Molecular Ecology (2007) **16**, 2768–2781 doi: 10.1111/j.1365-294X.2007.03348.x

The phylogeographical clade trade: tracing the impact of human-mediated dispersal on the colonization of northern Europe by the oak gallwasp *Andricus kollari*

GRAHAM N. STONE,* RICHARD J. CHALLIS,* RACHEL J. ATKINSON,† GYÖRGY CSÓKA,‡ ALEX HAYWARD,§ GEORGE MELIKA,¶ SERAP MUTUN,** SONJA PREUSS,* ANTONIS ROKAS,†† EBRAHIM SADEGHI‡‡ and KARSTEN SCHÖNROGGE§§

**Institute of Evolutionary Biology, University of Edinburgh, School of Biological Sciences, The King's Buildings, West Mains Road, Edinburgh, EH9 3JT, UK, †Charles Darwin Research Station, Puerto Ayora, Santa Cruz, Galapagos Islands, Ecuador, ‡Hungarian Forest Research Institute, Mátrafüred Research Station, 3232 Mátrafüred, Hungary, §Department of Biology, South Parks Road, Oxford OX1 3PS, UK, ¶Systematic Parasitoid Laboratory, Plant Protection and Soil Conservation Service of County Vas, Ambrozy Setany 2, 9762 Tanakajd, Hungary, **Faculty of Science & Arts, Department of Biology, Abant Izzet Baysal University, 14280 Bolu, Turkey, ††The Broad Institute of MIT & Harvard, 320 Charles Street, Cambridge, MA 02141, USA, ‡‡Research Institute of Forest & Rangelands, PO Box13185-116, Teheran, Iran, §§Centre for Ecology and Hydrology, Benson Lane, Crowmarsh Gifford, Wallingford OX10 888, UK*

Abstract

Human dispersal of organisms is an important process modifying natural patterns of biodiversity. Such dispersal generates new patterns of genetic diversity that overlie natural phylogeographical signatures, allowing discrimination between alternative dispersal mechanisms. Here we use allele frequency and DNA sequence data to distinguish between alternative scenarios (unassisted range expansion and long range introduction) for the colonization of northern Europe by an oak-feeding gallwasp, *Andricus kollari*. Native to Mediterranean latitudes from Portugal to Iran, this species became established in northern Europe following human introduction of a host plant, the Turkey oak *Quercus cerris*. Colonization of northern Europe is possible through three alternative routes: (i) unassisted range expansion from natural populations in the Iberian Peninsula; (ii) unassisted range expansion from natural populations in Italy and Hungary; or (iii) descent from populations imported to the UK as trade goods from the eastern Mediterranean in the 1830s. We show that while populations in France were colonized from sources in Italy and Hungary, populations in the UK and neighbouring parts of coastal northern Europe encompass allozyme and sequence variation absent from the known native range. Further, these populations show demographic signatures expected for large stable populations, rather than signatures of rapid population growth from small numbers of founders. The extent and spatial distribution of genetic diversity in the UK suggests that these *A. kollari* populations are derived from introductions of large numbers of individuals from each of two genetically divergent centres of diversity in the eastern Mediterranean. The strong spatial patterning in genetic diversity observed between different regions of northern Europe, and between sites in the UK, is compatible with leptokurtic models of population establishment.

Keywords: *Andricus*, cynipid, introduction, invasion, lifecycle, oak

Received 26 September 2006; revision received 13 February 2007; accepted 20 March 2007

Correspondence: Dr Graham N. Stone, Fax: 0131 6506564; E-mail: graham.stone@ed.ac.uk

Graham Stone and Richard Challis contributed equally to this manuscript.

Introduction

Human-assisted dispersal can be a powerful force in shaping the genetic structure of animal populations, superimposing new signatures on existing natural phylogeographical

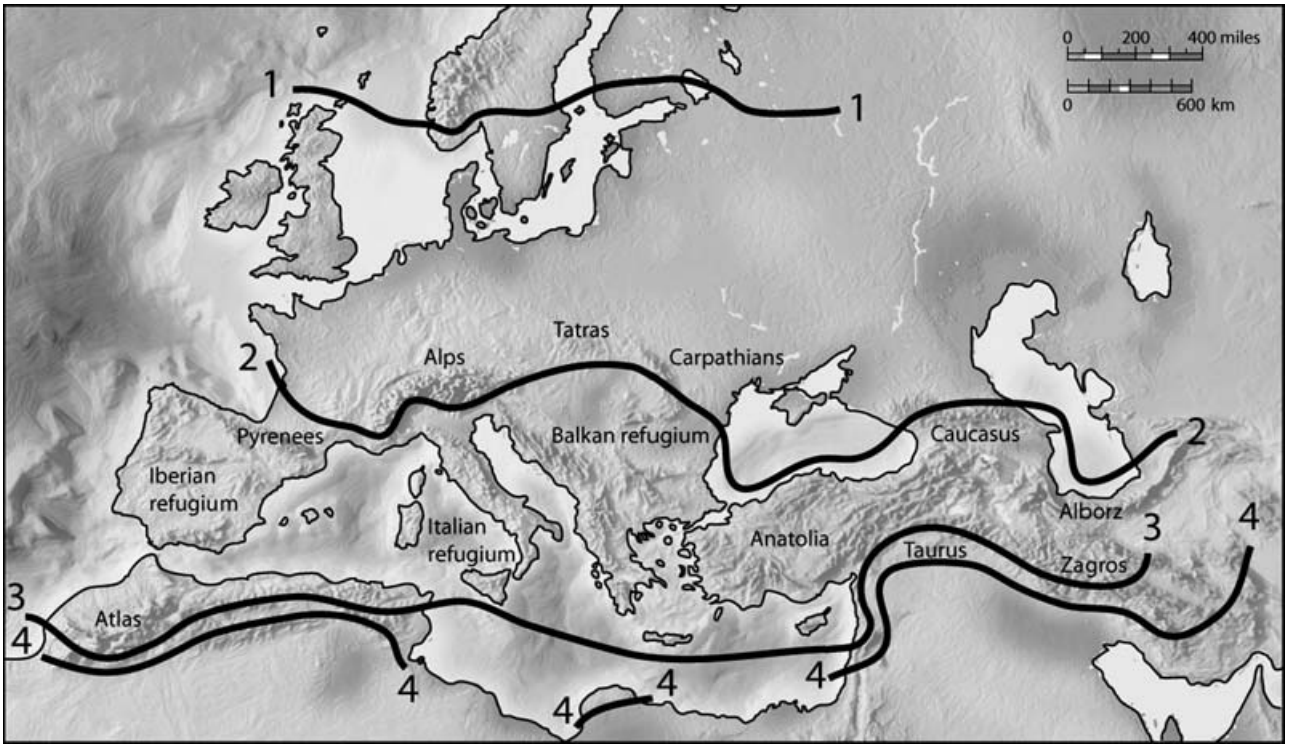

Fig. 1 Regions of the western Palaearctic occupied by the oak sections *Quercus sensu stricto* and *Cerris*. Line 1 represents the northern limit of oaks in the section *Quercus*, and of all oaks in the western Palaearctic. Only section *Quercus s.s.* oaks are native between lines 1 and 2. Line 2 represents the northern limit of the natural distribution of section *Cerris* oaks. Oaks in sections *Quercus s.s* and *Cerris* are naturally found together between lines 2 and 3. Line 3 represents the southern limit of oaks in the section *Quercus s.s.* Only section *Cerris* oaks are found between lines 3 and 4. Line 4 represents the southern limit of section *Cerris* oaks. Distribution sources are as follows: Europe (Tutin *et al*. 1993), Turkey (Yaltirik 1982; Davis 1965–85), the former USSR (Konarov 1936), Iran (Browicz & Menitsky 1971), Iraq (Townsend & Guest 1980), Palestine (Zohary 1966).

patterns. An understanding of the historical importance of organisms that have been widely traded, or are successful hitchhikers, is essential to unravelling their phylogeographical history. Human trade provides a dispersal mechanism most infamously exploited by the plague vector, *Xenopsylla cheopis* (the Oriental rat flea), and phylogeographical methods have been used extensively in analyses of the origin(s) and dispersal routes of pests (e.g. Grapputo *et al*. 2005; Navia *et al*. 2005), disease organisms (e.g. Morgan *et al*. 2005; Lehrmann *et al*. 2006), disease vectors (e.g. Mousson *et al*. 2005; Fonseca *et al*. 2006) and domestic animals (e.g. Morii *et al*. 2002; Haag *et al*. 2004; Meadows *et al*. 2005). Here we distinguish the consequences of two human impacts (dispersal of a host plant potentiating unassisted range expansion, and trade in the organism itself) on the population genetic structure of an herbivorous insect, the oak marble gallwasp *Andricus kollari*.

Andricus kollari is one of 11 oak gallwasps (Hymenoptera: Cynipidae) whose distributions in the western Palaearctic have been influenced by human activity (Stone & Sunnucks 1993; Stone *et al*. 2001, 2002; Csóka *et al*. 2005). Gallwasps are obligate parasites that develop within a sealed gall on the host plant. In common with most oak gallwasps, *A. kollari* has a cyclically parthenogenetic lifecycle involving strict alternation between a sexual spring generation and a parthenogenetic summer generation that overwinters in its gall. *A. kollari* belongs to a western Palaearctic clade of *Andricus* in which these two generations develop in galls on different groups of host oaks. The sexual generation gall develops on Cerris oaks (*Quercus* subgenus *Quercus* section *Cerris*) and the parthenogenetic generation gall develops on white oaks (section *Quercus sensu stricto*) (Cook *et al*. 2002). The natural distribution of such host-alternating gallwasps is limited to a latitudinal band in which oaks from both host sections grow together (Fig. 1). North of this band, the distributions of host-alternating gallwasps are constrained by the absence of section *Cerris* oaks, while south of this band they are constrained by the absence of section *Quercus s.s.* oaks. *A. kollari* has a natural distribution extending across this latitudinal band from Morocco through southern Europe and Turkey to Iran, and, in common with other host-alternating gallwasps,

shows a longitudinal sequence of sexual generation oak hosts. In the Iberian Peninsula, the sexual generation galls develop on cork oak, *Quercus suber*, from southeastern France to eastern Turkey, they develop on Turkey oak, *Q. cerris*, and in regions of Lebanon and Iran from which *Q. cerris* is absent, the probable host is the Lebanon oak, *Quercus libani* (Nieves Aldrey 2001; Stone *et al*. 2001; Hayward & Stone 2006; Challis *et al*. 2007).

The distribution of host-alternating gallwasps in Europe has been extended dramatically by the planting of *Q. cerris*. Native to regions south of the Alps and Carpathians, this oak has been dispersed far outside its native range over the last 500 years and is now found as far north as Scotland and Denmark and as far west as the Atlantic coast of France (Stone & Sunnucks 1993; Stone *et al*. 2001; Walker *et al*. 2002). Human introduction of *Q. cerris* to an area with widespread natural populations of white oaks (particularly sessile oak, *Quercus petraea*, and pedunculate oak, *Quercus robur*) has allowed nine host-alternating *Andricus* species to invade northern Europe, of which seven (*A. aries*, *A. corruptrix*, *A. lignicolus*, *A. kollari*, *A. lucidus*, *A. grossulariae* and *A. quercuscalicis*) have become established in the UK (Stone *et al*. 1995, 2002; Schönrogge *et al*. 1995, 1996a, b; Walker *et al*. 2002). Of these, *A. kollari* is the most widespread, reaching northern Scotland (Schönrogge *et al*. 1998, 1999; Walker *et al*. 2002).

For all of these species except for *A. kollari*, the invasion of northwestern Europe is thought to have occurred entirely without direct human assistance. Work on three species native to Italy and the Balkans shows a gradual and continuous loss of population genetic diversity away from the native range (Stone & Sunnucks 1993; Csóka *et al*. 1998). Populations farther along the invasion route show a subset of the allelic diversity found in populations closer to the native range. There is no evidence of the discontinuous distribution of genetic diversity predicted for large-scale introductions ahead of the invasion front, and no evidence that the English Channel presented any more of a barrier to dispersal than an equivalent distance over land. Previous work on *A. kollari* shows that the establishment of some populations in northwestern continental Europe can similarly be explained by range expansion by populations native to Italy and the Balkans, for which the introduced *Q. cerris* is also the natural sexual generation host. In contrast, Iberian populations have failed to expand beyond the northern limits of the natural distribution of cork oak in the Gironde, southwestern France (Stone *et al*. 2001; Hayward & Stone 2006).

However, a compatible alternative explanation exists for *A. kollari* populations in the UK. Uniquely among the range-expanding gallwasps, parthenogenetic generation galls of this species were deliberately imported by sea to the southwest of the UK in the 1830s (Walker *et al*. 2002). Oak cynipid galls contain high levels of tannins, and have been traded in the Mediterranean region for at least two and a half thousand years, since Theophrastus (371–286 BC) recorded the use of species including *A. kollari* in medicine, tanning of leather and the manufacture of inks in his *Historia Plantarum* (Beavis 1988). The scale of this trade is illustrated by Niebuhr's (1776–1780) account of a caravan of 1000 camels exporting an estimated 10 million oak galls (many of which will have still contained gallwasps) westwards from Syria. The *A. kollari* galls imported to the UK originated from a currently unknown region in the eastern Mediterranean. *A. kollari* adults escaped from the imported galls, and after initial establishment in southern England they reached Scotland within 50 years (Walker *et al*. 2002). This import of *A. kollari* predated the arrival in the UK and northern continental Europe of naturally invading gallwasps by a century (Stone & Sunnucks 1993).

This study has two parallel aims. First, we use sequence and allele frequency data to reconstruct the native range phylogeography of *A. kollari*, extending previous work to include sampling in possible eastern sources of the galls imported to the UK. In particular, we ask whether data for *A. kollari* support an eastern origin within the western Palaearctic, as inferred for other host-alternating *Andricus* species (Rokas *et al*. 2003; Challis *et al*. 2007).

Second, we attempt to distinguish between three alternative hypotheses for the origin of current British *A. kollari* populations, that they are: (i) primarily derived from a continuation of the unassisted range expansion inferred for the invasion of continental Europe, (ii) primarily the descendants of the 19th century introductions from an area east of these sources, or (iii) a combination of these. These scenarios are expected to leave different phylogeographical signatures in *A. kollari*. In addition to the patterns in genetic diversity described above, unassisted range expansion is predicted to result from rapid population growth from small numbers of founders. Sequences sampled from the invaded range are expected to represent a subset of ancestral native range polymorphism, and any clades restricted to the invaded range are expected to have a very young most recent common ancestor (MRCA). In contrast, large-scale introduction is expected to import much of the genetic diversity present in the source population to a new location, creating a new area of high diversity in the introduced range that lacks a signature of rapid population growth from small numbers of founders. If eastern Mediterranean source populations are genetically distinct from those in the Balkans and Iberia, we expect British populations to be genetically distinct from those already examined in northern France (Stone *et al*. 2001; Hayward & Stone 2006). The MRCA of such introduced sequences is expected to long predate the 1830s introduction events. Establishment of a large population at the periphery of the invaded range also has the potential to result in dispersal from the UK into continental Europe, against the prevailing direction of unassisted range expansion.

Materials and methods

Sample sites

We sampled the full longitudinal range of the natural distribution of *Andricus kollari*, from Portugal to Iran. No galls were found in Lebanon despite 14 man-days of sampling, and although recorded from Israel (Sternlicht 1968), this species is now extremely rare throughout the Levant. The only region where *A. kollari* could potentially be found that we have not sampled extends from the extreme southeast of Turkey into the Kurdish autonomous region of Iraq, and is currently inaccessible. Asexual generation females were reared from their galls under quarantine in Edinburgh. Wasps were stored at −80 °C until required, homogenized in allozyme extraction buffer (Peakall & Beattie 1991), and screened immediately as described below. Samples were subsequently stored at –80 °C until required for DNA extraction.

Analyses of allozyme data include 2092 individuals from 69 sites. Of these, data for 1457 individuals from 46 sites in continental Europe (Spain, France, Germany, Holland, Italy, and Hungary) and Turkey were presented in Stone *et al*. (2001). Allele frequency summaries for these sites (23–49 and 51–69 in on-line supplementary material Appendix 1) are available at http://www.blackwellpublishing.com/products/journals/suppmat/mec/mec1211/mec1211sm.htm. Here we present data for a further 562 individuals from 20 sites in the UK, 59 from two sites in the Republic of Ireland, and 14 from one site in Belgium (see Appendices I and II, Supplementary material). The locations of all sites used in allozyme analyses are shown in Fig. 3.

Sequences for fragments of the mitochondrial cytochrome b (*cytb*) gene and the D2 expansion region of the nuclear 28S gene (see below) were obtained for 160 individuals from 53 sites and 56 individuals from 27 sites, respectively (see Supplementary material Appendix III). Analyses of *cytb* data include sequences for 54 individuals published previously in Stone *et al*. (2001; GenBank Accession nos AF242739–AF242762 and AF242764–AF242766) and Hayward & Stone (2006; DQ925335–DQ925361). Our final *cytb* data set was distributed across sampled countries as follows (with numbers of individuals sequenced in parentheses): Spain (20), Portugal (1), France (52), UK (19), Ireland (3), Holland (6), Germany (1), Italy (16), Hungary (15), Turkey (22), Iran (3). A full alignment of all *cytb* haplotypes used in our analysis is available online in nexus format (see Appendix IV, Supplementary material).

Allozyme screening

Individual wasps were scored for allelic variants at 13 polymorphic loci using cellulose acetate electrophoresis (Zip-zone, Helena Laboratories) and substrate-staining protocols described by Richardson *et al*. (1986) and as used in Stone & Sunnucks (1993) and Stone *et al*. (2001). AK, αGPD1, αGPD2 and PEP-b were run on 40 mM sodium phosphate pH 6.3 (Stone & Sunnucks 1993), GOT-s, GOT-m, GPI, MDH-s, MDH-m, ME, 6PGD were run on 0.1 M Tris-EDTA-maleate-$MgCl_2$ pH 7.6 (Richardson *et al*. 1986; buffer F) and HK and PGM were run on 25 mM Tris-Glycine pH 8.5 (Richardson *et al*. 1986; buffer I).

Analyses of allele frequency data

Genotypic data were tested for deviations from Hardy–Weinberg equilibrium using the approach of Guo & Thompson (1992) incorporated in ARLEQUIN 2.0 (Schneider *et al*. 2000) with default settings, and for linkage equilibrium using the permutation procedure in GENETIX 4.0 (Belkhir 1999). Significance levels in both cases were adjusted for multiple tests using a Bonferroni correction (corrected threshold P value $= 1 - (1 - a)^{1/k}$ where k is the number of tests and a is the desired threshold value of 5%, as in Stone *et al*. (2001).

Relationships between populations were inferred using neighbour-joining analysis (NJBP, Jean-Marie Cornuet, INRA Laboratoire de Modélization et Biologie Évolutive, Montpellier) of Cavalli-Sforza's chord distance (Cavalli-Sforza & Edwards 1967) to allow comparison with previous work on this system (Stone *et al*. 2001). Bootstrapping was carried out over both populations and loci, with 1000 bootstrap replicates. Pairwise F_{ST} values between populations were calculated in ARLEQUIN 2.0 (Schneider *et al*. 2000) and their significance assessed using the permutation procedures in this programme.

We also determined the number of discrete populations (genotype pools) in *A. kollari* using the program STRUCTURE (Pritchard *et al*. 2000). STRUCTURE assumes a model in which a specified number of populations are characterized by a set of allele frequencies derived from multilocus genotype data. Individuals are assigned probabilistically to populations under the assumptions of Hardy–Weinberg and linkage equilibrium (shown to apply for our data below) using Markov chain Monte Carlo (MCMC) simulation. The simulation is rerun for models specifying different numbers of populations (K) and the posterior probabilities for each simulation are compared to infer the number of discrete populations best supported by the data. Simulations for K = 1–6 were run for 1×10^6 generations with a burn-in of 1×10^5 generations, and convergence in the estimated parameter values was checked over two independent runs at each K. STRUCTURE allows fitting of models with or without admixture, the latter allowing genotypes to arise through mating between individuals derived from different populations. The results with and without admixture were indistinguishable (as indicated by the small value of the Dirichlet parameter), and we report only the results without admixture.

DNA extraction, amplification and sequencing

DNA was extracted using the DNeasy Tissue Kit (QIAGEN), following the manufacturer's protocol for insect DNA extraction. A 433-bp fragment of the mitochondrial *cytb* gene was amplified using the primers CB1 (forward) 5′-TATGTACTACCATGAGGACAAATATC-3′ and CB2 (reverse) 5′-ATTACACCTCCTAATTTATTAGGAAT-3′ (Jermiin & Crozier 1994; Stone *et al*. 2001). A 489-bp fragment of the nuclear 28S D2 expansion region was amplified using the primers D2 forward 5′-CGTGTTGCTTGATAGTGCAGC-3′ and D2 reverse 5′-TCAAGACGGGTCCTGAAAGT-3′ (28S D2, Hancock *et al*. 1988; Rokas *et al*. 2002). Polymerase chain reactions (PCRs) of 25 μL were carried out in a PTC-200 DNA Engine (MJ Research) using 1 U *Taq* polymerase (Invitrogen or Promega), 2.5 μL 10 × *Taq* buffer, 1.5 μL $MgCl_2$ (25 mM), 0.5 μL dNTPs (10 mM), 0.35 μL primers (20 pmol), 1.0 μL template DNA and 18.85 μL distilled water. PCR products were purified using shrimp alkaline phosphatase (USB). PCR products were sequenced directly using ABI BigDye Terminator chemistry on ABI automated sequencers (Applied Biosystems), and in both directions to minimize PCR artefacts, ambiguities and base-calling errors. Sequences were a constant length for each locus and were analysed using SEQUENCHER 4.1 (Gene Codes). All new sequences are deposited in GenBank (Accession nos EF030046–EF030047 for 28S D2, EF031335–EF031457 for *cytb*), and a nexus format alignment for the cytochrome *b* sequences is available online (see Appendix IV, Supplementary material).

Phylogenetic analysis

Because only two 28S D2 haplotypes were found over the full range of *A. kollari*, formal phylogenetic analysis was only carried out on the *cytb* sequences. We assessed the validity of applying tree-based analyses to this data set using the method of Huson & Bryant (2006), which assesses the extent to which the data are compatible with a bifurcating tree. We first generated a 95% confidence phylogenetic network (1000 bootstrap replicates) in SPLITSTREE 4.4 (Huson & Bryant 2006). We used distances estimated under the HKY + I + G model of evolution using PAUP* 4b10 (Swofford 2001) and model parameters estimated using MODELTEST 3.6 (Posada & Crandall 1998). Base frequencies: A = 0.3743, C = 0.1167, G = 0.0913, T = 0.4177; ti/tv ratio = 5.7454, I = 0.5302; Gamma distribution shape parameter = 0.5510). The network was constructed using the NeighbourNet (Bryant & Moulton 2004) distances transformation and equal angle splits transformation (Dress & Huson 2004). The hypothesis that the data originated on a tree is accepted if the 95% confidence network contains a tree and rejected if it does not. The hypothesis was rejected for the *cytb* data, and we therefore represent relationships between haplotypes using the SPLITSTREE network. Splits computed from the data are represented as parallel edges rather than single branches, allowing visualization of ambiguous and conflicting signals in the data set (Huson & Bryant 2006). The network provides an implicit representation of evolutionary history (Huson & Bryant 2006). Before inferring relationships and demographic histories with these data, we confirmed the absence of any signature of selection using the McDonald–Kreitman test in DNASP 4.10 (Rozas *et al*. 2003).

Inference of population demographic history

Population demographic history was assessed using pairwise mismatch distributions (Harpending 1994) using ARLEQUIN (Schneider *et al*. 2000). Rapid population growth produces a unimodal mismatch distribution, whereas a large and stable population is expected to show a multimodal distribution (Slatkin & Hudson 1991). Mismatch distributions were tested for unimodality using the sum of squared deviations (SDD) test of Schneider & Excoffier (1999). For lineages showing significantly unimodal distributions, we used ARLEQUIN to estimate relative population sizes ($2\mu N$) before (θ_0) and after (θ_1) population growth, and the relative time since the onset of population expansion ($\tau = 2\mu t$). For these same lineages, we also calculated Fu's *F* statistic (Fu 1997), which under population growth has a large negative value due to the excess of singleton mutations (Fu 1997).

Results

Analyses of allozyme allele frequency data

Hardy–Weinberg and linkage equilibria. Over all 69 populations, all 13 of the allozyme loci were polymorphic, yielding a total of 64 alleles. Allele frequencies for sites in the UK and Ireland (sites 1–22) and Belgium (site 50) are available online (see Appendix II Supplementary material). Over all 69 populations, there were small numbers of significant departures from Hardy–Weinberg equilibrium (HWE: 14 of 484 tests) and linkage equilibrium (74 of 5850 pairwise tests). As in prior analyses of these loci in this species (Stone *et al*. 2001), there were no consistent departures from these equilibria either across loci within a given population, or across populations for a given locus, and both equilibria have been assumed in the STRUCTURE analysis.

Geographical patterns in allele frequencies. Variation in allele frequencies resolves *Andricus kollari* populations into four groups (Fig. 2). The native distribution of *A. kollari* resolves into two groups: (i) a strongly supported group containing all populations from the native range of *Quercus suber* in

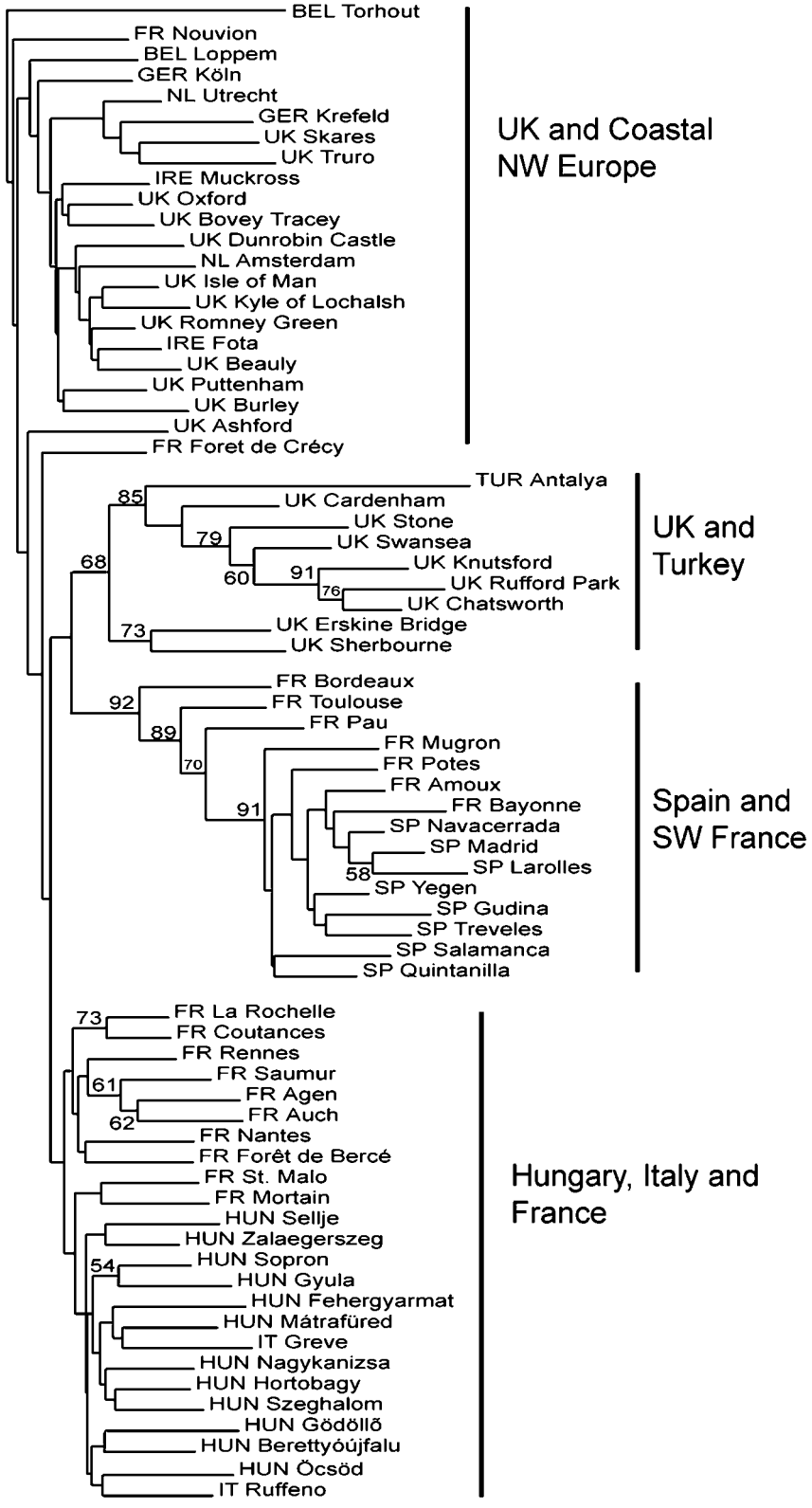

Fig. 2 A phylogram of relationships based on allele frequency data among *A. kollari* populations. Relationships are generated by neighbour joining of Cavalli-Sforza and Edward's chord distance. Numbers at nodes are bootstraps, over individuals and populations, expressed as a percentage of 1000 replicates. Nodes without bootstraps are supported by less than 50% of replicates.

Table 1 (a) Pairwise F_{ST} between sites in the two groups of UK populations, shown for four representative sites in each group. UK Sites whose allele frequencies cluster them with Turkey in Fig. 2 and whose individuals are predominantly in STRUCTURE population 3 are Chatsworth, Knutsford, Rufford Park and Swansea. UK sites affiliated to populations in northern coastal Europe and whose individuals are predominantly in STRUCTURE population 5 are Beauly, Puttenham Common, Oxford and Skares. All values in bold type are significantly different from zero at $P < 0.05$ using the permutation procedure in ARLEQUIN

	STRUCTURE population 3			STRUCTURE population 5			
Site name	Chatsworth	Knutsford	Rufford	Swansea	Beauly	Puttenham	Oxford
Knutsford	0.014						
Rufford	**0.020**	**0.055**					
Swansea	**0.048**	**0.064**	**0.067**				
Beauly	**0.379**	**0.361**	**0.449**	**0.329**			
Puttenham	**0.297**	**0.276**	**0.364**	**0.237**	0.025		
Oxford	**0.333**	**0.311**	**0.406**	**0.277**	**0.030**	0.003	
Skares	**0.369**	**0.331**	**0.446**	**0.328**	0.018	**0.025**	0.037

(b) Mean pairwise F_{ST} ± standard error (sample size) for the same two sets of UK populations and representative populations in continental northern coastal Europe (Torhout, Utrecht, Amsterdam, Köln), northwestern France (Bercé, Coutances, Crécy, St. Malo and Rennes), Hungary (Fehérgyarmat, Hortobágy, Mátrafüred, Nagykanizsa and Zalaegerszeg) and Turkey (Antalya). All values between sets of populations are significantly nonzero at $P < 0.01$ using the permutation procedure in ARLEQUIN

Region	UK sites in genotype pool 3	UK sites in genotype pool 5
UK Sites in population 3	0.045 ± 0.008 (6)	
UK Sites in population 5	0.360 ± 0.015 (12)	0.023 ± 0.004 (6)
Continental northern coastal Europe	0.261 ± 0.019 (16)	0.105 ± 0.020 (16)
Northwestern France	0.181 ± 0.011 (20)	0.115 ± 0.010 (20)
Hungary	0.200 ± 0.009 (20)	0.114 ± 0.008 (20)
Turkey	0.169 ± 0.016 (4)	0.375 ± 0.020 (4)

Spain and southwest France; and (ii) a grouping comprising southern central Europe (Italy and Hungary in the native range of *Quercus cerris*) and most populations in northwestern France. UK populations are divided between two further groupings: (iii) a group comprising eight UK populations and the Turkish population of Antalya; and (iv) 12 UK populations with neighbouring regions of northern coastal Europe [Republic of Ireland, Belgium, Holland, Germany and the two French sites of Crécy (49) and Nouvion (47) on the Channel coast].

While the two groups of UK populations are not geographically separated (both groupings comprise sites distributed from southern England to Scotland), they are genetically very divergent (Table 1a). While pairwise F_{ST} between sites within each group is low and sometimes not significantly different from 0, values between sites in the two groups have a mean of 0.36. Pairwise F_{ST} values between regions (Table 1b) show that the UK populations grouping with Turkey in Fig. 2 are relatively divergent from all regional groupings of populations, including Turkey. Those UK populations clustering with neighbouring regions of northern coastal Europe in Fig. 2 are very divergent from the Turkish population, but relatively less divergent from those in Hungary and Italy.

Both groups of UK populations are significantly differentiated from populations representing possible sources for unassisted range expansion from Iberia or central Europe. The main genetic divide between the UK and continental Europe is at the Channel coast, and is apparent in the distribution of regionally private alleles. Sites at the Channel coast of France have four alleles that are shared with Italy and Hungary (αGPD2 allele 1, GOT-m allele 5, PEP-b allele 8, PGM allele 5) but absent from the UK and northern coastal Europe. This northwards loss of alleles is potentially compatible with founder effects during unassisted range expansion. However, UK and Irish populations also possess six alleles that are absent from northwestern France (GOT-s allele 1, αGPD1 allele 1, HK allele 3, AK alleles 1, 2, 4). Significantly, AK allele 1 (found at 9 sites in the UK and Ireland, reaching a frequency of 40% at Knutsford) and AK allele 4 (only found as six copies in Cardinham, UK) are otherwise absent from the entire sampled range of *A. kollari*.

Analyses of individual multilocus genotypes. STRUCTURE analysis strongly supports the existence of five populations in *A. kollari* (Fig. 3). The posterior probability for $k = 5$ is ~1 and for all other values of k is ~0. Native range *A. kollari* are allocated predominantly to populations 1 (Iberia

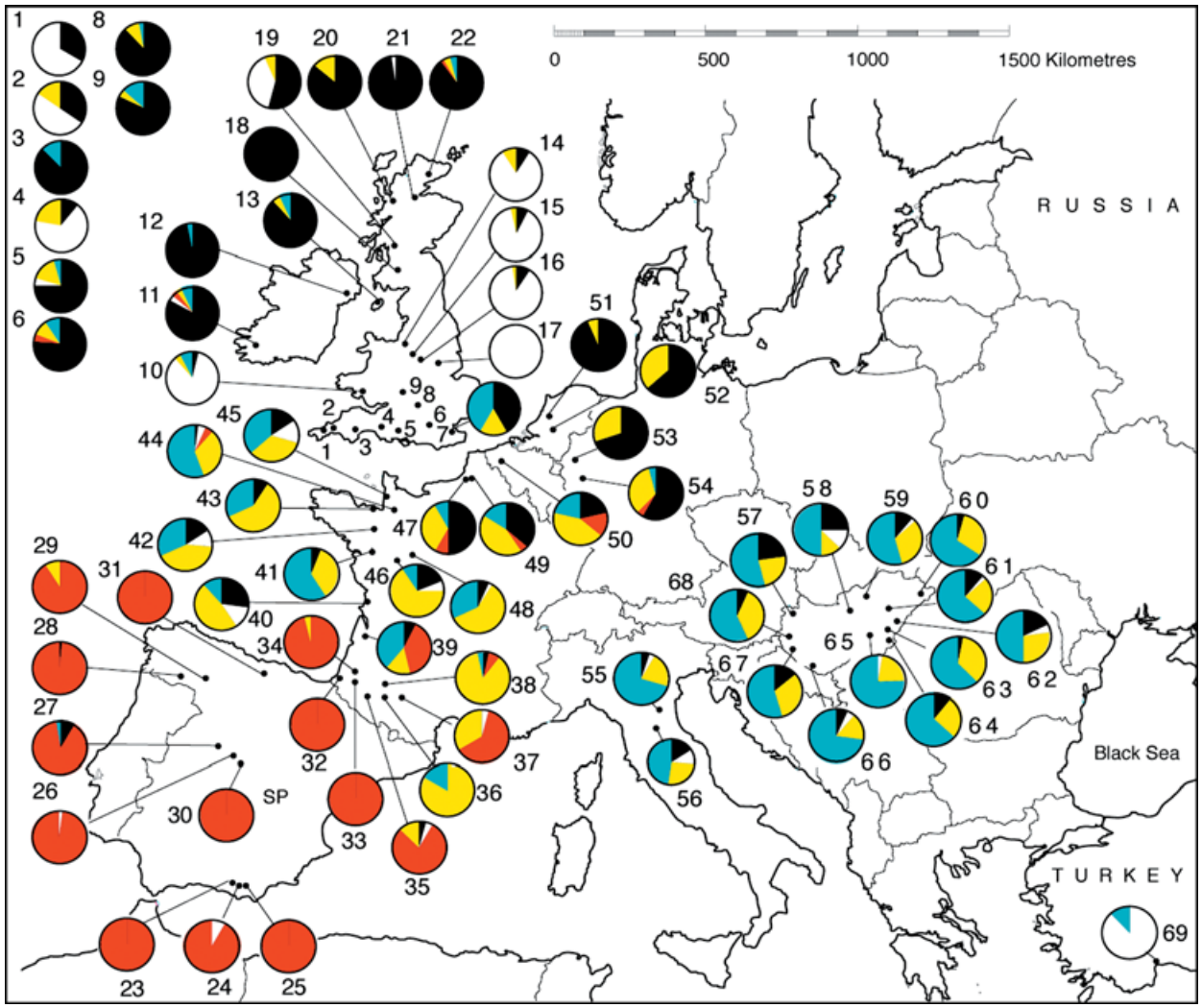

Fig. 3 The proportions of individuals in each population allocated to the 5 populations identified by STRUCTURE (1 = red, 2 = blue, 3 = white, 4 = yellow, 5 = black). Site numbers correspond to those in Table 1.

and southwestern France), 2 (Italy and Hungary) and 3 (Turkey). Population 4 contains individuals from sites throughout the native and introduced range of *Quercus cerris*. Population 5 primarily contains individuals from the UK and neighbouring regions of northern coastal Europe (Ireland, Belgium, the Netherlands and Germany), but is also represented in northern France, Italy and Hungary (Fig. 3).

This approach shows the strong genetic similarity between populations in northwestern France and the native range of *Q. cerris* in Italy and Hungary. Very few individuals sampled outside the native range of *Q. suber* are allocated to the Iberian population: of 359 individuals sampled in northern France, Belgium, Holland and northern Germany, only eight (2.2%) were allocated to the Iberian population 1, and these had low allocation probabilities (mean 0.51, range 0.37–0.71). Dispersal from the native and introduced ranges of *Q. cerris* into the Iberian Peninsula is also inferred to be very rare: of 284 individuals sampled in Spain, only 10 were not allocated to the Iberian population 1 (four, three, two and one individuals were allocated to populations 3, 2, 4 and 5, respectively, with probabilities ranging from 0.30 to 0.80).

UK sites fall into three groups on the basis of their STRUCTURE allocation (Fig. 3). One set (eight sites) contains sites for which a majority of individuals are allocated to the Turkish population 3. A second (13 sites) contains sites for which a majority of individuals are allocated to the fifth population. This division corresponds to the contrast between two groups of UK sites discussed above. The third group consists of Ashford in southeastern England, which is exceptional in having a high proportion of individuals allocated to population 2, thus resembling sites in northwestern France. Across the UK, only small minorities of individuals were allocated to the Iberian population 1

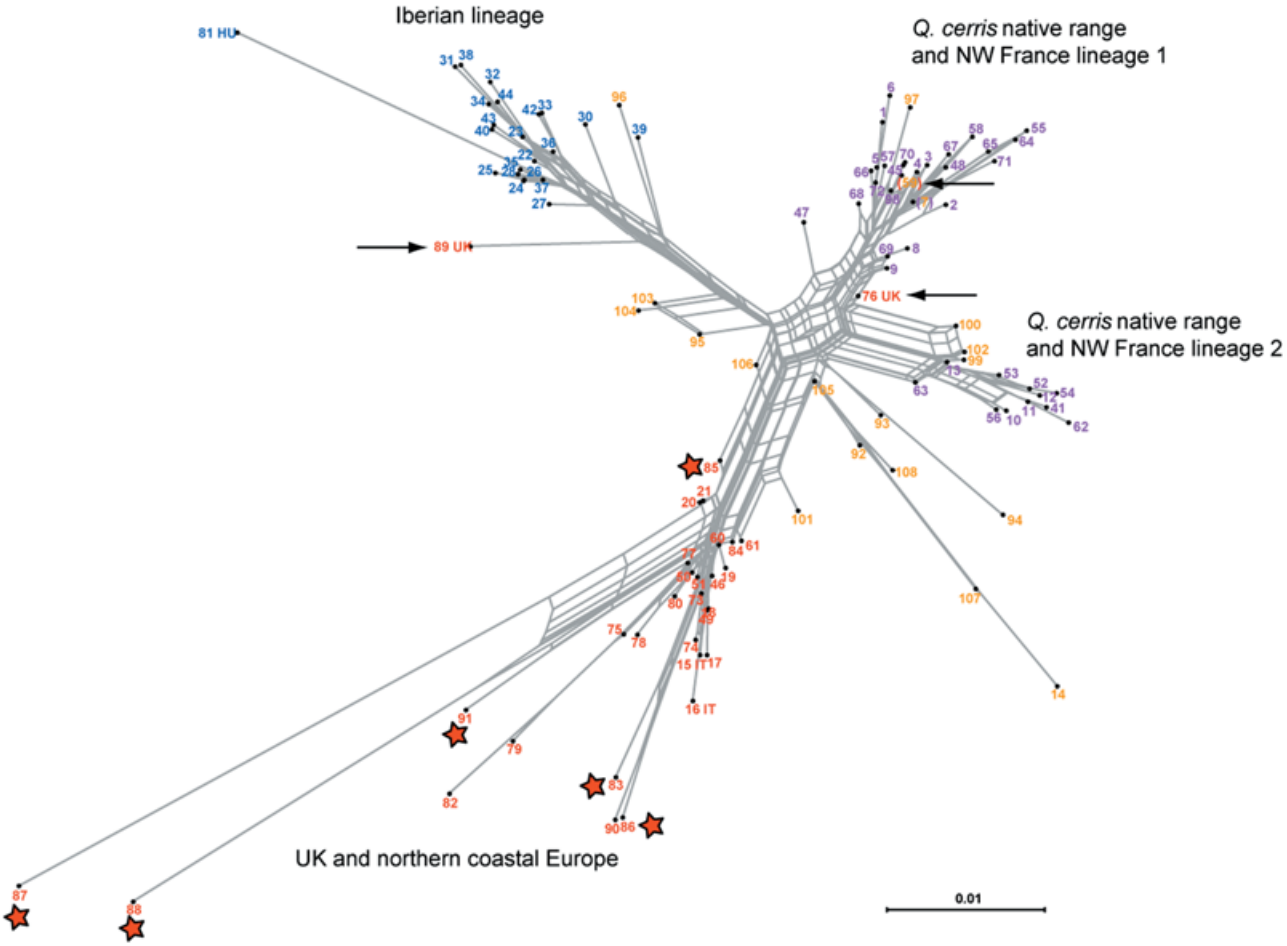

Fig. 4 A Neighbour-Net network (Bryant & Moulton 2004) of *cytb* haplotypes. The network was constructed using an equal angle splits transformation (Dress & Huson 2004) of distances under the HKY model of evolution in SplitsTree 4.4 (Huson & Bryant 2006). Geographical regions represented in each lineage are colour coded as follows: native range of *Q. suber* (blue), native range of *Q. cerris* in Europe and the introduced range of *Q. cerris* in France (purple), Turkey and Iran (orange) and the UK and northern coastal Europe (red). Haplotypes marked with a star indicate individuals whose multilocus allozyme genotypes were allocated to STRUCTURE population 3. Haplotypes marked with an arrow indicate individuals sampled in the UK and northern coastal Europe that fall outside the main lineage for this region.

(four individuals, 0.7%), and populations 2 (23 individuals, 4%) and 4 (37 individuals, 6.6%) common in the European native range of *Q. cerris*.

Sites in neighbouring regions of northern coastal Europe showed strong similarity to those UK populations dominated by the fifth population, but contained very few individuals allocated to Turkish population 3 (Fig. 3).

Analyses of sequence data

Phylogeographical patterns in cytochrome b *and 28S D2.* Our *cytb* sequencing added 62 new haplotypes (Genbank Accession nos EF031335–EF031457) to those recorded by Stone *et al.* (2001) and Hayward & Stone (2006), giving a total of 108 *A. kollari* haplotypes. The haplotypes recorded at each site and an alignment of all 108 sequences is provided in Appendices III and IV, Supplementary material. The entire data set of 108 discrete haplotypes has 94 polymorphic sites, 49 of which are parsimony informative.

The *cytb* network (Fig. 4) shows four major lineages diverging from an unresolved polytomy. All haplotypes from the native range of *Q. suber* form a single discrete lineage, which also contains three haplotypes sampled outside Iberia — one from Hungary (haplotype 81 from Gyula), one from the UK (haplotype 89 from Swansea) and one from Turkey (haplotype 96 from Bolu). Haplotypes from the native range of *Q. cerris* lie predominantly in two major lineages, each encompassing Central Europe and Turkey. The native range haplotypes not in these three major groups are from Turkey and Iran, and diverge independently from the central polytomy in the network. Across the native range, nucleotide diversity is greatest for Turkey (0.0200), intermediate in Italy (0.0151) and Hungary (0.0136), and lowest in Spain (0.0068).

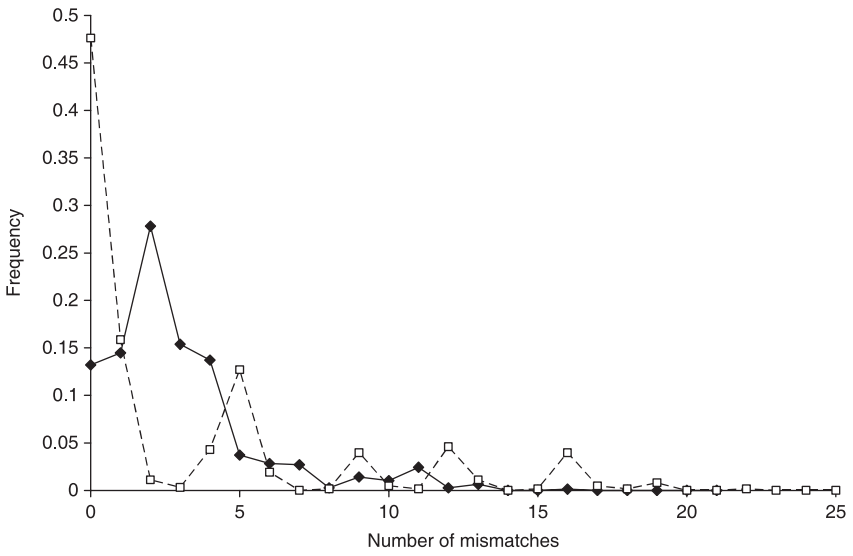

Fig. 5 Pairwise mismatch plots for *cytb* sequences in the Iberian (filled diamonds) and UK and northern coastal Europe (open squares) groups in Fig. 4.

The introduced range of *Q. cerris* in France contains haplotypes from both of the lineages representing the native range of this oak, implying independent colonization of northwestern Europe by members of each lineage. Contributions from both *Q. suber*- and *Q. cerris*-associated lineages mean that France contains the highest nucleotide diversity (0.0224) found in any single region.

With three exceptions, haplotypes from sites in the UK and northern coastal Europe (Ireland, Belgium, Holland and northern Germany) form a discrete lineage with high nucleotide diversity (0.0170). This lineage also includes two haplotypes sampled from sites in northern France (haplotypes 15 from Crécy and 82 from Nantes). Though the UK sequences form a discrete group relative to those sampled elsewhere, there is evidence that they comprise two genetically divergent populations. Individuals whose allozyme genotypes (comprising 11 nuclear and two mitochondrial loci) are allocated to STRUCTURE population 3 are significantly over-represented among the most divergent haplotypes (marked with stars in Fig. 4), while individuals allocated to STRUCTURE population 5 are significantly over-represented among the less divergent sequences in this lineage (*G*–test for association using Williams' correction, $G_{adj} = 8.00$ on 1 degree of freedom, $P < 0.005$; Sokal & Rohlf 1981).

Of the three UK and northern coastal Europe sequences lying outside the UK + northern coastal Europe lineage (marked with arrows in Fig. 4), one falls in the Iberian lineage (haplotype 89 from Swansea, UK), and two are associated with *Quercus cerris* native range lineage 1 (haplotypes 59 from Utrecht, Holland, and 76 from Cardinham, UK).

The D2 region of 28S revealed only two haplotypes, one present throughout the native range of *Q. suber* in Spain and southwestern France, and one shared by all other individuals. The individuals from Gyula and Bolu whose cytochrome *b* haplotypes lie in the Iberian lineage both had non-Iberian D2 haplotypes while the individual from Swansea had one copy of each of the D2 haplotypes. The two of these three individuals also genotyped for allozymes had non-Iberian multilocus genotypes: the Swansea individual was allocated to population 2, and the Gyula individual to population 5, rather than Iberian population 1.

Population demographic history. Of the lineages shown in Fig. 4, two showed a significantly unimodal pairwise mismatch distribution indicating rapid population expansion (see Fig. 5 for Iberia): for *Q. cerris* native range lineage 2 SDD = 0.0455, $P = 0.05$; for Iberia, SSD = 0.0916, $P < 0.001$. If we assume that μ is equivalent in the two lineages, the population expansion in the native range of *Q. cerris* was more dramatic ($\theta_0 = 0.025$, $\theta_1 = 3.63$, for the Iberian lineage $\theta_0 = 0.388$, $\theta_1 = 1.159$) and twice as long ago ($\tau = 2.559$, for the Iberian lineage $\tau = 1.248$). Significant population growth for both of these lineages was also inferred by Fu's *F* statistic test: $F_S = -8.143$, $P < 0.001$ for *Q. cerris* native range lineage 2, and $F_S = -10.674$, $P < 0.001$ for the Iberian lineage.

The UK and northern coastal Europe lineage has a multipeaked mismatch distribution more characteristic of a large and stable population (Fig. 5). No signature of population expansion was detected ($F_S = -2.278$, $P = 0.169$). The same was true for this clade when the more divergent sequences corresponding to STRUCTURE population 3 (identified in Fig. 4) were excluded ($F_S = -0.555$, $P = 0.35$).

Discussion

Native range phylogeography of Andricus kollari

Data for two other host-alternating *Andricus* species (*A. coriarius* and *A. quercustozae*) support the hypothesis

that European populations are derived from lineages originating in Turkey or Iran, leading to the hypothesis of an eastern cradle for the host-alternating clade of *Andricus* gallwasps (Rokas *et al*. 2003; Challis *et al*. 2007). The lack of resolution of the relationships between lineages in the *Andricus kollari* network mean that we cannot confirm or reject a similar 'out of the east' signature in *A. kollari*. However, the high nucleotide diversity in Turkey and Iran, and the starlike distribution of eastern sequences around the central polytomy in the haplotype network are both compatible with an eastern origin. The signature of population growth in one of the lineages containing the European native range of *Quercus cerris* matches that demonstrated for *Andricus coriarius* (Challis *et al*. 2007). In *A. coriarius*, this signature was inferred to be the result of westwards range expansion 1.6 million years ago. The signature in *A. kollari* is compatible with a similar event.

All of the approaches used here confirm earlier works (Stone *et al*. 2001; Hayward & Stone 2006) showing a major divide between Iberian populations galling cork oak, *Quercus suber* and those populations galling turkey oak, *Q. cerris*. In *A. coriarius* and *A. quercustozae*, Iberian haplotypes (and hence a lifecycle involving *Q. suber*) comprise a monophyletic clade derived from more eastern lineages associated with *Q. cerris* (Rokas *et al*. 2003; Challis *et al*. 2007). This relationship is compatible with the network obtained for *A. kollari*, but is not strongly supported by it. The low nucleotide diversity and signature of recent population growth detected in Iberian haplotypes could date from colonization of *Q. suber* from *Q. cerris* by a small number of founders, followed by population expansion over a newly available resource. Alternatively, it could indicate recovery from a subsequent population bottleneck.

The origins of UK Andricus kollari *and cross-channel gene flow*

Almost all *cytb* sequences for UK samples belong to a single lineage that includes samples from neighbouring coastal Europe but excludes all other sampled regions. This lineage has a pairwise mismatch distribution characteristic of a large and stable population, a signature that can only be explained by large-scale introduction from somewhere in the native range of this species. Extensive differentiation between UK sites and those in France, Iberia and southern central Europe is also supported by the allozyme data. This pattern contrasts strongly with three other host-alternating *Andricus* gallwasps (*A. corruptrix*, *A. lignicolus* and *A. quercuscalicis*) that colonized the UK without direct human assistance by dispersal from France (Stone & Sunnucks 1993; Csóka *et al*. 1998). The allozyme data for *A. kollari* suggest some affinity between a subset of UK populations and the single Turkish site sampled for these markers, but pairwise F_{ST} between these populations is still substantial. All of these lines of evidence suggest that UK populations are primarily derived from one or more large and genetically diverse introductions from unsampled parts of the eastern native range of *A. kollari*.

In addition to this primary route, haplotype sequences and STRUCTURE analysis provide evidence of rare gene flow into the UK across the English Channel from both *Q. suber*-(haplotype 89, Swansea) and *Q. cerris*-associated lineages (haplotype 76, Cardinham) in continental Europe. These haplotypes probably reached the UK by continuation of the unassisted range expansion process that brought related haplotypes from Italy and Hungary into north-western France. Direct dispersal from Spain is unlikely for the Swansea sample because although its mitochondrial haplotype is Iberian, one of its alleles for 28S D2 and its allozyme genotype are strongly diagnostic of *Q. cerris* regions outside Iberia (the allocation probability of this individual to the central European population 2 by STRUCTURE is 0.966). This suggests that this mitochondrial haplotype escaped Iberia by hybridization near where the distributions of *Q. suber* and planted *Q. cerris* meet in southern France, followed by multiple generations of backcrossing to central European wasps to result in a non-Iberian nuclear genotype. The existence of a UK site showing substantial similarity in population allele frequencies to sites in France (Ashford in southeastern England, Figs 2 and 3) suggests that successful colonization from France does occur, as inferred for three other host-alternating gallwasps (Stone & Sunnucks 1993; Csóka *et al*. 1998). Cytochrome *b* sequence and allele frequency data show that from the UK, *A. kollari* spread westwards into Ireland and eastwards into Belgium, Holland and northern Germany. Dispersal from the UK southwards into France appears to have been much more restricted, and only two of 30 haplotypes sampled from northern France fell within the UK and northern coastal Europe clade.

Genetic discontinuity across the English Channel and between France and northern coastal Europe is thus maintained despite bidirectional gene flow. Genetic discontinuities are not present in the same area in other closely related and ecologically similar invading gallwasps (Stone & Sunnucks 1993; Csóka *et al*. 1998), suggesting that the pattern in *A. kollari* is unlikely to be a consequence either of its biology or the distribution of its oak hosts. The pattern is consistent with predictions of models in which long-range dispersal ahead of an invasion front establishes spatial patterns in genetic diversity that are resistant to perturbation by subsequently arriving genotypes (Ibrahim *et al*. 1996; Lee & Hastings 2006). Even with rare long-range dispersal events, the patterns established in this way can persist for hundreds of generations. In the case of *A. kollari*, the number of introduced individuals was very large, and also (if we take other *Andricus* invasions as

a guide) occurred around 100 years before the unassisted arrival of central European genotypes. Both of these aspects of the introduction of *A. kollari* should extend the persistence of the resulting spatial patterns in genetic diversity.

How many sources?

The membership of a single lineage by all but two of the UK-sampled haplotypes for *cytb* suggests a single origin for most UK individuals. However, the correlation between nuclear allozyme genotypes and mitochondrial haplotypes within this lineage suggests that the UK *A. kollari* are derived from two related sources between which there has been limited gene flow in the past. Mean pairwise F_{ST} values between the two main populations within the UK and northern coastal Europe clade (0.36) are greater than those between Hungary and Spain (0.32, Stone *et al*. 2001), indicating differentiation at the level of discrete glacial refugia. There are two possible unsampled or little-sampled regions that could harbour such diversity, and yet be related genetically. One is south of the Taurus Mountains in Turkey and Syria (Fig. 1). The Taurus Mountains are part of a major faunistic and floristic divide termed the Anatolian Diagonal (Davis 1965–85, 1971; Çiplak 2003, 2004), which is also associated with major genetic divides in *Andricus* gallwasps (Rokas *et al*. 2003). Our sampling from this divide eastwards for *A. kollari* is limited to 10 haplotypes from sites near Kayseri, Bitlis and Mus (haplotypes 99–108 in Fig. 4). None of these haplotypes fall within the UK and northern coastal Europe lineage, but haplotype 104 (from Gevas, near Bitlis) is the most similar in sequence to this lineage of any native range samples. A second possible source is Lebanon and Israel, whose highlands support endemic flora (Zohary 1966; Shmida 1984) and which were probably isolated enough during Pleistocene ice ages to constitute discrete refugia for oaks, and hence associated gallwasps. We were unable to find any *A. kollari* in this region, but the possibility remains that they were present in the 19th century. Oak woodlands are under severe threat through degradation for charcoal production and by livestock throughout this region (Pons & Quézel 1985), and it is possible that one or both of the related sources for the UK *A. kollari* may be extinct in their region of origin.

How many introductions?

Unless gallwasps from the two inferred sources are able to mate assortatively after their introduction to the UK, we would expect associations between mitochondrial haplotypes and multilocus nuclear genotypes to have broken down rapidly through interbreeding. The fact that this has not occurred suggests that initial colonization of the UK was by discrete pulses of individuals, each dominated by one of the two sources. Variation in the source responsible for colonizing specific regions of the UK early in the invasion process could have established spatial patterning in the genetic make up of populations that has persisted to the present day. Such patchiness in the source of founding individuals is unlikely to have occurred were the two sources mixed in an individual shipment, and so spread together after their escape. It would be more likely if the source of gallwasps released in the UK varied between years. The trade in galls in the eastern Mediterranean was conducted through major trading centres such as Aleppo (Halab) in Syria that served a wide hinterland in Turkey, Syria, Iraq and Lebanon (Niebuhr 1776–1780). It is likely that the source of shipments exported through such centres varied from year to year.

Acknowledgements

We thank those who have helped in the collection of samples of *Andricus kollari*, particularly Gordon Brown, Roger Folliot, Ric Paul, Alex Rowe and Olivier Plantard.

References

Beavis IC (1988) *Insects and Other Invertebrates in Classical Antiquity* University of Exeter Press, Exeter, UK.

Belkhir K (1999) GENETIX, version 4.0. *A Windows Program for Population Genetic Analysis*. Laboratoire Génome, Populations, Interactions, Université Montpellier, Montpellier, France.

Browicz K, Menitsky GL (1971) Fagaceae. In: *Flora Des Iranischen Hochlandes und der Umrahmenden Gebirge*, Vol. 77 (ed. Rechinger KH), Akademische Druck-u.-Verlagsanstalt, Graz, Austria.

Bryant D, Moulton V (2004) NeighborNet: and agglomerative algorithm for the construction of planar phylogenetic networks. *Molecular Biology and Evolution*, **21**, 255–265.

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

Challis RJ, Stone GN, Mutun S *et al*. (2007) Longitudinal range expansion and cryptic eastern species in the western Palaearctic oak gallwasp *Andricus coriarius*. *Molecular Ecology*, **16**, 2103–2114.

Çiplak B (2003) Distribution of Tettigoniinae (Orthoptera, Tettigoniidae) bush crickets in Turkey: the importance of the Anatolian Taurus mountains in biodiversity and implications for conservation. *Biodiversity and Conservation*, **12**, 47–64.

Çiplak B (2004) Biogeography of Anatolia: the marker group Orthoptera. *Memorie della Società Entomologica Italia*, **82**, 357–372.

Cook JM, Rokas A, Pagel M, Stone GN (2002) Evolutionary shifts between host oak sections and host plant organs in *Andricus* gallwasps. *Evolution*, **56**, 1821–1830.

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. USDA, St Paul, Minnesota.

Csóka G, Stone GN, Melika G (2005) The biology, ecology and evolution of gallwasps. In: *Biology, ecology and evolution of gall-inducing anthropods* (eds Raman A, Schaeffer CW, Withers TM), pp. 573–642. Science Publishers Inc., Enfield, New Hampshire.

Davis PH (ed.) (1965–85) *Flora of Turkey and the East Aegean Islands*, Vol. 1–9. Edinburgh University Press, Edinburgh, UK.

Davis PH (1971) Distribution patterns in Anatolia with particular reference to endemism. In: *Plant Life of Southwest Asia* (eds Davis PH, Harper PC, Hedge IC), pp. 15–27. Botanical Society of Edinburgh, Edinburgh, UK.

Dress AWM, Huson DH (2004) Constructing splits graphs. *IEEE Transactions on Computational Biology and Bioinformatics*, **1**, 109–115.

Fonseca DM, Smith JL, Wilkerson RC, Fleischer RC (2006) Pathways of expansion and multiple introductions illustrated by large genetic differentiation among worldwide populations of the southern house mosquito. *American Journal of Tropical Medicine and Hygiene*, **74**, 284–289.

Fu YX (1997) Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics*, **141**, 915–925.

Grapputo A, Boman S, Lindstrom L, Lyytinen A, Mappes J (2005) The voyage of an invasive species across continents: genetic diversity of North American and European Colorado potato beetle populations. *Molecular Ecology*, **14**, 4207–4219.

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

Haag KL, Ayala FJ, Kamenetzky L, Gutierrez AM, Rosenzvit M (2004) Livestock trade history, geography and parasite strains: the mitochondrial genetic structure of *Echinococcus granulosus*. Argentina. *Journal of Parasitology*, **90**, 234–239.

Hancock JM, Tautz D, Dover GA (1988) Evolution of the secondary structures and compensatory mutations of the ribosomal RNAs of *Drosophila melanogaster*. *Molecular Biology and Evolution*, **5**, 393–414.

Harpending H (1994) Signature of ancient population growth in a low-resolution mitochondrial DNA mismatch distribution. *Human Biology*, **66**, 591–600.

Hayward A, Stone GN (2006) Comparative phylogeography across two trophic levels: the oak gall wasp *Andricus kollari* and its chalcid parasitoid *Megastigmus stigmatizans*. *Molecular Ecology*, **15**, 479–489.

Huson DH, Bryant D (2006) Application of phylogenetic networks to evolutionary studies. *Molecular Biology and Evolution*, **23**, 254–267.

Ibrahim KM, Nichols RA, Hewitt GM (1996) Spatial patterns of genetic variation generated by different forms of dispersal during range expansion. *Heredity*, **77**, 282–291.

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.

Konarov VL, ed. (1936) *Flora of the USSR*. Izdatel'stva Akademii Nauk SSSR. Moscow/Leningrad, USSR.

Lee CT, Hastings A (2006) Non-equilibrium genetic structure is insensitive to the shape of the dispersal distribution. *Evolutionary Ecology Research*, **8**, 279–293.

Lehrmann T, Marcet PL, Graham DH, Dahl ER, Dubey JP (2006) Globalization and the population structure of *Toxoplasma gondii*. *Proceedings of the National Academy of Sciences, USA*, **103**, 11 423–11 428.

Meadows JRS, Li K, Kantanen J *et al*. (2005) Mitochondrial sequence reveals high levels of gene flow between breeds of domestic sheep from Asia and Europe. *Journal of Heredity*, **96**, 494–501.

Morgan JAT, Dejong RJ, Adeoye GO *et al*. (2005) Origin and diversification of the human parasite *Schistosoma mansoni*. *Molecular Ecology*, **14**, 3889–3902.

Morii Y, Ishiguro N, Watanabe T *et al*. (2002) Ancient DNA reveals genetic lineage of *Sus scrofa* among archaeological sites in Japan. *Anthropological Science*, **110**, 313–328.

Mousson L, Dauga C, Garrigues T, Schaffner F, Vazeile M, Failloux AB (2005) Phylogeography of *Aedes* (*Stegomyia*) *aegypti* (L.) and *Aedes* (Stegomyia) *albopictus* (Skuse) (Diptera: Culicidae) based on mitochondrial DNA variations. *Genetical Research*, **86**, 1–11.

Navia D, de Moraes GJ, Roderick G, Navajas M (2005) The invasive coconut mite *Aceria guerreronis* (Acari: Eriophyidae): origin and invasion sources inferred from mitochondrial (16S) and nuclear (ITS) sequences. *Bulletin of Entomological Research*, **95**, 505–516.

Niebuhr C (1776–1780). *Reize Naar Arabië En Andere Omliggende Landen*. J.S. Baalde, Amsterdam.

Nieves Aldrey JL (2001) Fauna Iberica. *Hymenoptera, Cynipidae. Museo Nacional de Ciencias Naturales Consejo Superior de Investigaciones Cientificas, Madrid*, **16**, 1–636.

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

Pons AO, Quézel P (1985) The history of the flora and vegetation and past and present human disturbance in the Mediterranean region. In: *Plant Conservation in the Mediterranean Area* (ed. Gomez-Campo C), pp. 25–43. Dr W. Junk, Dordrecht, The Netherlands.

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

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

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

Rokas A, Atkinson RJ, Webster L, Csóka G, Stone GN (2003) Out of Anatolia: longitudinal gradients in genetic diversity support an eastern origin for a circum-Mediterranean oak gallwasp *Andricus quercustozae*. *Molecular Ecology*, **12**, 2153–2174.

Rokas A, Nylander JAA, Ronquist F, Stone GN (2002) A maximum likelihood analysis of eight phylogenetic markers using gallwasps (Hymenoptera: Cynipidae); implications for insect phylogenetic studies. *Molecular Phylogenetics and Evolution*, **22**, 206–219.

Rozas J, Sanchez-DeI, Barrio JC, Messeguer X, Rozas R (2003) DNASP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics*, **19**, 2496–2497.

Schneider S, Excoffier L (1999) Estimation of demographic parameters from the distribution of pairwise differences when the mutation rates vary among sites: application to human mitochondrial DNA. *Genetics*, **152**, 1079–1089.

Schneider S, Roessli D, Excoffier L (2000) ARLEQUIN, *Version 2.0: a Software for Population Genetics Data Analysis Genetics and Biometry Laboratory*. University of Geneva, Switzerland.

Schönrogge K, Walker P, Crawley MJ (1998) Invaders on the

move: parasitism in the galls of four alien gall wasps in Britain (Hymenoptera, Cynipidae). *Proceedings of the Royal Society of London. Series B, Biological Sciences*, **256**, 1643–1650.
Schönrogge K, Stone GN, Crawley MJ (1995) Spatial and temporal variation in guild structure: parasitoids and inquilines of *Audricus quercuscalicis* (Hymenoptera: Cynipidae) in its native and alien ranges. *Oikos*, **72**, 51–60.
Schönrogge K, Stone GN, Crawley MJ (1996a) Alien herbivores and native parasitoids: rapid development of guild structure in an invading gall wasp, *Andricus quercuscalicis* (Hymenoptera: Cynipidae). *Ecological Entomology*, **21**, 71–80.
Schönrogge K, Stone GN, Crawley MJ (1996b) Abundance patterns and species richness of the parasitoids and inquilines of the alien gall former *Audricus quercuscalicis* Burgsdorf (Hymenoptera: Cynipidae). *Oikos*, **77**, 507–518.
Schönrogge K, Walker P, Crawley MJ (1999) Complex life cycles in *Andricus kollari* (Hymenoptera, Cynipidae) and their impact on associated parasitoid and inquiline species. *Oikos*, **84**, 293–301.
Shmida A (1984) Endemism in the flora of Israel. *Botanische Jahrbucher für Systematik, Pflanzengeschichte und Pflanzen-Geographyraphie*, **104**, 537–567.
Slatkin M, Hudson RR (1991) Pairwise comparisons of mitochondrial DNA sequences in stable and exponentially growing populations. *Genetics*, **129**, 555–562.
Sokal RR, Rohlf FJ (1981) *Biometry*. W.H. Freeman, New York.
Sternlicht M (1968) The oak galls of Israel (*Quercus calliprinos* Webb. and *Quercus Ithaburensis* Decne.). *Israel Journal of Entomology*, **3**, 17–58.
Stone G, Atkinson R, Rokas A, Csóka G, Nieves-Aldrey JL (2001) Differential success in northwards range expansion between ecotypes of the marble gallwasp *Andricus kollari*: a tale of two lifecycles. *Molecular Ecology*, **10**, 761–778.
Stone GN, Schönrogge K, Atkinson RJ, Bellido D, Pujade-Villar J (2002) The population biology of oak gallwasps (Hymenoptera: Cynipidae). *Annual Review of Entomology*, **47**, 633–668.
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 P (1993) Genetic consequences of an invasion through a patchy environment — the cynipid gallwasp *Andricus quercuscalicis* (Hymenoptera, Cynipidae). *Molecular Ecology*, **2**, 251–268.
Swofford DL. (2001) *PAUP* Phylogenetic Analysis Using Parsimony (*and Other Methods)*. Sinauer Associates, Sunderland, Massachusetts.
Townsend CC, Guest ER (eds) (1980) *Flora of Iraq*, **4**, part 1: Cornaceae to Rubiaceae. Ministry of Agriculture and Agrarian Research, Baghdad, Iraq.
Tutin TG, Heywood VH, Burges NA *et al*. (1993) *Flora Europaea*. Cambridge University Press, Cambridge, UK.
Walker P, Leather SR, Crawley MJ (2002) Differential rates of invasion in three related alien oak gall wasps (Cynipidae: Hymenoptera). *Diversity and Distributions*, **8**, 335–349.
Yaltirik F (ed.) (1982) Fagaceae. *Flora of Turkey and the eastern Aegean Islands*. Edinburgh University Press, Edinburgh, UK.
Zohary M (1966) *Flora Palestina*, Vol. 1 The Israel Academy of Sciences and Humanities, Goldberg Press, Jerusalem.

This work forms part of an ongoing series of projects on the impact of range expansion on the community structure of oak gallwasps, led by Graham Stone (population genetics, phylogenetics) and Karsten Schönrogge (community dynamics). Richard Challis, Alex Hayward and Sonja Preuss apply molecular phylogenetic approaches to a wide range of within-and between species questions, ranging from detection of selection in *Plasmodium Var* genes to the evolution of Strepsiptera. Serap Mutun is using molecular approaches to study population structure, particularly of insects in Anatolia. György Csóka works on all aspects of the health of forests in Hungary, with particular interest in insect pests. George Melika has a long-term interest in the biology and taxonomy of gallwasps. Ebrahim Sadeghi works on applied entomology issues in forests and rangelands in Iran.

Supplementary material

The following supplementary material is available for this article:

Appendix SI Sample sites and sample sizes for 13 polymorphic allozyme loci in *Andricus kollari*. Latitude and longitude are given in decimal degrees, with a negative sign for western coordinates. Data for the other sites in our analyses are available online at http://www.blackwellpublishing.com/products/journals/suppmat/mec/mec1211/mec1211sm.htm

Appendix SII Allele frequencies by site for UK and Irish *Andricus kollari*. Site details are provided by site number in the online Supplementary material Appendix I.

Appendix SIII Sample sites, sample sizes and haplotype numbers (with numbers of copies when > 1 in brackets) for mitochondrial *cytb* and nuclear 28S D2 sequences. *Cytb* haplotype numbers are consistent with Hayward & Stone (2006), and numbers of contributing sequences previously published in Stone *et al*. (2001) (total 27 sequences, GenBank Accession nos AF242739–AF242762 and AF242764–AF242766) and Hayward & Stone (2006) (total 27 sequences, GenBank Accession nos DQ925335–DQ925361) are given in brackets in the *cytb* sample size column. Latitude and longitude are given in decimal degrees, with a negative sign for western coordinates. Site numbers refer to the map in Fig. 3.

Appendix SIV A nexus format file for the cytochrome *b* haplotypes of *Andricus kollari*. The first six sequences are outgroups in the genus *Andricus*: Acurva, *A. curvator*; Aconif, *A. conficus*; Acaput, *A. caputmedusae*; Asecko, *A. seckendorffi*; Asolit, *A. solitarius*; Acongl, *A. conglomeratus*. Genbank Accession numbers are given in Methods.

This material is available as part of the online article from:
http://www.blackwell-synergy.com/doi/abs/10.1111/j.1365-294X.2007.03348.x
(This link will take you to the article abstract).

Please note: Blackwell Publishing are not responsible for the content or functionality of any supplementary materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.