

Fig. 4 The genetic divide between eastern and western native distributions in *Andricus kollari*, as demonstrated by allozyme allele frequency data. The phylogram shown is generated by neighbour-joining, using Cavalli-Sforza and Edward's chord distance. Numbers at nodes are bootstraps, over individuals, expressed as percentages for 1000 replicates. Values in parentheses are bootstraps for the same node obtained by maximum likelihood. Nodes without bootstraps are supported by less than 50% of replicates.

Postglacial range expansion by A. kollari. Sites north of the glacial refuges for *A. kollari* cluster into two clear groups on the basis of allozyme allele frequency data (Fig. 5), supported by bootstrapping for ML and NJ trees, and across both loci and individuals for NJ trees. Populations in the south-west of France (Bordeaux, Toulouse, Pau, Mugron, Amoux, Bayonne) cluster with Spain, while all other sites in northern Europe are more closely associated with Hungarian and Italian sites. The same result is shown by pairwise F_{ST} calculations among groups of populations (Fig. 3). All populations north of the Pyrenees and Alps, bar those in south-western France, show only slight (but nevertheless significant) genetic differentiation from Hungary/Italy, but major differentiation from Spain (Fig. 3). The only populations in northern Europe to show more substantial differentiation from Italy and Hungary are those in Germany, north-central France, and the Low Countries.

Table 1 Tests of linkage disequilibrium for each pair of loci over all populations (global tests) and within each population. Results are given for permutation tests in GENETIX and for Fisher's exact tests in GENEPOP. Named populations show significant results in permutation tests, and if also significant in exact tests are followed by an asterisk (*)

Hungarian populations (n = 12)				
Locus 1	Locus 2	Global permutation	Global exact test	Individual populations
GOT-s	GOT-m	No	No	Mátrafüred, Nagykanizsa
GOT-s	PGM	No	No	Mátrafüred
GOT-m	HK	No	No	Hortobagy
GOT-m	PEP-b	No	$P < 0.01$	Mátrafüred(*), Ocsod
GOT-m	6PGD	No	$P < 0.01$	Hortobagy
GOT-m	PGM	No	$P < 0.05$	
GPI	MDH-s	No	No	Szeghalom
GPI	6PGD	No	No	Ocsod, Szeghalom
HK	PGM	No	No	Mátrafüred
ME	PGM	$P < 0.05$	No	Fehergyarmat
PEP-b	6PGD	No	$P < 0.05$	Hortobagy
PGM	6PGD	No	No	Mátrafüred
Spanish populations (n = 8)				
aGPD1	PGM	$P < 0.05$	No	
aGPD1	6PGD	$P < 0.05$	No	
GOT-m	PEPb	No	No	Madrid
GOT-m	PGM	$P < 0.05$	No	Salamanca
GOT-m	6PGD	$P < 0.05$	No	Salamanca
PGM	6PGD	$P < 0.05$	No	Quintanilla

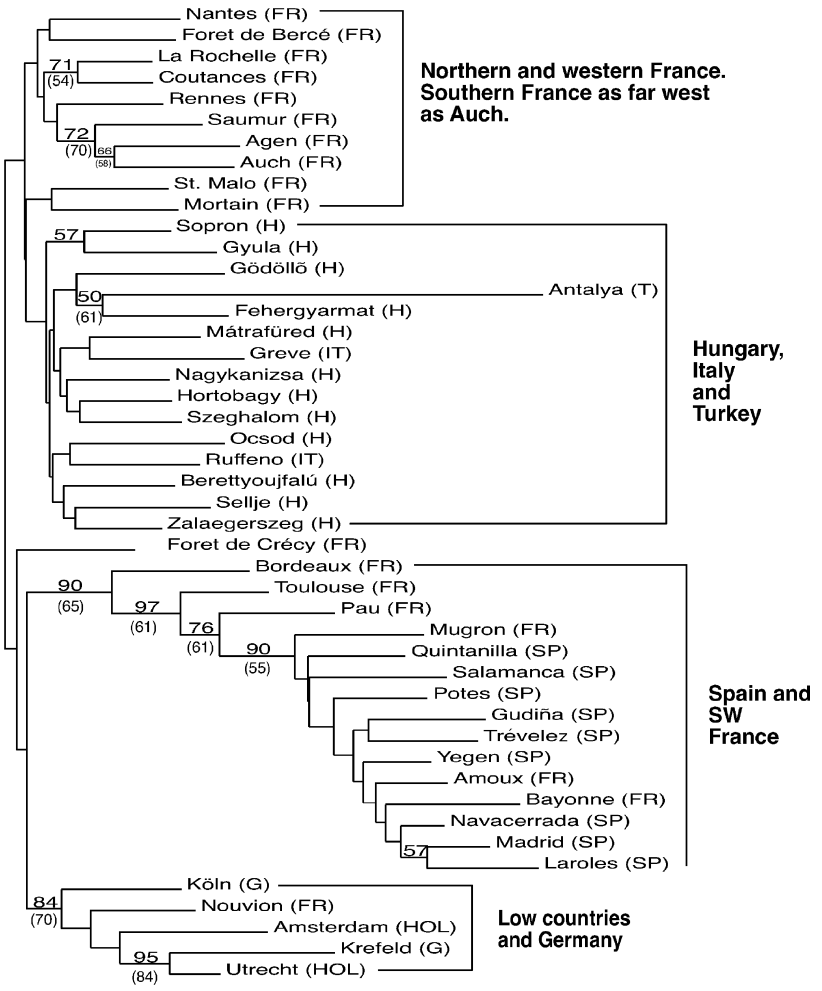

Fig. 5 A phylogram of relationships, based on allele frequency data, among *Andricus kollari* populations from both native and recently invaded regions. Relationships are generated by neighbour joining of Cavalli-Sforza and Edward's chord distance. Numbers at nodes are bootstraps, over individuals, expressed as a percentage of 1000 replicates. Values in parentheses are bootstraps for the same node obtained by maximum likelihood. Nodes without bootstraps are supported by less than 50% of replicates.

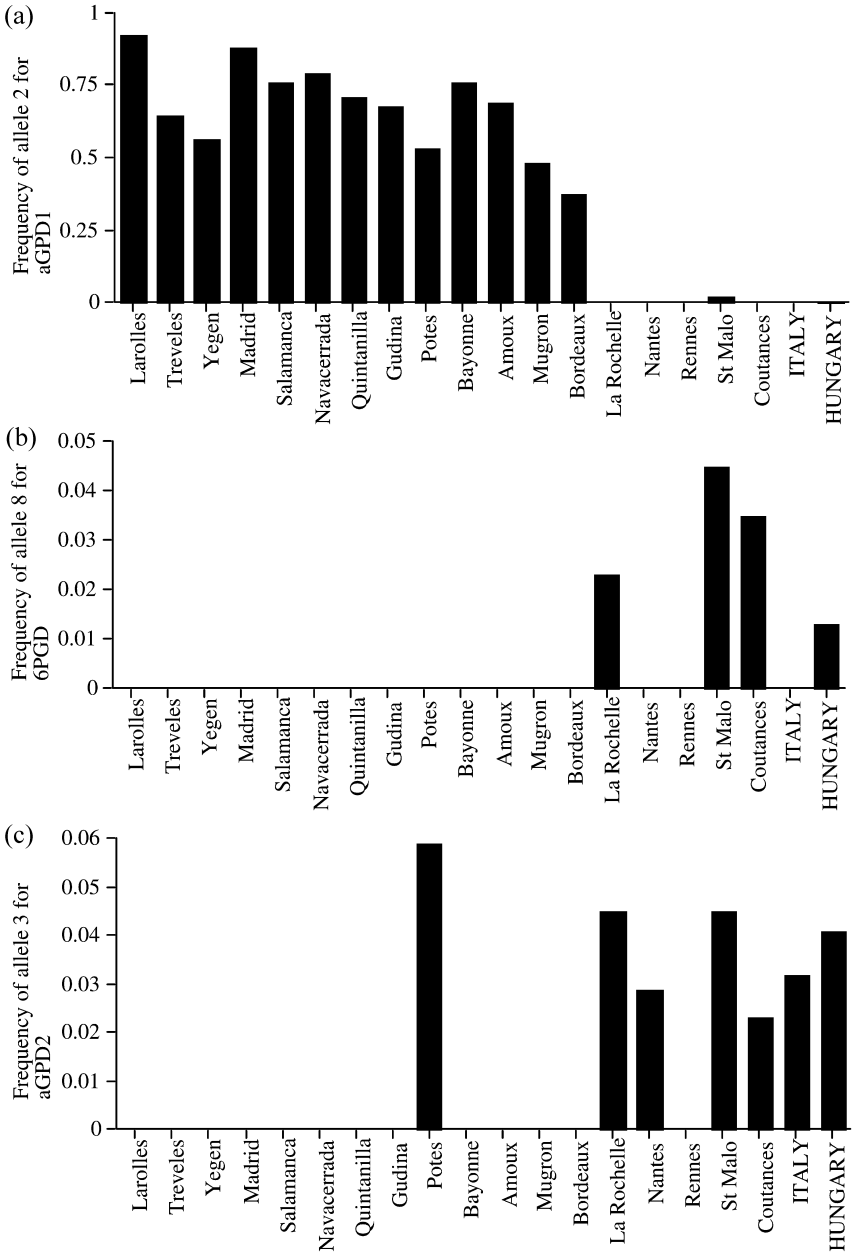

Fig. 6 Changes in the frequencies of three regionally private alleles indicating the genetic division between western and eastern gene pools in *Andricus kollari*. Sites are organized from south to north in a transect running from the extreme south of Spain to the channel coast of France (see Fig. 2 for locations).

These form a clade with high bootstrap support (Fig 5), and experience lower estimated gene flow with Italy and Hungary (pairwise F_{ST} 0.108 and 0.083, respectively) than populations in western or southern central France (Fig. 3).

The sharpness of the spatial transition from populations with Iberian affinities to those closer to central and eastern Europe is striking. Populations in north-western France, and north of the central Pyrenees (Auch, Agen and Toulouse) are all substantially differentiated from sites at the west of the Pyrenees (F_{ST} = 0.22–0.23) (Fig. 3). While French populations north of the central Pyrenees show little apparent gene flow with Spain (F_{ST} = 0.29), those at the west of the Pyrenees are closely allied to Spain (F_{ST} = 0.025). The differentiation is due mainly to abrupt changes in the frequency of several regionally private alleles, illustrated along a south-to-north series of sites in Fig. 6. In particular, allele 2 at αGPD1 falls from a high frequency (60–80%) in Spain and south-west France to zero at sites less than 100 km to the north.

Analyses of mitochondrial DNA (mtDNA) sequence data

Of the 433 bp in the amplified fragment of cytochrome *b*, 375 were constant across all sequences. Of the 58 variable

Table 2 Haplotypes for the 433 bp fragment of cytochrome *b* amplified in *Andricus kollari*. Parsimony-informative sites are indicated by an asterisk in the first row of the table

						11	111	111	111	122	222	222	222	222	223	333	333	333	333	333	4
	13	445	667	778	811	222	344	789	900	012	334	447	788	990	000	222	233	477	899	2
	702	695	790	392	625	147	025	213	625	840	281	454	706	282	458	235	967	069	279	1
Parsimony informative?	**	*	*	**	**		**	***	*	***	**	**	*		**	**	***	*	*	*	*	*	*
Haplotype 1	AAA	AAG	ATT	AAA	TTA	CTA	TAT	TAA	ATT	AAT	CTT	TCA	ATT	TTT	ATT	CCT	AGA	CTT	TAG	A
Haplotype 2	...	.G.	...	...	...	...	...	...	G..	...	...	...	...	...	...	...	...	...	...	.
Haplotype 3	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	T..	...	.
Haplotype 4	...	...	...	...	...	...	...	...	.C.	.G.	...	...	...	...	...	...	?..	T..	...	.
Haplotype 5	.G.	..A	...	...	...	...	...	.G.	...	.G.	...	...	...	...	...	..C	...	.?.	...	.
Haplotype 6	.G.	G.A	...	...	...	TC.	...	..T	G..	...	...	...	...	...	...	T.C	...	T..	CG.	G
Haplotype 7	.G.	G.A	...	...	...	TC.	...	..T	G..	...	...	...	...	...	...	T.C	...	T..	CG.	.
Haplotype 8	.G.	..A	.C.	..G	...	...	...	...	GCC	.G.	...	...	...	C..	...	TTC	G..	T..	...	.
Haplotype 9	.G.	..A	.C.	..G	...	...	...	...	GCC	.G.	...	...	...	C..	...	TTC	G..	T..	...	.
Haplotype 10	.G.	..A	.C.	..G	C..	...	...	...	GCC	.G.	...	...	...	C..	...	TTC	G..	T..	...	.
Haplotype 11	.GG	..A	.C.	..G	...	...	...	...	GCC	.G.	...	...	...	C..	...	T.C	G..	T..	...	.
Haplotype 12	.G.	..A	.C.	..G	...	...	...	...	GCC	.G.	...	...	...	C..	...	T.C	G..	T..	...	.
Haplotype 13	.G.	..A	.C.	...	...	...	...	...	GCC	.G.	...	...	...	C..	...	T.C	G..	T..	...	.
Haplotype 14	.G.	..A	.C.	...	...	...	...	...	.CC	.G.	...	...	...	C..	...	T.C	G..	T..	...	.
Haplotype 15	...	..A	.C.	...	.C.	..T	...	...	.CC	.G.	...	.T.	...	CC.	.CC	T.C	GA.	T..	..A	.
Haplotype 16	GG.	..A	...	.G.	...	...	C..	...	..C	.G.	.CC	C..	...	...	..C	T..	...	T..	...	.
Haplotype 17	.G.	..A	...	.G.	...	...	C..	...	..C	.G.	.CC	C..	...	...	..C	T..	...	T..	...	.
Haplotype 18	GG.	..A	...	.G.	...	...	C..	...	..C	.G.	.CC	...	...	...	..C	T..	...	...	...	.
Haplotype 19	GG.	..A	...	.GG	...	...	C..	...	..C	.G.	.CC	...	...	...	..C	T..	...	...	...	.
Haplotype 20	GG.	..A	...	.G.	...	...	C..	...	..C	.G.	.CC	...	...	...	..C	T..	...	T..	...	.
Haplotype 21	GG.	..A	...	.G.	...	...	C..	...	G.C	.G.	.CC	...	...	...	..C	T..	...	T..	...	.

Haplotype locations: haplotypes 1–15 are from the eastern distribution and associated sites in north-west France, haplotypes 16–21 are from Spain and associated sites in south-west France. 1 – Gödöllõ (H); 2 – Greve (I); 3 – Gödöllõ (H), St. Malo 37 (NW Fr); 4 – Foret de Crécy (NW Fr); 5 – Gödöllõ (H); 6 – Antalya (T); 7 – Antalya (T); 8 – Foret de Crécy (NW Fr); 9 – Ruffeno (I); 10 – Ruffeno (I); 11 – Foret de Crécy (NW Fr); 12 – Nantes (NW Fr); 13 – Foret de Crécy, Mortain (NW Fr); 14 – Mortain, St. Malo (NW Fr); 15 – Nantes (NW Fr); 16 – Potes (SP), Mugron, Amoux, Bayonne (SW Fr); 17 – Mugron (SW Fr); 18 – Amoux (SW Fr); 19 – Quintanilla (SP); 20 – Salamanca (SP); 21 – Potes (SP).

sites, 35 were informative in parsimony reconstruction. The 27 individuals sequenced yielded 20 haplotypes (Table 2).

Relationships between the eastern and western native distributions. Both tree-building approaches generate best fit trees with the same topology (Fig. 7), with the four Spanish sequences placed as a monophyletic clade within a paraphyletic group of eastern distribution sequences. This nesting of western within eastern distribution sequences is not, however, supported by bootstrap analysis (Fig. 7). While four lineages retain high bootstrap support (one contains the Spanish sequences, one contains the two Turkish sequences, and the remaining two contain sequences from Italy and Hungary), it is not possible to say whether eastern or western lineages are closer to the outgroup. If the HKY + G + I model is assumed, the level of sequence divergence between eastern and western distribution sequences ranges between 4.0 and 10.4%, close to that observed among the eastern lineages of 4.1–7.4% between any two of Ruffeno (Italy), Antalya (Turkey) and Gödöllõ (Hungary). If we use an estimate of the rate of mitochondrial sequence divergence, based on insects and other invertebrates, of 2.3% per million years (Brower 1994), these four lineages diverged from each other between 1.75 and 4.5 million years ago. Use of Kimura 2-parameter distances (a simpler model rejected for our data by the Modeltest program) gives a sequence divergence between the four lineages of 2.8–3.8%, and an estimated time since divergence of between 1.2 and 1.6 million years, still long before the end of the Pleistocene.

Postglacial range expansion by A. kollari. Sequences from sites north of putative glacial refuges fall into three of the four lineages established for the native range sequences. For both ML and MP methods, all sequences from sites in northern France lie within one of the two lineages containing sequences from Italy and Hungary (left-hand figure, Fig. 8), while sequences from south-west France cluster with those from Spain. Thus, the DNA results agree with site groupings

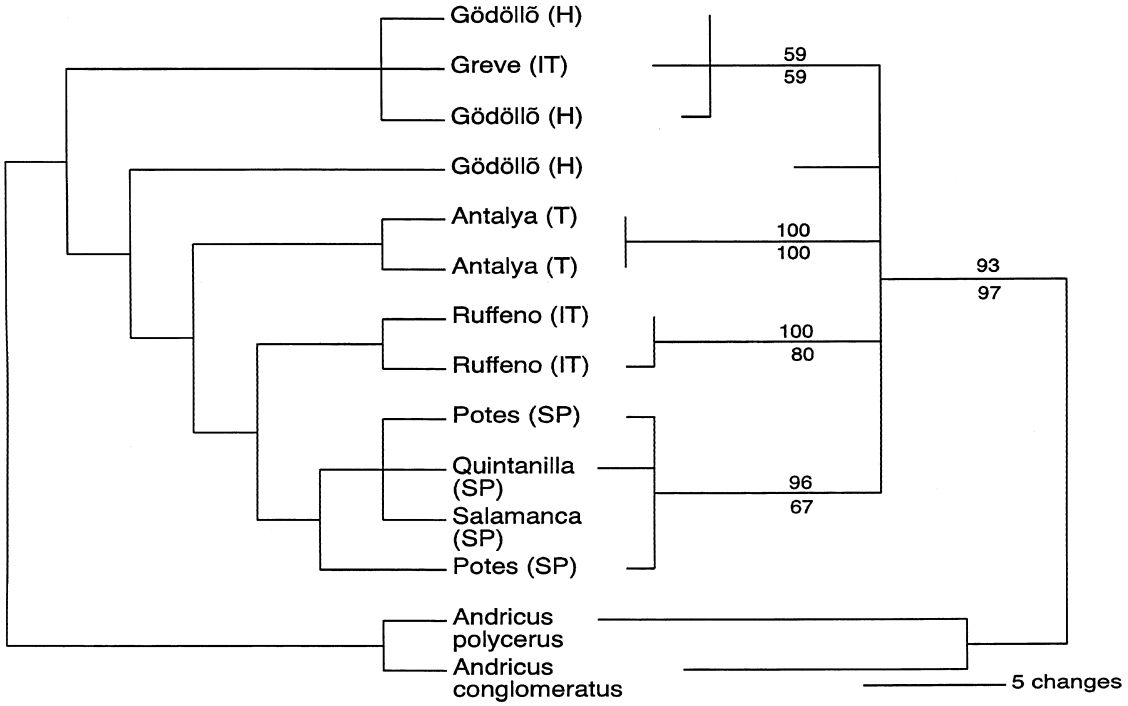

Fig. 7 Relationships, based on cytochrome *b* sequence data, among eastern and western distribution *Andricus kollari*. The left-hand cladogram represents the strict consensus of the two shortest trees generated using maximum parsimony in PAUP (length 69 steps, CI = 0.797, RI = 0.816). The best tree generated using maximum likelihood has almost exactly the same topology (–log likelihood = 902.4), with reversal of the relative positions of two clades (one containing Godollo and Greve, the second containing the two sequences from Ruffeno). The right-hand phylogram shows the majority rule consensus tree for 100 bootstrap replicates generated by maximum likelihood (–log likelihood = 903.93). Maximum parsimony returns an identical bootstrap consensus tree (length 75 steps, CI = 0.733, RI = 0.737). Numbers on the cladogram are bootstrap percentages, for maximum parsimony above the branch and maximum likelihood below.

on the basis of allele frequency data. Although this larger dataset again suggests that eastern sequences may be basal to western sequences (left-hand phylogeny in Fig. 8), bootstrap analysis shows membership of the main lineages to be only weakly supported for some northern French sequences, and the branching order of the main lineages is unresolved (right-hand phylogeny in Fig. 8). Similarity between Spanish sequences and those from south-western France remains strongly supported. Thus, addition of further sequences does not improve our ability to infer which lifecycle is ancestral and which derived.

Can eastern and western A. kollari *hybridize?*

Although the genetic transition between eastern and western distributions in south-western France is relatively abrupt (Fig. 6), allozyme data identified no hybrids near the contact area, and no hybrid zone. Only a single individual (from Mortain, north-west France) could be identified as a likely hybrid on the basis of allozyme data. This individual possessed alleles diagnostic of both east and west, combining allele 2 at αGPD1 (diagnostic of the western distribution) with allele 6 at 6PGD (diagnostic of the eastern distribution). A further three individuals from north-west France (one from each of St. Malo and Forêt de Crécy, and a second individual from Mortain) also possessed allele 2 for αGPD1. If we assume that this allele evolved only once, these individuals must be the result of long-range gene flow from the western distribution. The western distribution alleles could be present either in the offspring of purely Spanish immigrants, or in hybrids. Two possible hybrids (the two individuals from Mortain) were sequenced for the cytochrome *b* fragment. Each possessed a haplotype (numbers 13 and 14, Table 2) indicating origin in the eastern distribution. Because their haplotype is maternally inherited, these individuals must have acquired their western distribution αGPD1 allele by mating of a western distribution male with an eastern distribution female.

Discussion

Genetic relationships among native range populations of Andricus kollari

Allozyme and DNA sequence data suggest that eastern and western distributions of *Andricus kollari* have been

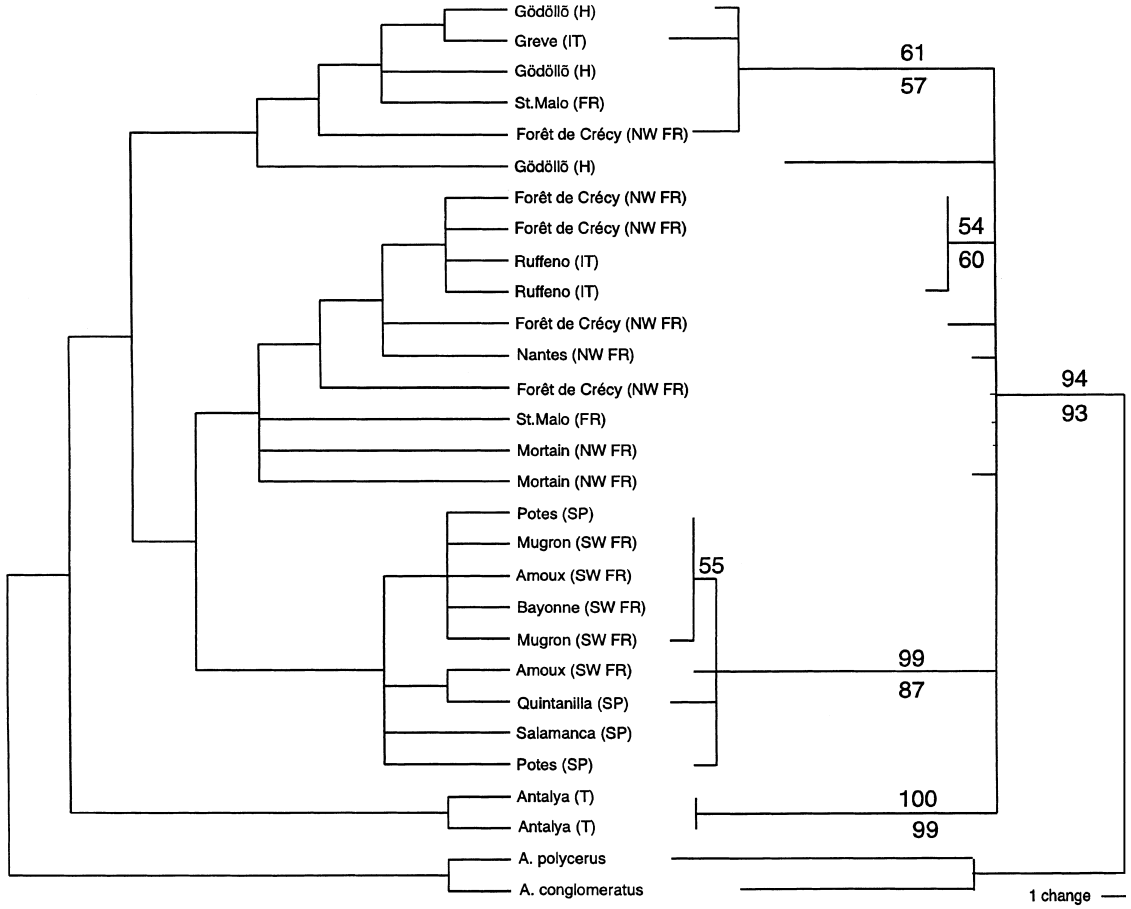

Fig. 8 Relationships, based on cytochrome *b* sequence data, among *Andricus kollari* populations from both native and recently invaded regions. The left-hand cladogram is the strict consensus of 55 shortest trees generated by maximum parsimony (length 73 steps, CI = 0.767, RI = 0.902). With the exception of a single trivial rearrangement, the topology of the best maximum likelihood tree is identical (–log likelihood = 934.41). The right-hand phylogram shows the majority rule consensus tree for 100 bootstrap replicates generated by maximum likelihood (–log likelihood = 959.46). The bootstrap topology returned by maximum parsimony is identical (length 91 steps, CI = 0.615, RI = 0.798). Numbers on the cladogram are bootstrap percentages, for maximum parsimony above the branch and maximum likelihood below.

genetically separated for a substantial period of time. Both regions contain private allozyme alleles, and each consists of a mutually exclusive set of cytochrome *b* haplotypes. It is probable, therefore, that the Iberian peninsular has acted as a discrete glacial refuge for several glacial cycles prior to the present interglacial. The small number of sequences sampled to date further suggest that three lineages diverged in the eastern distribution at approximately the same time that Iberian populations became isolated, perhaps corresponding to simultaneously isolated refugial populations in Italy, the Balkans and Turkey.

The complete separation of mtDNA lineages between east and west suggests greater genetic subdivision than is suggested by F_{ST} values calculated from allozyme data. While this difference may in part be due to the difference in sampling effort between mitochondrial and nuclear markers, it may also reflect the larger effective population size, and longer coalescence times, of nuclear alleles relative to mitochondrial markers. This means that allozyme alleles common to both regions (which tend to reduce F_{ST}-based estimates of differentiation) may represent ancestral polymorphisms retained in both regions, rather than the results of contemporary gene flow.

Population genetic evidence for the generality of a sexual generation in the lifecycle of A. kollari

Many European oak gallwasps appear to lack sexual generations in their lifecycle (Ambrus 1974; Nieves-Aldrey 1987; Csóka 1997). In most cases, inference of a purely parthenogenetic lifecycle is due to ignorance of any sexual generation, rather than demonstrated persistence of purely parthenogenetic reproduction. Painstaking rearing experiments have shown that the asexual generation of at least one oak gallwasp, *A. quadrilineatus*, includes both females able to give rise directly to a second generation of parthenogenetic

females, and females giving rise to a sexual generation (Folliot 1964). Such intraspecific variance in reproductive mode makes it possible that a species that has a sexual generation in one part of its range could lack it in another (geographical parthenogenesis). Geographical parthenogenesis has been proposed for three oak gallwasps (Yasumatsu 1951; Bailey & Stange 1966; Abe 1986), and is known to occur in rose gallwasps (Stille 1985a,b; Plantard *et al*. 1998, 1999). Though geographical parthenogenesis was proposed as a possibility for Iberian *A. kollari* (Csóka *et al*. 1998), allele frequency data for Spanish populations strongly support Pujade-Villar's (1991, 1992) conclusion that *A. kollari* has a sexual generation in Iberia. Many other gallwasps are known only from asexual generations, and application of a similar population genetic approach should reveal which are true purely parthenogenetic species, and which are half of a cyclically parthenogenetic lifecycle (Atkinson 2000).

Which lifecycle is ancestral?

A taxonomically diverse set of gallwasp species in the genera *Andricus*, *Callirhytis*, *Neuroterus* and *Synophrus* exploit *Quercus cerris* in the east and *Q. suber* in the west (Trotter & Cecconi 1904; Dalla Torre & Kieffer 1910; Nieves-Aldrey 1987, 1992). This raises two general questions. First, which represents the ancestral sexual generation host for each of these species? Do they share a common ancestral host? Second, did the shifts from one host to another occur at more or less the same time for the whole set, perhaps driven by a common set of circumstances, or have they occurred at different times in different species? Our analysis of *A. kollari* is a first step towards answering these broader questions of the origin of host association in oak cynipids.

Our cytochrome *b* data cannot resolve which of eastern or western distributions gave rise to the other. Although both MP and ML analyses suggest that Spanish cytochrome *b* sequences are derived with respect to an eastern distribution ancestor (left-hand figures in Figs 7, 8), there is no bootstrap support for this conclusion. What is clear is that lineages currently associated with *Q. suber* and *Q. cerris* diverged 1–2 million years ago, and thus the host shift, whichever its direction, may be relatively ancient.

An alternative approach to the question of which sexual generation host is ancestral is to take a phylogenetic view of the occurrence of *Andricus* lifecycles involving the two oak species. At least five host-alternating *Andricus* species have a sexual generation on *Q. cerris* where this oak is native. Three are members of a single closely related clade (*A. corruptrix*, *A. kollari*, and *A. lignicola*), while two are members of more distantly related species groups (*A. quercuscalicis* and *A. gemmea*) (Stone & Cook 1998; Cook *et al*. 1998). Of these six, only two (*A. kollari* and *A. gemmea*) are found in Iberia. In contrast, no host-alternating *Andricus* species are yet known that are restricted to Iberia. Taken at face value, this distribution would suggest that host alternation evolved in the native range of *Q. cerris*, followed by a host shift to *Q. suber* by one (or possibly two) species. However, this pattern may merely reflect a lack of knowledge of lifecycles in Iberian cynipids, and further work on this issue is required before conclusions can be drawn.

Origins of populations of A. kollari *north of glacial refuges*

Allozyme and sequence data show that populations north of the Pyrenees, Alps and Tatras are divided into two sets. We suggest that these sets are probably associated with escape from refuges on two different timescales.

Populations in south-western France. Populations in the extreme south-west of France are genetically indistinguishable from Spanish populations, sharing the same private allozyme alleles and cytochrome *b* haplotypes. This region of south-west France corresponds closely to the distribution of *Q. suber* at the western end of the Pyrenees (Fig. 1), and is compatible with a dependence of western distribution *A. kollari* on this host. As described above, *Q. suber* occupied a refuge in Iberia during the last glacial period, and populations in southern France are the result of natural range expansion from Iberia following retreat of the ice. Because the level of sequence divergence between Iberia and the other refuges suggests divergence long before the current interglacial, *A. kollari* has probably been associated with *Q. suber* in Iberia for multiple glacial cycles. Therefore, populations of *A. kollari* in south-western France probably represent the results of natural postglacial range expansion by Iberian *A. kollari* already dependent on *Q. suber*, 8000–10 000 years ago. The northern limit to this range expansion was probably imposed by the inability of *Q. suber* to spread beyond the geographical limits of the hot humid or subhumid Mediterranean climates it requires (Camus 1938).

Populations across the rest of northern Europe. All other populations in northern Europe show close affinity to populations in Italy and Hungary. Haplotypes from populations in north-western France lie in both of the lineages found in Italy and Hungary, and pairwise F_{ST} values indicate high gene flow to these refuges (Fig. 3). Exploitation of planted *Q. cerris* has apparently been entirely by *A. kollari* migrants from areas in which this oak is the natural host of the sexual generation. This confirms that, as for other oak gallwasps (Stone & Sunnucks 1993; Sunnucks & Stone 1996; Csóka *et al*. 1998), any geographical variation across northern continental Europe in available genotypes of section Quercus hosts has not prevented range expansion from eastern refuges.

Are there host races in A. kollari*?*

If Iberian *A. kollari* spread into south-western France following natural dispersal of cork oak, they were there for thousands of years before human introduction of *Q. cerris* and would have been exposed to Turkey oak immediately following its introduction. In contrast, immigrants from the east could only reach south-western France following dispersal across other *Q. cerris* patches, resulting in a substantial time lag (Stone & Sunnucks 1993). Despite a head start, western distribution *A. kollari* have remained restricted to the distribution of *Q. suber*, suggesting that Iberian asexual generation females may be unable to exploit *Q. cerris*.

Exploitation of novel *Q. cerris* requires: (ii) that asexual generation females adapted to location of *Q. suber* can locate *Q. cerris*, and receive suitable oviposition stimuli; (ii) that host tissues are located in the narrow time window during which they are developmentally susceptible to gall induction; and (iii) that their sexual generation offspring can successfully induce galls on the novel host (Craig *et al*. 1993; Brown *et al*. 1996; Abrahamson & Weis 1997). Inability of western *A. kollari* to exploit *Q. suber* could result from failure in any one of these requirements. The significance of host genotype for gallwasps remains little understood. Studies of oak hybrid zones have shown that hostplant genotypes have major effects on the abundance of specific oak gallwasp species (Boecklen & Spellenberg 1990; Aguilar & Boecklen 1991), but whether the causes of such variation act before or after oviposition remains unknown. Host races have been demonstrated in at least two gallwasps, one on oaks (Abe 1988, 1991) and one on roses (Stille 1985b). In both systems, host races are characterized by demonstrable divergence in the cues which elicit oviposition. Rearing experiments by Folliot (1964, personal communication) and J. Pujade-Villar (personal communication) suggest that similar divergence has evolved between eastern and western *A. kollari*. They found that asexual female *A. kollari* from Brittany (almost certainly of the eastern race) oviposited on *Q. cerris*, but never *Q. suber*, while females from Catalonia (almost certainly of the western race) would lay eggs in *Q. suber* but never *Q. cerris*. The split between *Q. cerris* and *Q. suber* is relatively ancient (Manos *et al*. 1999), and the two species may possess quite different profiles of the host secondary metabolites often used by phytophagous insects as host location and oviposition cues (Thompson 1994; Abrahamson & Weis 1997). As for *A. mukaigawae* (Abe 1988, 1991), asexual generation *A. kollari* adapted to one of these hosts may not be able to recognize the alternative host as an oviposition site. Although it has not yet been investigated for *A. kollari*, it is possible that even were oviposition on the alternative host to occur, gall induction might not be successful.

Hybridization between eastern and western races of A. kollari

Although our data show the two lifecycle races of *A. kollari* to meet at a relatively clearly defined boundary in south-western France, we found no individuals in this region combining alleles characteristic of eastern and western distributions. The absence of hybrid heterozygotes despite the high frequency and diversity of refuge-specific private alleles suggests that hybrids are either genuinely rare, or extremely local, or both. Because both eastern and western races share some alleles, allozyme data will underestimate the frequency of hybridization; analysis of haplotype distributions is a far more powerful tool, and is still at an early stage in this system. The identification of two individuals in north-western France with Iberia-specific alleles and mitochondrial haplotypes characteristic of the eastern distribution shows that there is no absolute barrier to introgression. Several aspects of the biology of oak gallwasps, however, may contribute to partial prezygotic barriers.

First, sexual generations of *Andricus* and of other gall-inducers mate as the sexual females emerge from their galls (G. Stone, unpublished data; Stille 1985b; Waring *et al*. 1990). Because the sexual generations of the two races of *A. kollari* develop on different host plants, mating is thus likely to be highly assortative. Host-mediated assortative mating is well-known in phytophagous insects, and among gall-inducers has been demonstrated to contribute to genetic isolation of host races in a japanese oak gallwasp *A. mukaigawae* (Abe 1991), and in the the tephritid gall fly *Eurosta solidaginis* (Abrahamson & Weis 1997). To overcome such assortment, either: (i) asexual generation females must induce galls on the unfamiliar sexual generation host, and their sexual offspring mate with members of the other race; or (ii) sexual individuals must disperse from *Q. cerris* to *Q. suber*, or vice versa, and mate. As discussed above, asexual generation females may be unable to exploit nonfamiliar sexual generation hosts, and dispersal of sexual adults seems a more likely source of hybrids. Although small cynipids can be carried for long distances in aerial currents (Hardy & Cheng 1986; Ros-Farré & Pujade-Villar 1998), any dispersing sexual adults would still have to recognize and land on the nonfamiliar oak host, and locate a mate.

Second, the sexual generation wasps are very short-lived (less than a week for *A. quercuscalicis*) and a small difference in the phenology of sexual generation galls on their respective hosts would effectively prevent interbreeding (allochronic isolation). Phenological differences are known to be important in the restriction of genetic exchange between races of other herbivorous insects (e.g. Mitter *et al*. 1979; Feder *et al*. 1993), including an oak gallwasp (Abe 1991). The phenology of sexual generations of *A. kollari* on *Q. cerris* and *Q. suber* remains unknown, but it is quite possible that they do not overlap.

Given that hybridization occurs despite such considerations, a hybrid zone between the host races of *A. kollari* may exist in south-western France. The resulting asexual generation females are the first to experience any detrimental effects of interracial matings. Unlike the sexual generation females, which gall the same section Quercus hosts regardless of their refugial origin, the asexual generation females must return either to *Q. suber* or to *Q cerris* to lay their eggs. Any host-specific differences associated with host location or oviposition (as suggested by the experiments described above for *A. kollari*) may be disrupted by hybridization. Even if oviposition does occur, studies in another galling system showed that hybrid eggs either fail to induce galls, or that resulting larvae die at an early stage of development (Abrahamson & Weis 1997). If there is similarly strong negative selection against hybrid *A. kollari*, only first generation hybrids are expected. Because hybrids are probably generated only when sexual adults disperse between the two sexual generation hosts, any hybrid zone is expected to track closely those areas in which *Q. suber* and *Q. cerris* come into close contact and to have a breadth determined by the dispersal distance of sexual generation females. South-western France is a hotspot of hybrid zones in a range of taxa (Taberlet *et al*. 1998; Hewitt 1999). If it exists, this *A. kollari* hybrid zone would be unusual in that its position can be explained by the postglacial fortunes of the two sexual generation hosts.

Q. suber and *Q. cerris* are naturally sympatric in Italy (Fig. 1), and though the gallwasp fauna is generally less well-studied in Italy than it is in either Iberia or the Balkans, *A. kollari* is present in the area of overlap (Trotter & Cecconi 1904; Dalla Torre & Kieffer 1910). This raises the question of whether the sexual generation of *A. kollari* might also be found on *Q. suber* in this region. The limited data we have on Italian *A. kollari* link populations from the north of this refuge with the Balkans, but the lifecycle and genetic make-up of populations in the south remains completely unknown. *Q. suber* populations in North Africa and Italy are thought to have been derived by range expansion from an Iberian refuge one or more interglacial periods ago (Reille *et al*. 1996; Toumi & Lumaret 1998). If the eastern and western lifecycles of *A. kollari* had already diverged by this time, it is possible that western race migrants accompanied *Q. suber* across Africa and into Italy. Here they would have encountered eastern race populations already exploiting *Q. cerris*. This raises the possibility of a second, more ancient, contact zone between two host-races of *A. kollari* in Italy.

Acknowledgements

Graham Stone was supported by grants from the NERC EDGE thematic programme (GR9/03553), the Royal Society (574006), the British Ecological Society (Small Ecological Project Grant number 1462) and The E.P. Abraham Research Fund of Oxford University. José-Luis Nieves-Aldrey was supported by a grant from the DGES (PGC, PB97-1241). We would like to thank all of those who helped at various times in the collection of samples, particularly Olivier Plantard, Alex Rowe and an army of Hungarian Foresters. We thank Susan Haley and Gordon Brown for their help in the allozyme electrophoresis, Jill Lovell for her help with sequencing, and Sylvain Piry for his patient, long-range help with NJBP.exe. We thank Roger Folliot and Juli Pujade-Villar for their many comments on the biology of Spanish *Andricus kollari.* We thank Karsten Schönrogge, Paul Sunnucks and an anonymous referee for their constructive comments on the submitted manuscript.

References

Abe Y (1986) Taxonomic status of the *Andricus mukaigawae* complex and its speciation with geographic parthenogenesis. *Applied Entomology and Zoology*, **21**, 436–447.

Abe Y (1988) Two host races in *Andricus mukaigawae* (Mukaigawa) (Hymenoptera: Cynipidae). *Applied Entomology and Zoology*, **23**, 382–387.

Abe Y (1991) Host race formation in the gall wasp *Andricus mukaigawae*. *Entomologia Experimentalis et Applicata*, **58**, 15–20.

Abrahamson WG, Weis AE (1997) *Evolutionary Ecology Across Three Trophic Levels: Goldenrods, Gallmakers and Natural Enemies*. Princeton Monographs in Population Biology Number 29. Princeton University Press, Princeton, USA.

Aguilar JM, Boecklen WJ (1991) Patterns of herbivory in the *Quercus grisea* × *Quercus gambelii* species complex. *Oikos*, **64**, 498–504.

Ambrus B (1974) Cynipida-gubacsok-cecidia cyniparum. In: *Fauna Hungariae*, Vol. 12 *(Hymenoptera 2), Part 1/a (Serial Number 116)*. Academic Press, Budapest, Hungary.

Askew RR (1984) The biology of gallwasps. In: *The Biology of Galling Insects* (ed. Ananthakrishnan TN), pp. 223–271. Oxford and IBH Publishing Co, New Delhi.

Atkinson R (2000) The genetic analysis of natural history, reproductive strategy and population structure in European oak gallwasps (Hymenoptera: Cynipidae). PhD Thesis. University of Edinburgh, UK.

Bailey SF, Stange LA (1966) The twig wasp of Cork Oak — Its biology and control. *Journal of Economic Entomology*, **59**, 663–668.

Beijerinck MW (1902) Über die sexuelle Generation von *Cynips kollari*. *Marcellia*, **1**, 13–20.

Belkhir K (1999) GENETIX, *Version 4.0. A windows programme for population genetic analysis*. Laboratoire Génome, Populations, interactions UPR 9060 du CNRS, Université Montpellier 2, Place E. Bataillon, 34095 Montpellier Cedex 05, France.

Bensasson D, Zhang DX, Hewitt GM (2000) Frequent assimilation of mitochondrial DNA by grasshopper nuclear genomes. *Molecular Biology and Evolution*, **17**, 406–415.

Bernatchez L, Wilson CC (1998) Comparative phylogeography of nearctic and palearctic fishes. *Molecular Ecology*, **7**, 431–452.

Boecklen WJ, Spellenberg R (1990) Structure of herbivore communities in two oak (*Quercus* spp.) hybrid zones. *Oecologia*, **85**, 92–100.

Brower AVZ (1994) Rapid morphological radiation and convergence among races of the butterfly *Heliconius erato* inferred from patterns of mitochondrial DNA evolution. *Proceedings of the National Academy of Sciences of the USA*, **91**, 6491–6495.

Brown JM, Abrahamson WG, Way PA (1996) Mitochondrial DNA phylogeography of host races of the goldenrod ball gallmaker, *Eurosta solidaginis* (Diptera: Tephritidae). *Evolution*, **50**, 777–786.

Camus A (1938) *Les Chênes,* Vol. 3. Monographie du genre *Quercus*. Lechevallier, Paris.

Cavalli-Sforza LL, Edwards AWF (1967) Phylogenetic analysis: models and estimation procedures. *American Journal of Human Geneticsi*, **19**, 233–257.

Cook JM, Stone GN, Rowe JA (1998) Patterns in the evolution of gall structure and life cycles in oak gall wasps (Hymenoptera: Cynipidae). In: *The Biology of Gall-Inducing Arthropods* (eds Csóka G, Mattson W, Stone GN, Price PW), pp. 261–279. General Technical Report NC 199. North Central Research Station USDA, St. Paul, Minnesota.

Craig TP, Itami JK, Horner JD, Abrahamson WG (1993) behavioural evidence for host-race formation in *Eurosta solidaginis*. *Evolution*, **47**, 1696–1710.

Csóka G (1997) *Plant Galls*. Agroinform Publishing, Budapest, Hungary.

Csóka G, Stone GN, Atkinson RJ, Schönrogge K (1998) The population genetics of postglacial invasions of northern Europe by cynipid gallwasps (Hymenoptera: Cynipidae). In: *The Biology of Gall-Inducing Arthropods* (eds Csóka G, Mattson W, Stone GN, Price PW), pp. 280–294. General Technical Report NC 199. North Central Research Station USDA, St. Paul, Minnesota.

Dalla Torre KW, Kieffer JJ (1910) Cynipidae. In: *Das Tierreich. Eine Zusammenstellung und Kennzeichnung der Rezenten Tierformen*, pp. 1–891. Der Deutsche Zoologischen Geselschaft, Berlin.

Docters van Leuwen WM (1959) Generatiewisseling en wisseling van waardplant bij vier Nederlandse galwespen. *De Levende Natuur*, **2**, 149–161.

Dumolin-Lapègue S, Demesure B, Fineschi S, Le Corre V, Petit RJ (1997) Phylogeographic structure of White oaks throughout the European continent. *Genetics*, **146**, 1475–1487.

Feder JL, Hunt TA, Bush L (1993) The effects of climate, host plant phenology and host fidelity on the genetics of apple and hawthorn infesting races of *Rhagoletis pomonella*. *Entomologia Experimentalis et Applicata*, **69**, 117–135.

Felsenstein J (1993) PHYLIP *(Phylogenetic inference package), Version 3.5c*. Department of Genetics, University of Washington, Seattle, USA.

Ferris C, Oliver RP, Davy AJ, Hewitt GM (1993) Native oak chloroplasts reveal an ancient divide across Europe. *Molecular Ecology*, **2**, 337–344.

Ferris C, Oliver RP, Davy AJ, Hewitt GM (1995) Using chloroplast DNA to trace postglacial migration routes of oaks into Britain. *Molecular Ecology*, **4**, 731–738.

Folliot R (1964) *Contribution à l'étude de la biologie des cynipides gallicoles (Hymenoptera, Cynipoidea)*. PhD Thesis. University of Rennes, France.

Garnier-Géré P, Dillman C (1992) A computer program for testing pairwise linkage disequilibria in subdivided populations. *Journal of Heredity*, **83**, 239.

Gómez A, Carvalho GR (2000) Sex, parthenogenesis and genetic structure of rotifers: microsatellite analysis of contemporary and resting egg bank populations. *Molecular Ecology*, **9**, 203–214.

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

Guo SW, Thompson EA (1992) Performing the exact test of Hardy–Weinberg proportions for multiple alleles. *Biometrics*, **48**, 361–372.

Hails RS, Crawley MJ (1991) The population dynamics of an alien insect: *Andricus quercuscalicis* (Hymenoptera: Cynipidae). *Journal of Animal Ecology*, **60**, 545–562.

Hardy AC, Cheng L (1986) Studies in the distribution of insects by aerial currents.3. Insect drift over the sea. *Ecological Entomology*, **1**, 283–290.

Hasegawa M, Kishino H, Yano T (1985) Dating the human-ape splitting by a molecular clock of mitochondrial DNA. *Journal of Molecular Evolution*, **22**, 160–174.

Hewitt GM (1996) Some genetic consequences of ice ages, and their role in divergence and speciation. *Biological Journal of the Linnean Society*, **58**, 247–276.

Hewitt GM (1999) Postglacial re-colonisation of European biota. *Biological Journal of the Linnean Society*, **68**, 87–112.

Houard C (1912) Les zoocécidies do nord de l'Afrique. *Annales de la Societé Entomologique de France*, **81**, 1–236.

Huelsenbeck JP, Hillis DM (1993) Success of phylogenetic methods in the 4 taxon case. *Systematic Biology*, **42**, 247–264.

Huntley BH (1990) European vegetation history: palaeovegetation maps from pollen data — 13 000 years BP to present. *Journal of Quaternary Science*, **5**, 103–122.

Jalas J, Suominen J (1987) *Atlas Florae Europeae. Distribution of Vascular Plants in Europe*, Vol. 2. Cambridge University Press, Cambridge.

Jermiin LS, Crozier RH (1994) The cytochrome-*b* region in the mitochondrial DNA of the ant *Tetraponera rufoniger* — sequence divergence in Hymenoptera may be associated with nucleotide content. *Journal of Molecular Evolution*, **38**, 282–294.

Manos PS, Doyle JJ, Nixon KC (1999) Phylogeny, biogeography and processes of molecular differentiation in *Quercus* subgenus *Quercus* (Fagaceae). *Molecular Phylogenetics and Evolution*, **12**, 333–349.

Marsden-Jones EM (1953) A study of the lifecycle of *Adleria kollari* Hartig, the marble or Devonshire gall. *Transactions of the Royal Entomological Society of London*, **104**, 195–222.

Mitter C, Futuyma DJ, Schneider JC, Hare DJ (1979) Genetic variation and host-plant relations in a parthenogenetic moth. *Evolution*, **33**, 777–790.

Nei M (1978) Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics*, **89**, 583–590.

Nei M, Maruyama T, Chakraborty R (1975) The bottleneck effect and genetic variability in populations. *Evolution*, **29**, 1–10.

Nichols RA, Hewitt GM (1994) The genetic consequences of long distance dispersal during colonisation. *Heredity*, **72**, 312–317.

Nieves-Aldrey JL (1982) Contribución al conocimiento de los cinípidos gallícolas (Hym. Cynipidae) de la encina y el alcornoque en la provincia de Salamanca. *Boletín de la Asociación Española de Entomología*, **5**, 59–74.

Nieves-Aldrey JL (1987) Estado actual del conocimiento de la subfamilia Cynipidae (Hymenoptera, Parasitica, Cynipidae) en la Península Ibérica. *Eos*, **63**, 179–195.

Nieves-Aldrey JL (1992) Revisíon de las especies europeas del género *Callirhytis* (Hymenoptera, Cynipidae). *Graellsia*, **48**, 171–183.

Peakall R, Beattie AJ (1991) The genetic consequences of worker ant pollination in a self-compatible clonal orchid. *Evolution*, **45**, 1837–1848.

Petit RJ, Pineau E, Demesure B, Bacilieri B, Ducousso A, Kremer A (1997) Chloroplast DNA footprints of postglacial recolonisation by oaks. *Proceedings of the National Academy of Sciences of the USA*, **94**, 9996–10001.

Plantard O, Rasplus J-Y, Mondor G, Le Clainche I, Solignac M (1998) Wolbachia-induced thelytoky in the rose gallwasp *Diplolepis spinosissimae* (Giraud) (Hymenoptera: Cynipidae), and its consequences on the genetic structure of its host. *Proceedings of the Royal Society of London B*, **265**, 1075–1080.

Plantard O, Rasplus J-Y, Mondor G, Le Clainche I, Solignac M (1999) Distribution and phylogeny of *Wolbachia*-induced thelytoky in Rhoditini and 'Aylacini' (Hymenoptera: Cynipidae). *Insect Molecular Biology*, **8**, 185–191.

Posada D, Crandall KA (1998) Modeltest: testing the models of DNA substitution. *Bioinformatics*, **14**, 817–818.

Pujade-Villar J (1991) *Contribució al coneixement dels cinípids cecidógens dels arbres i arbursts de Catalunya, dels cinípids associats a aquests i dels seus parásits*. Unpublished PhD Thesis. University of Barcelona, Spain.

Pujade-Villar J (1992) *Andricus kollari* (Hartig) (Insecta: Hymenoptera: Cynipidae) Part II: Consideracions sobre el seu cicle biológic. *La Sitja Del Llop*, **3**, 12.

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

Reille M, Andrieu V, de Beaulieu J-L (1996) Les grands traits de l'histoire de la végétation des montagnes méditerranéennes occidentales. *Ecologie*, **27**, 153–169.

Richardson BJ, Baverstock PR, Adams M (1986) *Allozyme Electrophoresis; a Handbook for Animal Systematics and Population Studies*. Academic Press, New York.

Ros-Farré P, Pujade-Villar J (1998) Estudio mediante una trampa malaise de la comunidad de cinipidos cecidogenos e inquilinos de Santa Coloma, Andorra (Hymneoptera, Cynipidae). *Ecología*, **12**, 441–454.

Saitou N, Nei M (1987) The neighbour-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution*, **4**, 1406–1425.

Schönrogge K, Walker P, Crawley MJ (1998) Invaders on the move: parasitism in the galls of four alien gallwasps in Britain (Hymenoptera: Cynipidae). *Proceedings of the Royal Society of London, B*, **256**, 1643–1650.

Simon J-C, Baumann S, Sunnucks P *et al.* (1999) Reproductive mode and population genetic structure of the ceral aphid *Sitobion avenae* studied using phenotypic and microsatellite markers. *Molecular Ecology*, **8**, 531–545.

Stille B (1985a) Population genetics of the parthenogenetic gall wasp *Diplolepis rosae* (Hymenoptera: Cynipidae). *Genetica*, **7**, 145–151.

Stille B (1985b) Host plant specificity and allozyme variation in the parthenogenetic gallwasp *Diplolepis mayri* and its relatedness to *D. rosae* (Hymenoptera: Cynipidae). *Entomologia Generalis*, **10**, 87–96.

Stone GN, Cook JM (1998) The structure of cynipid oak galls: patterns in the evolution of an extended phenotype. *Proceedings of the Royal Society of London, B*, **265**, 979–988.

Stone GN, Schönrogge K, Crawley MJ, Fraser S (1995) Geographic variation in the parasitoid community associated with an invading gallwasp, *Andricus quercuscalicis* (Hymenoptera: Cynipidae). *Oecologia*, **104**, 207–217.

Stone GN, Sunnucks PJ (1993) The population genetics of an invasion through a patchy environment: the cynipid gallwasp *Andricus quercuscalicis*. *Molecular Ecology*, **2**, 251–268.

Sunnucks PJ, de barro PJ, Lushai G, MacLean N, Hales D (1997) Genetic structure of an aphid using microsatellites: cyclic parthenogenesis, differentiated lineages, and host specialisation. *Molecular Ecology*, **7**, 339–353.

Sunnucks PJ, Hales DF (1996) Numerous transposed sequences of mitochondrial cytochrome oxidase I–II in aphids of the genus *Sitobion* (Hemiptera: Aphididae). *Molecular Biology and Evolution*, **13**, 510–524.

Sunnucks PJ, Stone GN (1996) Genetic structure of invading insects and the case of the knopper gallwasp. In: *Frontiers of Population Ecology* (eds Floyd RB, Sheppard AW, De Barro PJ), pp. 485–496. CSIRO, Canberra, Australia.

Suomalainen E, Saura A, Lokki J (1987) *Cytology and Evolution in Parthenogenesis*. CRC Press, Boca Raton, Florida, USA.

Swofford DL (1998) PAUP* 4.0 — Phylogenetic Analysis Using Parsimony (*and other methods). Sinauer Associates, Sunderland, MA, USA.

Swofford DL, Olsen GJ, Waddell PJ, Hillis DM (1996) Phylogenetic inference. In: *Molecular Systematics* (eds Hillis DM, Moritz C, Mable BK), pp. 407–514. Sinauer, Sunderland, Massachusetts, USA.

Swofford DL, Selander RB (1981) BIOSYS 1: a FORTRAN program for the comprehensive analysis of data in population genetics and systematics. *Journal of Heredity*, **72**, 281–283.

Taberlet P, Fumagalli L, Wust-Saucy A-G, Cosson J-F (1998) Comparative phylogeography and postglacial colonisation routes in Europe. *Molecular Ecology*, **7**, 453–464.

Thompson JN (1994) *The Coevolutionary Process*. University of Chicago Press, Chicago USA.

Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL-W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research*, **22**, 4673–4680.

Toumi L, Lumaret R (1998) Allozyme variation in cork oak (*Quercus suber* L.): the role of phylogeography and genetic introgression by other Mediterranean oak species and human activities. *Theoretical and Applied Genetics*, **97**, 647–656.

Trotter A, Cecconi G (1904) Cecidotheca italica. *Marcellia*, **3**, 76–81.

Waring GL, Abrahamson WG, Howard DJ (1990) Genetic differentiation among host-associated populations of the gallmaker *Eurosta solidaginis* (Diptera: Tephritidae). *Evolution*, **44**, 1648–1655.

Weir BS (1991) *Genetic Data Analysis*. Sinauer, Sunderland, Massachussetts, USA.

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

Yaltirik F (1984) *Türkiye Meseleri Teshis kilavuzu (The oak trees of Turkey)*. Yenilik Basimevi, Istanbul.

Yang Z (1996) Among-site variation and its impact on phylogenetic analyses. *Trends in Ecology and Evolution*, **11**, 367–372.

Yasumatsu K (1951) A new Dryocosmus injurious to chestnut trees in Japan (Hym. Cynipidae). *Mushi*, **22**, 89–93.

Zhang DX, Hewitt GM (1996) Highly conserved nuclear copies of the mitochondrial control region in the desert locust *Schistocerca gregaria*: some implications for population studies. *Molecular Ecology*, **5**, 295–300.

Graham Stone has a long-standing interest in the phylogeography and evolution of oak cynipids, and in particular in the consequences of human dispersal of Turkey oak for the distributions and population genetic structure of host-alternating species. Rachel Atkinson and Antonis Rokas are both graduate students in the gallwasp group in Edinburgh. Rachel is working on interspecific comparisons of population structure in oak gallwasps, ranging from broad studies of phylogeography to population genetic analyses of the number of females laying eggs in a single oak bud. Antonis Rokas is working on a sequence-based phylogeny of oak cynipids, and has a special interest in the phylogenetic distribution of *Wolbachia* symbionts in gallwasps. Gyuri Csóka has a lifelong interest in all aspects of the biology of gallwasps, particularly in patterns of host oak exploitation. José-Luis Nieves-Aldrey has published many taxonomic papers on gallwasps and their associated communities, and has just completed a major monograph on this group in the Iberian Peninsular.

Supplementary material

The following material is available from http://www.blackwell-science.com/products/journals/suppmat/mec/mec1211/mec1211sm.htm

Appendix I

Sampling effort and allele frequencies at the 13 polymorphic loci for populations of *Andricus kollari*. Degrees latitude and longitude are decimalized