

# *Cardinium* symbiosis as a potential confounder of mtDNA based phylogeographic inference in *Culicoides imicola* (Diptera: Ceratopogonidae), a vector of veterinary viruses

Jack Pilgrim (✉ [j.pilgrim@liverpool.ac.uk](mailto:j.pilgrim@liverpool.ac.uk))

University of Liverpool Institute of Infection and Global Health <https://orcid.org/0000-0002-2941-1482>

Stefanos Siozios

University of Liverpool Institute of Infection, Veterinary and Ecological Sciences

Matthew Baylis

University of Liverpool Institute of Infection, Veterinary and Ecological Sciences

Gert Venter

ARC Onderstepoort Veterinary Research Campus: Agricultural Research Council Onderstepoort  
Veterinary Research

Claire Garros

CIRAD

Gregory D. D. Hurst

University of Liverpool Institute of Infection, Veterinary and Ecological Sciences

---

## Short report

**Keywords:** Cardinium, symbiont, symbiosis, COI, Culicoides

**Posted Date:** September 30th, 2020

**DOI:** <https://doi.org/10.21203/rs.3.rs-83200/v1>

**License:**   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

**Version of Record:** A version of this preprint was published at Parasites & Vectors on February 8th, 2021.  
See the published version at <https://doi.org/10.1186/s13071-020-04568-3>.

# Abstract

**Background:** *Culicoides imicola* (Diptera: Ceratopogonidae) is an important Afrotropical and Palearctic vector of disease, transmitting viruses of animal health and economic significance. The apparent incursions of *C. imicola* into mainland Europe via wind-movement events has made it important to trace this species to better predict new areas of arbovirus outbreaks. A widely used method for tracking dispersal patterns of *C. imicola* employs a phylogeographic approach anchored on the mtDNA marker *COI* (cytochrome c oxidase subunit I). However, a problem with this approach is that maternally-inherited symbiotic bacteria can alter the frequency of *COI* mitochondrial haplotypes (mitotypes), masking the true patterns of movement and gene flow.

**Methods:** In this study, the mtDNA structure of *C. imicola* in relation to *Cardinium* infection status was investigated through haplotype network analysis. *COI* Sanger sequences from infected and uninfected individuals were first compared, before extending the haplotype network to include mitotypes from a geographic range where *Cardinium* infection has previously been investigated.

**Results:** The mitotype network of a South African population, containing both *Cardinium*-infected and uninfected individuals, demonstrated the presence of two broad mitotype groups. All *Cardinium*-infected specimens fell into one group (Fisher's exact test,  $P < 0.001$ ) demonstrating a linkage disequilibrium between symbiont and mitochondria. Furthermore, by extending this haplotype network to include other *C. imicola* populations from the Mediterranean basin, we revealed mitotype variation between the Eastern and Western Mediterranean basins (EMB and WMB) mirrored *Cardinium*-infection heterogeneity.

**Conclusions:** These observations suggest that the linkage disequilibrium of *Cardinium* and mitochondria reflects symbiont gene flow within the Mediterranean basin but may not assist in elucidating host gene flow. Subsequently, we urge caution on the single usage of the *COI* marker to determine population structure and movement in *C. imicola*, and instead suggest the complementary utilisation of additional molecular markers.

## Background

*Culicoides* spp. (Diptera: Ceratopogonidae) are a group of tiny biting midges responsible for the spread of several pathogens to animals, including African horse sickness, Schmallenberg and bluetongue viruses. An important vector species, *Culicoides imicola*, was initially considered to be limited to Afrotropical regions but recent genetic analysis suggests this species has been present in the Mediterranean basin since the Late Pleistocene or Early Holocene (1). The movement of *C. imicola* at the northern fringes of the Mediterranean basin is likely driven by wind-events which have led to recurrent incursions of the species into mainland Europe (2). Thus, the tracking of *C. imicola* through population genetics and climatic modelling is important in establishing where potential future outbreaks of midge-borne viruses may occur.

Mitochondrial DNA (mtDNA) markers have previously been used in phylogeographic studies to detect biodiversity and dispersal events of *C. imicola* populations (1–3). However, mtDNA based inferences can be confounded by the linkage disequilibrium (co-inheritance) of symbionts and host mtDNA. As symbionts enter a naïve population, selective sweeps of the bacteria can occur and with it the homogenisation of linked mtDNA haplotypes, leading to the apparent absence of biodiversity (4). Conversely, variation in symbiont presence over space can lead to a perceived genetic population structure when none exists. Subsequently, inferences from these studies can only be interpreted in the context of any history of infection with symbiotic bacteria.

In this study, we examine the relationship between mtDNA in *C. imicola* and the heritable symbiont *Candidatus Cardinium hertigii* (Bacteroidetes). *Cardinium* is present in several biting midge species, including *C. imicola*, although *Cardinium's* significance in *Culicoides* biology is unknown (5–7). As *Cardinium* infection could have implications for inferring movement of *C. imicola* into naïve geographic areas through mtDNA marker analysis, we analysed associations between mitochondrial haplotypes (mitotypes) and *Cardinium* infection in this important vector species.

## Methods

106 *C. imicola* individuals were collected between 2015–2016 using light traps from two countries: France (location: Corsica) and South Africa (location: Onderstepoort). Amplification of the host COI gene (8) and the *Cardinium* Gyrase B (GyrB) gene (7) was undertaken by conventional PCR assay. DNA extracts which passed quality control through successful COI amplification but were negative for GyrB were then screened with a more sensitive nested PCR assay (cycling conditions and primer sequence information in Table S1). PCR assays consisted of a total of 15  $\mu$ L per well, comprising of 7.5  $\mu$ L GoTaq® Hot Start Polymerase (Promega), 5.1  $\mu$ L nuclease free water, 0.45  $\mu$ L forward and reverse primers (concentration 10 pmol/ $\mu$ L) and 1.5  $\mu$ L DNA template. PCR products were separated on 1% agarose gels stained with Midori Green Nucleic Acid Staining Solution (Nippon Genetics Europe). COI amplicons identified by gel electrophoresis were purified enzymatically (ExoSAP) before being sent for sequencing through the Sanger method (GATC Biotech AG, Konstanz, Germany). Chromatograms were then visually assessed for correct base calls. Accession numbers for COI sequences generated in this study are LR877434–LR877461. We then investigated possible associations of the symbiont *Cardinium* with *C. imicola* mitotype distribution through haplotype network analysis. To this end, haplotype networks were constructed, with analysis and visualisation undertaken using the TCS haplotype network algorithm generated in PopART v1 (9). The networks included *C. imicola* mitotypes from this study (South Africa and Corsica), as well as mitotypes detected from the Iberian peninsula and Israel [(3); Portugal n = 12, Israel n = 29; Accession numbers: AF078098–AF078100, AF080531–AF080535, AJ549393–AJ549426]. The Iberian population is known to be free from *Cardinium*, whereas that of Israel is known to carry *Cardinium* (5, 6).

## Results

To confirm individuals in this study were not cryptic species of *C. imicola* missed by morphological identification, a distance estimation of mtDNA barcodes was assessed giving a minimum identity of > 98%, consistent with all individuals belonging to the same species. *Cardinium* infection was observed in the South African population (10/33 individuals) but not in Corsican populations (0/73 individuals) (Table 1). The additional nested screening revealed no evidence of low-titre infections.

Table 1  
*Cardinium* screening results using conventional and nested *GyrB* PCR assays.

Species	Collection site	Year	Number of <i>Cardinium</i> positive individuals by conventional PCR	Number of <i>Cardinium</i> positive individuals by nested PCR
<i>C. imicola</i>	Corsica (Site 1)	2015	0/46	0/46
	Corsica (Site 2)		0/27	0/27
	South Africa (Onderstepoort)	2016	10 /33	10 /33

To investigate associations between symbiont and mitotype, mtDNA haplotypes and the *Cardinium* infection status of *C. imicola* individuals within and between populations were compared. A *C. imicola* mitotype network (Fig. 1A) of the single South African population, containing both *Cardinium*-infected and uninfected individuals showed the presence of seven mtDNA haplotypes split into two broad mitotype groups separated by three SNPs. *Cardinium* infection status and these general mitotype groupings were associated, with all *Cardinium*-infected specimens falling in one group (Fisher's exact test,  $P < 0.001$ ).

Extending the haplotype network to include *C. imicola* populations from the Western Mediterranean Basin (WMB; Corsica and Portugal) and the Eastern Mediterranean Basin (EMB; Israel), led to 12 haplotypes being observed with 0-1.77% range of divergence over a 451 bp region. Further to this, 24/29 Israeli (EMB) haplotypes clustered within the main *Cardinium* infected mitotype observed in South Africa, whereas the Portuguese (WMB) haplotypes all clustered within an uninfected Corsican mitotype.

## Discussion

Insect symbionts are generally considered in the context of driving diverse and extreme host phenotypes, such as reproductive manipulations and "virus-blocking" (10, 11). However, their role in confounding interpretations of mtDNA structure is also of importance (4). As both symbiont and mitochondria are maternally inherited together, a selective sweep of *Cardinium* is likely to have led to our observed concordance of *C. imicola* mitotype and *Cardinium* (Fig. 1A). Null models of symbiont-mtDNA dynamics predict three phases. First, the symbiont sweep, during which the symbiont associated mitotype increases in frequency, and the infected and uninfected individuals have different associated mitotypes. This is followed by the mitotype of infected individuals becoming increasingly represented in uninfected

individuals where the symbiont fails to transmit from mother to offspring. In the final phase, mutation re-establishes diversity in the mitotype pool. Although there is a structure observed between infection status and mtDNA haplotype, the presence of two uninfected individuals in the main infected haplotype in *C. imicola* from South Africa is consistent with a symbiont showing imperfect vertical symbiont transmission, generating uninfected individuals from lineages with the dominant *Cardinium* infected mitotype.

Although the infection status of the Portuguese and Israeli individuals included in the extended haplotype network (Fig. 1B) have not been directly assessed, the recent Pagès (6) screening for *Cardinium* found no indication of *C. imicola* infection in the Iberian Peninsula and we observed no infections in our Corsican population. Despite this, others (5) have observed the symbiont to be common in Israel, indicating infection heterogeneity exists between the EMB and WMB. In light of this, mitotype variation between the EMB and WMB mirrors infection heterogeneity between *C. imicola* populations from the EMB and WMB. Thus, mtDNA structure may be explained by the dispersal of *Cardinium* infected individuals exclusively into the EMB.

The matrilineal subdivision between *Cardinium* infected and uninfected *C. imicola* (Fig. 1B) is similar to observations by previous studies (3, 12) which noted a remarkable divergence between *C. imicola* mitotypes from the EMB and WMB. This observation suggests that the linkage disequilibrium of *Cardinium* and mitochondria reflects symbiont gene flow within the Mediterranean basin but may not assist in elucidating host gene flow. For example, if dispersal events occur involving *Cardinium*-infected *C. imicola*, selective sweeps of the symbiont may erase any previous biodiversity in this marker. This is reminiscent of similar patterns observed in other insects with symbiont infections: the mosquito vector, *Culex pipiens* (13), the parasitoid wasp, *Nasonia vitripennis* (14) and the ladybird, *Adalia bipunctata* (15). In all these cases, the frequency of mtDNA haplotypes are more closely associated with the symbionts in a population rather than geography.

## Conclusions

The haplotype networks produced in this study (Fig. 1) suggest *Cardinium* infection in *Culicoides* is another example where the presence of a symbiont impacts mtDNA based inference of population history. Previous work had indicated Israeli and Southern Africa samples grouped together on analysis of mtDNA variation but were distinct on analysis of microsatellite variation (1). We conclude that association of *Cardinium* and mtDNA provides a likely explanation for this discordance, and these data support the general importance of using multiple loci (e.g. microsatellites and nuclear markers) in phylogeographic analysis.

## List Of Abbreviations

EMB=Eastern Mediterranean Basin

WMB=Western Mediterranean Basin

COI=cytochrome c oxidase subunit I

## Declarations

**Ethics approval and consent to participate:** Not applicable

**Consent for publication:** Not applicable

**Availability of data and materials:** Accession numbers for *COI* sequences generated in this study are LR877434-LR877461.

**Competing interests:** Not applicable

**Funding:** This work was supported by a BBSRC Doctoral Training Partnership studentship (BB/M011186/1) awarded to JP and a Marie Curie Individual Fellowship (H2020-MSCA-IF-2014) grant awarded to SS.

**Authors' contributions:** JP, GDDH, SS and MB and assisted in the conception and design of the study. Field work was undertaken by JP, CG and GV. Lab work was undertaken by JP. Analyses and interpretation of the data were undertaken by JP, GDDH, MB, SS, CG and GV, as well as drafting of the manuscript.

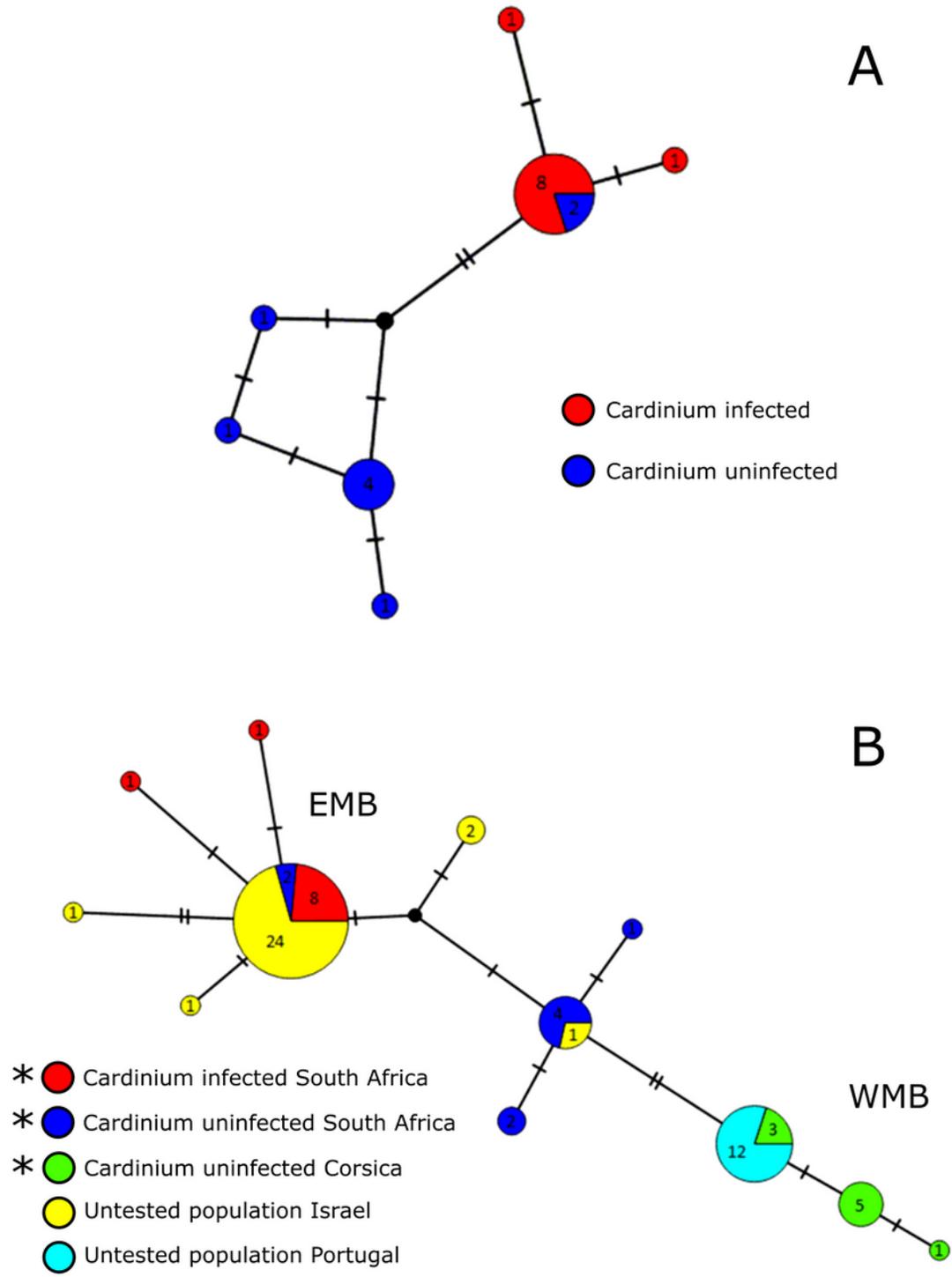
**Acknowledgements:** The authors would like to acknowledge the Onderstepoort Veterinary Institute, Agricultural Research Council, Pretoria, South Africa for allowing the undertaking of field work at their site.

## References

1. Jacquet S, Garros C, Lombaert E, Walton C, Restrepo J, Allene X, et al. Colonization of the Mediterranean basin by the vector biting midge species *Culicoides imicola*: an old story. *Mol Ecol*. 2015;24(22):5707–25.
2. Jacquet S, Huber K, Pagès N, Talavera S, Burgin LE, Carpenter S, et al. Range expansion of the Bluetongue vector, *Culicoides imicola*, in continental France likely due to rare wind-transport events. *Sci Rep*. 2016;6(1):27247.
3. Dallas JF, Cruickshank RH, Linton YM, Nolan D V., Patakakis M, Braverman Y, et al. Phylogenetic status and matrilineal structure of the biting midge, *Culicoides imicola*, in Portugal, Rhodes and Israel. *Med Vet Entomol*. 2003;17(4):379–87.
4. Hurst GDD, Jiggins FM. Problems with mitochondrial DNA as a marker in population, phylogeographic and phylogenetic studies: the effects of inherited symbionts. *Proc Biol Sci*. 2005;272(1572):1525–34.

5. Morag N, Klement E, Saroya Y, Lensky I, Gottlieb Y. Prevalence of the symbiont *Cardinium* in *Culicoides* (Diptera: Ceratopogonidae) vector species is associated with land surface temperature. *FASEB J.* 2012;26(10):4025–34.
6. Pagès N, Muñoz-Muñoz F, Verdún M, Pujol N, Talavera S. First detection of *Wolbachia*-infected *Culicoides* (Diptera: Ceratopogonidae) in Europe: *Wolbachia* and *Cardinium* infection across *Culicoides* communities revealed in Spain. *Parasit Vectors.* 2017;10(1):582.
7. Lewis SE, Rice A, Hurst GDD, Baylis M. First detection of endosymbiotic bacteria in biting midges *Culicoides pulicaris* and *Culicoides punctatus*, important Palaearctic vectors of bluetongue virus. *Med Vet Entomol.* 2014;28(4):453–6.
8. Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek R. DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Mol Mar Biol Biotechnol.* 1994;3(5):294–9.
9. Leigh JW, Bryant D. Popart: Full-Feature Software for Haplotype Network Construction. *Methods Ecol Evol.* 2015;6(9):1110–6.
10. Werren JH, Baldo L, Clark ME. *Wolbachia*: master manipulators of invertebrate biology. *Nat Rev Microbiol.* 2008;6(10):741–51.
11. Bourtzis K, Dobson SL, Xi Z, Rasgon JL, Calvitti M, Moreira LA, et al. Harnessing mosquito-*Wolbachia* symbiosis for vector and disease control. *Acta Trop.* 2014;132(1).
12. Calvo JH, Calvete C, Martinez-Royo A, Estrada R, Miranda MA, Borrás D, et al. Variations in the mitochondrial cytochrome c oxidase subunit I gene indicate northward expanding populations of *Culicoides imicola* in Spain. *Bull Entomol Res.* 2009;99(6):583–91.
13. Rasgon JL, Cornel AJ, Scott TW. Evolutionary history of a mosquito endosymbiont revealed through mitochondrial hitchhiking. *Proc R Soc B Biol Sci.* 2006;273(1594):1603–11.
14. Raychoudhury R, Grillenberger BK, Gadau J, Bijlsma R, van de Zande L, Werren JH, et al. Phylogeography of *Nasonia vitripennis* (Hymenoptera) indicates a mitochondrial-*Wolbachia* sweep in North America. *Heredity.* 2010;104(3):318–26.
15. Schulenburg JHGVD, Hurst GDD, Tetzlaff D, Booth GE, Zakharov IA, Majerus MEN. History of infection with different male-killing bacteria in the two-spot ladybird beetle *Adalia bipunctata* revealed through mitochondrial DNA sequence analysis. *Genetics.* 2002;160(3):1075-86.

## Figures



**Figure 1**

*Culicoides imicola* mtDNA haplotype networks A) mtDNA haplotype network of *Cardinium* infected and uninfected *Culicoides imicola* from a single site at ARC-OVI, Onderstepoort, South Africa; based on a 517 bp COI sequence. B) mtDNA haplotype network of *Culicoides imicola* from sites spanning South Africa and the Mediterranean basin; based on a 451 bp COI sequence. Haplotype networks were generated using the TCS network algorithm in PopART v1.7 (Leigh & Bryant 2015). Numbers within circles represent the

numbers of individuals designated to each haplotype. The numbers of substitutions separating haplotypes are indicated by dashes. EMB=Eastern Mediterranean Basin, WMB=Western Mediterranean Basin \*Indicates mtDNA haplotypes generated in this study.