Bulletin of Entomological Research (2009) **99**, 583–591
doi:10.1017/S0007485309006622

First published online 19 February 2009

Variations in the mitochondrial cytochrome c oxidase subunit I gene indicate northward expanding populations of *Culicoides imicola* in Spain

J.H. Calvo[1]*, C. Calvete[2], A. Martinez-Royo[1], R. Estrada[3], M.A. Miranda[4], D. Borras[4], V. Sarto I Monteys[5], N. Pages[5], J.A. Delgado[6], F. Collantes[6] and J. Lucientes[3]

[1]Unidad de Tecnología en Producción animal, CITA-ARAID, Zaragoza, Spain: [2]Unidad Sanidad Animal, CITA, Zaragoza. Spain: [3]Departamento de Patología Animal, Universidad de Zaragoza, Zaragoza, Spain: [4]Laboratorio de Zoología, Universidad de las Islas Baleares, Palma de Mallorca, Spain: [5]Unidad de Entomología. Fundación CReSA, UAB Bellaterra, Spain: [6]Departamento de Zoología y Antropología Física, Universidad de Murcia, Murcia, Spain

Abstract

Culicoides imicola is the main vector for bluetongue (BT) and African horse sickness (AHS) viruses in the Mediterranean basin and in southern Europe. In this study, we analysed partial mitochondrial cytochrome c oxidase subunit I (*COI*) gene to characterize and confirm population expansion of *Culicoides imicola* across Spain. The data were analysed at two hierarchical levels to test the relationship between *C. imicola* haplotypes in Spain ($n = 215$ from 58 different locations) and worldwide ($n = 277$). We found nineteen different haplotypes within the Spanish population, including 11 new haplotypes. No matrilineal subdivision was found within the Spanish population, while western and eastern Mediterranean *C. imicola* populations were very structured. These findings were further supported by median networks and mismatch haplotype distributions. Median networks demonstrated that the haplotypes we observed in the western Mediterranean region were closely related with one another, creating a clear star-like phylogeny separated only by a single mutation from eastern haplotypes. The two, genetically distinct, sources of *C. imicola* in the Mediterranean basin, thus, were confirmed. This type of star-like population structure centred around the most frequent haplotype is best explained by rapid expansion. Furthermore, the proposed northern expansion was also supported by the statistically negative Tajima's D and Fu's Fs values, as well as predicted mismatch distributions of sudden and spatially expanding populations. Our results thus indicated that *C. imicola* population expansion was a rapid and recent phenomenon.

*Author for correspondence
Fax: +34 976716335
E-mail: jhcalvo@aragon.es

Keywords: *Culicoides imicola*, African horse sickness, bluetongue, cytochrome c oxidase I, population expansion, mtDNA

(Accepted 14 October 2008)

Introduction

Bluetongue (BT) and African horse sickness (AHS) are two non-contagious, arthropod-borne viral diseases affecting ruminants and horses, respectively. Bluetongue disease is caused by the bluetongue virus (BTV), the prototype of the *Orbivirus* genus within the Reoviridae family, which also includes AHSV (Holmes *et al*., 1995). Both viruses are transmitted by species of biting midges belonging to the *Culicoides* genus (Diptera: Ceratopogonidae) (Du Toit, 1944; Mellor & Pitzolis, 1979; Mellor *et al*., 1990) and are maintained naturally through a series of alternative cycles of replication between *Culicoides* vectors and susceptible hosts (Takamatsu *et al*., 2003).

In the Mediterranean basin and in southern Europe, the main vector for both viruses is *Culicoides imicola* (Mellor, 1996) although other Paleartic *Culicoides* species, including *C. obsoletus*, *C. scoticus*, *C dewulfi* and *C. chiopterus*, can also serve as main BTV vectors (De Liberato *et al*., 2005; Mehlhorn *et al*., 2007; Meiswinkel *et al*., 2007, 2008; Calvete *et al*., 2008). In northern Europe, species of *Culicoides obsoletus* and *scoticus*, and the Avaritia subgenus (*Culicoides dewulfi* and chiopterus), are thought to be involved in bluetongue pothology.

In Europe, several BTV episodes during the past few decades have been reposted, with the first of these occurring between 1956 and 1960 in the southwestern region of the Iberian Peninsula (Manso-Ribeiro *et al*., 1957). Since then, sporadic occurrences have been documented across Mediterranean Europe (Mellor & Pitzolis, 1979; Jennings *et al*., 1983), with a large BT epizootic in 1998 expanding northward across Europe. During this epidemic, several BTV serotypes were postulated to affect many European countries, including several that had not previously been affected by the virus (Mellor & Wittmann, 2002; Mehlhorn *et al*., 2007).

Between 1956 and 1990, one BT and two AHS epizootics broke out in Spain, concentrating mainly in Andalusia (southern Spain) (Mellor & Pitzolis, 1979; Jennings *et al*., 1983; Ortega *et al*., 1999); and, more recently, a BT epizootic in 2000 occurred in the Balearic Islands of Majorca and Menorca (Office International des Epizooties, 2000). Subsequent trapping revealed that *C. imicola* was the most likely vector involved, with *Culicoides obsoletus* and *scoticus* also abundantly present on these two islands (Miranda *et al*., 2003). *C. imicola* was first detected in 2000 on the Mediterranean Islands of Sardinia, Sicily and Corsica, located east of the Balearic Islands, as well as in the Italian mainland regions of Calabria and Tuscany (Wittmann *et al*., 2001). BTV-4 outbreaks were reported in Menorca in 2003 and in Majorca by 2004, which then subsequently spread during the following year to other southern and mainland municipalities. In 2007, another BTV-1 outbreak emerged on the southern coast of the Iberian Peninsula, most likely transported by a viraemic animal or infected *Culicoides* specimen from northern Africa. The distribution of BTV-1 spread was similar to that of *C. imicola* but also spread to more northern areas where this species is absent (http://www.oie.int).

This northeastward expansion of the BT virus across Europe appears to be partially linked to the recent trend of *C. imicola* across Mediterranean Europe. Extensive surveys estimating current *C. imicola* distributions have been carried out in Portugal, Italy and Spain (Calistri *et al*., 2003; Capela *et al*., 2003; Calvete *et al*., 2006). Sarto I Monteys *et al*. (2005) have suggested that *C. imicola* is currently expanding into the Catalonian region (northeastern Spain). Furthermore, Calvete *et al*. (2006) detected a *C. imicola* presence throughout Spain, but the considerable place-to-place and year-to-year variation of *C. imicola* populations (Rawlings *et al*., 1997), in addition to the lack of prior extensive surveys, does not allow for an accurate assessment of whether this species is indeed expanding into the remaining areas of Spain.

Understanding population genetic structures of insect vectors has important implications for the spatial scales over which vector-borne diseases spread (Tabachnick & Black, 1995). Dipteran disease vectors differ enormously in the degree of matrilineal subdivision. *C. imicola* carried by midges from Portugal, Rhodes, Israel and South Africa belong to the same phylogenetic clade of mitochondrial cytochrome c oxidase subunit I (*COI*) and are phylogenetically distinct from the four other species of the *C. imicola* species-complex (Dallas *et al*., 2003). This suggests that the South African *C. imicola* is phylogenetically distinct (Linton *et al*., 2002) and that the level of the *C. imicola* matrilineal subdivision between the eastern and western ends of the Mediterranean basin is remarkably high, and is comparable to the geographical structure of BTV serotypes observed in outbreaks around the Mediterranean between 1998–2001 (Dallas *et al*., 2003). Recent phylogenetic studies of the *Culicoides* species in France and Italy have been performed based on nuclear ITS1-rDNA and ITS2-rDNA sequences (Perrin *et al*., 2006; Gomulski *et al*., 2005, 2006; Nolan *et al*., 2007).

The aim of the present study was to characterize the *C. imicola* population in Spain and its relationship with other *C. imicola* populations in other areas. By comparing differences in *mitochondrial* cytochrome c oxidase subunit I (*COI*) gene, we suggest a hypothesis regarding current *C. imicola* expansion in Spain.

Materials and methods

Insect samples

Trapping was carried out on farms using a 4 W ultraviolet light trap fitted with a suction fan (Miniature Blacklight Model 1212, John Hock Company, Gainesville, FL, USA). Each trap was placed outside selected farms within 30 m of livestock and 1.7–2.5 m from ground level. All collected insects were transported to the laboratory and preserved in 70% ethanol. Upon examination, species of *Culicoides* were separated from other insects, and identification of *C. imicola* was achieved based on previously described wing patterns (Rawlings, 1996). Between summer of 2005 and summer of 2006, 215 insects were randomly selected from 58 sites in Spain (fig. 1, table 1). Insects from

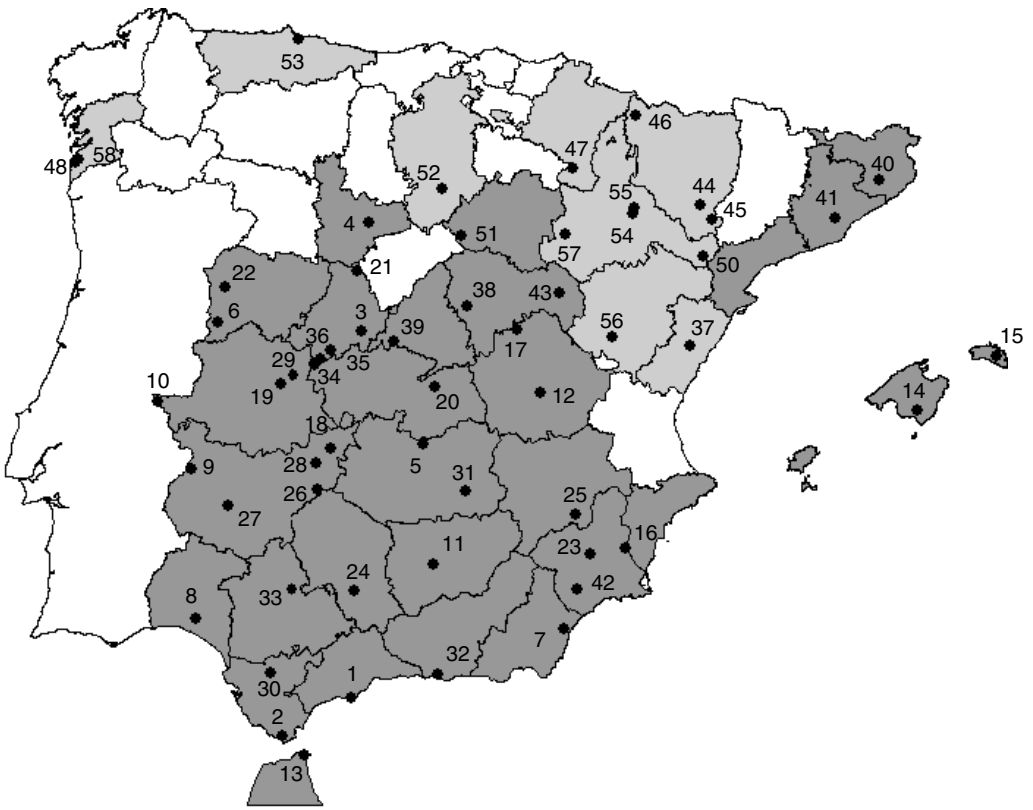

Fig. 1. Location of farms sampled under the Spanish National Surveillance Programme for the detection of *Culicoides imicola* in the present study. Dark grey: provinces in which *C. imicola* had been previously reported; light grey: provinces in which *C. imicola* was reported for the first time. Population abbreviations are as follows: Mijas (01); Tarifa (02); Navaluenga (03); Tudela de Duero (04); Fuente del Fresno (05); Ituero de Azaba (06); Vera (07); Moguer (08); Badajoz (09); Cedilla (10); Begijar (11); Valera de Abajo (12); Ceuta (13); Mallorca (14); Sant Lluis (15); Orihuela (16); Villar del Infantado (17); Herrera del Duque (18); Serrejón (19); Huerta de Valdecarabanos (20); Arévalo (21); Banobarez (22); Mula (23); Montemayor (24); Hellín (25); Cabeza de Buey (26); Fuente del Maestre (27); La Puebla de Alcocer (28); Navalmoral de la Mata (29); Arcos de la Frontera (30); Alcubillas (31); Salobrena (32); Lora del Río (33); Calzada de Oropesa (34); Arenas de San Pedro (35); Candeleda (36); Ibarzos (37); Guadalajara (38); Aldea del Fresno (39); Caldes de Malavella (40); Piera (41); Lorca (42); Molina de Aragón (43); Albalate de Cinca (44); Fraga (45); Hecho (46); Castejón (47); Gándara (48); Estivada (49); Fabara (50); Torremocha de Ayllón (51); Tejada (52); Asturias (53); Zaragoza (54); Montanana (55); San Blas (56); Villarroya (57); and Iglesia (58).

regions of the most recent BT outbreaks and where *C. imicola* had been previously found were sampled along with samples from sites where *C. imicola* had been absent (sites 47–58). In places with four insects or less, this number was the maximum number of insects trapped. Sample sites in Spain were approximately 6–1100 km apart, with the pairwise distance between the centre of the sampling area approximately 290 km.

DNA extraction, polymerase chain reaction (PCR) amplification and DNA sequencing

Total DNA was extracted from single midges using a commercial kit (Nucleo Spin Tissue, Mancherey-Nagel, Düren, Germany). The mitochondrial *COI* gene was amplified using the following primers: 5′-*GGAGGATTTGGAAATTGATTAGT*-3′ and 5′- *CAGGTAAAATTAAAATATAAACTTCTGG*-3′ (Dallas *et al*., 2003). PCRs were carried out in a final volume of 25 µl containing 50 ng of total DNA, 5 pmol of each primer, 200 µM dNTPs, 2 mM $MgCl_2$, 50 mM KCl, 10 mM Tris-HCl, 0.1% Triton X-100 and 0.5 U *Taq* polymerase (*Taq* polymerase, Biotools, Spain). Thirty cycles were performed with the following step-cycle program: strand denaturation at 94°C for 30 s, primer annealing at 54°C for 30 s, and primer extension at 72°C for 30 s. A 523 bp fragment of the *COI* gene was amplified from DNA extracts of individual midges (1–10 midges from each site), and PCR products were directly sequenced using the primers described above. All amplified products were sequenced from both strands.

Analytical methodologies

Sequence traces of each PCR fragment were edited in BioEdit v7.1.0 (Isis Pharmaceuticals, Inc, USA) to retain the maximum length sequence reads in both forward and reverse directions. Data were aligned without gaps and inspected using MEGA 2.0 (Kumar *et al*., 2001) to discriminate between singletons (polymorphisms that appeared in only one insect) and parsimony informative sites. *COI* reading frames were detected by alignment with other *Culicoides* *COI* sequences (table 2). Indices of sequence variation, and

https://doi.org/10.1017/S0007485309006622 Published online by Cambridge University Press

Table 1. Haplotype sequences and distribution of Spanish *Culicoides imicola* populations. IMICOI indicate the *COI C. imicola* haplotype. IMICOI 24-IMICOI 34 represent new haplotypes described in this study. Vertical numbers indicate the SNP position relative to the DQ868882 reference sequence. Population abbreviations are as follows: Mijas (01); Tarifa (02); Navaluenga (03); Tudela de Duero (04); Fuente del Fresno (05); Ituero de Azaba (06); Vera (07); Moguer (08); Badajoz (09); Cedilla (10); Begijar (11); Valera de Abajo (12); Ceuta (13); Mallorca (14); Sant Lluis (15); Orihuela (16); Villar del Infantado (17); Herrera del Duque (18); Serrejón (19); Huerta de Valdecarabanos (20); Arévalo (21); Banobarez (22); Mula (23); Montemayor (24); Hellín (25); Cabeza de Buey (26); Fuente del Maestre (27); La Puebla de Alcocer (28); Navalmoral de la Mata (29); Arcos de la Frontera (30); Alcubillas (31); Salobrena (32); Lora del Río (33); Calzada de Oropesa (34); Arenas de San Pedro (35); Candeleda (36); Ibarzos (37); Guadalajara (38); Aldea del Fresno (39); Caldes de Malavella (40); Piera (41); Lorca (42); Molina de Aragón (43); Albalate de Cinca (44); Fraga (45); Hecho (46); Castejón (47); Gándara (48); Estivada (49); Fabara (50); Torremocha de Ayllón (51); Tejada (52); Asturias (53); Zaragoza (54); Montanana (55); San Blas (56); Villarroya (57); and Iglesia (58).

			561113334444444	POPULATION																												
			392370380011255	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29
			5695891703825																													
IMICOI	1	GCTCCTGCAAGACGC		2	5	4	3	3	4	5	4	3	3	9	3	3	10	5	2	1		2	2	2	3	3	4	4	4	3	3	4
IMICOI	14	ACTCCTGCAAGACGC														1																
IMICOI	15	GCCCCTGCAAAACGC				1										2																
IMICOI	16	GCTCCCGCAAGACGC		1											1					1							1					1
IMICOI	17	GCTCCTACAAGACGC			2			1																								
IMICOI	18	GCTCCTGCAAAACGC																														
IMICOI	19	GCTCCTGCAAGGCGC					1					1	1								1											
IMICOI	20	GCTCTTGCAAGACGC											1																			
IMICOI	24	GCTCCTGTAAGACGC		1																												
IMICOI	25	GCTCCTGCAGGGCGC		1																												
IMICOI	26	GCTCCTACAAGATGC			1							1																			1	
IMICOI	27	GCTCCTGCAAGACAC					1									1														2	1	
IMICOI	28	GCTCCTGCAAGACGT						1	1						1																	
IMICOI	29	GCTTCTGCAAGACGC									1																					
IMICOI	30	GCTCCTATAAGACGC														1																
IMICOI	31	GCACCTGCAAGACGC														1																
IMICOI	32	GCTCCTGCGAGACGC																											1			
IMICOI	33	ACTCCTGCAAGACGT																														
IMICOI	34	GATCCTGCAAGGCGC																														
				5	8	5	5	5	5	5	5	5	5	9	5	9	10	5	2	2	1	2	2	2	3	5	4	5	5	5	5	5

			561113334444444	POPULATION																												
			392370380011255	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58
			5695891703825																													
IMICOI	1	GCTCCTGCAAGACGC		2	4	4	4	4	4	4	3	3	4	4	5	4	1	1	1	1	1			1	1		1	1	1	1	1	1
IMICOI	14	ACTCCTGCAAGACGC																														
IMICOI	15	GCCCCTGCAAAACGC				1										1										1						
IMICOI	16	GCTCCCGCAAGACGC						1			1	1										1										
IMICOI	17	GCTCCTACAAGACGC																														
IMICOI	18	GCTCCTGCAAAACGC							1																							
IMICOI	19	GCTCCTGCAAGGCGC		1			1																									
IMICOI	20	GCTCTTGCAAGACGC																														
IMICOI	24	GCTCCTGTAAGACGC																														
IMICOI	25	GCTCCTGCAGGGCGC																														
IMICOI	26	GCTCCTACAAGATGC																														
IMICOI	27	GCTCCTGCAAGACAC			1																						1					
IMICOI	28	GCTCCTGCAAGACGT									1	1																				
IMICOI	29	GCTTCTGCAAGACGC																														
IMICOI	30	GCTCCTATAAGACGC																														
IMICOI	31	GCACCTGCAAGACGC																														
IMICOI	32	GCTCCTGCGAGACGC												1																		
IMICOI	33	ACTCCTGCAAGACGT									1																					
IMICOI	34	GATCCTGCAAGGCGC																					1									
				3	5	5	5	5	5	4	5	5	5	5	5	5	1	1	1	1	1	1	1	1	1	1	2	1	1	1	1	1

haplotype structure and neutrality tests were then calculated using DnaSP 4.0 (Rozas *et al*., 2003). Tajima's D (Tajima, 1989) tests the difference of θ_π and θ_w, both diversity estimators of the neutral parameter 4 Neμ. θ_π is the average heterozygosity per nucleotide site obtained from the mean number of pairwise differences among sample sequences (Nei & Li, 1979), and θ_w is estimated from the expected nucleotide diversity and number of segregating sites (Watterson, 1975). If the population sample fit the infinite sites model, θ_π and θ_w would have the same expected value and $D = 0$. Under non-neutral evolution, since θ_w is calculated only using the number of segregating sites, it is more affected than θ_π by low-frequency variants calculated using the frequency of each variant. Therefore, an excess of singletons (as expected in cases of recent selective sweeps or population expansion) will produce negative Tajima's D values (Braverman *et al*., 1995; Aris-Brosou & Excoffier, 1996). Fu's Fs tests whether the excess of rare alleles is significant in relation to the expected value under a mutation-drift model. The D* statistic is based on the difference

https://doi.org/10.1017/S0007485309006622 Published online by Cambridge University Press

Table 2. World *Culicoides imicola* haplotypes (IMICOI) and their distribution within the Mediterranean basin and in Europe. In this study, Spain 2006 refers to the *Culicoides imicola* survey. GenBank references are as follows: Spain 2006 haplotypes: DQ868882, DQ868883, DQ868886-DQ868890, DQ868892, DQ868894, DQ871030; Spain[1]: AJ867227, AJ867228; Italy[1]: AJ877224, AJ877225, AJ867232-AJ867234; Morocco[1]: AJ867226; Portugal[2]: AJ549415-AJ549426; Portugal[1]: AJ867229-AJ867231; Rhodes[2]: AJ549388-AJ549392; Israel[1]: AF078097; Israel[2]: AJ549393-AJ549414, AF080531, AF080532, AF080534, AF080535, AF078098-AF078100, South Africa[3]: AF069231-AF069233, AF069249.

IMICOI	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34
Localization																																		
Spain 2006 ($n=215$)	165													1	6	9	3	1	6	1				1	1	6	7	5	1	1	1	2	1	1
Spain[1] ($n=2$)																1	1																	
Italy[1] ($n=5$)													1	1							1	1	1											
Morocco[1] ($n=1$)															1																			
Portugal[2] ($n=12$)	12																																	
Portugal[1] ($n=3$)																		1	1	1														
Rhodes[2] ($n=5$)		3	2																															
Israel[1] ($n=1$)												1																						
Israel[2] ($n=29$)			2	21	2	1	1	1	1																									
SouthAfrica[3] ($n=4$)										3	1																							

[1] Haplotypes corresponding to GenBank sequences.
[2] Haplotypes corresponding to sequences isolated in Dallas *et al*. (2003).
[3] Haplotypes corresponding to sequences isolated in Linton *et al*. (2002).

between the number of singletons and total number of segregating positions, while the F* statistic is based on the difference between the number of singletons and θ_π. Both the D* and F* statistics test whether mutations are selectively neutral. The mismatch distribution was calculated using Arlequin (http://anthropologie.unige.ch/arlequin/) and represents the number of differences that exist between every pairwise combination of haplotypes plotted as a function of frequency. Mismatch distribution and the raggedness index of the observed distribution defined by Harpending (1994) were also calculated to estimate recent demographic (Rogers & Harpending, 1992; Schneider & Excoffier, 1999; Excoffier, 2004) and spatial expansions (Ray *et al*., 2003; Excoffier, 2004). Significance levels of all test statistics were evaluated by randomization tests as described in Arlequin. Median networks using Network version 4.1.1.2 (http://www.fluxusengineering.com) were constructed to investigate the relationship between haplotypes. A minimum spanning network was built based on the absolute number of nucleotide differences. Data were analysed at two different hierarchical levels to test the relationship between Spanish *C. imicola* (entire dataset isolated in this study) and worldwide *C. imicola* haplotypes (entire dataset isolated in this study plus those deposited in GenBank from other countries and locations) (table 2). Non-Spanish samples were Genbank sequences obtained from Italy, Morocco, Portugal, Rhodes, Israel and South Africa.

Results

COI *sequences of* Culicoides imicola

A 455 bp mitochondrial sequence was collected from each of 215 *C. imicola* samples from 58 sites in Spain (fig. 1, table 1). The Spanish *C. Imicola* population contained 19 haplotypes defined by 15 polymorphic sites (four polymorphic sites were singletons), including two non-synonymous substitutions at positions 69 and 136 and three nucleotide variants in one polymorphic site at position 125 (SNP (Single Nucleotide polymorphim) position is relative to the DQ868882 reference sequence). Considering worldwide *C. imicola* populations, 34 haplotypes were defined by 27 mutations, including five non-synonymous substitutions at positions 46, 69, 115, 136 and 253. A total of 11 previously undescribed haplotypes were found among the Spanish *C. imicola* populations (GenBank Accession Nos. DQ868882, DQ868883, DQ868886-DQ868890, DQ868892, DQ868894 and DQ871030). The 11 phylogenetically informative sites were used to calculate nucleotide diversity ($\pi = 0.00132 \pm 0.17 \times 10^{-3}$) and haplotype diversity ($Hd = 0.428 \pm 0.043$) within the Spanish population. When separately considering each sampling location, the maximum nucleotide diversity and haplotype diversity were located in Ceuta (13) ($\pi = 0.00405$, $Hd = 0.888$), Tarifa (02) ($\pi = 0.00299$, $Hd = 0.607$) and Mijas (01) ($\pi = 0.00352$, $Hd = 0.900$) and the minimum only where the IMICO 1 haplotype was found (table 1). The 21 phylogenetically informative sites (six were singletons) in the worldwide *C. imicola* populations were used to calculate nucleotide ($\pi = 0.00344 \pm 0.14 \times 10^{-3}$) and haplotype diversity ($Hd = 0.596 \pm 0.035$).

Population characteristics of Culicoides imicola *within the Mediterranean basin*

To investigate haplotype structure, differences were calculated between each pairwise combination of the 19 Spanish and 34 worldwide haplotypes. The distribution of

https://doi.org/10.1017/S0007485309006622 Published online by Cambridge University Press

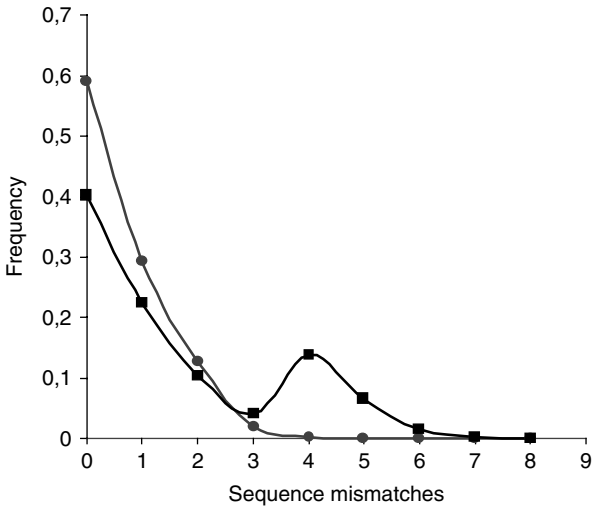

Fig. 2. Frequency distribution of the number of sequence mismatches between pairwise combinations of *Culicoides imicola* mtDNA haplotypes in Spain and worldwide. Distinct peaks indicate two haplotypes groups (–●–, Spain; –■–, world).

these differences indicated one peak in Spainish *C. imicola* haplotypes while two peaks were found worldwide, indicating the presence of divergent haplotype groups (fig. 2). This finding was supported by the construction of median networks (fig. 3), containing one median vector (mv1), which was missing one haplotype between western and eastern Mediterranean populations. Mv1 represents possible unsampled sequences or extinct ancestral sequences. It is clear when studying only the torso of the network, that haplotypes were also grouped in western Mediterranean (IMICOI 1, 14, 17, 24, 28, 31, 33) and eastern Mediterranean (IMICOI 2, 5, 6, 9, 13), showing in the latter two groups: Central Mediterranean, with samples of Rhodes, Sicily and one of Israel, and eastern Mediterranean. The last median vector (mv2) is a haplotype that connects this group to the second one, formed from Israel haplotypes. Fifteen haplotypes radiated from the western torso of the haplotype network, creating a clear star-like phylogeny, and samples from South Africa were situated outside of the torso of the network.

Neutrality tests and mismatch distributions

Neutrality tests were applied to determine whether neutrality could be discarded and to verify that the Spanish population was indeed expanding. Segregating sites displayed an irregular frequency distribution pattern, which represented an excess of variants that are both non- and highly polymorphic. Fu's Fs was -21.57196 ($P < 0.001$), indicating a clear deviation from the expected allele frequency spectrum in the case of an excess of rare alleles. Nevertheless, the observed excess of rare alleles does not imply that *COI* variation is high. In fact, nucleotide variation was markedly reduced in our study. Despite the low number of segregating sites due to the excess of low frequency variants, the number of alleles was actually higher than expected. Tajima's D statistic was -1.99 (significant, $P < 0.05$), while the D* and F* statistics were -1.353 (not significant, $P > 0.10$) and -1.921 (not significant, $0.10 > P > 0.05$), respectively. Negative values of neutrality estimators indicate an excess of low frequency variants. Fu's Fs was statistically significant, providing evidence for an excess of rare alleles produced by recent mutations. Tajima's D and Fu's Fs could indicate a recent population expansion, as there is an excess of low-frequency haplotypes with all new and rare mutations. Since Fu and Li's and Tajima's tests combine non-synonymous and synonymous mutations, it is impossible to separate effects of heterogeneous (purifying selection) and homogeneous (population expansion) processes. An excess of low-frequency haplotypes is expected under purifying selection, as new mutations with deleterious effects are removed (Hahn *et al*., 2002). Pairwise mismatch distributions were calculated to distinguish whether Tajima's and Fu's Fs parameters could indicate a population expansion. Pairwise mismatch distributions were unimodal for Spanish *COI* sequences within Spanish *C. imicola* populations, as expected after a sudden increase in population size. The distribution of mtDNA did not differ significantly from those expected from a sudden expansion model (sum of squared deviations under the least-squares approach = 0.00038, $P = 0.650$; Harpending's raggedness index = 0.12, $P = 0.800$) (fig. 4). In the same way, the haplotype distribution did not differ significantly from those expected from a spatial expansion model (sum of squared deviations under the least-squares approach = 0.00038, $P = 0.800$; Harpending's raggedness index = 0.12, $P = 0.700$) (fig. 4). Harpending's raggedness index was very low, indicating that the observed distribution was unimodal. This index takes larger values for multimodal distributions than unimodal or smoother distributions commonly found in a stationary population. Thus, analysis of mismatch distributions was consistent with the hypothesis that the sampled Spanish *C. imicola* population had undergone a recent expansion.

Discussion

In the present work, we characterized a population of *C. imicola* in Spain to detect putative expansion across Spain by comparing differences in the mitochondrial cytochrome c oxidase subunit I (*COI*) gene. In the Spanish *C. imicola* population, a large number of rare haplotypes separated by single nucleotide differences were revealed. The distribution patterns of DNA variation, genotypes and haplotypes represent the first evidence of rapid *C. imicola* population expansion in recent times.

The Mijas (01), Tarifa (02) and Ceuta (13) sampling sites exhibit the maximum haplotype and nucleotide diversities. Ceuta is a north African Spanish town on the border of Morocco, while the other two are located along the southern coast of the Iberian Peninsula near Morocco. This could indicate that these places are a point of incursion of *C. imicola* in the Iberian Peninsula, as haplotypes described in Morocco were also found in these locations. This is in accord with the points of the BTV-4 outbreak in 2004 (Gomez-Tejedor, 2004) and the BTV-1 outbreak in 2007.

Mismatch distributions of various haplotypes were consistent with results from Dallas *et al*. (2003), who demonstrated that *COI* haplotypes in *C. imicola* were remarkably high in the eastern and western Mediterranean basin, concordant with the geographical structure of BTV serotypes in outbreaks reported around the area. Two peaks were identified, indicating the presence of divergent haplotypes groups (fig. 2). One peak contained no mismatches and the

https://doi.org/10.1017/S0007485309006622 Published online by Cambridge University Press

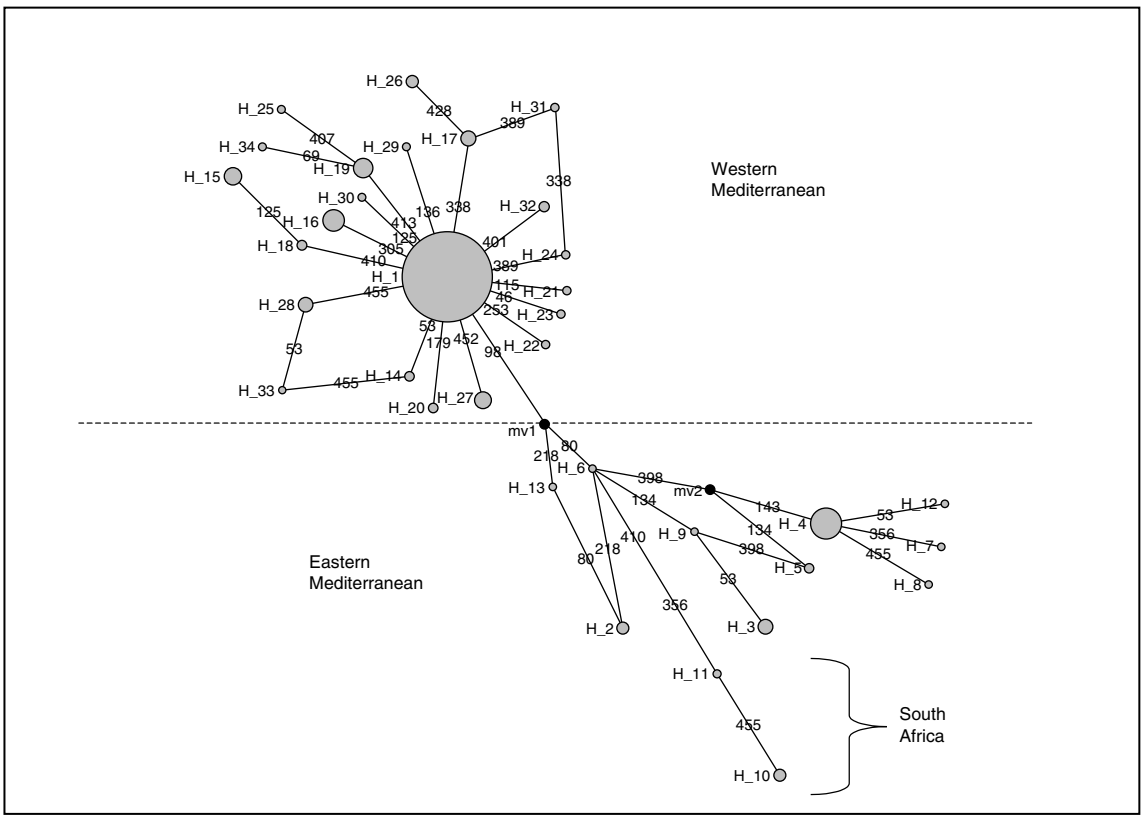

Fig. 3. Median joining network diagrams show the approximate location of the two major geographic regions from which *Culicoides imicola* haplotypes were isolated. The diagram illustrates both the relationship between haplotypes and their observed frequencies. Haplotype number corresponds to IMICOI haplotype. The diameter of H_6 (IMICOI 6) represents one animal. Mutated nucleotides position (relative to the reference sequence) is indicated with numbers.

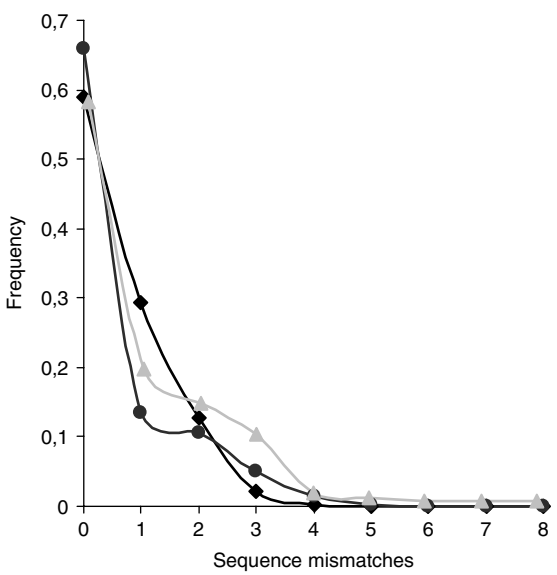

Fig. 4. Frequency distribution of sequence mismatches between pairwise combinations of observed and estimated Spainish *Culicoides imicola* mtDNA haplotypes under the sudden and spatial population expansion models (–◆–, observed; –●–, sudden expansion model; —▲—, spatial expansion model).

second had four. The first peak corresponded to haplotypes described in western Mediterranean while the second one corresponded to eastern Mediterranean haplotypes, indicating two genetically distinct sources better than genetic drift associated with dependence on fragmented area. This was confirmed by median joining networks that showed one median vector, 'mv1', (one missing haplotype) between western and eastern populations. The median network (fig. 3) also demonstrated that the haplotypes observed in the western Mediterranean region were closely related with one another, creating a clear star-like phylogeny centred around the most frequent haplotype. This type of population structure is best explained by rapid population expansion after a bottleneck, when molecular markers may show quite uniform structures over large areas. In such cases, the haplotype and nucleotide diversities can be low, even if the population *de facto* has some breeding structure (Walker *et al*., 1998). Another explanation for the observed star-like population structure is a selective sweep, where one genotype is favoured by natural selection and has recently replaced all others (Rich *et al*., 1998). Median joining networks confirmed previous data obtained by Dallas *et al*. (2003), who suggested that gene flow between western and eastern areas in the Mediterranean area was restricted to the central Mediterranean region because insects isolated in Sicilia grouped with western and eastern specimens. The postulated northern expansion was further supported by the

https://doi.org/10.1017/S0007485309006622 Published online by Cambridge University Press

statistically significantly negative Tajima's D and Fu's Fs values, as well as expected mismatch distributions of expanding populations (fig. 4). Furthermore, D* and F* were negative, which again reflects population growth with no background selection. Samples drawn from populations at demographic equilibrium usually exhibit a multimodal distribution, as it reflects the highly stochastic shape of gene trees, but the distribution becomes unimodal in populations that have passed through a recent demographic expansion (Slatkin & Hudson, 1991; Rogers & Harpending, 1992) or range expansion with high migration rates between neighbouring populations (Ray *et al*., 2003; Excoffier, 2004). A population generally undergoes spatial expansion if the population range is initially restricted to a very small area, then increases over time and space. Based on simulations from Ray *et al*. (2003), a large spatial expansion in a mismatch distribution often produces the same signal as for a purely demographic expansion in a panmictic population, but only if neighbouring subpopulations exchange many migrants (50 or more). These results, in conjunction with the presence of one common haplotype in almost all sample populations, support the recent spatial and demographic population expansion of *C. imicola* into northern regions (fig. 1). When western Mediterranean and worldwide *imicola* populations were combined, significant *C. imicola* expansion was similary found. However, no significance was seen when only eastern haplotypes were considered, possibly due to small sample sizes for sites in the eastern Mediterranean region. We would like to note, however, that results from worldwide *C. imicola* analysis should be interpreted with caution, as non-Spanish insects samples were collected in different years and, thus, could produce a bias in genetic analysis.

Previous data, as well as new data from this study, demonstrate that *C. imicola* is indeed expanding northwards. Expansion was limited to the southwest region of the Iberian Peninsula prior to 2000 and, then, was observed in more northern areas, such as Catalonia, the Balearic Islands and, recently, in the Spanish Basque country (Miranda *et al*., 2003; Sarto I Monteys *et al*., 2005; Calvete *et al*., 2006, 2008). Our results imply that *C. imicola* expansion was a rapid and recent phenomenon. Ecosystem modifications depend on complex relationships established between the environment and the biology and physiology of the vector, virus and vertebrate hosts, providing a possible reason for the rapid expansion of *C. imicola*. Future research is necessary to assess the potential for other *Culicoides* species to act as vectors, such as the *Culicoides pulicaris*, and the putative competency of *C. imicola*, *C. obsoletus group* (*C. obsoletus* and *scoticus*), as well as species of the Avaritia subgenus (*C. dewulfi* and chiopterus).

Acknowledgements

The authors thank the veterinary services of regional governments for their efforts in carrying out the samplings. The national survey was supported by the Spanish Ministry of Agriculture, Fisheries and Food.

References

Aris-Brosou, S. & Excoffier, L. (1996) The impact of population expansion and mutation rate heterogeneity on DNA sequence polymorphism. *Molecular Biology Evolution* **13**, 494–504.

Braverman, J.M., Hudson, R.R., Kaplan, N.L., Langley, C.H. & Stephan, W. (1995) The hitchhiking effect on the site frequency-spectrum of DNA polymorphisms. *Genetics* **140** (2), 783–796.

Calistri, P., Goffredo, M., Caporale, V. & Meiswinkel, R. (2003) The distribution of *Culicoides imicola* in Italy: application and evaluation of current Meditteranean models based on climate. *Journal of Veterinary Medicine B* **50**, 132–138.

Calvete, C., Miranda, M.A., Estrada, R., Borras, D., Sarto, I., Monteys, V., Collantes, F., Garcia-de-Francisco, J.M., Moreno, N. & Lucientes, J. (2006) Spatial distribution of *Culicoides imicola*, the main vector of bluetongue virus, in Spain. *Veterinary Record* **158**(4), 130–131.

Calvete, C., Calvo, J.H., Calavia, R., Miranda, M.A., Borrás, D., Estrada, R., Lucientes, J., Mañuz, B. & Romero, L. (2008) *Culicoides* species and transmission of bluetongue virus in Spain. *Veterinary Record* **162**, 255.

Capela, R., Purse, B.V., Pena, I., Wittmann, F.J., Margarita, Y., Capela, M., Romao, L., Mellor, P.S. & Baylis, M. (2003) Spatial distribution of *Culicoides* species in Portugal in relation to the transmisión of African horse sickness and bluetongue viruses. *Medical and Veterinary Entomology* **17**, 165–177.

Dallas, J.F., Cruickshank, R.H., Linton, Y.M., Nolan, D.V., Patakakis, M., Braverman, Y., Capela, R., Capela, M., Pena, I., Meiswinkel, R., Ortega, M.D., Baylis, M., Mellor, P.S. & Mordue, A.J. (2003) Phylogenetic status and matrilineal structure of the biting midge, *Culicoides imicola*, in Portugal, Rhodes and Israel. *Medical and Veterinary Entomology* **17**, 379–387.

De Liberato, C., Scavia, G., Lorenzetti, R., Scaramozzino, P., Amaddeo, D., Cardeti, G., Scicluna, M., Ferrari, G. & Autorino, G.L. (2005) Identification of *Culicoides obsoletus* (Diptera: Ceratopogonidae) as a vector of bluetongue virus in central Italy. *The Veterinary Record* **156**, 301–304.

Du Toit, R.M. (1944) The transmission of blue-tongue and horsesickness by *Culicoides*. *Onderstepoort Journal of Veterinary Science and Animal Industry* **19**, 7–16.

Excoffier, L. (2004) Patterns of DNA sequence diversity and genetic structure after a range expansion: lessons from the infinite-island model. *Molecular Ecology* **13**(4), 853–864.

Gomez-Tejedor, C. (2004) Brief overview of the bluetongue situation in Mediterranean Europe, 1998–2004. *Veterinaria Italiana* **40**(3), 57–60.

Gomulski, L.M., Meiswinkel, R., Delecolle, J.C., Goffredo, M. & Gasperi, G. (2005) Phylogenetic relationships of the subgenus *Avaritia* Fox, 1955 including *Culicoides obsoletus* (Diptera, Ceratopogonidae) in Italy based on internal transcribed spacer 2 ribosomal DNA sequences. *Systematic Entomology* **30**(4), 619–631.

Gomulski, L.M., Meiswinkel, R., Delecolle, J.C., Goffredo, M. & Gasperi, G. (2006) Phylogeny of the subgenus *Culicoides* and related species in Italy, inferred from internal transcribed spacer 2 ribosomal DNA sequences. *Medical and Veterinary Entomology* **20**(2), 229–238.

Hahn, M.W., Rausher, M.D. & Cunningham, C.W. (2002) Distinguishing between selection and population expansion in an experimental lineage of bacteriophage T7. *Genetics* **161**(1), 11–20.

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

Holmes, I.H., Boccardo, G., Estes, M.K. & Furuichi, M.K. (1995). Family Reoviridae. pp. 208–239 *in* Murphy, F.A.,

Fauquet, C.M., Bishop, D.H.L. & Ghabrial, S.A. (*Eds*) *Virus taxonomy, classification and nomeclature of viruses*. Sixth report of the international committee on taxonomy of viruses. Archives of Virology, Suppl. 10, Wien, New York, Springer.

Jennings, D.M., Boorman, J.P.T. & Ergün, H. (1983) *Culicoides* from Western Turkey in relation to bluetongue disease of sheep and cattle. *Revue d'Elevage et de Mèdecine Vètérinaire des Pays Tropicaux* **36**, 67–70.

Kumar, S., Tamura, K., Jakobsen, I.B. & Nei, M. (2001) MEGA2: molecular evolutionary genetics analysis software. *Bioinformatics* **17**(12), 1244–1245.

Linton, Y.M., Mordue, A.J., Cruickshank, R.H., Meiswinkel, R., Mellor, P.S. & Dallas, J.F. (2002) Phylogenetic analysis of the mitochondrial cytochrome oxidase subunit I gene of five species of the *Culicoides imicola* species complex. *Medical and Veterinary Entomology* **16**(2), 139–146.

Manso-Ribeiro, J.J., Rosa-Azevedo, A., Noroña, F.O., Blanco-Forte, M.C., Grave-Pereira, C. & Vasco-Fernandez, M. (1957) Fiebre catarrhale du mouton (Bluetongue). *Bulletin de l'Office International des Épizooties* **48**, 350–367.

Mehlhorn, H., Walldorf, V., Klimpel, S., Jahn, B., Jaeger, F., Eschweiler, J., Hoffmann, B. & Beer, M. (2007) First occurrence of *Culicoides obsoletus*-transmitted bluetongue virus epidemic in Central Europe. *Parasitology Research* **101**, 219–228.

Meiswinkel, R., van Rijn, P., Leijs, P. & Goffredo, M. (2007) Potential new *Culicoides* vector of bluetongue virus in northern Europe. *Veterinary Record* **161**, 564–565.

Meiswinkel, R., Baldet, T., de Deken, R., Takken, W., Delecolle, J.C. & Mellor, P.S. (2008) The 2006 outbreak of bluetongue in northern Europe: The entomological perspective. *Preventive Veterinary Medicine* **87**(1–2), 55–63.

Mellor, P.S. (1996) *Culicoides*: vectors, climate change and disease risk. *Veterinary Bulletin* **66**, 301–306.

Mellor, P.S. & Pitzolis, G. (1979) Observations on breeding sites and light trap collections of *Culicoides* during an outbreak of bluetongue in Cyprus. *Bulletin of Entomological Research* **69**, 229–234.

Mellor, P.S. & Wittmann, E.J. (2002) Bluetongue virus in the Mediterranean basin 1998–2001. *Veterinary Journal* **164**, 20–37.

Mellor, P.S., Boned, J., Hamblin, C. & Graham, S. (1990) Isolations of African horse sickness virus from vector insects made during the 1988 epizootic in Spain. *Epidemiology and Infection* **105**, 447–454.

Miranda, M.A., Borras, D., Rincon, C. & Alemany, A. (2003) Presence in the Balearic Islands (Spain) of the midges *Culicoides imicola* and *Culicoides obsoletus* group. *Medical and Veterinary Entomology* **17**, 52–54.

Nei, M. & Li, W.H. (1979) Mathematical model for studying genetic-variation in terms of restriction endonucleases. *Proceedings of the National Academy of Sciences of the United States of America* **76**(10), 5269–5273.

Nolan, D.V., Carpenter, S., Barber, J., Mellor, P.S., Dallas, J.F., Mordue, A.J., Mordue, A.J. & Piertney, S.B. (2007) Rapid diagnostic PCR assays for members of the *Culicoides obsoletus* and *Culicoides pulicaris* species complexes, implicated vectors of bluetongue virus in Europe. *Veterinary Microbiology* **124**(1–2), 82–94.

Office International des Epizooties (OIE) (2000) Bluetongue in Spain: in the Balearic Islands. Follow-up report No. 1. *Disease Information* **13**, 224.

Ortega, M.D., Holbrook, F.R. & Lloyd, J.E. (1999) Seasonal distribution and relationship to temperature and precipitation of the most abundant species of *Culicoides* in five provinces of Andalusia, Spain. *Journal of the American Mosquito Control Association* **15**(3), 391–399.

Perrin, A., Cetre-Sossah, C., Mathieu, B., Baldet, T., Delecolle, J.C. & Albina, E. (2006) Phylogenetic analysis of *Culicoides* species from France based on nuclear ITS1-rDNA sequences. *Medical and Veterinary Entomology* **20**(2), 219–228.

Rawlings, P. (1996) A key, based on wing patterns of biting midges (genus *Culicoides* Latreille – Diptera: Ceratopogonidae) in the Iberian Peninsula, for use in epidemiological studies. *Graellsia* **52**, 57–71.

Rawlings, P., Pro, M.J., Peña, L., Ortega, M.D. & Capela, R. (1997) Spatial and seasonal distribution of *Culicoides imicola* in Iberia in relation to the transmission of African horse sickness virus. *Medical and Veterinary Entomology* **11**, 49–57.

Ray, N., Currat, M. & Excoffier, L. (2003) Intra-Deme Molecular Diversity in Spatially Expanding Populations. *Molecular Biology and Evolution* **20**(1), 76–86.

Rich, S.M., Licht, M.C., Hudson, R.R. & Ayala, F.J. (1998) Malaria's Eve: Evidence of a recent population bottleneck throughout the world population of *Plasmodium falciparum*. *Proceedings of the National Academy of Sciences of the United States of America* **95**, 4425–4430.

Rogers, A.R. & Harpending, H. (1992) Population growth makes waves in the distribution of pairwise genetic differences. *Molecular Biology and Evolution* **9**, 552–569.

Rozas, J., Sanchez-DelBarrio, J.C., Messeguer, X. & Rozas, R. (2003) DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics* **19**(18), 2496–2497.

Sarto I Monteys, V., Ventura, D., Pages, N., Aranda, C. & Escosa, R. (2005) Expansion of *Culicoides imicola*, the main bluetongue virus vector in Europe, into Catalonia, Spain. *Veterinary Record* **156**, 415–417.

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.

Slatkin, M. & Hudson, R.R. (1991) Pairwise comparisons of mitochondrial-DNA sequences is stable and sponentially growing populations. *Genetics* **129**(2), 555–562.

Tabachnick, W.J. & Black, W.C. (1995) Making a case for molecular population genetic-studies of arthropod vectors. *Parasitology Today* **11**(1), 27–30.

Tajima, F. (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123**, 585–595.

Takamatsu, H., Mellor, P.S., Mertens, P.P.C., Kirkham, P.A., Burroughs, J.N. & Parkhouse, R.M.E. (2003) A possible overwintering mechanism for bluetongue virus in the absence of the insect vector. *Journal of general Virology* **84**, 227–235.

Walker, D.E., Moler, P.E., Buhlmann, K.A. & Avise, J.C. (1998) Phylogeographic uniformity in mitochondrial DNA of the snapping turtle (*Chelydra serpentina*). *Animal Conservation* **1**, 55–60.

Watterson, G.A. (1975) Number of segregating sites in genetic models without recombination. *Theorical Population Biology* **7**(2), 256–276.

Wittmann, E.J., Mellor, P.S. & Baylis, M. (2001) Using climate data to map the potential distribution of *Culicoides imicola* (Diptera: Ceratopogonidae) in Europe. *Revue Scientifique et Technique de l'Office International des Epizooties* **20**, 731–740.