Heredity **87** (2001) 294–304 *Received 3 November 2000, accepted 1 February 2001*

Understanding patterns of genetic diversity in the oak gallwasp *Biorhiza pallida*: demographic history or a *Wolbachia* selective sweep?

ANTONIS ROKAS*, RACHEL J. ATKINSON, GORDON S. BROWN, STUART A. WEST & GRAHAM N. STONE

Institute of Cell, Animal and Population Biology, Ashworth Laboratories, University of Edinburgh, King's Buildings, West Mains Road EH9 3JT, U.K.

The endosymbiont *Wolbachia* can be responsible for selective sweeps on mitochondrial DNA variability within species. Similar signals can also result from demographic processes, although crucially the latter affect nuclear as well as mitochondrial loci. Here we present data on *Wolbachia* infection status and phylogeographic patterning for a widely distributed insect host, the oak gallwasp *Biorhiza pallida* (Hymenoptera: Cynipidae). Two hundred and eighteen females from eight European countries were screened for *Wolbachia*. All individuals from Hungary, Italy, France, U.K., Ireland, Switzerland, Sweden, and northern and southern Spain were infected with a single group A strain of *Wolbachia*, while populations in central Spain were not infected. A mitochondrial marker (cytochrome *b*) shows low variation and departure from neutrality in infected populations, but greater variation and no deviation from neutrality in *Wolbachia*-free populations. This pattern is compatible with a *Wolbachia*-induced selective sweep. However, we also find parallel differences between infected and uninfected populations for nuclear markers (sequence data for ITS1 and ITS2). All markers support the existence of a deep split between populations in Spain (some free of *Wolbachia*), and those in the rest of Europe (all infected). Allelic variation for five allozyme loci is also consistent with the Spain–rest of Europe split. Concordant patterns for nuclear and mitochondrial markers suggest that differences in the nature and extent of genetic diversity between these two regions are best explained by differing demographic histories (perhaps associated with range expansion from Pleistocene glacial refugia), rather than a *Wolbachia*-associated selective sweep.

Keywords: cytochrome *b*, gallwasp, phylogeography, range expansion, selective sweep, *Wolbachia*.

Introduction

Bacteria of the genus *Wolbachia* are intracellular microorganisms that infect the reproductive tissues of arthropods and nematodes (O'Neill *et al*., 1997; Werren, 1997; Stouthamer *et al*., 1999). They are inherited cytoplasmically (i.e. passed from mother to daughter) and alter reproduction in their arthropod hosts in a number of ways, including cytoplasmic incompatibility, male killing, feminization and imposition of parthenogenesis (see O'Neill *et al*., 1997 and chapters therein). *Wolbachia* are extremely common, infecting 16–22% of insects (Werren *et al*., 1995; West *et al*., 1998; Werren & Windsor, 2000), with a recent study indicating that the percentage might be even higher (Jeyaprakash & Hoy, 2000). In addition to immediate reproductive modifications, *Wolbachia* infection has a range of longer term evolutionary impacts on host taxa (O'Neill *et al*., 1997; Werren, 1997; Stouthamer *et al*., 1999).

Most research on *Wolbachia*, particularly for insect hosts, has concentrated on understanding the phylogenetic distribution and extent of infection (Werren *et al*., 1995; West *et al*., 1998; Werren & Windsor, 2000), using a small number of individuals for each species. The few studies examining geographical variation in levels of *Wolbachia* infection of a single host have revealed spatial patterns in the presence/absence of the bacterium and the occurrence of multiple infections (Turelli *et al*., 1992; Solignac *et al*., 1994; Plantard *et al*., 1998; Malloch *et al*. 2000). Studies that combine analyses of

*Correspondence. E-mail: a.rokas@ed.ac.uk

spatial variation in *Wolbachia* infection, and of genetic diversity in the host, are of particular relevance because of the potential impact of *Wolbachia* on the host's genetic structure (described below). In this paper, we address the potential impact of *Wolbachia* infection on large-scale genetic patterning in an insect host, the oak gallwasp *Biorhiza pallida*.

When an advantageous mutation is driven through a population to fixation (a process known as a selective sweep), much of the neutral variation at linked loci is eliminated during the process (Maynard Smith & Haigh, 1974). The genetic variants at the linked loci that are initially paired with the advantageous mutation, although neutral in themselves, 'hitchhike' to fixation. Spread of a specific *Wolbachia* strain through a host population can have a directly analogous effect on variation in other cytoplasmically inherited markers, such as mitochondrial DNA. Evidence from *Drosophila* suggests that a single mtDNA haplotype may become widespread in the host population through hitchhiking with a successful *Wolbachia* strain (Turelli *et al*., 1992; Ballard *et al*., 1996). Selective sweeps on mtDNA not only reduce haplotype diversity but also cause the remaining set of host haplotypes to deviate from predictions based on neutrality (Johnstone & Hurst, 1996).

These potential impacts of *Wolbachia* infection on mitochondrial markers are particularly important because the latter are often employed as phylogeographic markers under the assumption of neutral evolution (Johnstone & Hurst, 1996). Low variability in mtDNA, and departure from neutrality, can also be caused by demographic processes such as range expansions or population bottlenecks (Avise, 2000). Demographic and *Wolbachia*-related explanations for spatial patterning in mtDNA variation can be distinguished by comparing patterns seen for nuclear and mitochondrial markers. Unlike mitochondrial markers, nuclear markers are typically inherited in a Mendelian fashion and are not expected to show any change in nucleotide diversity or deviation from neutrality in response to *Wolbachia* infection, although this depends on the phenotypic effect that *Wolbachia* has on its host. More specifically, only if sexual reproduction is involved can this assumption hold. If the phenotypic effect of *Wolbachia* on its host is the induction of parthenogenesis (and assuming that genetic exchange with sexual relatives is rare), concordance between nuclear and mitochondrial markers of the host may be expected.

In contrast to a *Wolbachia*-induced selective sweep, demographic processes cause changes in the extent and nature of variation for both mitochondrial and nuclear markers, although mitochondrial markers are expected to show stronger responses due to their lower effective population size (Avise, 2000). This difference is the basis for demonstrating *Wolbachia*-associated selective sweeps using comparative studies of sequence variation in mitochondrial and nuclear markers (for example, Turelli *et al*., 1992; Ballard *et al*., 1996). For hosts infected with *Wolbachia*, concordance in spatial patterning of genetic diversity in nuclear and mitochondrial markers argues against a significant causative role for *Wolbachia* infection.

Undetected selective sweeps on mtDNA variation due to *Wolbachia* infection can thus generate at least two types of artefact if interpreted in a phylogeographic context. Firstly, loss of mtDNA diversity in part of the host species' range could be attributed to a demographic effect having a similar impact, such as a population bottleneck. Here, a process that in fact affects only cytoplasmically inherited markers could be mistaken for a demographic process affecting both mitochondrial and nuclear markers. Secondly, patterning in mtDNA variation generated by spatial patterning in *Wolbachia* infection could be interpreted mistakenly as indicative of the phylogeographic history of the host. These potential pitfalls are highly relevant to studies of insect phylogeography, because of the high proportion of insect species infected with *Wolbachia* (Werren *et al*., 1995; West *et al*., 1998; Jeyaprakash & Hoy, 2000; Werren & Windsor, 2000) and the widespread use of mitochondrial markers in phylogeographic reconstruction (Avise, 2000).

In this paper, we analyse spatial patterns of variation in *Wolbachia* infection, and mitochondrial and nuclear markers for a widespread European phytophagous insect, the oak-apple gallwasp *Biorhiza pallida* (Olivier 1791). *B. pallida* induces galls on oaks in the genus *Quercus* (Csóka, 1997). *B. pallida is* extremely widely distributed in the western Palaearctic, extending from Morocco in the west to Georgia in the east, and as far north as Sweden. *Wolbachia* infection was detected in central European populations of *B. pallida* as part of a broader phylogenctic survey of *Wolbachia* in oak gallwasps (Rokas, unpublished data). These two features of *B. pallida* make it a suitable taxon within which to examine large scale spatial patterns in *Wolbachia* infection, and any associated impacts on variation at both mitochondrial and nuclear markers.

Oak gallwasps are obligate parasites of oak trees, and their spatial patterns of genetic variation are expected to reflect to an extent the phylogeographic history of their oak hosts. During the Pleistocene (1.8 mya until 0.01 mya), cycles of glacials and interglacials led to repeated range contraction and expansion of many taxa across Europe, with many species surviving the glacials in refugial areas in the southern of Europe (Italy, the Balkans and Spain) (Huntley & Webb, 1989; Hewitt,

1999). The retreat of the last ice sheet across Europe at the end of the Pleistocene era was followed by northern range expansion by many organisms from one or more of these refugia, resulting in the distribution pattern we see today. A growing body of work on oak gallwasps, covering 12 species in the genera *Andricus* and *Cynips*, shows that their current distributions and geographical patterns of genetic variation are determined largely by two factors: (a) the number and location of regions that acted as glacial refugia, and (b) the extent to which alternative refugia have contributed colonists to post-glacial range expansion (Stone & Sunnucks, 1993; Sunnucks & Stone, 1996; Atkinson, 2000; Stone *et al.* 2001). Consistent features of all species studied include (a) the existence of refuge-specific allozyme alleles and mtDNA haplotypes, and (b) a decline in genetic diversity with increasing latitude and distance from refugia. Where both have been studied, patterns of genetic variation north of glacial refugia are similar for mitochondrial and nuclear markers (Stone *et al.* 2001), and are consistent with genetic subsampling of neutral variation associated with the range expansion process (Stone & Sunnucks, 1993; Sunnucks & Stone, 1996; Atkinson, 2000). Preliminary sampling of the species whose phylogeography has been studied also suggests that, unlike *B. pallida*, they are free of *Wolbachia* infection. Although precise phylogeographic scenarios vary among species, these studies indicate that in the absence of any impact of *Wolbachia* infection, similar concordance in spatial patterning in nuclear and mitochondrial markers represents a qualitative null expectation for *B. pallida*. In contrast, discordant patterning in nuclear and mitochondrial markers, and concordance in patterning of *Wolbachia* infection and mtDNA diversity would suggest a significant impact of *Wolbachia*.

We established spatial patterns in *Wolbachia* infection by using a polymerase chain reaction (PCR)-based screening technique for 218 *B. pallida* individuals sampled from 46 localities in eight countries across Europe (Fig. 1). Strain diversity of *Wolbachia* was assessed by sequencing a fragment of the *wsp* gene. Host genetic diversity was analysed using a mitochondrial sequence (a fragment of the cytochrome *b* gene) and two nuclear sequences (the internal transcriber regions ITS1 and ITS2). We extended the diversity of nuclear markers sampled by screening 270 individuals for five polymorphic allozyme loci. We use these data to address the following questions: (a) how many strains of *Wolbachia* are present in this host, and how many infection events have occurred?; (b) is infection with

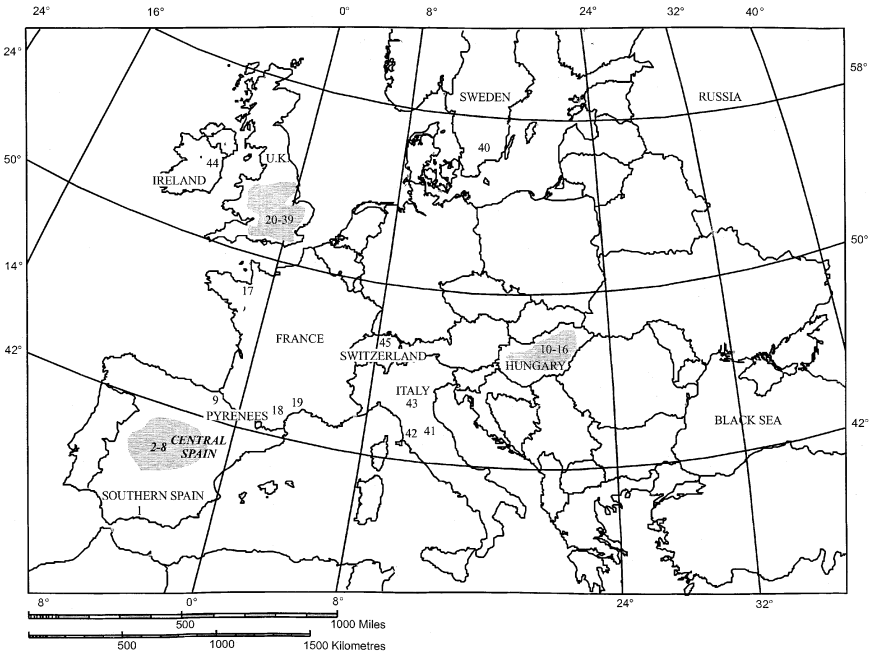

Fig. 1 Map of collection sites of populations of *Biorhiza pallida*. Spain: (1) Prado del Rey (2–8) Guadalix de la Sierra; Los Molinos; Zarzalejo; Avila-El Escorial; Cercedilla; Soto del Real; Villaviciosa. Northern Spain: (9) Puerto de Velate. Hungary: (10–16) Tiszaigar; Mátrafüred, Szentendre; Karcag; Bajna; Szeghalom; Visegrad. France: (17) Rennes; (18) St Jean pied de Porte; (19) Clermont-l'Hérault. U.K. (20–39) Cambridge; Hertford; Birnwood Forest; Oxford; Isle of Wight; Hampstead Heath; Fakenham; Broughton; SW Lincoln; Thetford; Elsfield; London; Ascot; Cawood; Chatham; West Shropshire; Grace Dieu Wood; Shorne; Hertingfordbury; Stoughton. Sweden: (40) Uppsala. Italy: (41) Chianti; (42) Volterra; (43) Casina. Ireland: (44) Dublin. Switzerland: (45) Luin. Populations in ***bold italics*** in central Spain (2–8) indicate absence of *Wolbachia* infection.

Wolbachia associated with lower sequence diversity, and departure from neutrality, for the mitochondrial marker?; (c) do the mitochondrial and nuclear markers show similar or discordant patterns of variation?; and (d) can we discriminate between demographic processes and a *Wolbachia*-associated selective sweep as possible causes of observed variation in host mitochondrial DNA?

Materials and methods

Collection and DNA extraction

Sexual generation galls of *B. pallida* were collected from 46 localities in eight European countries (Fig. 1, Table 1). *B. pallida* galls are multilocular (more than one offspring emerges from a single gall). To minimize screening of siblings we used one female from each gall, except where there was just a single gall from a particular location, when two females were screened. DNA was extracted from 218 female wasps as described by Stone & Cook (1998). To avoid contamination, each female wasp was soaked in 5% bleach and then serially rinsed in drops of sterile water prior to DNA extraction. With each DNA extraction three control extractions were performed using a *Nasonia Wolbachia*-positive strain, a *Nasonia Wolbachia*-negative strain and a no-DNA sample.

Wolbachia *screening*

Screening for *Wolbachia* was performed by PCR using *Wolbachia*-specific primers for the *ftsZ* cell-cycle gene (Werren *et al*., 1995). These and all other PCRs were performed in a PTC-200 DNA engine (MJ Research Waltham, MA, USA). The forward primer was ftsZF1 (Werren *et al*., 1995) and a new reverse primer WOLG-R was designed based on sequences available in GenBank. The sequence of WOLG-R (26 nucleotides) is 5′-GCA GVA TCA ACY TCA AAY ARA GTC AT-3′ (V = G/A/C, Y = C/T, *R* = A/G). The ftsZF1–WOLG-R pair amplifies the A, B and C groups of *Wolbachia*. Screening PCRs were attempted for sample DNA extractions at dilutions ranging from 1/10 to 1/100. Control PCRs were always performed. The PCR cycle for *ftsZ* was: one cycle of 94°C for 3 min, 55°C for 90 s and 72°C for 5 min, followed by 35 cycles of 94°C for 30 s, 55°C for 90 s, 72°C for 5 min and a final extension step at 72°C for 5 min. The PCRs were performed in 25 μL volumes, consisting of 1 μL DNA sample, 2.5 μL 10 × PARR buffer (Hybaid Ashford, UK), 1 μL $MgCl_2$ (25 mM), 0.5 μL dNTPs (10 mM), 0.35 μL of each primer (20 mM), 0.25 μL *Taq* (Promega Madison, WI, USA) and 19.05 μL of distilled, deionized H_2O. A 1% ethidium bromide-stained agarose gel was used for electrophoresis, and loaded with 15 μL of each reaction. To check that any *Wolbachia*-negative samples were not artefactual because of (a) failed DNA extraction; (b) presence of PCR inhibitors; or (c) incorrect DNA concentration, control PCRs with the general eukaryotic 28S rDNA primers 28Sf and 28Sr were performed as described in Werren *et al*. (1995). Of the 218 individuals screened for *Wolbachia* infection, 12 did not amplify for 28S rDNA and were discarded.

Table 1 Sampling regions of *Biorhiza pallida*. + *indicates* infection with *Wolbachia* − indicates absence of infection

Geographical region	Number of individuals tested	*Wolbachia*
Central Spain	34	−
Southern Spain	5	+
Northern Spain	5	+
Hungary	26	+
France	17	+
U.K.	95	+
Sweden	2	+
Italy	12	+
Ireland	5	+
Switzerland	5	+

PCR amplification and sequencing

All sequencing reactions were carried out at least twice (either with the forward and reverse primers for *wsp* and cytochrome *b* or twice with the forward primers for the ITS fragments) to minimize PCR artefacts, ambiguities and base-calling errors. Sequencing was carried out using the a sequencing kit and an automated sequencer (BigDye Terminator kit and ABI 377 sequencer; Perkin-Elmer Foster City, CA, USA). *Wolbachia* diversity in infected populations was assessed by PCR and sequencing of a fragment of the *wsp* gene (nine specimens were analysed: two from Hungary; one from southern Spain; two from France, southern and central; one from Switzerland; one from the U.K.; one from Ireland; and one from Italy). This is the most polymorphic gene so far isolated from *Wolbachia* (Zhou *et al*., 1998) and hence, the most likely to distinguish between two closely related *Wolbachia* strains. *Wolbachia*-infected individuals were sequenced for *wsp* using the 81F and 691R primers following the methods described by Zhou *et al*. (1998). The total volume of three PCR reactions for each individual wasp was electrophoresed on a 1% agarose gel. The expected bands were cut from the gel and cleaned with a DNA extraction kit (QIAQuick gel extraction kit, no. 28704; Qiagen, Crawley, UK), and the clean DNA fragment was quantified and sequenced. Twenty-nine individuals (10 from central Spain; five

from southern Spain; three from the U.K.; three from southern France; four from Hungary; one from northern Spain; one from Switzerland; one from Ireland; and one from Italy) were sequenced for a 433-base pair (bp) fragment of cytochrome *b* showing 13 distinct haplotypes (Table 2). The cytochrome *b* fragment was amplified using the primers CB1 and CB2 as described by Stone & Cook (1998), and purified and sequenced as described above for *wsp*. Internal transcriber regions were amplified using the universal primers ITS4 and ITS5 (White *et al*., 1990). The amplified fragment consisted of the internal transcriber regions, ITS1 and ITS2, and the 5.8S rDNA region of the rDNA array. Fifty-four clones of a 474-bp ITS2 fragment were sequenced for 13 individuals (three from central Spain, one from southern Spain, one from northern Spain, two from France, two from Italy, one from Hungary, one from Switzerland, one from Ireland and one from the U.K.) and 28 clones of a 635-bp fragment of ITSl for six (of the 13) individuals (see below). The PCR cycle consisted of an initial denaturation step of 94°C for 2 min, followed by 35 cycles of 94°C for 15 s, 55°C for 60 s and 72°C for 2 min, and a final extension step of 15 min at 72°C. Reaction conditions were as for *Wolbachia* screening, except that 1.5 μL $MgCl_2$ and 18.55 μL of deionized, distilled H_2O were used. The rDNA array is present in multiple copies in the typical eukaryote genome and undergoes concerted evolution (Hillis & Dixon, 1991). However, for fast-evolving regions of the array, concerted evolution is not always perfect, resulting in intraindividual variation. To cheek for this variation, PCR products were cloned using a cloning kit (TOPO TA cloning kit no. 4500–01; Invitrogen Groningen, Netherlands) and 2–6 clones from each specimen were subsequently sequenced. Plasmids containing the fragment of interest were isolated using a commercial kit (QIAprep Spin miniprep kit, no. 27104; Qiagen). Plasmid DNA was subsequently quantified and sequenced.

Allozyme screening

Two hundred and seventy individuals were screened at five variable allozyme loci using cellulose–acetate gel electrophoresis (Zip Zone equipment; Helena Laboratories Gatshead, UK) as described in Stone & Sunnucks (1993) and Stone *et al*. (2001). Of an original set of 13 loci (see Stone *et al*. 2001) screened in *B. pallida*, five systems were found to be polymorphic: peptidase b (PEPb) (EC 3.4.11); aspartate aminotransferase (GOTm) (EC 2.6.1.1); 6-phosphogluconate dehydrogenase (6PGD) (EC 1. 1. 1.44); malate dehydrogenase (MDHs) (EC 1.1.1.37); and glucose-phosphate isomerase (GPI) (EC 5.3.1.9). PEPb, GOTm, 6PGD and MDHs were run on a sodium phosphate buffered gel (pH 6.3), and GPI was run using a Tris-EDTA–maleate–$MgCl_2$ buffered gel (pH = 7.6).

Table 2 List of haplotypes for the 433-bp fragment of *cytochrome b* from *Biorhiza pallida*. Parsimony-informative sites are indicated by an asterisk (*) and singletons by a full stop (.)

			1	112	222	233	333	334	4
	111	566	990	170	124	902	224	771	1
	034	912	020	535	499	321	250	394	8
	**.	**.	.**	***	*..	***	***	.**	.
Haplotype 1	GTA	GTA	GAT	GTT	TAC	ATT	TGC	ATC	T
Haplotype 2	...	..G	...	...	.G.	...	...	...	.
Haplotype 3	...	...	...	...	...	...	...	...	C
Haplotype 4	...	...	...	...	...	...	...	G..	.
Haplotype 5	...	...	...	...	..T	...	.A.	...	.
Haplotype 6	...	...	...	...	...	...	...	..T	.
Haplotype 7	...	...	...	A..	...	...	..T	...	.
Haplotype 8	...	A..	...	A.C	C..	GAC	.A.	.AT	.
Haplotype 9	...	A..	A..	A.C	C..	GAC	.A.	.AT	.
Haplotype 10	...	A..	.G.	A.C	C..	GAC	.A.	.AT	.
Haplotype 11	A..	.A.	.G.	A.C	C..	.AC	CA.	..T	.
Haplotype 12	A.G	.A.	.G.	A.C	C..	.AC	CA.	..T	.
Haplotype 13	..C.	...	.GC	ACC	...	GAC	C.T	..T	.

Haplotype 1: St Jean pied de Porte, France; Clermont-l'Hérault, France; Szeghalom, Hungary; Broughton, U.K.; Luin, Switzerland; Szentendre, Hungary; Oxford, U.K.; Cambridge, U.K. Haplotype 2: Puerto de Velate, Northern Spain. Haplotype 3: Mátrafüred, Hungary. Haplotype 4: St Jean pied de Porte, France. Haplotype 5: Bajna, Hungary. Haplotype 6: Dublin, Ireland. Haplotype 7: Casina, Italy. Haplotype 8: Avila-El Escorial, central Spain. Haplotype 9: Guadalix de la Sierra, central Spain. Haplotype 10: Guadalix de la Sierra, central Spain; Zarzalejo, central Spain; Villaviciosa, central Spain; Los Molinos, central Spain; Soto del Real, central Spain. Haplotype 11: Cercedilla, central Spain; Zarzalejo, central Spain. Haplotype 12: Cercedilla, central Spain. Haplotype 13: Prado del Rey, southern Spain (five individuals).

Analysis of sequence data

Sequences were aligned by CLUSTALW (Thompson *et al*., 1994) using the default settings. ITS1 and ITS2 alignments were manually checked to verify that there were no ambiguities. Departures from neutrality were tested using Tajima's *D* (Tajima, 1989) and Fu and Li's *D** and *F** (Fu & Li, 1993) statistics incorporated into the software program (DNASP, version 3.0) (Rozas & Rozas, 1999). For ITS1 and ITS2, sites with alignment gaps were excluded from the neutrality tests. All generated sequences were used (including the multiple clones from each individual).

Phylogenies for cytochrome *b* were generated using parsimony (MP) and maximum likelihood (ML) in PAUP* (version 4.0b3) (Swofford, 1999). MP was performed for 1000 bootstraps replicated with all sites equally weighted, using the tree bisection-rooting (TBR) option in a heuristic search. All ML analyses were heuristic searches using the TBR option. To reduce computational time, only 100 bootstrap replications were performed for each ML analysis. A hierarchical series of increasingly complex models of sequence evolution was employed to identify the model that made the data most likely using likelihood ratio test (LRT) statistics (reviewed by Huelsenbeck & Rannala, 1997). We tested, singly and together, the effects of unequal base frequencies, different rates between transitions and transversions (ti/tv) and rate variation over nucleotide sites (Hasegawa *et al*., 1985). The assumption of among-site rate heterogeneity (Yang, 1993) and the enforcement of a molecular clock were also tested for the best-fitting model. The shape parameter α of the gamma distribution and the ti/tv ratio were calculated from a 50% majority rule consensus tree using unweighted parsimony. Whenever unequal base frequencies were employed, we used the empirical frequencies of the nucleotides (Yang *et al*., 1994).

For ITS, phylogenies were constructed using combined data for ITS1 and ITS2. Each gap was coded as missing and as a fifth nucleotide character. Trees were constructed using MP (as above) and ML, using PUZZLE (Strimmer & von Haeseler, 1996). MP bootstraps were carried out as above for 100 replicates. ML was performed using the quartet puzzling algorithm with 10 000 puzzling steps, the HKY85 model of DNA substitution (Hasegawa *et al*., 1985) and a gamma-shaped distribution to account for rate heterogeneity. All parameters were estimated from the dataset. A parsimony analysis was also performed using the gap insertions as unique indel (insertion/deletion) events. They were treated separately from the sequence dataset because no data exist concerning their frequency of substitution.

Results

Wolbachia *screening and diversity*

Of the 206 *B. pallida* successfully screened, 85% were infected with *Wolbachia*. Over the entire sampled range only populations in central Spain were not infected (Fig. 1, Table 1). All sequences were 564-bp long (GenBank accession number AF339629), with only two nucleotide positions polymorphic. Position 218 was polymorphic for T/C and position 230 was polymorphic for A/G in all the individual sequences. As the sequencing was done from a PCR fragment, it could not be determined whether these sites are genuinely polymorphic (not fixed yet, four alleles) or represent two different strains (two alleles). The low *wsp* diversity suggests that all infected European populations carry the same strain of *Wolbachia*.

Mitochondrial phylogeography

The 29 individuals sequenced for the cytochrome *b* fragment were polymorphic for 25 positions, 18 of which were parsimony informative (4.18%). This variation yielded a total of 13 discrete haplotypes (see Table 2). Likelihood ratio tests showed that the best-fitting model for the *B. pallida* data used empirical base frequencies, a ti/tv ratio equal to 10.44 and variation in rate among sites (using a gamma distribution with $\alpha = 0.0057$ and four rate categories). The estimated value $\alpha = 0.0057$ denotes a very strong rate variation (see Yang *et al*., 1994). The assumption of a molecular clock could not be rejected at the 0.05% probability level (unconstrained model *versus* model with molecular clock enforced, $-\Delta L = 44.42$, d.f = 27, $0.01 < P < 0.05$).

All MP and ML analyses reveal a deep and well-supported split (based on 10 changes) between populations in central and southern Spain *versus* those in northern Spain and the rest of Europe (Fig. 2). Sequence diversity was higher in *Wolbachia*-free populations in central and southern Spain ($P = 0.0082$) than in the *Wolbachia*-infected populations throughout northern Spain and the rest of Europe ($P = 0.0029$). The 14 sequences from the populations in northern Spain and the rest of Europe showed significant deviation from predictions under neutrality (Tajima's $D = -2.09$, $P < 0.05$; Fu and Li's $D^* = -2.73$, $P < 0.05$ and $F^* = -2.93$, $P < 0.05$). In contrast, central and southern Spanish sequences, from either the 10 uninfected central Spanish individuals (Tajima's $D = 0.52$, $P > 10$; Fu and Li's $D^* = 0.62$, $P > 0.10$ and $F^* = 0.67$, $P > 10$), or the central and southern Spanish individuals combined (15 sequences) (Tajima's $D = 1.20$, $P > 0.10$; Fu and Li's

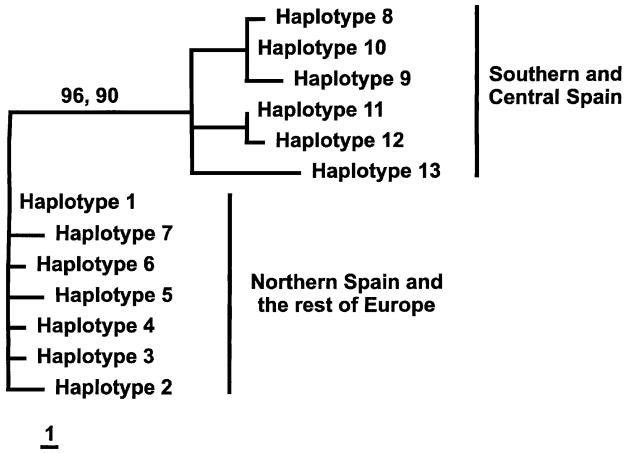

Fig. 2 Consensus phylogram (50% majority rule) for cytochrome *b* haplotypes using an ML model that accounts for unequal base frequencies, a different ti/tv ratio and rate variation among sites. See Table 2 for haplotype information. Values above the branches denote bootstrap support under ML and MP, respectively. The tree is 35 steps long with a consistency index (CI) of 0.714 and a retention index (RI) of 0.931. GenBank accession numbers: AF339616–AF339628.

$D^* = 0.89$, $P > 0.10$ and $F^* = 1.13$, $P > 0.10$), showed no significant deviation from neutrality.

ITS phylogeography

The 13 individuals sequenced for ITS2 and the six individuals sequenced for ITS1 and ITS2 revealed 43 parsimony-informative positions (Table 3). The length of the fragments varied between 456 and 467 bp for ITS2 and between 626 and 632 bp for ITS1. Length differences from clones from the same individual were usually much smaller or absent. The ML tree for ITS constructed with the quartet puzzling algorithm, and the MP tree coding alignment gaps as missing information, were both poorly resolved. However, MP analysis with gaps coded as a fifth base strongly supports the split of the population from central and southern Spain *versus* that from northern Spain and the rest of Europe revealed by the mitochondrial cytochrome *b* data (Fig. 3). MP analysis of the ITS indel dataset (with 10 of 12 indels being parsimony informative) revealed the same topology, although with lower bootstrap support.

Deviation from neutrality was tested separately for ITS1 and ITS2. Given that these are nuclear loci, within which recombination might be occurring, the estimates for deviation from neutrality are conservative. Significantly, regional variation in departures from neutral expectations parallel those seen in cytochrome *b*. For both loci, Fu and Li's D^* and F^* and Tajima's D showed significant departure from neutrality for populations outside Spain (ITS1: Tajima's $D = -1.92$, $P < 0.05$; Fu and Li's $D^* = -2.77$, $P < 0.05$ and $F^* = -2.92$, $P < 0.05$. ITS2: Tajima's $D = -1.98$, $P < 0.05$; Fu and Li's $D^* = 4.15$, $P < 0.05$ and $F^* = 4.05$, $P < 0.05$) but were nonsignificant for central Spanish specimens (ITS1: Tajima's $D = -1.01$, $P > 0.10$; Fu and Li's $D^* = -0.99$, $P > 0.10$ and $F^* = -1.07$, $P > 0.10$. ITS2: Tajima's $D = -0.77$, $P > 0.10$, Fu and Li's $D^* = 0.86$, $P > 0.10$ and $F^* = -0.95$, $P > 0.10$).

Allozyme variability

The sample sizes obtained from each location (maximum 12 galls, and so 12 females screened) were too small to allow meaningful analysis of allele frequencies. We thus limit our interpretation of the data to the presence/absence of specific alleles, as summarized in Table 4. Four alleles were found only in Spain. Of these, two were present only in central Spain; allele 1 at 6PGD was present only in one population, whereas allele 2 at 6PGD was present in four populations. The third allele was present in the only southern Spanish population (allele 4 at GPI) and one was present in all 10 Spanish populations (allele 1 at MDHs). Populations in the rest of Europe possessed two alleles absent from Spain. One was from a single individual from Stoughton, U.K. (allele 2 at GPI) and the other was very common in individuals from Phoenix Park, Ireland (allele 4 at MDHs). These regional differences in locally restricted alleles support the substantial genetic divergence between central and southern Spain *versus* northern Spain and the rest of Europe implied by the cytochrome *b* and ITS sequence analyses.

Discussion

Geographic variation in Wolbachia *infection*

Extensive screening revealed that, with the exception of central Spain, all European populations of *B. pallida* sampled are infected with *Wolbachia*. The apparent single strain infecting *B. pallida* appears to be relatively cosmopolitan; it is shared with three other European gallwasp species (Rokas, unpublished data), and database searches suggest that it also infects tse-tse flies. There are three possible explanations for existence of infected populations of *B. pallida* both north and south of a *Wolbachia*-free region in central Spain: (a) there was a single infection event and central Spanish populations have lost their infection; (b) there was a single infection in a common ancestor of the southern Spanish, northern Spanish and the rest of Europe *B. pallida* populations not shared with populations in central Spain; or (c) infection has occurred independently in southern Spain

Table 3 Parsimony-informative sites for ITS1 and ITS2. A dash (–) denotes a gap and a full stop (.) denotes an identical site. Individuals with the same parsimony-informative sites were grouped together (see below.) Numbers indicate number of individual and number of clone, respectively. All numbers refer to the individuals in Fig. 3. An asterisk (*) indicates individuals that were sequenced for both ITS1 and ITS2. Individuals without an asterisk were sequenced only for ITS2. Groups – **1:** 123.5; **2:** 140.10*; **3:** 117.3*, 123.11, 123.13, 48.4, 48.11, 53.4, 212.11, 35.3, 48.12, 53.11, 96.20*, 53.2, 48.7, 35.1, 48.3; **4:** 212.3, 212.9, 212.7, 212.8, 96.13*; **5:** 96.10*, 96.18*, 96.16*; **6:** 117.15*; **7:** 117.2*; **8:** 123.1; **9:** 141.6*, 140.5*, 141.10*, 141.3*, 48.5; **10:** 96.19*; **11:** 140.13*; **12:** 141.5*; **13:** 141.12*; **14:** 147.5*; **15:** 147.14*; **16:** 147.15*; **17:** 157.6, 157.16; **18:** 157.13, 157.7, 172.5; **19:** 147.19*; **20:** 147.13*; **21:** 32.5*; **22:** 32.7*, 32.3*, 32.8*; **23:** 32.10*; **24:** 147.3*; **25:** 123.3

ITS1

			2	2	2	3	4	4	4	4	5	5	5	5	5
	8	9	4	7	7	5	5	7	8	8	0	0	1	1	1
Group	5	4	5	6	7	5	2	2	3	4	5	6	3	4	5
1	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?
2	A	A	A	–	–	T	A	A	–	–	T	A	A	T	T
3	A	A	A	–	–	T	A	A	–	–	T	A	A	T	T
4	.	.	.	A	A	C	.	T	–	–	.	.	.	.	.
5	.	.	.	A	A	.	.	T	–	–	.	.	.	.	.
6	.	.	.	–	–	.	.	T	–	–	.	.	–	–	–
7	.	.	.	–	–	.	.	.	–	–	.	.	.	.	.
8	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?
9	.	.	.	–	–	.	G	T	–	–	.	.	–	–	.
10	.	.	.	A	A	.	.	T	–	–	.	.	.	.	.
11	.	.	.	–	–	.	.	.	–	–	.	.	.	.	.
12	.	.	.	–	–	C	G	T	–	–	.	.	–	–	–
13	.	.	.	–	–	.	G	T	–	–	.	.	–	–	–
14	.	.	–	–	–	.	.	T	T	A	–	–	–	–	–
15	.	.	–	–	–	.	.	T	T	A	–	–	–	–	–
16	.	.	–	–	–	.	.	T	T	A	–	–	–	–	–
17	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?
18	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?
19	.	T	–	–	–	.	.	T	T	A	–	–	–	–	–
20	.	T	–	–	–	.	.	T	T	A	–	–	–	–	–
21	C	.	.	–	–	.	.	T	–	–	–	–	–	–	–
22	C	.	.	–	–	.	.	T	T	A	–	–	–	–	–
23	C	.	.	–	–	.	.	T	T	A	–	–	–	–	–
24	.	.	–	–	–	.	.	T	T	A	–	–	–	–	–
25	?	?	?	?	?	?	?	?	?	?	?	?	?	?	?

ITS2

				1	1	1	2	2	2	2	2	2	2	2	3	3	3	3	3	3	3	4	4	4	4	4	4	4
	1	3	4	3	6	6	0	1	1	1	1	4	4	4	1	4	4	5	5	8	9	0	0	0	0	0	1	2
Group	5	8	8	9	1	2	6	0	1	2	3	6	7	8	6	8	9	0	1	7	9	5	6	7	8	9	6	2
1	T	–	–	A	–	–	T	T	A	–	–	–	–	–	A	G	T	–	–	C	T	A	–	–	–	–	–	T
2	.	–	–	.	–	–	.	.	.	–	–	–	–	–	.	.	.	G	T	.	.	.	–	–	–	–	A	.
3	.	–	–	.	–	–	.	.	.	–	–	–	–	–	.	.	.	–	–	.	.	.	–	–	–	–	A	.
4	C	–	C	.	–	–	G	.	.	–	–	–	–	–	.	.	.	–	–	.	.	.	–	–	–	–	A	.
5	C	–	C	.	–	–	G	.	.	–	–	–	–	–	.	.	.	–	–	.	.	.	–	–	–	–	A	.
6	.	–	C	G	–	–	G	.	.	–	–	–	–	–	.	.	.	–	–	.	.	.	–	–	–	–	A	.
7	.	–	–	G	–	–	.	.	.	–	–	–	–	–	.	.	.	–	–	.	.	.	–	–	–	–	A	.
8	.	–	–	.	–	–	.	.	.	T	A	C	A	T	.	.	.	G	T	.	A	.	–	–	–	–	A	.
9	C	–	–	.	G	T	.	–	–	–	–	C	A	T	.	.	.	G	T	.	A	.	–	–	–	–	A	.
10	C	–	–	.	–	–	.	.	.	T	A	C	A	T	.	.	.	G	T	.	A	.	A	–	–	–	A	.
11	C	–	–	.	G	T	.	–	–	–	–	C	A	T	.	.	.	G	T	.	A	.	A	–	–	–	A	.
12	C	–	–	.	G	T	.	–	–	–	–	C	A	T	.	.	.	G	T	.	A	–	–	–	–	–	A	.
13	C	–	–	.	G	T	.	–	–	–	–	C	A	T	.	.	.	G	T	.	A	–	–	–	–	–	A	.
14	C	T	–	.	G	T	.	–	–	–	–	–	–	–	.	–	–	–	–	.	A	.	A	–	–	–	A	.
15	C	T	–	.	G	T	.	–	–	–	–	–	–	–	.	–	–	–	–	.	A	.	A	–	–	A	A	.
16	C	T	–	.	G	T	.	–	–	–	–	–	–	–	.	–	–	–	–	.	A	.	–	–	–	–	A	.
17	C	A	–	.	G	T	.	–	–	–	–	–	–	–	.	–	–	–	–	T	A	G	A	A	A	A	T	C
18	C	A	–	.	G	T	.	–	–	–	–	–	–	–	.	–	–	–	–	.	A	G	A	A	A	A	T	.
19	C	A	–	.	G	T	.	–	–	–	–	–	–	–	.	–	–	–	–	.	A	–	–	–	A	A	A	.
20	C	A	–	.	G	T	.	–	–	–	–	–	–	–	.	–	–	–	–	.	A	A	–	–	–	–	A	.
21	C	T	–	.	G	T	.	–	–	–	–	–	–	–	G	–	–	–	–	.	A	.	A	–	–	–	A	.
22	C	T	–	.	G	T	.	–	–	–	–	–	–	–	G	–	–	–	–	.	A	.	–	–	–	–	A	.
23	C	T	–	.	G	T	.	–	–	–	–	–	–	–	G	–	–	–	–	.	A	–	–	–	–	–	A	.
24	C	T	–	.	G	T	.	–	–	–	–	–	–	–	G	–	–	–	–	.	A	–	–	–	–	–	A	.
25	C	–	–	.	G	T	.	–	–	–	–	C	A	T	.	.	.	–	–	.	.	.	–	–	–	–	A	.

and in Europe from the Pyrenees northwards. Too little is currently known about the mechanisms of *Wolbachia* transfer between lineages for these possibilities to be distinguished, although evidence for horizontal transfer of *Wolbachia* between and within species is accumulating (Vavre *et al*., 1999; Huigens *et al*. 2000). We also know very little about the phenotypic consequences of *Wolbachia* infection in *B. pallida*. *Wolbachia* has been shown to induce parthenogenesis in some non-oak cynipids (e.g. Plantard *et al*., 1999), but the presence of males in all our samples suggests that this does not occur in *B. pallida*. Additionally, the slow lifecycle of this host, and obligate development of the larva within oak tissues make the mating and curing experiments needed to examine *Wolbachia*-associated effects (such as for example, cytoplasmic incompatibility) extremely difficult (see also Plantard *et al*., 1998; Plantard *et al*., 1999).

Patterns of genetic diversity in B. pallida

Variability at the mitochondrial locus was higher in populations free of *Wolbachia* than in the northern Spanish and the rest of Europe populations infected with the symbiont. Similarly, infected northern Spanish and rest of Europe populations showed significant departures from neutrality, while central and southern Spanish populations showed no such signature. Such covariation between infection and sequence diversity is compatible with a *Wolbachia*-induced selective sweep.

Large-scale demographic processes such as range expansion can have a similar impact on mtDNA variation to that caused by a selective sweep in a population (Donnelly & Tavaré, 1995). Studies on other European oak gallwasps predict a significant impact of range expansion on spatial patterns in *B. pallida*, and should be regarded as a more parsimonious underlying cause if compatible with the data. As discussed above, demographic processes tend to generate qualitatively similar patterns at both mitochondrial and nuclear markers. This is exactly what we see in *B. pallida*, indicating that lower genetic diversity in central and northern Europe is far more likely to result from either historically low genetic diversity in nearby refugia, or loss of genetic diversity associated with range expansion, rather than a *Wolbachia*-associated selective sweep. *Wolbachia* infection may also generate similar patterns of genetic diversity in nuclear and mitochondrial markers, provided that the effect of *Wolbachia* on its host is to induce parthenogenesis. However, the presence of males in every population of *B. pallida* that we collected argues against a scenario of *Wolbachia*-induced parthenogenesis.

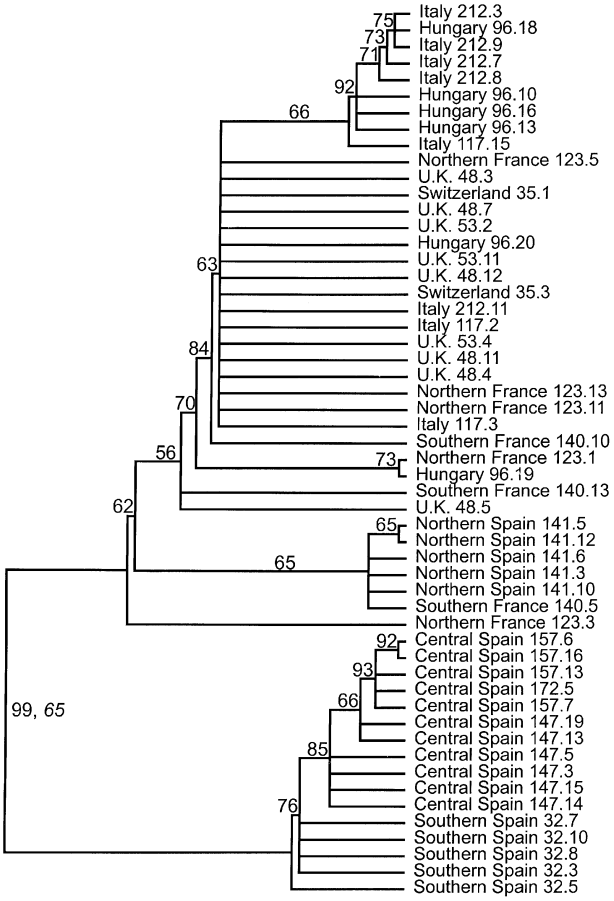

Fig. 3 Consensus tree (50% majority rule) for ITS using MP, with gaps coded as the fifth base. Numbers in branches denote bootstrap support under MP (alignment gaps as fifth base). For southern and central Spain *versus* northern Spain and the rest of Europe split, the bootstrap support given by MP analysis of the indel dataset is shown in ***italics***. Numbers next to countries' names indicate number of individual and number of clone, respectively. The tree is 141 steps long with a consistency index (CI) of 0.794 and a retention index (RI) of 0.925. GenBank accession numbers: ITS1, AF340069–AF340096; ITS2, AF340097–AF340150.

The dominant feature of genetic variation in *B. pallida* is the deep split between the central and southern Spanish population *versus* those of northern Spain and the rest of Europe, supported by all the *B. pallida* datasets. The division of these two groups is robustly supported by the cytochrome *b* (Fig. 2), and less robustly by the ITS (Fig. 3) analyses. The lower support observed in ITS is due to the fact that the informative sites are in the indel regions. When gaps are encoded as the fifth base or when they are treated as unique indel insertions, the split between central and southern Spain *versus* northern Spain and the rest of Europe is more strongly supported. Genetic discontinuities between areas north and south of the Pyrenees are known for many plants and animals, and many show hybrid zones at the Pyrenees (see reviews by Taberlet *et al*., 1998; Hewitt, 1999). For many species, including gallwasps (Stone & Sunnucks, 1993; Stone *et al*., 2001), expansion following retreat of the ice was principally from Italy and the Balkans, while Spanish populations failed to expand far into France. A similar scenario could explain the patterns seen in *B. pallida*. Region-specific haplotypes and nuclear alleles imply the existence of discrete refuge populations in Spain and in regions to the east. Some oak gallwasps show further differentiation between distinct refuges in Italy and the Balkans

Table 4 *Biorhiza pallida* allozyme alleles from *5* allozyme loci, from 4 areas of Europe. Locally restricted alleles are written in bold. Numbers in parentheses denote the number of individuals and number of populations screened, respectively

Area/allozyme	6PGD	GOTm	PEPb	MDHs	GPI
Central and southern Spain (43–10)	**1,2,**3,4	2,4	**1,**2,3,4	1,2,3	3,**4**
Southern France (20–3)	4	4	1,2,3,4	2	3
Italy, Hungary (51–11)	4	2,4	1,2,3,4	2	3
U.K., Eire, Northern France (156–30)	3,4	4	1,2,3,4	2,3,**4**	**2,**3

(Atkinson, 2000), but there is inadequate resolution in the *B. pallida* data to see if the same is true for this species. Range expansion from central or eastern Europe and associated rapid population growth would generate both the similarity among sites and the departure from expectations of neutrality seen outside Spain. Inability of Spanish populations to expand and spread beyond the Pyrenees would lead to absence of Spanish haplotypes or nuclear alleles from more northerly populations, and the absence of any deviation from neutrality. Inability to escape from the Spanish refuge has been clearly demonstrated in one other oak gallwasp, resulting from the evolution of oak-specific ecotypes whose hosts are essentially limited to Spain (Stone *et al*., 2001). It is interesting to note that patterns of post-Pleistocene range expansion for gallwasps and for their host oaks do not match. Molecular evidence for oaks suggests that individuals from all three refugia (Spain, Italy and the Balkans) contributed to the colonization of central Europe (Ferris *et al*., 1993; Dumolin-Lapègue *et al*., 1997). In contrast, gallwasps expanded principally from Italy and the Balkans (see above; Atkinson, 2000; Stone *et al*., 2001). This discordance may be explained by adaptation of gallwasp populations in the Spanish refuge to local endemic oak species that failed to expand further north (for more discussion see Atkinson, 2000; Stone *et al*., 2001).

This phylogeographic hypothesis is supported by dating estimates based upon a *B. pallida* mitochondrial molecular clock, assuming a 2.3% divergence of mtDNA sequences per million years (Brower, 1994). There are two issues for which timing estimates are meaningful. The first concerns the timing of the event leading to the current levels of genetic diversity among northern Spanish and rest of Europe populations. Assuming that the 0–0.9% divergence that is observed among these populations is the result of substitutions that have occurred just after that event, we get an estimation between 0 and 390 000 years ago. This estimation roughly encompasses the range expansion by the oak hosts of *B. pallida* following the end of the Pleistocene (Ferris *et al*., 1993; Dumolin-Lapègue *et al*., 1997). The second issue is the split between central and southern Spain *versus* northern Spain and the rest of Europe. Levels of divergence between the groups of around 2.5% suggest that this split is ancient, pointing to a separation long before the end of the Pleistocene. Such a division into long-standing eastern and western refugia is supported by similar data for other oak gallwasp species, such as *Andricus kollari* and *Andricus quercustozae* (Atkinson, 2000; Stone *et al*., 2001). Our study shows that while patterns of variation in mitochondrial sequence diversity in this system do not allow a *Wolbachia*-induced selective sweep to be discounted, consideration of nuclear marker diversity points to a demographic cause.

Acknowledgements

AR is partially funded by a NERC studentship and RJA holds a BBSRC studentship. GNS and GSB are supported by a grant from the NERC EDGE programme (GR9/03553), with additional support to GNS from the Royal Society (grant no. 574006), and the British Ecological Society (Small Ecological Project Grant, no. 1462). SAW holds a BBSRC fellowship. We especially thank Gil McVean and two anonymous referees for comments on the manuscript. Thanks to Mark Blaxter for unlimited access to Mac computers for running the ML analyses; to David Guiliano for providing DNA from *Wolbachia* strain C to test our WOLG-R primer; Jill Lovell for excellent technical support with the automated sequencing and John Werren for providing the *Nasonia* control strains. Special thanks to the members of the British Plant Gall Society, György Csóka, José-Luis Nieves-Aldrey, Felix Fontal-Cazalla, Gil McVean, Olivier Plantard and Jean-Yves Rasplus for help with gall collecting.

References

ATKINSON, R. J. 2000. *The Genetic Analysis of Natural History, Reproductive Strategy and Population Structure in European Oak Gallwasps (Hymenoptera: Cynipidae)*. PhD Thesis, University of Oxford, Oxford, UK.

AVISE, J. C. 2000. *Phylogeography: the History and Formation of Species*. Harvard University Press, Harvard, USA.

BALLARD, J. W. O., HATZIDAKIS, J., KARR, T. L. AND KREITMAN, M. 1996. Reduced variation in *Drosophila simulans* mitochondrial DNA. *Genetics*, **144**, 1519–1528.

BROWER, A. V. Z. 1994. Rapid morphological radiation and convergence among races of the butterfly *Heliconius erato* inferred from patterns of mitochondrial DNA evolution. *Proc. Natl. Acad. Sci. U.S.A.*, **91**, 6491–6495.

CSÓKA, G. 1997. *Plant Galls*. Agroinform Publishing, Budapest, Hungary.

DONNELLY, P. AND TAVARÉ, S. 1995. Coalescents and genealogical structure under neutrality. *A. Rev. Genet.*, **29**, 401–421.

DUMOLIN-LAPÈGUE, S., DEMESURE, B., FINESCHI, S. L. E., CORRE, V. AND PETIT, R. J. 1997. Phylogeographic structure of white oaks throughout the European continent. *Genetics*, **146**, 1475–1487.

FERRIS, C., OLIVER, R. P., DAVY, A. J. AND HEWITT, G. M. 1993. Native oak chloroplasts reveal an ancient divide across Europe. *Mol. Ecol.*, **2**, 337–344.

FU, Y.-X. AND LI, W.-H. 1993. Statistical tests of neutrality of mutations. *Genetics*, **133**, 693–709.

HASEGAWA, M., KISHINO, M. AND YANO, T. 1985. Dating of the human–ape splitting by a molecular clock of mitochondrial DNA. *J. Mol. Evol.*, **22**, 160174.

HEWITT, G. M. 1999. Post-glacial re-colonization of European biota. *Biol. J. Linn., Society*, **68**, 87–112.

HILLIS, D. M. AND DIXON, M. T. 1991. Ribosomal DNA: molecular evolution and phylogenetic inference. *Q. Rev. Biol.*, **66**, 411–453.

HUELSENBECK, J. P. AND RANNALA, B. 1997. Phylogenetic methods come of age: testing hypotheses in an evolutionary context. *Science*, **276**, 227–232.

HUIGENS, M. E., LUCK, R. F., KLAASSEN, R. H. G., MAAS, M. F. P. M., TIMMERMANS, M. J. T. N. AND STOUTHAMER, R. 2000. Infectious parthenogenesis. *Nature*, **405**, 178–179.

HUNTLEY, B. AND WEBB, T. 1989. Migration – species' response to climatic variations caused by changes in the earth's orbit. *J. Biogeogr.*, **16**, 5–19.

JEYAPRAKASH, A. AND HOY, M. A. 2000. Long PCR improves *Wolbachia* DNA amplification: *wsp* sequences found in 76% of sixty-three arthropod species. *Insect Mol. Biol.*, **9**, 393–406.

JOHNSTONE, R. A. AND HURST, G. D. D. 1996. Maternally inherited male-killing microorganisms may confound interpretation of mitochondrial DNA variability. *Biol. J. Linn. Soc.*, **58**, 453–470.

MALLOCH, G., FENTON, B. AND BUTCHER, B. D. J. 2000. Molecular evidence for multiple infections of a new subgroup of *Wolbachia* in the European raspberry beetle *Byturus tomentosus*. *Mol. Ecol.*, **9**, 77–90.

MAYNARD SMITH, J. AND HAIGH, J. 1974. The hitch-hiking effect of a favorable gene. *Genet. Res.*, **23**, 23–35.

O'NEILL, S. L., HOFFMANN, A. A. AND WERREN, J. H., eds. 1997. *Influential Passengers: Inherited Microorganisms and Arthropod Reproduction*. Oxford University Press, Oxford, UK.

PLANTARD, O., RASPLUS, J. Y., MONDOR, G. L. E., CLAINCHE, I. AND 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. *Proc. R. Soc. Lond. B*, **265**, 1075–1080.

PLANTARD, O., RASPLUS, J.-Y., MONDOR, G. L. E., CLAINCHE, I. AND SOLIGNAC, M. 1999. Distribution and phylogeny of *Wolbachia* inducing thelytoky in Rhoditini and 'Aylacini' (Hymenoptera: Cynipidae). *Insect Mol. Biol.*, **8**, 185–191.

ROZAS, J. AND ROZAS, R. 1999. DnaSP, version 3: an integrated program for molecular population genetics and molecular evolution analysis. *Bioinformatics*, **15**, 174–175.

SOLIGNAC, M., VAUTRIN, D. AND ROUSSET, F. 1994. Widespread occurrence of the proteobacteria *Wolbachia and* partial cytoplasmic incompatibility in *Drosophila melanogaster*. *C. R. Acad. Sci. Paris*, **III** (317), 461–470.

STONE, G. N., ATKINSON, R., ROKAS, A., CSÓKA, G. AND NIEVES-ALDREY, J.-L. 2001. Differential success in northwards range expansion between ecotypes of the marble gallwasp *Andricus kollari:* a tale of two lifecycles. *Mol. Ecol.*, **10**, 1–18.

STONE, G. N. AND COOK, J. M. 1998. The structure of cynipid oak galls: patterns in the evolution of an extended phenotype. *Proc. R. Soc. Lond. B*, **265**, 979–988.

STONE, G. N. AND SUNNUCKS, P. 1993. Genetic consequences of an invasion through a patchy environment – the cynipid gallwasp *Andricus quercuscalicis* (Hymenoptera: Cynipidae). *Mol. Ecol.*, **2**, 251–268.

STOUTHAMER, R., BREEUWER, J. A. J. AND HURST, G. D. D. 1999. *Wolbachia pipientis:* Microbial manipulator of arthropod reproduction. *Ann. Rev. Microbiol.*, **53**, 71–102.

STRIMMER, K. AND VON HAESELER, A. 1996. Quartet puzzling: a quartet maximum-likelihood method for reconstructing tree topologies. *Mol. Biol. Evol.*, **13**, 964–969.

SUNNUCKS, P. AND STONE, G. S. 1996. Genetic structure of invading insects and the case of the knopper gallwasp. In: Floyd, R. B., Sheppard, A. W. and De Barro, P. J., eds. *Frontiers of Population Ecology*, pp. 485–496. CSIRO Publishing, Australia.

SWOFFORD, D. L. 1999. *PAUP*: Phylogenetic Analysis Using Parsimony (and Other Methods)*. Sinauer Associates.

TABERLET, P., FUMAGALLI, L., WUST-SAUCY, A. G. AND COSSON, J. F. 1998. Comparative phylogeography and postglacial colonization routes in Europe. *Mol. Ecol.*, **7**, 453–464.

TARMA, F. 1989. Statistical method for testing the neutral ruination hypothesis by DNA polymorphism. *Genetics*, **123**, 585–595.

THOMPSON, J. D., HIGGINS, D. G. AND GIBSON, T. J. 1994. CLUSTALW: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucl Acids Res.*, **22**, 4673–4680.

TURELLI, M., HOFFMANN, A. A. AND MCKECHNIE, W. S. 1992. Dynamics of cytoplasmic incompatibility and mtDNA variation in natural *Drosophila simulans* populations. *Genetics*, **132**, 713–723.

VAVRE, F., FLEURY, F., LEPETIT, D., FOUILLET, P. AND BOULETREAU, M. 1999. Phylogenctic evidence for horizontal transmission of *Wolbachia* in host parasitoid associations. *Mol. Biol. Evol.*, **16**, 1711–1723.

WERREN, J. H. 1997. Biology of *Wolbachia*. *Ann. Rev. Entomol.*, **42**, 587–609.

WERREN, J. H. AND WINDSOR, D. M. 2000. *Wolbachia* infection frequencies in insects: evidence of a global equilibrium? *Proc. R. Soc. Lond. B.*, **267**, 1277–1285.

WERREN, J. H., WINDSOR, D. AND GUO, L. 1995. Distribution of *Wolbachia* among neotropical arthropods. *Proc. R. Soc. Lond. B*, **262**, 197–204.

WEST, S. A., COOK, J. M., WERREN, J. H. AND GODFRAY, C. J. 1998. *Wolbachia* in two insect host–parasitoid communities. *Mol. Ecol.*, **7**, 1457–1465.

WHITE, T. L., BRUNS, T., LEE, S. AND TAYLOR, J. 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: Innis, M. A. Gelfand, D. H. and Sninsky, J. J., eds. PCR. *Protocols. A Guide to Methods and Applications*, pp. 315–322. Academic Press, Inc., USA.

YANG, Z. 1993. Maximum-likelihood estimation of phylogeny from DNA sequences when substitution rates differ over sites. *Mol. Biol. Evol.*, **10**, 1396–1401.

YANG, Z., GOLDMAN, N. AND FRIDAY, A. 1994. Comparison of models for nucleotide substitution used in maximum-likelihood phylogenetic estimation. *Mol. Biol. Evol*, **11**, 316–324.

ZHOU, W., ROUSSET, F. AND O'NEILL, S. 1998. Phylogeny and PCR-based classification of *Wolbachia* strains using *wsp* gene sequences. *Proc. R. Soc. Lond. B*, **265**, 509–515.