

Molecular taxonomy and population structure of a *Culicoides* midge vector

D.V. Nolan, J.F. Dallas & A.J. Mordue (Luntz)

School of Biological Sciences, University of Aberdeen, United Kingdom

Correspondence: D.V. Nolan, Zoology Building, Tillydrone Avenue, Aberdeen AB24 2TZ, Scotland, United Kingdom

Summary

The biting midge *Culicoides imicola* Kieffer (Diptera: Ceratopogonidae) is the major Old World vector of the arboviruses that cause African horse sickness (AHS) and bluetongue (BT). Recently, the incidence and geographical scales of AHS and BT outbreaks in the Mediterranean Basin have increased, with serotype distribution in the BT outbreaks being geographically structured. The authors review molecular approaches for assessing the contribution of cryptic species and population subdivision in *C. imicola* to BT serotype structure in this region. No evidence was found for cryptic species. In contrast, evidence was found for marked matrilineal subdivision between the eastern and western parts of the Mediterranean Basin. This pattern is comparable to the geographic structure of BT serotypes, suggesting that subdivision in the insect vector potentially constrains serotype spread. The authors are presently testing this hypothesis.

Keywords

African horse sickness – Bluetongue – *Culicoides imicola* – Cytochrome oxidase subunit I – Europe – Mediterranean Basin – Mitochondrial DNA – Molecular population genetics.

Introduction

The biting midge *Culicoides imicola* Kieffer, 1913 (Diptera: Ceratopogonidae) is the most important Old World vector of the arboviruses that cause African horse sickness (AHS) and bluetongue (BT) (16). AHS is endemic in sub-Saharan Africa and BT is endemic in tropical latitudes worldwide. Before 1980, outbreaks of both diseases in the Mediterranean Basin were localised and sporadic. This pattern has changed recently, with AHS outbreaks in Iberia since the mid-1980s lasting longer (14, 15) and the largest epidemic of BT ever recorded in the Mediterranean Basin between 1998 and 2001 (1).

The most likely causes of these epidemiological changes are increases in the range and abundance of *C. imicola* (2, 15, 20, 21), and these changes might be responses to climate change (27). Although primarily an Afro-Asiatic species, *C. imicola* is reported widely throughout the Mediterranean Basin: from Algeria, Cyprus, Egypt, Israel, the islands of Lesbos and

Rhodes, Morocco, Portugal, Spain, Tunisia and Turkey (16), and most recently from Italy (7) and the Balearic Islands (18). Moreover, risk maps for predicting the distribution of *C. imicola* from climatic variables (10, 28) and vegetation indices (4) suggest that *C. imicola* is even more widespread in this region, and that its apparent absence potentially reflects a lack of sampling.

Taxonomic status in the North American BT virus vector *Culicoides variipennis* (Coquillett) has crucial implications for epidemiology and vector control. According to analyses of male morphology and genetic distance, *Culicoides variipennis* consists of three subspecies, only one of which, *C. v. sonorensis* Wirth and Jones, is an efficient BT virus vector (24, 25). Subdivision of insect vector populations has important implications for control measures and predicting disease spread (26). Despite this, the phylogenetic status and genetic structure of *C. imicola* in Mediterranean countries were unknown prior to recent work by our group at the University of Aberdeen.

Identification of insect species: a 'DNA barcoding' system?

It is now recognised that coherence in insect systematics will ultimately depend on having a large database of comparable DNA sequence data. Four genes presently stand out as well-surveyed and informative across a wide taxonomic range: the cytochrome oxidase subunit I (COI) and 16S genes of mitochondrial DNA (mtDNA), and the nuclear 18S and *EF-1 α* genes (9). Indeed, a DNA-based identification system based on COI can potentially provide a consistent means to resolve species diversity using 'DNA barcoding' (12). When fully developed, this approach is expected to provide a reliable, cost-effective and accessible solution to current problems of species identification.

Molecular taxonomy of the *Imicola* Complex in South Africa

Species identification in *Culicoides* is often problematic, owing to their small size and environment-dependent morphology, the prevalence of taxa with very similar morphologies and the expression of diagnostic traits in adult males only. Analyses of morphological characters of adult insects suggest that the *Imicola* complex in South Africa comprises at least ten species. Eight species are confirmed (*C. imicola sensu stricto*; *C. brevitarsis* Kieffer; *C. pseudopallidipennis* Clastrier; *C. nudipalpis* Delfinado; *C. bolitinos* Meiswinkel; *C. miombo* Meiswinkel; *C. loxodontis* Meiswinkel, *C. tuttifrutti* Meiswinkel, Cornet and Dyce,) and two are provisional (*C. kwagga* Meiswinkel, and *C. sp. #103* Meiswinkel). These species have distinct larval habitats, including rich damp clay soil (*C. imicola*), marshy areas (*C. miombo*), elephant dung (*C. loxodontis*), horse, rhinoceros and zebra dung (*C. kwagga*), buffalo and cattle dung (*C. bolitinos*), rotting fruits (*C. tuttifrutti*) and banana stumps (*C. pseudopallidipennis*). According to analysis of random amplification of polymorphic DNA (RAPD) markers, these species are genetically different but not phenetically distinct (22). The evolution of vector competence and transitions in larval habitat in the *Imicola* complex are best elucidated using phylogenetic approaches. The phylogenetic status of four confirmed species (*C. imicola s.s.*, *C. bolitinos*, *C. loxodontis* and *C. tuttifrutti*) and one provisional (*C. kwagga*) species of this complex was assessed, using a phylogenetic analysis of COI (13).

The *Culicoides* spp. sequences were aligned against the corresponding COI sequence of *Anopheles gambiae*, and base differences among the sequences were identified relative to the IMI 16 sequence of *C. imicola*. All the base substitutions within species

were synonymous, and several base differences between species were evident (Fig. 1).

A neighbour-joining tree (Fig. 2a) yielded consistently high bootstrap support for each separate species but consistently low support for clades containing more than one species. All parsimony analyses yielded the same group of four equally parsimonious trees. One of these trees had the same topology as the NJ tree, which also had the highest likelihood. The other trees differed in the placement of *C. bolitinos*, *C. loxodontis* or *C. tuttifrutti*. A maximum likelihood tree (Fig. 2b) yielded high quartet puzzling support for all species except *C. imicola* as distinct groups, and for the clade containing *C. bolitinos*, *C. loxodontis* and *C. tuttifrutti*.

Thus, the usefulness of COI was demonstrated for species identification in the *Imicola* Complex, and showed that five of its members are phylogenetically distinct.

Species identification and subdivision in *Culicoides imicola*

The species status and matrilineal subdivision in *C. imicola* in the Mediterranean Basin has been characterised using COI sequences (11). Partial COI sequences (472 bp without primers) were obtained from 19 sites in Portugal, Rhodes and Israel (Table I).

Phylogenetic trees were constructed from the COI sequences of *C. imicola* in the Mediterranean Basin and South Africa and of four other species of the *Imicola* Complex from southern Africa. The maximum likelihood tree (Fig. 3) contained five separate clades. Each clade represented a single species in the *Imicola* Complex, and each was supported by high (73-100) bootstrap values. All the *C. imicola* sequences grouped in the same clade. Phylogenetic substructure within the *C. imicola* clade was absent, according to the maximum likelihood bootstrap values of less than 60. Thus, midges of the morphospecies *C. imicola* from Portugal, Rhodes, Israel and South Africa belong to the same phylogenetic clade of COI, and are phylogenetically distinct from four other species of the *Imicola* Complex. These results are consistent with all *C. imicola* from the study areas being potentially competent AHS and BT vectors.

The samples from Portugal had zero values of haplotype and nucleotide diversity, being monomorphic for haplotype IMICOI 01. The samples from Israel had a haplotype diversity of 0.478 (SD 0.112) and a nucleotide diversity of 0.00246 (SD 0.00067). One haplotype (IMICOI 03) was shared between Israel and Rhodes, and the other

eight haplotypes were unique to each country. According to an analysis of molecular variance, genetic differentiation between the samples from Portugal and Israel was highly significant (permutation test, $P < 0.001$). The proportion of molecular variance within countries was 0.20 and between countries 0.80. The corresponding value of Φ_{ST} was 0.80, and this value was significantly

different from zero (permutation test, $P < 0.001$). Thus, the level of matrilineal subdivision in *C. imicola* between the eastern and western ends of the Mediterranean Basin is marked and is comparable to the highest levels known for Diptera.

Whether this result reflects genetic drift associated with dependence on fragmented habitat, or a recent

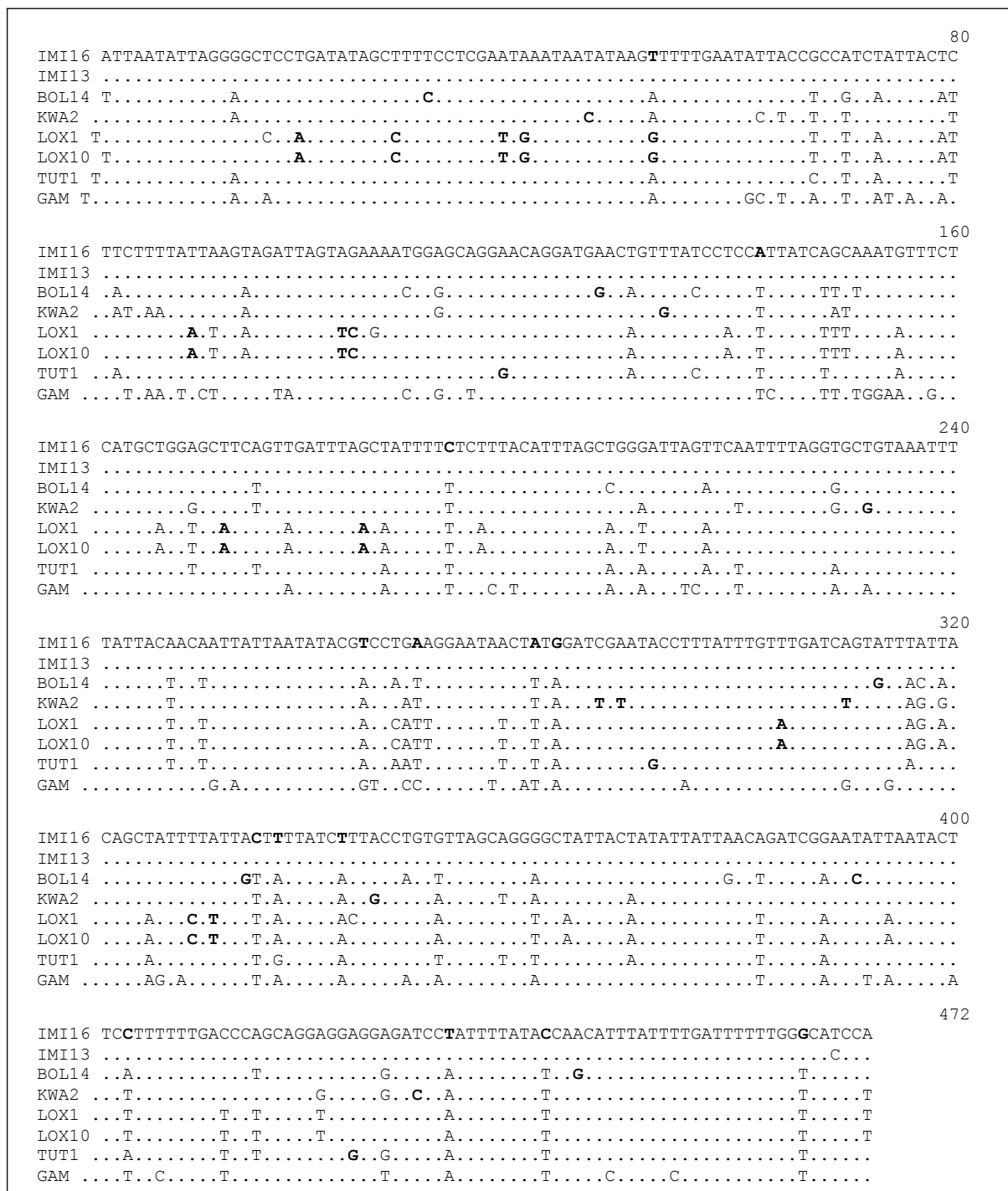
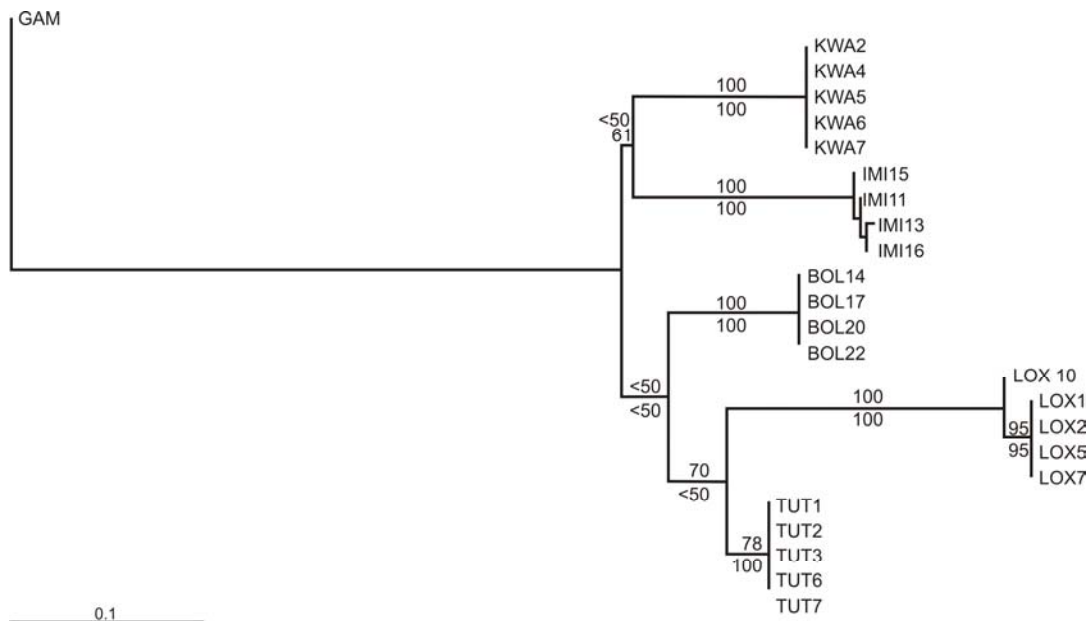


Figure 1 Alignment of nucleotide sequences (5'-3') of 472 bp (without primers) of the mitochondrial DNA cytochrome oxidase subunit I gene of *Culicoides imicola* (IMI), *C. bolitinos* (BOL), *C. kwagga* (KWA), *C. loxodontis* (LOX), *C. tuttifrutti* (TUT) from South Africa and *Anopheles gambiae* (GAM) Identical sequences are omitted, dots indicate identical bases, and species-diagnostic bases are shown in bold Redrawn from Linton *et al.* (13)

- a) Neighbour-joining tree constructed from Tamura-Nei distances with γ -distributed rates
Numbers on the nodes represent NJ (above) and parsimony (below) bootstrap proportions (10 000 replicates each)



- b) Maximum likelihood tree constructed using a GTR substitution model with γ -distributed rates
Numbers on the nodes represent quartet puzzling values (10 000 replicates). The full species names corresponding to the abbreviations are given in Figure 1

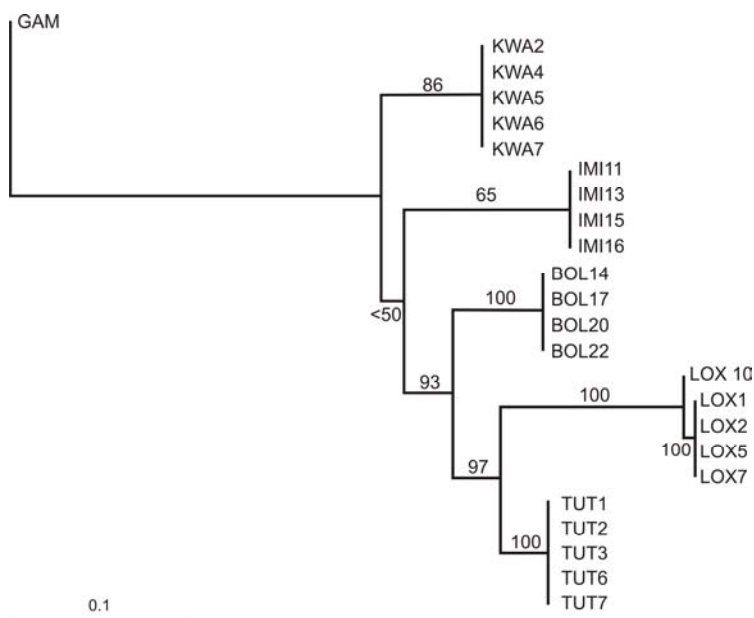


Figure 2
Phylogenetic relationships among partial nucleotide sequences of the mitochondrial DNA cytochrome oxidase subunit I gene of five members of the *Imicola* Complex
Redrawn from Linton *et al.* (13)

Table 1
Locations and years of sampling of *Culicoides imicola*, number of sequences and COI haplotypes obtained at each site

Country	Site	Latitude and longitude	Year	Site code	No.	IMICOI haplotypes
Portugal	Alcaíns	39°55'N 7°26'W	2001	PO1	2	01
	Avis	39°03'N 7°54'W	2001	PO2	2	01
	Castelo de Vide	39°26'N 7°28'W	2001	PO3	2	01
	Mourão	38°23'N 7°18'W	2001	PO4	2	01
	Pedrogão	38°46'N 8°19'W	2001	PO5	2	01
	Barca D'Alva	41°01'N 6°57'W	2001	PO6	2	01
Greece	Dimilia, Rhodes	36°16'N 28°02'E	2001	GR1	5	02, 03
Israel	Beit Dagan	32°00'N 34°50'E	1996	IBET	3	04, 08, 09
			2001	IS1	2	04, 05
	Newe Ya'ar	32°42'N 35°11'E	1996	INY	4	04
	Karmiyya	31°36'N 34°33'E	2001	IS2	2	04
	Kefar Malal	32°10'N 34°54'E	2001	IS3	2	04
	Kefar Daniyyel	31°56'N 34°56'E	2001	IS4	2	04
	Kefar Silver	31°41'N 34°32'E	2001	IS5	2	04
	Mishmar Hasharon	32°21'N 34°54'E	2001	IS6	2	04, 05
	Regba	32°59'N 35°06'E	2001	IS7	2	04
	Ma'ale Hahamisha	31°49'N 35°07'E	2001	IS8	2	04
	Merom Golan	33°10'N 35°42'E	2001	IS9	2	03
	Nahalal	32°41'N 35°12'E	2001	IS10	2	04
	Yotvata	29°53'N 35°03'E	2001	IS11	2	06, 07
	South Africa*	Onderstepoort	25°39'S 28°11'E	1996	IMI	4

* Sequences were reported in Linton *et al.* (13)

Redrawn from Dallas *et al.* (11)

colonisation of the Mediterranean Basin by two or more genetically distinct sources, is presently unclear. Seasonal airstreams are thought to mediate the movement of arbovirus-infected *C. brevitarsis* Kieffer in Australia (5, 19) and the immigration of BT virus (BTV)-infected *C. imicola* into Portugal (23) and Israel (6). Our result suggests that matrilineal gene flow between the latter areas in the past has been limited.

The matrilineal subdivision in *C. imicola* is comparable to the geographic structure of BTV serotypes in outbreaks around the Mediterranean during 1998-2001. BTV serotypes 4, 9 and 16 spread westwards from Israel and Turkey to Greece and Bulgaria, while BTV-2 spread northwards from Tunisia through Sardinia, Corsica, and Sicily to mainland Italy (17). Assuming that the *C. imicola* populations in these areas are equally competent BTV vectors, we hypothesise that the matrilineal subdivision reflects true population subdivision in *C. imicola*, and that this structure constrained the recent BT outbreaks in these areas to different geographical sources and routes of spread. This hypothesis has important implications for epidemiological monitoring and vaccination programmes for BT, and we are presently testing it using a phylogeographic analysis of a comprehensive sampling of *C. imicola* from this region.

Future work

Outbreaks of BT have occurred in areas where *C. imicola*, although looked for, has never been recorded (as far north as 44°N in Serbia and Bosnia-Herzegovina, north-west Greece, Bulgaria and west European Turkey) (3, 17). In these areas, novel *Culicoides* vectors, possibly members of the widespread *Obsoletus* and *Pulicaris* Complexes, are potential vectors (17). Moreover, in 2002, in Sicily, where outbreaks of BT occurred in the absence of *C. imicola*, bluetongue viral RNA was detected in wild-caught, non-blood-engorged, parous *C. pulicaris*, further suggesting that *C. pulicaris* can be a fully competent BT vector (8). Therefore, we plan to extend the molecular approaches described above to these novel vectors.

Analysis of specimens from mainland Italy, and of larger samples from the areas in the above studies, is ongoing. The utility of COI to detect genetic subdivision within countries such as Portugal is limited, however. We are therefore analysing mtDNA and nuclear genes more polymorphic than COI for a comprehensive characterisation of genetic subdivision in *C. imicola* in the Mediterranean Basin. The results of these molecular studies will be combined with GIS data and used to construct a risk map for bluetongue in Europe. This map will be

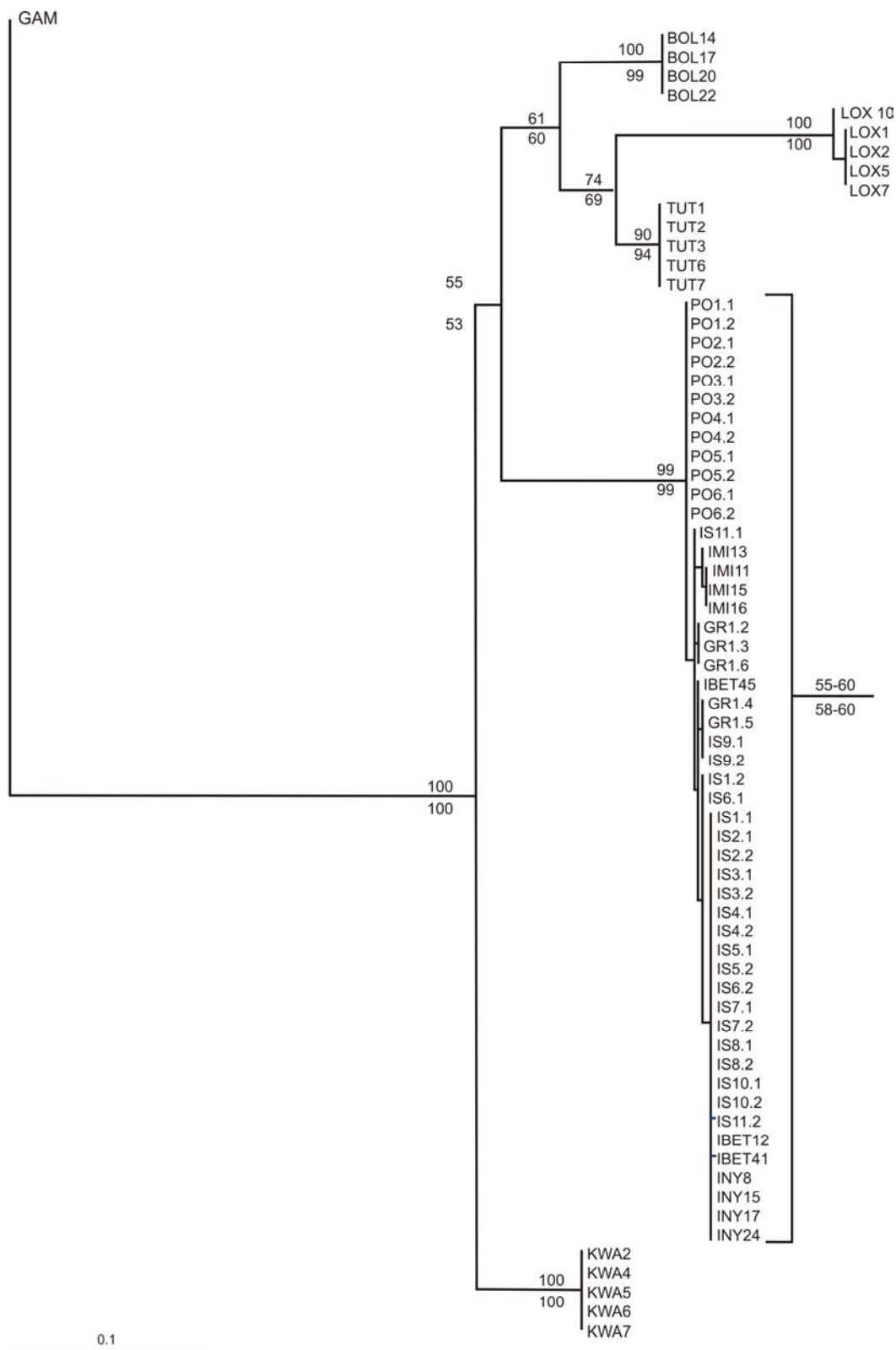


Figure 3
Phylogenetic relationships among COI haplotypes of *C. imicola* from Portugal, Rhodes, Israel and South Africa, and of four other species of the Imicola Complex

Maximum likelihood tree. Numbers are bootstrap percentages. Values above the nodes are for the GTR+I+ γ model and values below the nodes are for the GTR+ γ model. The details corresponding to the code of each *C. imicola* sample are given in Table I. The species names corresponding to the abbreviations used for the other species are given in Figure 1. Redrawn from Dallas *et al.* (13)

used to predict outbreaks and inform agricultural and veterinary practice in the participating countries and the EU as a whole. The combination of data on subdivision in vector species and on BT serotype distribution is expected to provide critical information in developing risk assessment relating to emerging diseases and will become more important as vector ecologists realise the full potential of population genetics and its application to medical veterinary entomology.

Acknowledgements

Part of this work was financed by the European Commission (Contract No. QLK2-CT-2000-00611).

References

- Baylis M. (2002). – The re-emergence of bluetongue. *Vet. J.*, **164**, 5-6.
- Baylis M., El-Hasnaoui H., Bouayoune H., Touti J. & Mellor P.S. (1997). – The spatial and seasonal distribution of African horse sickness and its potential *Culicoides* vectors in Morocco. *Med. Vet. Entomol.*, **11**, 203-212.
- Baylis M. & Mellor P.S. (2001). – Bluetongue around the Mediterranean in 2001. *Vet. Rec.*, **149**, 659.
- Baylis M., Mellor P.S., Wittmann E.J. & Rogers D.J. (2001). – Prediction of areas around the Mediterranean at risk of bluetongue by modelling the distribution of its vector using satellite imaging. *Vet. Rec.*, **149**, 639-643.
- Bishop A.L., Barchia I.M. & Spohr L.J. (2000). – Models for the dispersal in Australia of the arbovirus vector, *Culicoides brevitarsis* Kieffer (Diptera: Ceratopogonidae). *Prev. Vet. Med.*, **47**, 243-254.
- Braverman Y. & Chechik F. (1996). – Air streams and the introduction of animal diseases borne on *Culicoides* (Diptera, Ceratopogonidae) into Israel. *Rev. Sci. Tech. Off. Int. Épiç.*, **15**, 1037-1052.
- Calistri P., Goffredo M., Caporale V. & Meiswinkel R. (2003). – The distribution of *Culicoides imicola* in Italy: Application and evaluation of current Mediterranean models based on climate. *J. Vet. Med. B*, **50**, 132-138.
- Caracappa S., Torina A., Guercio A., Vitale F., Calabro A., Purpari G., Ferrantelli V., Vitale M. & Mellor P.S. (2003). – Identification of a novel bluetongue virus vector species of *Culicoides* in Sicily. *Vet. Rec.*, **153**, 71-74.
- Caterino M.S., Cho S. & Sperling F.A.H. (2000). – The current state of insect molecular systematics: a thriving Tower of Babel. *Ann. Rev. Entomol.*, **45**, 1-54.
- Conte A., Giovannini A., Savini L., Goffredo M., Calistri P. & Meiswinkel R. (2003). – The effect of climate on the presence of *Culicoides imicola* in Italy. *J. Vet. Med. B*, **50**, 139-147.
- Dallas J.F., Cruickshank R.H., Linton Y.-M., Nolan D.V., Patakakis M., Braverman Y., Capela R., Pena I., Meiswinkel R., Ortega M.D., Baylis M., Mellor P.S. & Mordue (Luntz) A.J. (2003). – Phylogenetic status and matrilineal structure of the biting midge, *Culicoides imicola*, in Portugal, Rhodes and Israel. *Med. Vet. Entomol.*, **17** (4), 379-387.
- Hebert P.D.N., Cywinska A., Ball S.L. & DeWaard J.R. (2003). – Biological identifications through DNA barcodes. *P. Roy. Soc. Lond. B-Bio.*, **270**, 313-321.
- 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. *Med. Vet. Entomol.*, **16**, 139-146.
- Mellor P.S. (1994). – Epizootiology and vectors of African horse sickness virus. *Comp. Immunol. Microbiol. Infect. Dis.*, **17**, 287-296.
- Mellor P.S. & Boorman J. (1995). – The transmission and geographical spread of African horse sickness and bluetongue viruses. *Ann. Trop. Med. Parasitol.*, **89**, 1-15.
- Mellor P.S., Boorman J. & Baylis M. (2000). – *Culicoides* biting midges: Their role as arbovirus vectors. *Ann. Rev. Entomol.*, **45**, 307-340.
- Mellor P.S. & Wittmann E.J. (2002). – Bluetongue virus in the Mediterranean Basin 1998-2001. *Vet. J.*, **164**, 20-37.
- Miranda M.A., Borrás D., Rincon C. & Alemany A. (2003). – Presence in the Balearic islands (Spain) of the midges *Culicoides imicola* and *Culicoides obsoletus* group. *Med. Vet. Entomol.*, **17**, 52-54.
- Murray M.D. (1995). – Influences of vector biology on transmission of arboviruses and outbreaks of disease – the *Culicoides brevitarsis* model. *Vet. Microbiol.*, **46**, 91-99.
- Rawlings P. & Mellor P.S. (1994). – African horse sickness and the overwintering of *Culicoides* spp in the Iberian Peninsula. *Rev. Sci. Tech. Off. Int. Épiç.*, **13**, 753-761.
- Rawlings P., Capela R., Pro M.J., Ortega M.D., Pena I., Rubio C., Gasca A. & Mellor P.S. (1998). – The relationship between climate and the distribution of *Culicoides imicola* in Iberia. *Arch. Virol.*, 93-102.
- Sebastiani F., Meiswinkel R., Gomulski L.M., Guglielmino C.R., Mellor P.S., Malacrida A.R. & Gasperi G. (2001). – Molecular differentiation of the Old World *Culicoides imicola* species complex (Diptera, Ceratopogonidae), inferred using random amplified polymorphic DNA markers. *Molec. Ecol.*, **10**, 1773-1786.
- Sellers R.F., Pedgley D.E. & Tucker M.R. (1979). – Possible windborne spread of bluetongue to Portugal, June-July 1956. *J. Hyg., Camb.*, **81**, 189-196.
- Tabachnick W.J. (1992). – Genetic differentiation among populations of *Culicoides variipennis* (Diptera,

- Ceratopogonidae), the North American vector of bluetongue virus. *Ann. Entomol. Soc. Am.*, **85**, 140-147.
25. Tabachnick W.J. (1996). – *Culicoides variipennis* and bluetongue virus epidemiology in the United States. *Ann. Rev. Entomol.*, **41**, 23-43.
 26. Tabachnick W.J. & Black W.C. (1995). – Making a case for molecular population genetic studies of arthropod vectors. *Parasitol. Today*, **11**, 27-30.
 27. Wittmann E.J. & Baylis M. (2000). – Climate change: effects on *Culicoides*-transmitted viruses and implications for the UK. *Vet. J.*, **160**, 107-117.
 28. Wittmann E.J., Mellor P.S. & Baylis M. (2001). – Using climate data to map the potential distribution of *Culicoides imicola* (Diptera: Ceratopogonidae) in Europe. *Rev. Sci. Tech. Off. Int. Épiçz.*, **20**, 731-740.