Controversial patterns of *Wolbachia* infestation in the social parasitic *Maculinea* butterflies (Lepidoptera: Lycaenidae)

Judit Bereczki[1,2], Rita Rácz[1,2], Zoltán Varga[1], János P. Tóth[3]

[1]Department of Evolutionary Zoology and Human Biology, Institute of Biology and Ecology, University of Debrecen, H-4032 Debrecen, Egyetem tér 1, Hungary, [2]MTA-DE 'Lendület' Behavioural Ecology Research Group, H-4032 Debrecen, Egyetem tér 1, Hungary, [3]Research Institute for Viticulture and Oenology, Tokaj (RIVOT), H-3915 Tarcal, Könyves Kálmán utca 54., Hungary.

Corresponding author: Judit Bereczki, Department of Evolutionary Zoology & Human Biology, Institute of Biology & Ecology, University of Debrecen, Debrecen, P.O.Box 3, H-4010, Hungary. Telephone: +36 52 512 900/62526. Fax: +36 52 512 941. E-mail: bereczki.judit@gmail.com

Abstract

Wolbachia is a common group of intracellular bacteria found in arthropods and filarial nematodes. Since the past decade they have attracted considerable interest owing to their various effects on hosts, which range from reproductive manipulation to mutualism. Moreover, they can influence the mitochondrial DNA pattern which do not reflect the real evolutionary history of the target species and may be incongruent with nuclear data. Previously *Wolbachia*-manipulated mtDNA patterns, namely mito-nuclear discordance and deep mitochondrial splits associated with specific *Wolbachia* infections have been also discovered in the genus *Maculinea*. Here we present a comprehensive study on *Wolbachia* infestation and the genetic diversity of all *Maculinea* species in the Carpathian Basin. The prevalence and the pattern of the infestation highly differ among *Maculinea* species. *M. alcon* and *M. arion* are infected in 100%, each of these species with a single strain, but the infection level of *M. nausithous* and *M. teleius* is much lower, additionally, they are infected with multiple strains. The genetic diversity of *Maculinea* species proved to be generally low, only *M. nausithous* showed geographic pattern based on mitochondrial sequences and allozymes. In contrast with the previous studies, we could not detect mito-nuclear discordance or find evidence for *Wolbachia*-induced selective sweep. Based on our results, we cannot hold only *Wolbachia* responsible for the restricted genetic diversity of *Maculinea* in the Carpathian Basin. Probably several factors shape together the level and pattern of genetic variability in *Maculinea* butterflies.

Key words: mitochondrial DNA, cytoplasmic incompatibility, selective sweep, *Wolbachia* MLST

Introduction

In the last half century, genetic markers have become widely used in species delimitation and in the investigation of the phylogenetic relationships among species due to the development of advanced molecular techniques. In the animal kingdom, mitochondrial DNA (mtDNA) is one of the most popular markers because of its many ideal properties (Galtier et al. 2009). Although mtDNA is proved to be suitable for barcoding animal life and highly informative in phylogenetic and phylogeographic studies, mtDNA patterns do not always reflect the real evolutionary history of the studied species and may be incongruent with nuclear data (Dupuis et al. 2012). Low intraspecific mtDNA variation can make it impossible to reveal the phylogeography of focal groups of organisms. Such reduced mtDNA diversity may be caused by an indirect selective sweep of a favoured mtDNA variant, e.g. by *Wolbachia*. These intracellular bacteria infect a wide variety of arthropods and filarial nematodes (Werren et al. 2008) and a recent meta-analysis estimated that more than 65% of insect species harbour *Wolbachia* (Hilgenboecker et al. 2008).

Usually, *Wolbachia* are vertically transmitted from mother to offspring while they can manipulate the host reproduction in ways that enhance their own transmission chance to future generations. Accordingly, *Wolbachia* infections are associated with a variety of phenotypic effects on the hosts, such as cytoplasmic incompatibility (CI, the sperm of infected males is incapable to fertilize the eggs of females which are uninfected or infected with a different strain), male-killing (MK), feminization of genetic males and induction of parthenogenesis (PI) (Werren et al. 2008). Regardless of the specific mechanism by which they manipulate host reproduction, the spread of *Wolbachia* will be accompanied with the concomitant spread of all maternally inherited organelles, including mitochondria, being present in the initially infected female (Turelli and Hoffmann 1991, 1995; Hoffmann et al. 1998). This means that one particular mitotype (and its mutational derivatives) will hitchhike through the infected population in association with the selectively-driven sweep of *Wolbachia* (Turelli and Hoffmann 1991, 1995; Turelli 1994). Consequently, *Wolbachia* may be expected to dramatically affect host mtDNA genome evolution.

Assuming the bacteria are transmitted with very high efficiency between generations, an infected population should have lower mtDNA diversity than an uninfected one since each *Wolbachia* sweep reduces the effective population size of mtDNA to one. At the same time, a single population may be infected with more than one strain which may inflate the mtDNA variability within a population given the association between the particular *Wolbachia* strain and the concomitant host mitotype.

Simultaneously, *Wolbachia* may affect levels of mtDNA differentiation. The mtDNA of infected and uninfected individuals or the individuals infected with different strains will evolve to be distinct, because there is little or no flow of mtDNA between them. Besides, the spread of *Wolbachia* through a single population may increase the mtDNA differentiation between populations if the bacteria are unable to invade neighbouring populations. As reported in the satyrine butterfly *Coenonympha tullia* (Kodandaramaiah et al. 2013) the two divergent mitochondrial clades found in North America are associated with two different *Wolbachia* strains leading to a 'two barcodes – one species' phenomenon. Neither genitalia morphology nor nuclear gene sequence supports cryptic species as an explanation, instead selective sweeps driven by *Wolbachia*. That is, *Wolbachia* can make one species appear as two due to the high intraspecific mtDNA diversity associated with possession of different bacterial strains. On the other hand, *Wolbachia* sweeps can result in minimal mtDNA differentiation if there is only a very small amount of interbreeding and parallel sweeps of mtDNA variants happen in different

populations. The butterfly species pair *Acraea encedon* and *A. encedana* bear the same male-killing *Wolbachia* strain which homogenizes haplotypes of the two species causing a 'one barcode – two species' phenomenon (Jiggins 2003). While these species are clearly distinct on morphological and nuclear DNA grounds, they appear identical on the basis of barcoding sequences of the infected individuals. The most likely explanation for this is that rare hybridization events open the door to the transfer of the male-killer and associated mitotype from species to species.

As a consequence of these widespread effects, *Wolbachia* may interfere with the phylogenetic and phylogeographic studies of hosts. While *Wolbachia* influence the mtDNA variability to a large extent, patterns of variation at nuclear genes should be largely unaffected by the presence of these bacteria leading to mito-nuclear discordance (Jiggins 2003; Shoemaker et al. 2003; Kodandaramaiah et al. 2013).

Wolbachia are also present in the social parasitic *Maculinea* butterflies (Sielezniew et al. 2012; Patricelli et al. 2013; Ritter et al. 2013; Bereczki et al. 2014). Reduced mitochondrial diversity refers to *Wolbachia*-induced selective sweeps in *M. alcon* ([Denis & Schiffermüller], 1775) and *M. arion* (Linnaeus, 1758) (Bereczki et al. 2014; Patricelli et al. 2013; Sielezniew et al. 2012). In addition, the possibility of CI also arises in *M. arion*. In the Carpathian Basin this species exists in two phenological forms ('spring and summer type' according to their flight periods) which co-occur in certain habitats. Bereczki et al. (2011, 2014) revealed that these forms differentiate neither on the basis of barcoding gene nor in allozymes but they are clearly distinct morphologically (based on wing and male genitalia traits) and all individuals were infected with *Wolbachia*. Based on the discordance between mitochondrial and morphological results, it might be assumed that *Wolbachia* make an effect on the speciation of *M. arion* causing a 'one barcode – two species' phenomenon (Bereczki et al. 2014). At the same time, inflated diversity and deep splits in mitochondrial sequences proved to be in correspondence with different *Wolbachia* strains in *M. nausithous* (Bergsträsser, 1779) and *M. teleius* (Bergsträsser, 1779) (Ritter et al. 2013), i.e. *Wolbachia* seems to have an influence on the mitochondrial DNA pattern of *Maculinea* hosts in different ways.

Here we present a comprehensive study on the *Wolbachia* infestation and the genetic diversity of all *Maculinea* species in the Carpathian Basin. Our main aim is to investigate the prevalence and infestation patterns of *Wolbachia* in each species and reveal their possible effects on hosts' genetic diversity and speciation using mitochondrial and nuclear gene sequences as well as allozyme loci. Therefore, our study provides new information to interpret the genetic patterns of these long-surveyed butterflies.

Materials and Methods

Samples

We sampled all *Maculinea* species in 18 geographical regions of the Carpathian Basin from 2001 to 2013 (Fig. 1a, b; Online Resource 1). The identification of the two ecotypes of *M. alcon* was carried out based on the initial food plants and the date of sampling. The 'spring' and 'summer type' of *M. arion* were also separated phenologically. In certain habitats two or even three *Maculinea* species were present. Imagoes were collected at the end of the egg laying period and stored at –80°C until molecular analyses.

Wolbachia screening and strain identification

DNA was extracted by homogenising the heads or thoraxes of butterflies following the protocol in Bereczki et al. (2014). Altogether 506 individuals were screened (1-11 individuals/population) from 18 geographical regions (Online Resource 1) by the amplification of the highly conservative 16S ribosomal RNA gene with *Wolbachia* specific primers W-Spec of Werren and Windsor (2000). Amplification from 1 µl of DNA extracts was carried out in 25 µl final reaction volumes containing 5X PCR buffer, 0.2 mM dNTPs, 2 mM $MgCl_2$, 0.5 mg/ml bovine serum albumin, 0.02 units/µl of Taq DNA polymerase (Phusion Hot Start II High-Fidelity, Thermo Scientific) and 0.3 µM of each primer. Amplification was carried out in an ABI Veriti thermal cycler programmed for: initial denaturation for 2 min at 98°C; 40 cycles of 15 s at 98°C, 45 s at 60°C, 1 min 30 s at 72°C; final elongation of 10 min at 72°C. Amplification procedure was repeated twice in *M. teleius* and *M. nausithous* since we could obtain faint bands as a result of the first 40 cycles. We used positive (surely infected samples) and negative controls (master mix without any DNA sample) in each reaction. The success of PCRs was checked by running 2 µl of product on 1% agarose gels stained with GelRed Nucleic Acid Stain (Biotium Inc.). The infection rates of the populations were plotted on geographic maps using QGIS v. 2.6 (2014). PCR-products originating from 70 individuals (Online Resource 2) were sequenced by commercial service provider Macrogen Inc. (Seoul, South-Korea). Sequences were edited and revised manually by Chromas Lite v. 2.01 and aligned by MEGA v. 6 (Tamura et al. 2013). Bayesian phylogenetic relationships were assessed in MrBayes v. 3.2.2 (Ronquist et al. 2012) as described below.

Wolbachia strain identification was carried out in *M. alcon* and *M. arion* according to *Wolbachia* MLST (=Multilocus Sequence Typing) system (Baldo et al. 2006) by the amplification of *Wolbachia* surface protein (WSP) and five conserved gene regions (gatB, coxA, hcpA, ftsZ, fbpA) following the PCR protocol mentioned above. The latter five genes determine together the sequence type (ST). WSP typing was achieved in a single individual each in *M. alcon* and *M. arion*. *Wolbachia* MLST was completed altogether in 32 specimens in the two species (Online Resource 2). After sequencing we defined the strains using the reference sequences of *Wolbachia* MLST database (http://pubmlst.org/wolbachia/) and searched the identified sequence types in other organisms in the mentioned database.

Phylogenetic analyses of mitochondrial and nuclear DNA

The same DNA extracts were used to amplify mitochondrial and nuclear genes as those in *Wolbachia* studies. We sequenced the mitochondrial cytochrome c oxidase subunit I gene (COI) together with the nuclear gyceraldehyde 3-phosphate dehydrogenase (GAPDH), malate dehydrogenase (MDH) and wingless (wg) (Online Resource 2) using commercial service provider Macrogen Inc. These genes were amplified by specific primers modified at their 5'-end to include the universal sequencing primer T7promoter (Wahlberg and Wheat 2008). We followed the above amplification protocol and the guidelines of the Nymphalidae Systematic Group (http://nymphalidae.utu.fi/). After the revision and alignment of sequences we estimated the level of genetic diversity by the following parameters: the number of variable and informative sites (V and PI), haplotype number and diversity (h and HD) (Librado and Rozas 2009), nucleotide diversity (π) (Nei and Li 1979; Nei 1987; Tajima 1983), the number of segregating sites (θ_w) (Watterson 1975; Nei 1987). In cases of nuclear genes PHASE haplotype reconstruction option was used because of the presence of ambiguous sites (Stephens et al.

2001; Stephens and Donnelly 2003). Neutrality tests were performed on all genes to determine the departure from the neutral model of molecular evolution. Under neutrality, both Tajima's D and Fu & Li's D are expected to be zero. Positive D values indicate an excess of intermediate-frequency variants and can be due to the operation of natural selection. On the contrary, a value significantly less than zero indicates a higher-than-expected number of low frequency variants and might be the consequence of a recent selective sweep or processes such as background selection (Tajima 1989; Aris-Brosou and Excoffier 1996). The parameters of gene diversity and neutrality tests were calculated using the program DnaSP v. 5.10.01 (Librado and Rozas 2009).

Based on the concatenated sequences, Bayesian analyses were carried out using MrBayes 3.2.2 (Ronquist et al. 2012). The combined data were partitioned by gene and analyzed as independent partitions. Several possible models of molecular evolution were sampled for each gene (both single and combined data) during the analysis using the model-jumping feature of MrBayes v. 3.2.2 applying the following sets "lset applyto = (all) nucmodel=4by4 nst=mixed rates=gamma covarion=no;". Two independent MCMC runs each with four simultaneous chains (one cold and three heated) were run for 10 million generations. The trees were sampled in every 1000 generations. Convergence of the two runs was determined by the stationary distribution plot of the log likelihood values against number of generations and confirmed by the average standard deviation of split frequencies, which were lower than 0.05 in all cases. The first 2,500,000 generations were discarded as burn-in and trees were summarized under the 50 percent majority rule method.

Haplotype networks were built in R computing environment (R Development Core Team 2014) with pegas package v. 0.6 (Paradis 2010) based on the mitochondrial COI sequences in *M. nausithous* and *M. teleius*. In this analysis our samples (55 individuals, see in Online Resource 2: HG326644, KM517275-KM517328) and the sequences from GenBank (JX311050-JX311307, Ritter et al. 2013) were used. First the original alignment was reduced to get the maximal overlapping sequences using MEGA v.6 (Tamura et al. 2013). Then we deleted those specimens from the alignment which had ambiguous sites in their sequences. After this process, 101 specimens with 984 bp remained in *M. nausithous* and 158 specimens with 850 bp in *M. teleius*. Haplotype networks were created using the reduced datasets in both species.

Allozyme studies

Allozyme polymorphism was studied at 11 loci by vertical polyacrylamide gel electrophoresis altogether in 1266 individuals of four *Maculinea* species (Online Resource 1). Thoraxes homogenized in 300 µl of extraction buffer were used to study *Gpdh*, *G6pgdh*, *Hk*, *Idh*, *Mdh*, *Pgi*, *Pgm* and *Sod*. Abdomens homogenized in 200 µl of extraction buffer were used to analyse *Acon*, *Aox* and *Est*. The extraction buffer, the electrophoresis buffer systems and running conditions, together with the staining solutions were applied as described in Bereczki et al. (2005).

Genotypes of the individuals were scored according to their enzyme pattern. Genotype and allele frequencies were calculated on the basis of banding patterns. Measures of genetic variation, i.e. average number of alleles per locus (N_a), the effective number of alleles (N_e), Shannon's information index (I), average observed heterozygosity (H_o) and proportion of polymorphic loci on the basis of the 95% criterion (P%) were calculated for each species using GenAlEx v. 6.41 (Peakall and Smouse 2006).

Allele frequencies were used to estimate Cavalli-Sforza and Edwards (1967) chord distances, and a UPGMA dendrogram (Sneath and Sokal 1973) was constructed on the basis of the distance matrix using Past v. 2.17

(Hammer et al. 2001). In the case of *M. alcon*, a principal component analysis (PCA) was conducted using the allele frequency data of the individuals to investigate the basis of the differentiation between the two types using different food plants ('cruciata and pneumonanthe type'). PCA analysis was performed using Past v. 2.17 (Hammer et al. 2001).

The population genetic structure was also analysed by Bayesian-clustering method (Pritchard et al. 2000). Here we estimated the most probable number of genetically differentiated groups (K) in our populations and assigned the individuals to these groups. Structure v. 2.3.4 was run to carry out these analyses with initial burn in 100,000 and running length 500,000. In the evaluation of the results ΔK was computed which indicates the change in log probability between successive K values (Evanno et al. 2005). Structure Harvester Web v. 0.6.93 (Earl and vonHoldt 2012) was used to compute the ΔK values.

Results

Wolbachia infection

The prevalence of *Wolbachia* was 100% both in *M. alcon* and *M. arion* irrespective of phenology or differential food plant usage while the infection level was 36.2% in *M. nausithous* and 14.4% in *M. teleius* (Online Resource 1; Fig. 1a, b). In *Wolbachia* infestation of the latter two species we could not detect any geographical pattern (Fig. 1b).

All *M. alcon* individuals were infected with a single *Wolbachia* strain which belonged to supergroup B (Table 1; Fig. 2). The WSP allele No. 575 and the sequence type (ST) No. 235 were identified in *M. alcon*. Similarly, all *M. arion* individuals were infected with only one strain which belonged to supergroup A (Table 1; Fig. 2). A new WSP allele was identified from this species and submitted to the *Wolbachia* MLST database (No. 685). The sequence type (ST) No. 403 was found in *M. arion*.

M. nausithous harboured various 16S rRNA sequences belonging to both *Wolbachia* supergroups A and B (Fig. 2; Table 2). Additionally, highly divergent sequences occurred even in a single population of this species. Similarly, *M. teleius* contained different 16S rRNA sequences belonging to both supergroups A and B (Fig. 2; Table 2). The parameters of 16S rRNA gene diversity were proved to be higher in *M. teleius* than in *M. nausithous* (Table 2). Additionally, neutrality tests indicated significantly positive departure from the neutral model of molecular evolution.

The genetic diversity of *Maculinea* species

Mitochondrial DNA

The final COI alignment from the Carpathian Basin contained 110 sequences (Table 2, Online Resource 2) with a total length of 1369 bases out of which 60 (4.4%) sites were variable and 45 (3.3%) were parsimony informative. The highest number of these sites was detectable in *M. nausithous*. In total, 29 unique haplotypes were observed out of which 11 were included in *M. teleius*. At the same time, *M. alcon* 'pneumonanthe type' involved only one mitotype similarly to *M. arion* 'spring type' which also included only a single mitochondrial haplotype. Nucleotide diversity and the number of segregating sites were very low and varied from 0 to 0.00476 and from 0 to 0.00671, respectively (Table 2). The highest measures were found in *M. nausithous*. Neutrality

tests showed negative D values almost in all cases but significant deviation from the neutral model of molecular evolution was experienced only in *M. nausithous* where Fu & Li's D was positive (Table 2).

Based on the mitochondrial COI sequences, not all *Maculinea* species differentiated from each other with high statistical support (Fig. 3a). Interestingly, the split of *M. nausithous* and *M. teleius* was practically without statistical support (posterior probability: 0.55). In *M. nausithous* strong geographical pattern was recognized. The differentiation of the three large regions (Western Hungarian, Transylvanian and Eastern Carpathian) was highly pronounced. Additionally, the populations of the Eastern Carpathian clade also diverged geographically but reduced mitochondrial diversity was experienced in the other two regions (Western Hungarian and Transylvanian). At the same time, the resolution of COI sequences in the other three *Maculinea* species proved to be very low. *M. nausithous* and *M. teleius* specimens infected with *Wolbachia* did not form separate clades.

The mitochondrial haplotype network which included also the sequences of Ritter et al. (2013) showed five main haplotype groups in *M. nausithous* (Online Resource 5a). The central group involved mostly Asian individuals and specimens from Poland and Bulgaria. The Hungarian, Slovenian and Transylvanian individuals as well as the majority of Croatian specimens formed a distinct haplotype group in the Carpathian Basin. The Western European individuals constituted two independent groups separated from each other by numerous mutation steps. One of these groups involved the majority of Western German specimens and certain French individuals (Western Europe 2 in Online Resource 5a) linked to the so-called *Wolbachia* clade although they were differentiated from it by numerous mutations, and Ritter et al. (2013) could not detect current infection in this group. *Wolbachia* clade included mostly but not exclusively infected individuals from Croatia, Russia and the Eastern Carpathian region of Romania but none of the specimens from Hungary and Slovenia despite the fact that these populations are mostly infected with *Wolbachia* (Fig. 1b). In the case of *M. teleius*, there is a central haplotype group which contained the majority of individuals both from Europe and Asia (Online Resource 5b). Certain individuals from various geographical regions were separated from this central group by one or two mutation steps (and three in a single case). Only the Chinese and the majority of Japanese specimens differed from the central set of haplotypes by numerous mutations. A *Wolbachia*-dominated clade which involved Asian specimens was also recognizable in this species although several infected individuals can be found in the central group as well.

Nuclear DNA

Altogether 1680 bases were studied from three different gene regions (GAPDH, MDH and wingless) in 58 individuals (Table 2, Online Resource 2). In general, the diversity of these DNA fragments was lower than that of mitochondrial sequences. The highest variability was detected in *M. arion* and the lowest in *M. teleius*. Neutrality tests did not indicate significant departure from the neutral model in any of the species.

Based on the concatenated nuclear sequences, all *Maculinea* species differed from each other perfectly (Fig. 3b). Some pattern was experienced only in *M. nausithous* but it did not correspond with geographical regions. The infected individuals clustered into separate clades neither in this case. The resolution of the different nuclear regions did not differ from each other considerably although wingless gene seemed to be the less suitable for the reconstruction of phylogeography (Online Resource 3 a, b, c).

Allozymes

The allozyme variability of *Maculinea* species was generally low. *M. arion* exhibited the highest level of polymorphism while *M. teleius* did the lowest level (Table 3).

All *Maculinea* species clearly differed from each other based on allozyme loci as well (Fig. 4a). Allozymes showed highly homogeneous pattern in *M. arion* and *M. teleius* (Online Resource 4). At the same time, both the dendrogram and the Bayesian analysis of the genetic structure indicated strong differentiation among the large geographical regions in *M. nausithous* similarly to mitochondrial sequences (Fig. 4a, Online Resource 4). The pattern emerging in *M. alcon* corresponded with different food plant usage but this correspondence was not entirely clear (certain 'pneumonanthe type' populations, i.e. ALpt-Dra and ALpt-Gyk clustered into the 'cruciata type' populations, see in Fig. 4a). Simultaneously, principal component analysis clearly showed that the 'cruciata and pneumonanthe type' populations were separated along the first axis which explained 46.9% of the total variance and the locus esterase showed the largest loading in this axis (Fig. 4b).

Discussion

In this paper, we provide a comprehensive view about *Wolbachia* infection and the genetic diversity of *Maculinea* butterflies in the Carpathian Basin. The prevalence and the pattern of the infestation highly differ among *Maculinea* species. *M. alcon* and *M. arion* are infected in 100%, each of these species with a single strain. Similarly, Sielezniew et al. (2012) reported a 100% prevalence of a single strain in the Polish and Lithuanian populations of *M. alcon*. Simultaneously, all studied individuals in the Polish and Italian populations of *M. arion* proved to be infected with one strain (Sielezniew et al. 2012; Patricelli et al. 2013). The *Wolbachia* 16S rRNA gene sequences were identical in all studied European populations in *M. alcon* and in *M. arion*, separately, but the two types of strains harboured by these species did not mix with each other in any particular case probably because there is cytoplasmic incompatibility between them. Interestingly, 19 populations of *Euphydryas aurinia* in the United Kingdom (Smee 2011) proved to be also infected in 100% with the same strain as *M. alcon* in our study. Nevertheless, the phenotype of this *Wolbachia* strain has remained unknown in both host species.

The establishment of the *Wolbachia*' effects on the hosts strongly requires experimental confirmation since one particular strain may cause the same (Dyson et al. 2002) or different (Fujii et al. 2001) reproductive alterations in different hosts. Namely, the phenotype seems to be dependent on interactions between the bacteria and the genomic make-up of the individual host. Without any experimental confirmation, we can only hypothesise the potential phenotypes of the strains based on the infection rate and the genetic diversity of host species. The infestation pattern, i.e. 100% prevalence of a single strain each in *M. alcon* and *M. arion* refers to the perfect vertical transmission of *Wolbachia* in these hosts. Vertically transmitted endosymbionts are generally considered to be evolved towards mutualism because their evolutionary fate is closely linked to that of their hosts (Fine 1975; Yamamura 1993; Zug and Hammerstein 2014). Probably these particular strains were able to overcome these hosts' immune system and rule out the other strains competitively through cytoplasmic incompatibility.

The effects of *Wolbachia* on these hosts' gene diversity seem to be variable in the different European regions. The Polish and Lithuanian populations of *M. alcon* have a complete lack of mitochondrial sequence variation.

Moreover, barcoding gene sequences proved to be identical from Western Europe to Kazakhstan and Kyrgyzstan (Sielezniew et al. 2012). The Polish populations of *M. arion* also have restricted mtDNA variation. However, mitochondrial haplotype diversity varies according to geographical regions in Italy. Generally, Italian populations share variable mitochondrial haplotypes apart from some Northern and Central Italian populations which have reduced diversity (Patricelli et al. 2013). In both species, the nuclear elongation factor 1α characterised by low substitution rate showed higher variability than the mitochondrial sequences whose evolutionary speed is commonly known to be high. On the contrary, in the Carpathian Basin nuclear variation based on 1680 bases from three different gene regions (GAPDH, MDH and wingless) was lower even than the restricted mitochondrial diversity thus the nuclear data were uninformative to reconstruct the phylogeography of these species. In addition, neutrality tests did not result in significantly negative values in any of these species. Therefore, we could not find evidence for *Wolbachia*-induced selective sweeps which arose in the previous studies in spite of the presence of a single strain and the highly reduced mitochondrial variation.

Additionally, it seems that not *Wolbachia* are responsible for the controversial patterns of mitochondrial and morphological variability of the different forms of *M. arion*. The results of the present study do not support the 'one barcode – two species' hypothesis. On the one hand, the allozymes do not show differentiation between the phenological forms. Although allozyme diversity of *Maculinea* species is generally low compared with other lycaenids (Schmitt et al. 2003; Schmitt and Hewitt 2004; Aagaard et al. 2002), it is suitable for species separation in this genus (Pecsenye et al. 2007). On the other hand, although these forms harbour the same strain, we could not detect mito-nuclear discordance and significant deviation from the neutral model of molecular evolution. One further possibility to explain the differentiation between the 'spring and summer type' of *M. arion* is their differential host ant usage but we do not possess any data about it in the Carpathian Basin. Thus, there is an urgent need of information about host ants on this geographic scale. The two forms of *M. alcon* seem to differentiate only based on their initial host plant usage since the 'cruciata and pneumonanthe type' populations show a strong trend to differ on the basis of esterase locus which is known to be related to adaptation to host plants. In *Colias* butterflies, esterase-D allozymes co-vary with food plant use, suggesting a role in detoxifying plants' chemical defences (Burns 1975).

The infection level of *M. nausithous* and *M. teleius* is much lower than in *M. alcon* and *M. arion*. On the one hand, this means lower infection rate in both species. On the other hand, probably fewer *Wolbachia* are present in one individual because we could amplify the highly conservative 16S rRNA gene only in double PCRs. Additionally, these hosts – even in a single population of *M. nausithous* – are infected with multiple strains belonging to supergroup A or B. This infestation pattern refers to incomplete vertical transmission of *Wolbachia* in these host species and extensive horizontal transfer in/between them. Although several potential horizontal transfer routes can exist in all *Maculinea* species, e.g. transfer by *Myrmica* host ants being in regular physical contact with caterpillars (ant workers feed them by regurgitation or caterpillars feed on the ant brood) or transfer by parasitoids (in certain cases the host's immune system may neutralise the parasitoid larva, see Davies and Vinson 1986; Hu et al. 2003), the lifecycle of *M. teleius* and *M. nausithous* is closely related to each other in many respects. Females in both species lay their eggs on *Sanguisorba officinalis*. Besides, they can use the same host ants and even the same ant nest may harbour *M. teleius* and *M. nausithous* individuals since they often occur in the same habitat (Tartally 2008). Additionally, caterpillars can be parasitized by the same wasp species, e.g. *Neotypus melanocephalus* which infects the young larvae on the food plant (Shaw et al. 2009). Namely,

extensive horizontal transfer routes can exist to mediate *Wolbachia* in both species (and between them) which corresponds with the presumably incomplete vertical transmission. In the conventional view, horizontal transmission favours parasitism (Anderson and May 1982) but we also need experimental studies to identify the phenotype of *Wolbachia* in these hosts.

As it is indicated by the significantly positive values of the neutrality tests in the case of *Wolbachia* 16S rRNA gene in these host species, there is likely to be strong selection on *Wolbachia* themselves, possibly exerted by host defensive mechanisms. *Wolbachia* strains can differ in their ability to transfect different hosts and not all host species are equally permissive (Werren et al. 2008). *M. teleius* and *M. nausithous* seem to be more resistant to *Wolbachia* than *M. alcon* and *M. arion* which is inferred from their much lower infection rate and the attempted invasion of multiple strains into these hosts. Probably different strains try to colonise these hosts and they could compete with each other.

Previously, cryptic speciation has been hypothesised for *M. teleius* and *M. nausithous*, based on deep mitochondrial split in each of these species (Ugelvig et al. 2011; Pech et al. 2004; Als et al. 2004). Ritter et al. (2013) tested the theory of cryptic speciation on a comprehensive sample across the Palaearctic ranges and revealed that deep mitochondrial divergence did not correspond with microsatellite data but was concordant with *Wolbachia* infection in both species. Haplotypes previously attributed to cryptic species were part of *Wolbachia*-infected clades thus deep intraspecific divergences found in DNA barcode studies coincide with specific infection patterns. In contrast with these results, in the Carpathian Basin mitochondrial haplotypes of specimens infected with *Wolbachia* formed separate clades neither in *M. nausithous* nor in *M. teleius*. Although the genetic diversity was the highest in *M. nausithous*, the phylogenetic and allozyme patterns mainly reflect the biogeographical history of the species since the great part of the variability arises from the differentiation of the large geographic regions which coincides the disjunct distribution of *M. nausithous*. Namely, *M. nausithous* only occurs on the Western and the Eastern edges of the Carpathian Basin with a large distributional gap in the centre where the species is missing. Interestingly, we could not detect this pattern in the nuclear genome maybe because of data deficiency (we could not amplify nuclear sequences from the Eastern Carpathian clade at all because of DNA degradation). The structured haplotype network reflects complex phylogeographical pattern in *M. nausithous*. Beside the various Asian haplotypes, the European populations are also divided to distinct haplogroups which suggest that the species survived during Pleistocene ice ages within different European glacial refugia. At the same time, one of the Western European haplogroups is interposed between the Asian and *Wolbachia*-dominated clade. Although, Ritter et al. (2013) could not detect current infection in this group (Western Europe 2), it is possible that the pronounced genetic distance between the two Western European haplogroups is the result of a historical infection. Therefore, it seems that biogeographical factors and *Wolbachia* infestation could shape together the genetic structure of *M. nausithous*. Simultaneously, the diversity of *M. teleius* was very low on the basis of all the studied genetic marker sets and neutrality tests did not lead to significantly negative results. The haplotype network that includes specimens from the whole Palearctic shows the higher diversity of the species in Eastern Asia. This elevated genetic variability refers to isolated refugial areas in this region owing to the complex topography. Simultaneously, the considerable differentiation of *Wolbachia*-dominated clade indicates the effect of the infestation on the genetic diversity of *M. teleius* although further studies need to discover the exact effects of *Wolbachia* on this host.

In summary, our research provides a comprehensive description of the *Wolbachia* infestation and the genetic diversity of *Maculinea* butterflies in the Carpathian Basin. Based on our results, we cannot hold *Wolbachia* responsible for the highly reduced genetic diversity of the target species exclusively. Probably several factors shape together the level and pattern of genetic variability in *Maculinea* butterflies, e.g. biogeography and/or population dynamics. *Wolbachia* are only one of these factors but their importance is very high not only in terms of evolutionary but also conservation biology. A case study about North American *Lycaeides* butterfly species emphasises that *Wolbachia* has to be considered before introducing individuals from one population to another because they may cause failure of introduction attempts due to CI (Nice et al. 2009). Introducing individuals from a source population infected with *Wolbachia* to another population infected with another strain can destroy the population that we wanted to save originally. *Maculinea* species are endangered and their reintroduction has been also achieved several times (Wynhoff 1998; Andersen et al. 2014; Thomas et al. 2009). Therefore, we cannot work out and execute reintroduction action plans responsibly without considering *Wolbachia* infection.

Acknowledgements

The study was supported by the Hungarian Scientific Research Fund (OTKA K109223 and K84071), János Bolyai Scholarship of the Hungarian Academy of Sciences, EU EESD-ENV-2OO2-NAS (Mac Man) project and Social Renewal Operational Programme (TÁMOP-4.2.2/B-10/1-2010-0024). Grateful acknowledgements are due to S. Szabó, A. Tartally, T. Korompai, Z. Ilonczay, A. Ambrus, L. Peregovits, L. Rákosy, F. Rebousek and R. Verovnik for collecting samples. The technical assistance of V. Mester, T. Tóth, K. Kecsmár, H. Boros and B. Prokaj in the molecular work is very much respected. The support of the Nature Conservation Authorities is also greatly appreciated.

Table 1 *Wolbachia* Multilocus Sequence Typing in *M. alcon* and *M. arion*. Genes: WSP – *Wolbachia* surface protein, gatB – aspartyl/glutamyl-tRNA(Gln) amidotransferase, subunit B, coxA – cytochrome c oxidase, subunit I, hcpA – conserved hypothetical protein, NCBI COG0217, ftsZ – cell division protein, fbpA – fructose-bisphosphate aldolase. The length of sequences is in parentheses. ST – sequence type based on five conserved gene regions.

	M. alcon	***M. arion***
WSP (465/513 bp)	575	685 (new allele)
gatB (369 bp)	9	14
coxA (402 bp)	144	203
hcpA (444 bp)	91	238
ftsZ (435 bp)	36	185
fbpA (429 bp)	221	394
ST	235	403

Table 2 – The parameters of DNA variability. N – the number of sequences; V – the number of variable sites; PI – the number of parsimony informative sites; h – the number of haplotypes; HD – haplotype (gene) diversity; π – nucleotide diversity (per site); θ_w – Theta (per site) from the number of variable sites; Neutrality tests – Tajima' D, Fu & Li's D, significance: * $p < 0.05$; ** $p < 0.02$.

	mitochondrial	nuclear (1680 bp)			*Wolbachia*
	COI (1369 bp)	**GAPDH** (663 bp)	**MDH** (648 bp)	**wg** (369 bp)	**16S rRNA** (382 bp)
Maculinea alcon					
N	27	30	38	38	20
V/PI	4/1	1/1	5/2	1/1	0/0
h/HD	5/0.442	2/0.129	6/0.471	2/0.102	1/0.000
π	0.00039	0.00019	0.00088	0.00028	0.00000
θ_w	0.00076	0.00038	0.00183	0.00065	0.00000
Tajima' D	-1.27896	-0.76373	-1.34375	-0.82551	-
Fu & Li's D	-1.90894	0.59448	-1.63150	0.56879	-
M. alcon cruciata type					
N	20	20	26	26	6
V/PI	4/1	1/1	2/1	1/0	0/0
h/HD	5/0.558	2/0.189	3/0.280	2/0.077	1/0.000
π	0.00051	0.00029	0.00045	0.00021	0.00000
θ_w	0.00082	0.00043	0.00081	0.00071	0.00000
Tajima' D	-1.11111	-0.59155	-0.95988	-1.15559	-
Fu & Li's D	-1.69308	0.64952	-0.68907	-1.63377	-
M. alcon pneumonanthe type					
N	7	10	12	12	14
V/PI	0/0	0/0	4/2	1/0	0/0
h/HD	1/0.000	1/0.000	5/0.758	2/0.167	1/0.000
π	0.00000	0.00000	0.00173	0.00045	0.00000
θ_w	0.00000	0.00000	0.00204	0.00090	0.00000
Tajima' D	-	-	-0.53993	-1.14053	-
Fu & Li's D	-	-	-0.45895	-1.32974	-
Maculinea arion					
N	25	18	14	18	16
V/PI	6/2	3/1	4/1	4/3	0/0
h/HD	5/0.477	3/0.307	5/0.505	5/0.673	1/0.000
π	0.00059	0.00065	0.00122	0.00271	0.00000
θ_w	0.00116	0.00132	0.00194	0.00315	0.00000
Tajima' D	-1.47260	-1.40138	-1.22200	-0.42024	-
Fu & Li's D	-1.74111	-1.19315	-1.41428	0.21103	-
M. arion spring type					
N	9	4	4	4	4
V/PI	0/0	0/0	0/0	2/0	0/0

h/HD	1/0.000	1/0.000	1/0.000	2/0.500	1/0.000
π	0.00000	0.00000	0.00000	0.00271	0.00000
θ_w	0.00000	0.00000	0.00000	0.00296	0.00000
Tajima' D	-	-	-	-0.70990	-
Fu & Li's D	-	-	-	-0.70990	-
M. arion summer type					
N	16	14	10	14	12
V/PI	6/2	3/1	4/1	4/2	0/0
h/HD	5/0.667	3/0.385	5/0.667	5/0.725	1/0.000
π	0.00087	0.00083	0.00164	0.00283	0.00000
θ_w	0.00132	0.00142	0.00218	0.00341	0.00000
Tajima' D	-1.16710	-1.27826	-0.94297	-0.56007	-
Fu & Li's D	-1.36562	-1.03687	-1.12706	-0.55476	-
Maculinea nausithous					
N	32	26	26	26	25
V/PI	37/37	5/1	1/1	0/0	24/14
h/HD	8/0.817	6/0.412	2/0.443	1/0.000	6/0.641
π	0.00476	0.00078	0.00068	0.00000	0.00775
θ_w	0.00671	0.00198	0.00040	0.00000	0.01293
Tajima' D	-1.05174	-1.70919	1.12990	-	-1.29057
Fu & Li's D	1.78939**	-2.24127	0.61208	-	1.72242**
Maculinea teleius					
N	26	26	34	34	9
V/PI	13/5	1/0	2/1	1/0	19/18
h/HD	11/0.735	2/0.077	3/0.169	2/0.059	6/0.784
π	0.00137	0.00012	0.00027	0.00016	0.02009
θ_w	0.00257	0.00040	0.00075	0.00066	0.01301
Tajima' D	-1.78318	-1.15559	-1.27659	-1.13783	2.09187*
Fu & Li's D	-1.53360	-1.63377	-0.77789	-1.72294	1.54015**
Total					
N	110	100	112	116	70
V/PI	60/45	29/23	23/18	13/11	43/32
h/HD	29/0.910	13/0.803	16/0.829	10/0.770	10/0.650
π	0.03350	0.01298	0.00821	0.00972	0.01627
θ_w	0.01866	0.00845	0.00670	0.00661	0.01241
Tajima' D	1.99896	1.62704	0.65680	1.24070	0.89898
Fu & Li's D	1.48974	-0.11654	-0.25446	0.28246	1.89932**

Table 3 – The parameters of allozyme variability based on 11 loci. N = sample size; N_a = average number of alleles per locus; N_e = the number of effective alleles = 1 / (Sum pi^2); I = Shannon's information index = -1* Sum (pi * Ln (pi)); H_o = observed heterozygosity = No. of Hets / N; P% = percentage of polymorphic loci.

	N	**N_a**	**N_e**	**I**	**H_o**	**P%**
Maculinea alcon	22.019	1.699	1.320	0.268	0.139	45.932
M. alcon cruciata type	23.063	1.776	1.331	0.281	0.143	48.252
M. alcon pneumonanthe type	19.758	1.530	1.294	0.241	0.128	40.907
Maculinea arion	15.397	1.959	1.379	0.350	0.199	60.333
M. arion spring type	16.500	2.023	1.349	0.346	0.210	61.365
M. arion summer type	14.766	1.922	1.396	0.353	0.193	59.743
Maculinea nausithous	17.936	1.786	1.333	0.281	0.148	45.988
Maculinea teleius	24.376	1.745	1.151	0.202	0.097	63.638
Total	20.295	1.780	1.293	0.270	0.142	52.790

Data Accessibility

DNA sequences: GenBank accessions: HG326619-HG326646, KM517249-KM517565
Sampling locations and data uploaded as online Supporting Information.
Allozyme genotypes: Dryad doi: 10.5061/dryad.v8q8r
Wolbachia WSP and MLST sequence information uploaded into *Wolbachia* MLST database.

Electronic Supplementary Material
Additional supporting information may be found in the online version of this article at the publisher's web-site.

Online Resource 1 Sampling sites. Abb. – abbreviations of localities. Region – the geographic region from which the samples originate. The table includes sampling time, the number of *Maculinea* individuals tested for *Wolbachia* in parentheses and the number of individuals used in allozyme studies in bold.

Online Resource 2 GenBank accession numbers of the sequenced mitochondrial and nuclear genes of *Maculinea* butterflies as well as 16S rRNA genes of *Wolbachia* found in *Maculinea* species. ID – individual identifiers of *Maculinea* specimens. The infected individuals are indicated with W. Specimens from which *Wolbachia* strain identification was carried out are indicated with '+'. WSP – *Wolbachia* surface protein, MLST – Multilocus Sequence Typing based on five genes (gatB, coxA, hcpA, ftsZ, fbpA).

Online Resource 3 Results of Bayesian inference analyses of single nuclear gene datasets in *Maculinea*. **(a)** GAPDH, **(b)** MDH, **(c)** wg.

Online Resource 4 Results of the Bayesian-clustering Structure analysis based on 11 allozyme loci. The most probable K values are indicated in the case of each *Maculinea* species. See the abbreviations of populations in Online Resource 1. *M. alcon*: 'pneumonanthe type' populations are in bold. *M. arion*: 'spring type' populations are in bold. Populations belonging to the same geographic region are joined together on the upper part of each barplot.

Online Resource 5 COI haplotype networks for *Maculinea nausithous* **(a)** and *M. teleius* **(b)** including the sequences of Ritter et al. (2013). Colours show the geographic regions from where haplotypes originate. Circle size is proportional to haplotypes frequency. Dots on lines linking haplotypes indicate the number of mutations. A separate *Wolbachia* clade is present in both species but there are numerous infected individuals outside this clade.

Figure Legends

Fig. 1 Sampling sites. See the abbreviations in Online Resource 1. **(a)** *M. alcon* and *M. arion*. The infection rate is 100% in both species. **(b)** *M. nausithous* and *M. teleius*. The infection rate is indicated with the shade of symbols representing populations. Darker symbols show higher infection rate.

Fig. 2 The result of the Bayesian inference analysis with posterior probability values based on *Wolbachia* 16S ribosomal RNA gene sequences obtained from different *Maculinea* species. ALct – *M. alcon* 'cruciata type', ALpt – *M. alcon* 'pneumonanthe type', ARsp – *M. arion* 'spring type', ARsu – *M. arion* 'summer type', NA – *M. nausithous*, TE – *M. teleius*. See the abbreviations in Online Resource 1 and 2.

Fig. 3 (a) The result of the Bayesian inference analysis with posterior probability values based on mitochondrial COI sequences, **(b)** Consensus phylogeny from the Bayesian inference analysis with posterior probability values based on the combined dataset of three nuclear regions (GAPDH, MDH and wg). ALct – *M. alcon* 'cruciata type', ALpt – *M. alcon* 'pneumonanthe type', ARsp – *M. arion* 'spring type', ARsu – *M. arion* 'summer type', NA – *M. nausithous*, TE – *M. teleius*. The infected individuals are indicated with W. See the abbreviations in Online Resource 1 and 2.

Fig. 4 (a) UPGMA dendrogram based on allozymes using Cavalli-Sforza & Edwards' chord distances. **(b)** The results of PCA in *M. alcon*. Each symbol represents one sample in the reduced space by variances. The first two axes explain 66.6% of the total variance. The 'cruciata and pneumonanthe type' of *M. alcon* clearly differ along the first axis determined by the allele 2 and 3 of esterase locus. ALct – *M. alcon* 'cruciata type', ALpt – *M. alcon* 'pneumonanthe type', ARsp – *M. arion* 'spring type', ARsu – *M. arion* 'summer type', NA – *M. nausithous*, TE – *M. teleius*. See the abbreviations in Online Resource 1.

References

Aagaard, K., Hindar, K., Pullin, A. S., James, C. H., Hammarstedt, O., Balstad, T., et al. (2002). Phylogenetic relationships in brown argus butterflies (Lepidoptera: Lycaenidae: Aricia) from northwestern Europe. *Biological Journal of the Linnean Society, 75*(1), 27-37, doi:10.1046/j.1095-8312.2002.00004.x.

Als, T. D., Vila, R., Kandul, N. P., Nash, D. R., Yen, S.-H., Hsu, Y.-F., et al. (2004). The evolution of alternative parasitic life histories in large blue butterflies. [10.1038/nature03020]. *Nature, 432*(7015), 386-390, doi:http://www.nature.com/nature/journal/v432/n7015/suppinfo/nature03020_S1.html.

Andersen, A., Simcox, D. J., Thomas, J. A., & Nash, D. R. (2014). Assessing reintroduction schemes by comparing genetic diversity of reintroduced and source populations: A case study of the globally threatened large blue butterfly (*Maculinea arion*). *Biological Conservation, 175*(0), 34-41, doi:http://dx.doi.org/10.1016/j.biocon.2014.04.009.

Anderson, R. M., & May, R. M. (1982). Coevolution of hosts and parasites. *Parasitology, 85*(02), 411-426, doi:doi:10.1017/S0031182000055360.

Aris-Brosou, S., & Excoffier, L. (1996). The impact of population expansion and mutation rate heterogeneity on DNA sequence polymorphism. *Molecular Biology and Evolution, 13*(3), 494-504.

Baldo, L., Dunning Hotopp, J. C., Jolley, K. A., Bordenstein, S. R., Biber, S. A., Choudhury, R. R., et al. (2006). Multilocus sequence typing system for the endosymbiont *Wolbachia pipientis*. *Applied and Environmental Microbiology, 72*(11), 7098-7110, doi:10.1128/aem.00731-06.

Bereczki, J., Pecsenye, K., Peregovits, L., & Varga, Z. (2005). Pattern of genetic differentiation in the *Maculinea alcon* species group (Lepidoptera, Lycaenidae) in Central Europe. *Journal of Zoological Systematics and Evolutionary Research, 43*(2), 157-165, doi:DOI 10.1111/j.1439-0469.2005.00305.157-165.

Bereczki, J., Tóth, J. P., Sramkó, G., & Varga, Z. (2014). Multilevel studies on the two phenological forms of Large Blue (*Maculinea arion*) (Lepidoptera: Lycaenidae). *Journal of Zoological Systematics and Evolutionary Research, 52*(1), 32-43, doi:10.1111/jzs.12034.

Burns, J. M. (1975). Isozymes in evolutionary systematics. In C. L. Markert (Ed.), *Isozymes IV: Genetics and Evolution* (pp. 49-62). New York: Academic Press.

Cavalli-Sforza, L. L., & Edwards, A. W. F. (1967). Phylogenetic analysis: models and estimation procedures. *Evolution, 21*, 550-570.

Davies, H. D., & Vinson, S. B. (1986). Passive evasion by eggs of braconid parasitoid *Cardiochiles nigriceps* of encapsulation in vitro by haemocytes of host *Heliothis virescens*. Possible role for fibrous layer in immunity. *Journal of Insect Physiology, 32*(12), 1003-1010, doi:http://dx.doi.org/10.1016/0022-1910(86)90119-8.

Dupuis, J. R., Roe, A. D., & Sperling, F. A. H. (2012). Multi-locus species delimitation in closely related animals and fungi: one marker is not enough. *Molecular Ecology, 21*(18), 4422-4436, doi:10.1111/j.1365-294X.2012.05642.x.

Dyson, E. A., Kamath, M. K., & Hurst, G. D. D. (2002). *Wolbachia* infection associated with all-female broods in *Hypolimnas bolina* (Lepidoptera: Nymphalidae): evidence for horizontal transmission of a butterfly male killer. *Heredity, 88*(3), 166-171.

Earl, D., & vonHoldt, B. (2012). STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources, 4*(2), 359-361, doi:10.1007/s12686-011-9548-7.

Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software structure: a simulation study. *Molecular Ecology, 14*(8), 2611-2620, doi:10.1111/j.1365-294X.2005.02553.x.

Fine, P. E. M. (1975). Vectors and vertical transmission: an epidemiologic perspective. *Annals of the New York Academy of Sciences, 266*(1), 173-194, doi:10.1111/j.1749-6632.1975.tb35099.x.

Fujii, Y., Kageyama, D., Hoshizaki, S., Ishikawa, H., & Sasaki, T. (2001). Transfection of *Wolbachia* in Lepidoptera: the feminizer of the adzuki bean borer *Ostrinia scapulalis* causes male killing in the Mediterranean flour moth *Ephestia kuehniella*. *Proceedings of the Royal Society of London. Series B: Biological Sciences, 268*(1469), 855-859, doi:10.1098/rspb.2001.1593.

Galtier, N., Nabholz, B., Glémin, S., & Hurst, G. D. D. (2009). Mitochondrial DNA as a marker of molecular diversity: a reappraisal. *Molecular Ecology, 18*(22), 4541-4550, doi:10.1111/j.1365-294X.2009.04380.x.

Hammer, Ø., Harper, D. A. T., & Ryan, P. D. (2001). PAST: Paleontological statistics software package for education and data analysis. *Paleontologica Electronica, 4*(1), 9.

Hilgenboecker, K., Hammerstein, P., Schlattmann, P., Telschow, A., & Werren, J. H. (2008). How many species are infected with *Wolbachia*? – a statistical analysis of current data. *Fems Microbiology Letters, 281*(2), 215-220, doi:10.1111/j.1574-6968.2008.01110.x.

Hoffmann, A. A., Hercus, M., & Dagher, H. (1998). Population dynamics of the *Wolbachia* infection causing cytoplasmic incompatibility in *Drosophila melanogaster*. *Genetics, 148*(1), 221-231.

Hu, J., Zhu, X.-X., & Fu, W.-J. (2003). Passive evasion of encapsulation in *Macrocentrus cingulum* Brischke (Hymenoptera: Braconidae), a polyembryonic parasitoid of *Ostrinia furnacalis* Guenée (Lepidoptera: Pyralidae). *Journal of Insect Physiology, 49*(4), 367-375, doi:http://dx.doi.org/10.1016/S0022-1910(03)00021-0.

Jiggins, F. M. (2003). Male-killing *Wolbachia* and mitochondrial DNA: selective sweeps, hybrid introgression and parasite population dynamics. *Genetics, 164*(1), 5-12.

Kodandaramaiah, U., Simonsen, T. J., Bromilow, S., Wahlberg, N., & Sperling, F. (2013). Deceptive single-locus taxonomy and phylogeography: *Wolbachia*-associated divergence in mitochondrial DNA is not reflected in morphology and nuclear markers in a butterfly species. *Ecology and Evolution, 3*(16), 5167-5176, doi:10.1002/ece3.886.

Librado, P., & Rozas, J. (2009). DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, *25*(11): 1451-1452, doi:10.1093/bioinformatics/btp187.

Nei, M. (1987). *Molecular evolutionary genetics*. New York: Columbia University Press.

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

Nice, C. C., Gompert, Z., Forister, M. L., & Fordyce, J. A. (2009). An unseen foe in arthropod conservation efforts: The case of *Wolbachia* infections in the Karner blue butterfly. *Biological Conservation, 142*(12), 3137-3146, doi:10.1016/j.biocon.2009.08.020.

Paradis, E. (2010). PEGAS: an R package for population genetics with an integrated-modular approach. *Bioinformatics*, *26*(3): 419-420, doi:10.1093/bioinformatics/btp696.

Patricelli, D., Sielezniew, M., Ponikwicka-Tyszko, D., Ratkiewicz, M., Bonelli, S., Barbero, F., et al. (2013). Contrasting genetic structure of rear edge and continuous range populations of a parasitic butterfly infected by *Wolbachia*. *BMC Evolutionary Biology, 13*(1), 14.

Peakall, R. O. D., & Smouse, P. E. (2006). Genalex 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes, 6*(1), 288-295, doi:10.1111/j.1471-8286.2005.01155.x.

Pech, P., Fric, Z., Konvička, M., & Zrzavý, J. (2004). Phylogeny of *Maculinea* blues (Lepidoptera: Lycaenidae) based on morphological and ecological characters: evolution of parasitic myrmecophily. *Cladistics, 20*(4), 362-375, doi:10.1111/j.1096-0031.2004.00031.x.

Pecsenye, K., Bereczki, J., Tihanyi, B., Tóth, A., Peregovits, L., & Varga, Z. (2007). Genetic differentiation among the *Maculinea* species (Lepidoptera: Lycaenidae) in eastern Central Europe. *Biological Journal of the Linnean Society*, *91*(1), 11-21.

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics, 155*(2), 945-959.

QGIS Development Team, 2014. QGIS Geographic Information System. Open Source Geospatial Foundation Project.

R Development Core Team (2014). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.

Ritter, S., Michalski, S. G., Settele, J., Wiemers, M., Fric, Z. F., Sielezniew, M., et al. (2013). *Wolbachia* infections mimic cryptic speciation in two parasitic butterfly species, *Phengaris teleius* and *P. nausithous* (Lepidoptera: Lycaenidae). *Plos One, 8*(11), e78107, doi:10.1371/journal.pone.0078107.

Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D. L., Darling, A., Höhna, S., et al. (2012). MrBayes 3.2: efficient bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology, 61*(3), 539-542, doi:10.1093/sysbio/sys029.

Schmitt, T., Giessl, A., & Seitz, A. (2003). Did *Polyommatus icarus* (Lepidoptera: Lycaenidae) have distinct glacial refugia in southern Europe? Evidence from population genetics. *Biological Journal of the Linnean Society, 80*(3), 529-538, doi:10.1046/j.1095-8312.2003.00261.x.

Schmitt, T., & Hewitt, G. M. (2004). The genetic pattern of population threat and loss: a case study of butterflies. *Molecular Ecology, 13*(1), 21-31, doi:10.1046/j.1365-294X.2004.02020.x.

Shaw, M. R., Stefanescu, C., & Van Nouhuys, S. (2009). Parasitoids of European butterflies. In J. Settele, T. G. Shreeve, M. Konvicka, & H. Van Dyck (Eds.), *Ecology of butterflies in Europe* (pp. 130-156). Cambridge: Cambridge University Press.

Shoemaker, D., Keller, G., & Ross, K. G. (2003). Effects of *Wolbachia* on mtDNA variation in two fire ant species. *Molecular Ecology, 12*(7), 1757-1771, doi:10.1046/j.1365-294X.2003.01864.x.

Sielezniew, M., Rutkowski, R., Ponikwicka-Tyszko, D., Ratkiewicz, M., Dziekańska, I., & Švitra, G. (2012). Differences in genetic variability between two ecotypes of the endangered myrmecophilous butterfly *Phengaris* (=*Maculinea*) *alcon* – the setting of conservation priorities. *Insect Conservation and Diversity, 5*(3), 223-236, doi:10.1111/j.1752-4598.2011.00163.x.

Smee, M. R. (2011). *Population ecology and genetics of the Marsh Fritillary Butterfly Euphydryas aurinia*. University of Exeter, Exeter (UK).

Sneath, P. H., & Sokal, R. R. (1973). *Numerical taxonomy* San Francisco: W. H. Freeman.

Stephens, M., & Donnelly, P. (2003). A comparison of bayesian methods for haplotype reconstruction from population genotype data. *The American Journal of Human Genetics, 73*(5), 1162-1169, doi:http://dx.doi.org/10.1086/379378.

Stephens, M., Smith, N. J., & Donnelly, P. (2001). A new statistical method for haplotype reconstruction from population data. *The American Journal of Human Genetics, 68*(4), 978-989, doi:http://dx.doi.org/10.1086/319501.

Tajima, F. (1983). Evolutionary relationship of DNA sequences in finite populations. *Genetics, 105*(2), 437-460.

Tajima, F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics, 123*(3), 585-595.

Tamura, K., Stecher, G., Peterson, D., Filipski, A., & Kumar, S. (2013). MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. *Molecular Biology and Evolution, 30*(12), 2725-2729, doi:10.1093/molbev/mst197.

Tartally, A. (2008). *Myrmecophily of Maculinea butterflies in the Carpathian Basin (Lepidoptera: Lycaenidae).* PhD, University of Debrecen,

Thomas, J. A., Simcox, D. J., & Clarke, R. T. (2009). Successful conservation of a threatened *Maculinea* butterfly. *Science, 325*(5936), 80-83, doi:10.1126/science.1175726.

Turelli, M. (1994). Evolution of incompatibility-inducing microbes and their hosts. *Evolution, 48*(5), 1500-1513, doi:10.2307/2410244.

Turelli, M., & Hoffmann, A. A. (1991). Rapid spread of an inherited incompatibility factor in California *Drosophila*. [10.1038/353440a0]. *Nature, 353*(6343), 440-442.

Turelli, M., & Hoffmann, A. A. (1995). Cytoplasmic incompatibility in *Drosophila simulans*: dynamics and parameter estimates from natural populations. *Genetics, 140*(4), 1319-1338.

Ugelvig, L. V., Vila, R., Pierce, N. E., & Nash, D. R. (2011). A phylogenetic revision of the *Glaucopsyche* section (Lepidoptera: Lycaenidae), with special focus on the *Phengaris-Maculinea* clade. *Molecular Phylogenetics and Evolution, 61*(1), 237-243, doi:10.1016/j.ympev.2011.05.016.

Wahlberg, N., & Wheat, C. W. (2008). Genomic outposts serve the phylogenomic pioneers: designing novel nuclear markers for genomic DNA extractions of Lepidoptera. *Systematic Biology, 57*(2), 231-242, doi:10.1080/10635150802033006.

Watterson, G. A. (1975). On the number of segregating sites in genetical models without recombination. *Theoretical Population Biology, 7*(2), 256-276, doi:http://dx.doi.org/10.1016/0040-5809(75)90020-9.

Werren, J. H., Baldo, L., & Clark, M. E. (2008). *Wolbachia*: master manipulators of invertebrate biology. [10.1038/nrmicro1969]. *Nat Rev Micro, 6*(10), 741-751.

Werren, J. H., & Windsor, D. M. (2000). *Wolbachia* infection frequencies in insects: evidence of a global equilibrium? *Proceedings of the Royal Society of London. Series B: Biological Sciences, 267*(1450), 1277-1285, doi:10.1098/rspb.2000.1139.

Wynhoff, I. (1998). Lessons from the reintroduction of *Maculinea teleius* and *M. nausithous* in the Netherlands. *Journal of Insect Conservation, 2*(1), 47-57, doi:10.1023/A:1009692723056.

Yamamura, N. (1993). Vertical transmission and evolution of mutualism from parasitism. *Theoretical Population Biology, 44*(1), 95-109, doi:http://dx.doi.org/10.1006/tpbi.1993.1020.

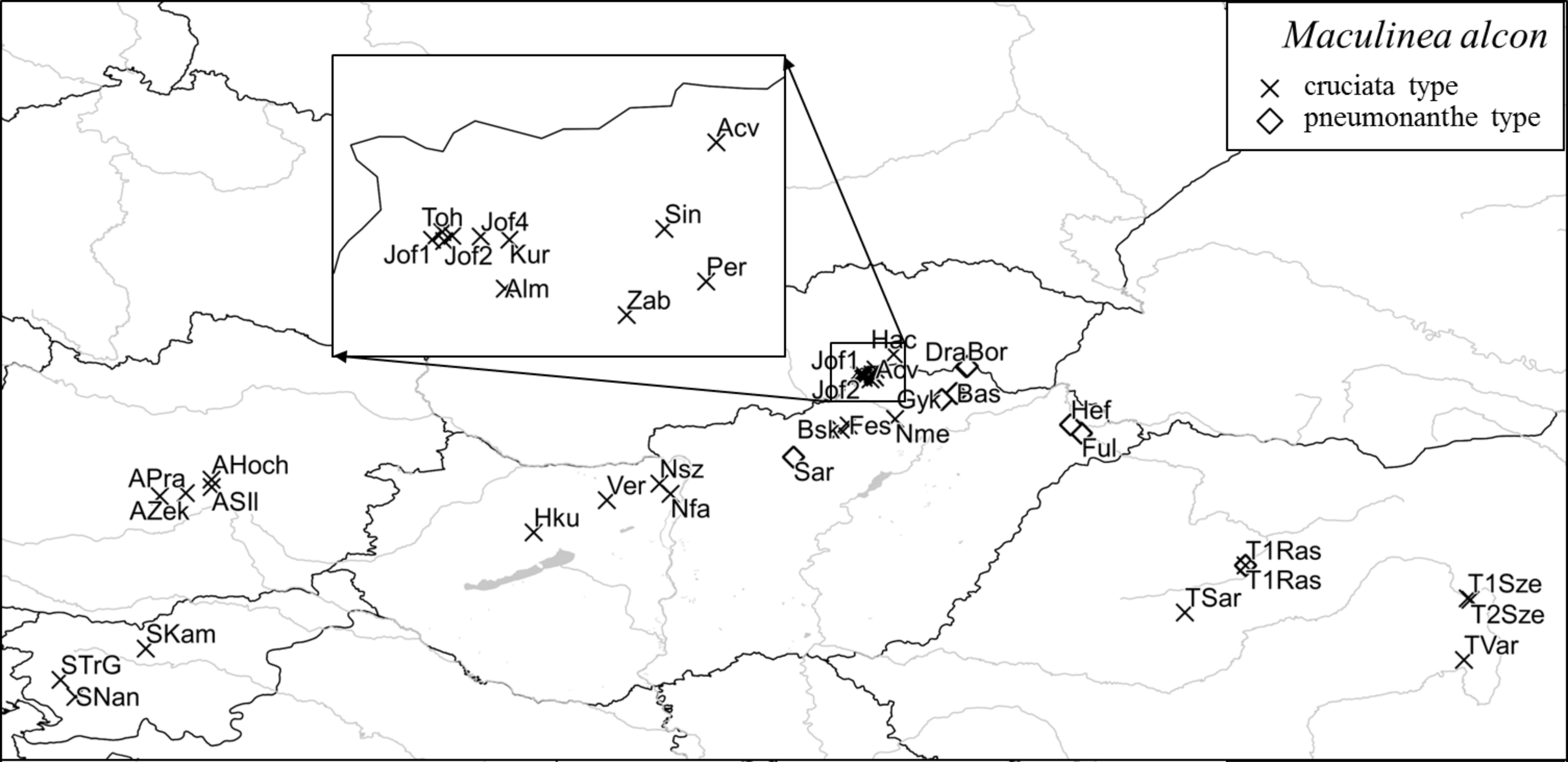
Zug, R., & Hammerstein, P. (2014). Bad guys turned nice? A critical assessment of *Wolbachia* mutualisms in arthropod hosts. *Biological Reviews*, doi:10.1111/brv.12098.

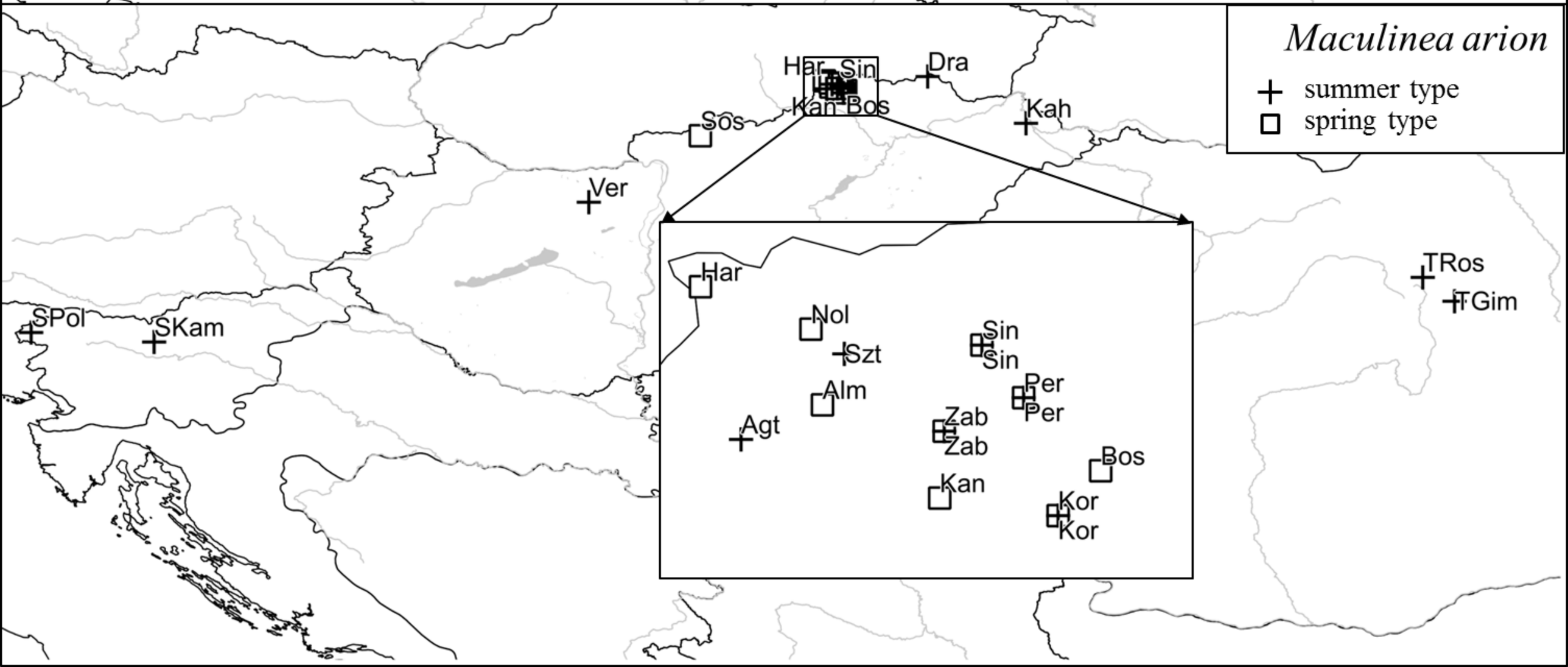
Controversial patterns of *Wolbachia* infestation in the social parasitic *Maculinea* butterflies (Lepidoptera: Lycaenidae)

Organisms, Diversity and Evolution

Judit Bereczki, Rita Rácz, Zoltán Varga, János P. Tóth

Corresponding author: Judit Bereczki, Department of Evolutionary Zoology & Human Biology, Institute of Biology & Ecology, University of Debrecen, Debrecen, P.O.Box 3, H-4010, Hungary. Telephone: +36 52 512 900/62526. Fax: +36 52 512 941. E-mail: bereczki.judit@gmail.com


Maculinea alcon
cruciata type
pneumonanthe type
Acv
Toh
Jof4
Sin
Jof1
Jof2
Kur
Per
Alm
Zab
Hac
DraBor
Jof1
Acv
Jof2
Gyk
Bas
Hef
Bsk
Fes
Nme
Ful
Sar
APra
AHoch
AZek
ASll
Nsz
Ver
Nfa
Hku
T1Ras
T1Ras
TSar
T1Sze
T2Sze
TVar
SKam
STrG
SNan


Maculinea arion
summer type
spring type
Har
Sin
Dra
Kah
Bos
Sos
Ver
TRos
TGim
SPol
SKam
Har
Nol
Sin
Szt
Sin
Per
Alm
Per
Zab
Agt
Zab
Bos
Kan
Kor
Kor

Maculinea nausithous

Hha Ket Api Osp Orf Szf Szo Mfa

Urh Nab

Nos Nag Cst

Hha Ket Mfa

SCel

BSu BRa BRu

T2Ras T1Ras

TFan

Infection rate

0%

100%

Bor Dra Tsj Agt RaM Tbv Bas Gyk

Agt Tsj Bor Dra RaM Tbv Gyk

Ful

Nab Hid Him

Csh Kun

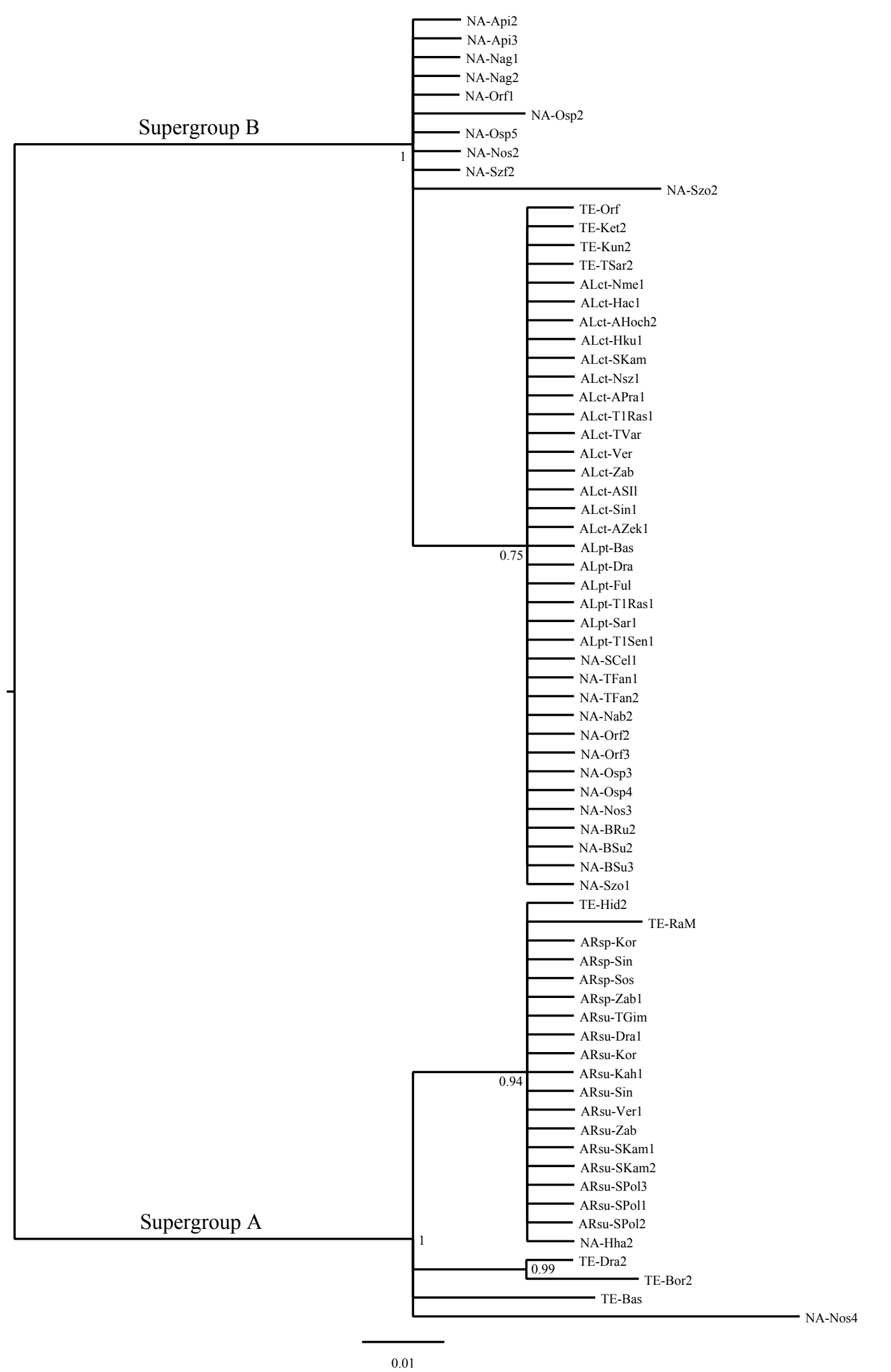
Orf Api Ket Mfa

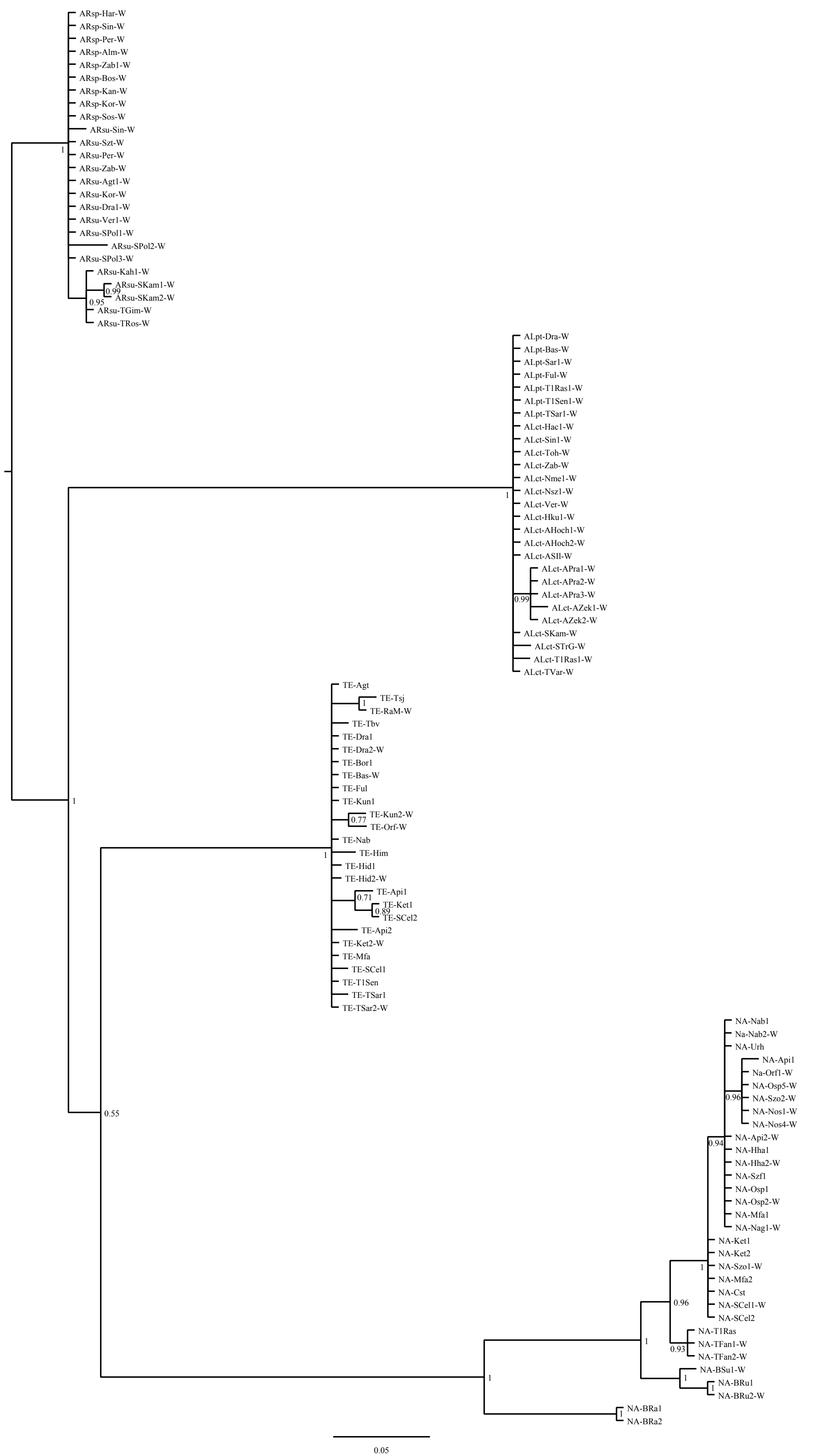
SCel

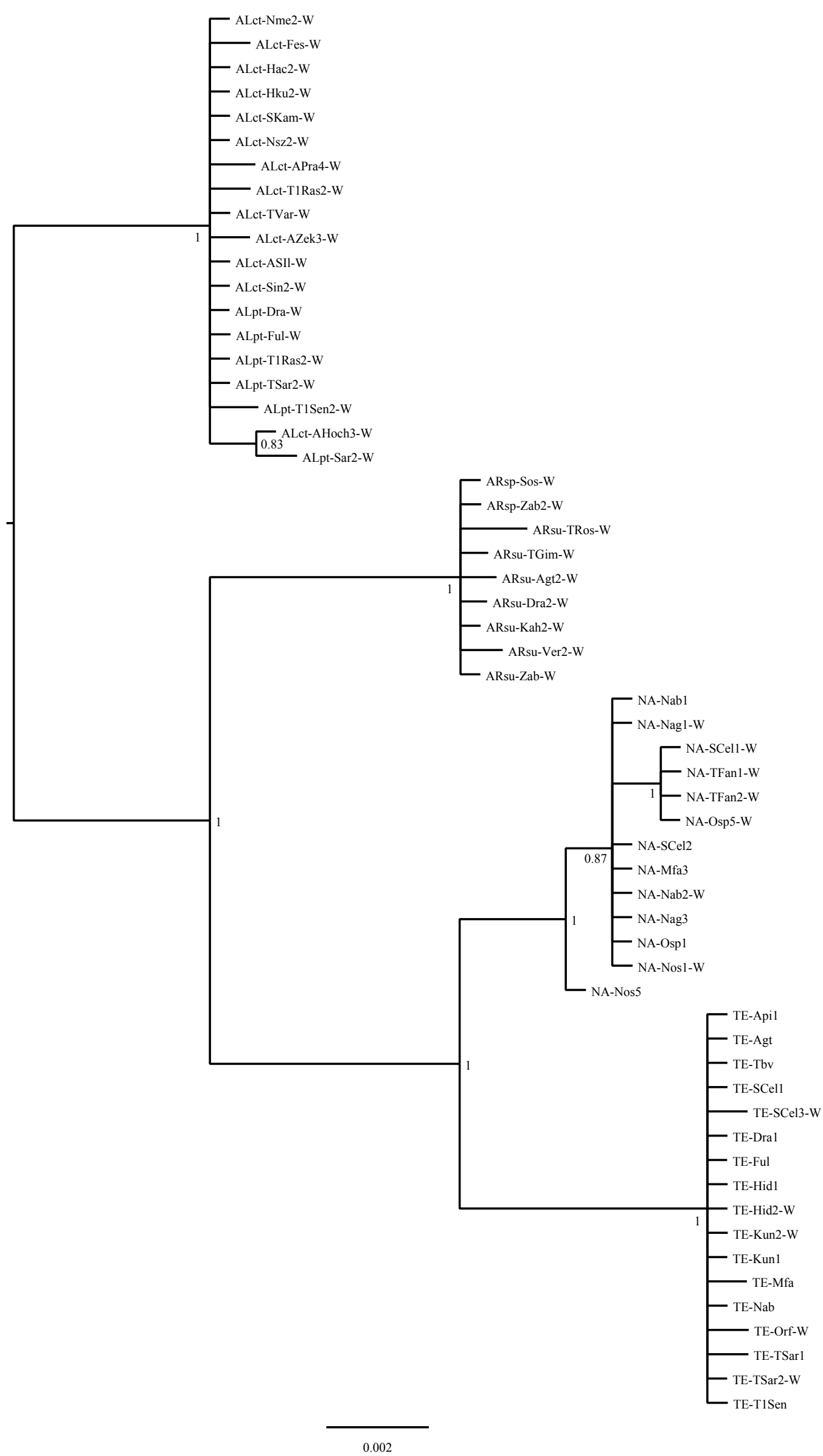
TSar

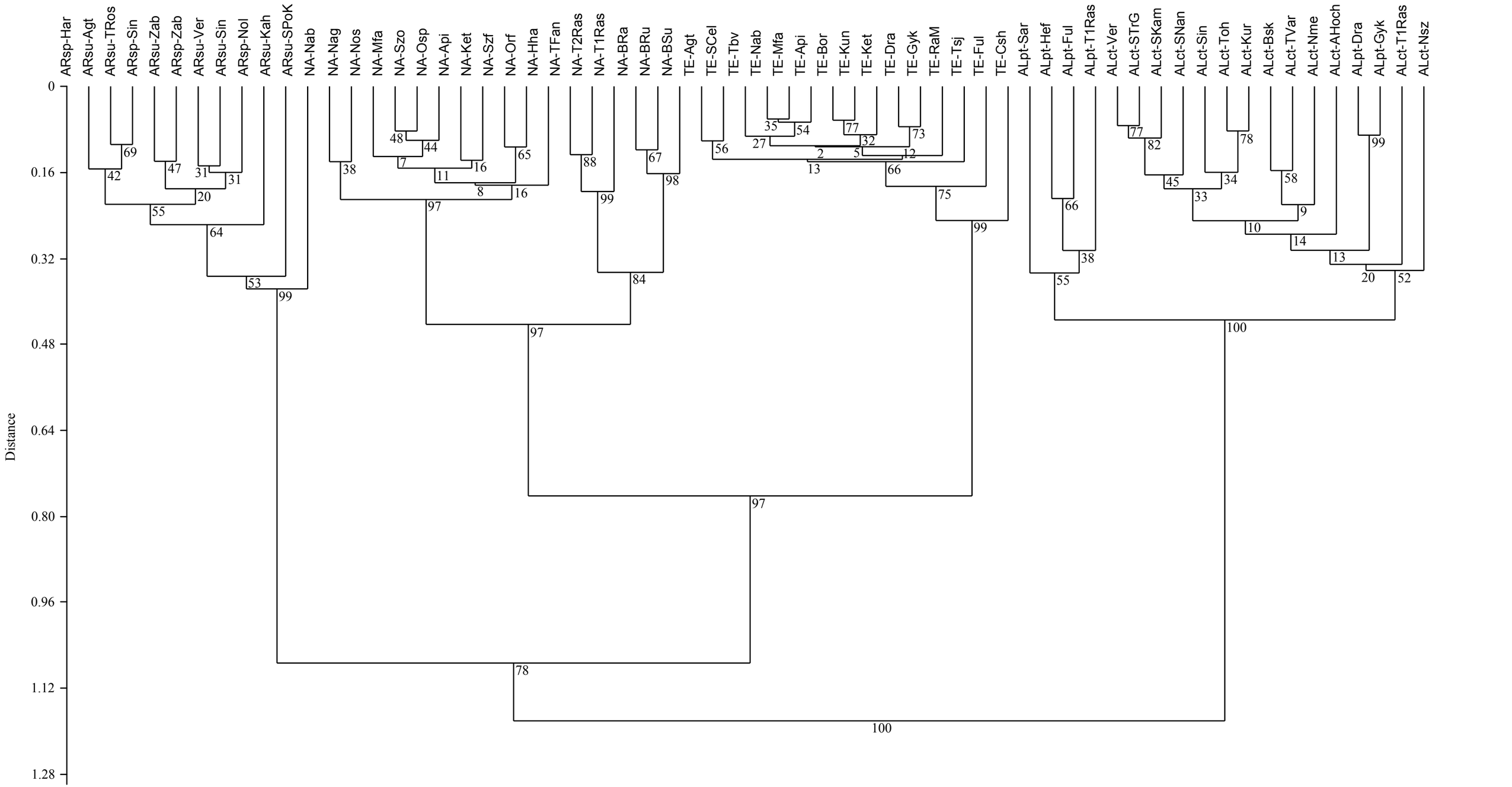
T1Sen

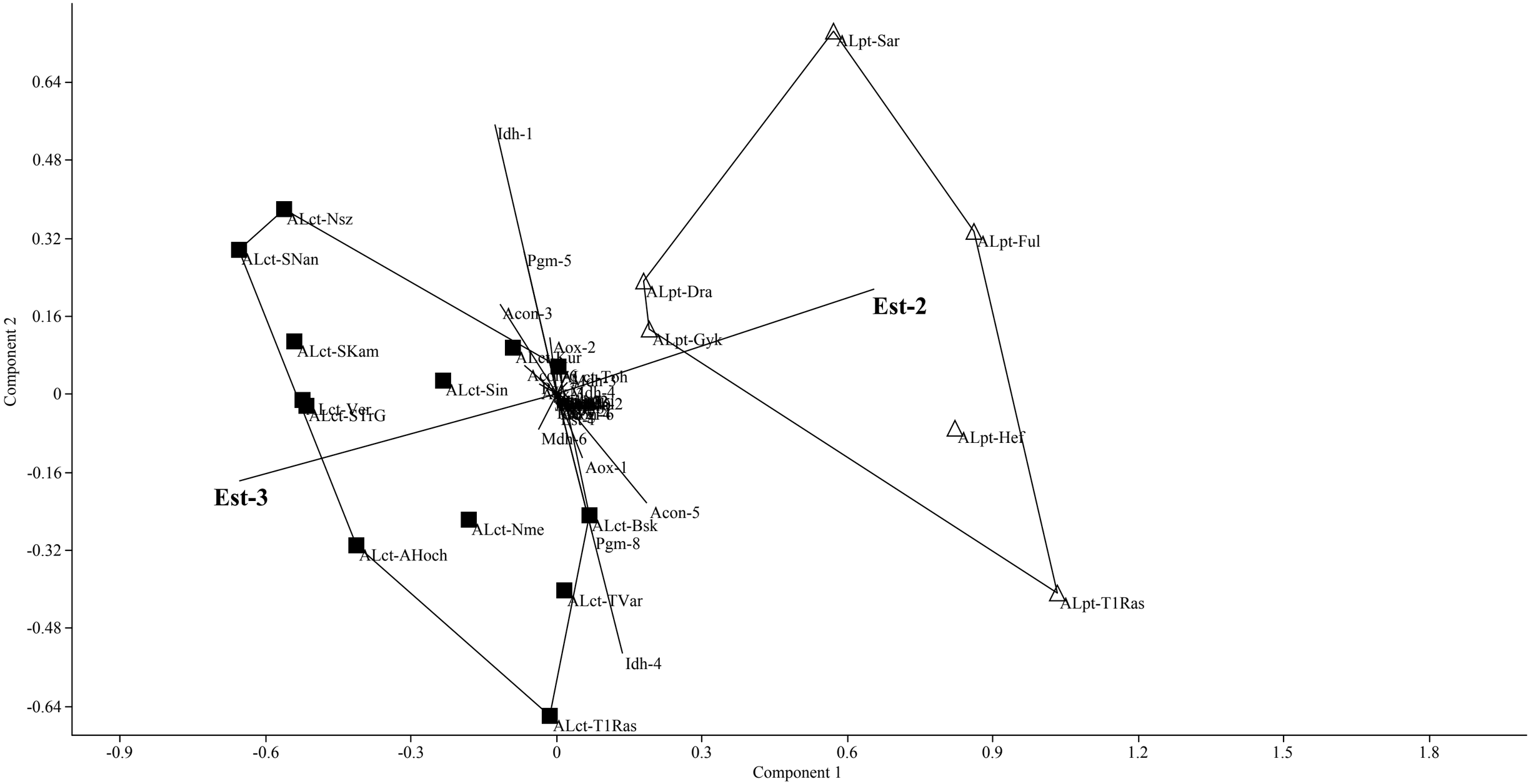
Maculinea teleius


Supergroup B
Supergroup A
NA-Api2
NA-Api3
NA-Nag1
NA-Nag2
NA-Orf1
NA-Osp2
NA-Osp5
NA-Nos2
NA-Szf2
NA-Szo2
TE-Orf
TE-Ket2
TE-Kun2
TE-TSar2
ALct-Nme1
ALct-Hac1
ALct-AHoch2
ALct-Hku1
ALct-SKam
ALct-Nsz1
ALct-APra1
ALct-T1Ras1
ALct-TVar
ALct-Ver
ALct-Zab
ALct-ASIl
ALct-Sin1
ALct-AZek1
ALpt-Bas
ALpt-Dra
ALpt-Ful
ALpt-T1Ras1
ALpt-Sar1
ALpt-T1Sen1
NA-SCel1
NA-TFan1
NA-TFan2
NA-Nab2
NA-Orf2
NA-Orf3
NA-Osp3
NA-Osp4
NA-Nos3
NA-BRu2
NA-BSu2
NA-BSu3
NA-Szo1
TE-Hid2
TE-RaM
ARsp-Kor
ARsp-Sin
ARsp-Sos
ARsp-Zab1
ARsu-TGim
ARsu-Dra1
ARsu-Kor
ARsu-Kah1
ARsu-Sin
ARsu-Ver1
ARsu-Zab
ARsu-SKam1
ARsu-SKam2
ARsu-SPol3
ARsu-SPol1
ARsu-SPol2
NA-Hha2
TE-Dra2
TE-Bor2
TE-Bas
NA-Nos4
1
0.75
0.94
1
0.99
0.01


ARsp-Har-W
ARsp-Sin-W
ARsp-Per-W
ARsp-Alm-W
ARsp-Zab1-W
ARsp-Bos-W
ARsp-Kan-W
ARsp-Kor-W
ARsp-Sos-W
ARsu-Sin-W
ARsu-Szt-W
ARsu-Per-W
ARsu-Zab-W
ARsu-Agt1-W
ARsu-Kor-W
ARsu-Dra1-W
ARsu-Ver1-W
ARsu-SPol1-W
ARsu-SPol2-W
ARsu-SPol3-W
ARsu-Kah1-W
ARsu-SKam1-W
ARsu-SKam2-W
ARsu-TGim-W
ARsu-TRos-W
1
0.99
0.95
ALpt-Dra-W
ALpt-Bas-W
ALpt-Sar1-W
ALpt-Ful-W
ALpt-T1Ras1-W
ALpt-T1Sen1-W
ALpt-TSar1-W
ALct-Hac1-W
ALct-Sin1-W
ALct-Toh-W
ALct-Zab-W
ALct-Nme1-W
ALct-Nsz1-W
ALct-Ver-W
ALct-Hku1-W
ALct-AHoch1-W
ALct-AHoch2-W
ALct-ASll-W
ALct-APra1-W
ALct-APra2-W
ALct-APra3-W
ALct-AZek1-W
ALct-AZek2-W
ALct-SKam-W
ALct-STrG-W
ALct-T1Ras1-W
ALct-TVar-W
1
0.99
TE-Agt
TE-Tsj
TE-RaM-W
TE-Tbv
TE-Dra1
TE-Dra2-W
TE-Bor1
TE-Bas-W
TE-Ful
TE-Kun1
TE-Kun2-W
TE-Orf-W
TE-Nab
TE-Him
TE-Hid1
TE-Hid2-W
TE-Api1
TE-Ket1
TE-SCel2
TE-Api2
TE-Ket2-W
TE-Mfa
TE-SCel1
TE-T1Sen
TE-TSar1
TE-TSar2-W
1
0.77
1
0.71
0.89
1
0.55
NA-Nab1
Na-Nab2-W
NA-Urh
NA-Api1
Na-Orf1-W
NA-Osp5-W
NA-Szo2-W
NA-Nos1-W
NA-Nos4-W
NA-Api2-W
NA-Hha1
NA-Hha2-W
NA-Szf1
NA-Osp1
NA-Osp2-W
NA-Mfa1
NA-Nag1-W
NA-Ket1
NA-Ket2
NA-Szo1-W
NA-Mfa2
NA-Cst
NA-SCel1-W
NA-SCel2
NA-T1Ras
NA-TFan1-W
NA-TFan2-W
NA-BSu1-W
NA-BRu1
NA-BRu2-W
NA-BRa1
NA-BRa2
0.96
0.94
1
0.96
1
0.93
1
1
1
1
0.05


ALct-Nme2-W
ALct-Fes-W
ALct-Hac2-W
ALct-Hku2-W
ALct-SKam-W
ALct-Nsz2-W
ALct-APra4-W
ALct-T1Ras2-W
ALct-TVar-W
1
ALct-AZek3-W
ALct-ASII-W
ALct-Sin2-W
ALpt-Dra-W
ALpt-Ful-W
ALpt-T1Ras2-W
ALpt-TSar2-W
ALpt-T1Sen2-W
ALct-AHoch3-W
0.83
ALpt-Sar2-W
ARsp-Sos-W
ARsp-Zab2-W
ARsu-TRos-W
ARsu-TGim-W
1
ARsu-Agt2-W
ARsu-Dra2-W
ARsu-Kah2-W
ARsu-Ver2-W
ARsu-Zab-W
NA-Nab1
NA-Nag1-W
NA-SCel1-W
NA-TFan1-W
1
NA-TFan2-W
NA-Osp5-W
NA-SCel2
0.87
NA-Mfa3
NA-Nab2-W
1
NA-Nag3
NA-Osp1
NA-Nos1-W
NA-Nos5
1
TE-Api1
TE-Agt
TE-Tbv
TE-SCel1
TE-SCel3-W
TE-Dra1
TE-Ful
TE-Hid1
TE-Hid2-W
1
TE-Kun2-W
TE-Kun1
TE-Mfa
TE-Nab
TE-Orf-W
TE-TSar1
TE-TSar2-W
TE-T1Sen
1
0.002


ARsp-Har
ARsu-Agt
ARsu-TRos
ARsp-Sin
ARsu-Zab
ARsp-Zab
ARsu-Ver
ARsu-Sin
ARsp-Nol
ARsu-Kah
ARsu-SPoK
NA-Nab
NA-Nag
NA-Nos
NA-Mfa
NA-Szo
NA-Osp
NA-Api
NA-Ket
NA-Szf
NA-Orf
NA-Hha
NA-TFan
NA-T2Ras
NA-T1Ras
NA-BRa
NA-BRu
NA-BSu
TE-Agt
TE-SCel
TE-Tbv
TE-Nab
TE-Mfa
TE-Api
TE-Bor
TE-Kun
TE-Ket
TE-Dra
TE-Gyk
TE-RaM
TE-Tsj
TE-Ful
TE-Csh
ALpt-Sar
ALpt-Hef
ALpt-Ful
ALpt-T1Ras
ALct-Ver
ALct-STrG
ALct-SKam
ALct-SNan
ALct-Sin
ALct-Toh
ALct-Kur
ALct-Bsk
ALct-TVar
ALct-Nme
ALct-AHoch
ALpt-Dra
ALpt-Gyk
ALct-T1Ras
ALct-Nsz
Distance
0
0.16
0.32
0.48
0.64
0.80
0.96
1.12
1.28


ALpt-Sar
ALpt-Ful
Idh-1
Pgm-5
ALct-Nsz
ALct-SNan
ALpt-Dra
Est-2
Acon-3
ALpt-Gyk
Aox-2
ALct-SKam
ALct-Kur
ALct-Toh
ALct-Sin
ALct-Ver
ALct-STrG
ALpt-Hef
Mdh-6
Aox-1
Est-3
Acon-5
ALct-Nme
ALct-Bsk
Pgm-8
ALct-AHoch
ALct-TVar
ALpt-T1Ras
Idh-4
ALct-T1Ras
0.64
0.48
0.32
0.16
0
-0.16
-0.32
-0.48
-0.64
-0.9
-0.6
-0.3
0
0.3
0.6
0.9
1.2
1.5
1.8
Component 2
Component 1

Controversial patterns of *Wolbachia* infestation in the social parasitic *Maculinea* butterflies (Lepidoptera: Lycaenidae)

Organisms, Diversity and Evolution

Judit Bereczki, Rita Rácz, Zoltán Varga, János P. Tóth

Corresponding author: Judit Bereczki, Department of Evolutionary Zoology & Human Biology, Institute of Biology & Ecology, University of Debrecen, Debrecen, P.O.Box 3, H-4010, Hungary. Telephone: +36 52 512 900/62526. Fax: +36 52 512 941. E-mail: bereczki.judit@gmail.com

Sampling site	**Abbr.**	**Region**	***M. alcon* cruciata type (ALct)**	***M. alcon* pneumonanthe type (ALpt)**	***M. arion* spring type (ARsp)**	***M. arion* summer type (ARsu)**	***M. nausithous* (NA)**	***M. teleius* (TE)**
Aggtelek	Agt	Aggtelek Karst (Hungary)				14/7/2011 (5) **23** 27/7/2003 (5) 6/8/2002 (1)		6/8/2002 (5) **34** 29/7/2005 (1)
Acskó-völgy	Acv		29/6/2011 (5)					
Almás-tető	Alm		29/6/2011 (5)		26/5/2011 (6)			
Boszorkány-völgy	Bos				25/5/2011 (1)			
Haragistya	Har				31/5/2011 (6) **28**			
Jósvafő 1	Jof1		22/6/2011 (5)					
Jósvafő 2	Jof2		23/6/2011 (5)					
Jósvafő 3	Jof3		23/6/2011 (5)					
Jósvafő 4	Jof4		23/6/2011 (5)					
Kánó	Kan				24/5/2011 (1)			
Korlát-hegy	Kor				25/5/2011 (6)	3/8/2011 (1)		
Kuriszlán	Kur		23/6/2011 (5) **22**					
Nagyoldal	Nol				28/5/2003 (5) **15**			
Perkupa	Per		22/6/2011 (5)		28/5/2011 (6)	4/8/2011 (1)		
Szin	Sin		23/6/2011 (5) **20**		25/5/2011 (6) **12**	2/8/2011 (6) **11**		
Szőlőtető	Szt					14/7/2011 (6)		
Tohonya-hát	Toh		22/6/2011 (5) **35** 23/6/2011 (1)					
Zabanyik	Zab		22/6/2011 (5)		24/5/2011 (6) **13**	29/6/2011 (2) **6** 14/7/2011 (3) 3/8/2011 (1)		
Cserszegtomaj	Cst	Balaton Uplands (Hungary)					31/7/2003 (4)	
Nagygörbő	Nag						2/8/2003 (5) **18**	
Hárskút	Hku		27/6/2002 (5)					
Noszlop	Nos						18/8/2002 (5) **6**	
Nagy-Szénás	Nsz	Budapest (Hungary)	27/6/2003 (3) **11**					
Normafa	Nfa		9/7/2001 (5)					

Sampling site	**Abbr.**	**Region**	***M. alcon* cruciata type (ALct)**	***M. alcon* pneumonanthe type (ALpt)**	***M. arion* spring type (ARsp)**	***M. arion* summer type (ARsu)**	***M. nausithous* (NA)**	***M. teleius* (TE)**
Bükkszentkereszt	Bsk	Bükk Mountains (Hungary)	10/7/2001 (5) **24**					
Feketesár	Fes		8/7/2003 (5)					
Nagymező	Nme		7/7/2011 (5) **28**					
Rakacaszend-Meszes	RaM	Cserehát (Hungary)						25/7/2002 (5) **35**
Tornabarakonyi-völgy	Tbv							25/7/2002 (5) **33**
Tornaszentjakab	Tsj							1/8/2002 (5) **10**
Hidegség	Hid	Fertő-Hanság (Hungary)						30/8/2013 (5)
Himod	Him							30/8/2013 (5)
Csévharaszt	Csh	Kiskunság (Hungary)						9/8/2001 (5) **32**
Kunpeszér	Kun							17/8/2002 (5) **21**
Sárhegy	Sar	Mátra (Hungary)		8/7/2013 (5) **20**				
Sóshartyán	Sos				19/5/2002 (3)			
Apátistvánfalva	Api	Őrség (Hungary)					8/15/2002 (5) **7**	15/8/2002 (5) **10**
Hármashatár	Hha						8/14/2002 (5) **8**	
Kétvölgy	Ket						8/14/2002 (5) **20** 1/8/2003 (1)	14/8/2002 (5) **23** 18/7/2013 (5)
Magyarszombatfa	Mfa						7/13/2002 (5) **25**	17/8/2002 (5) **18**
Orfalu	Orf						8/15/2002 (5) **12**	15/8/2002 (5)
Őriszentpéter	Osp						1/8/2003 (5) **24**	
Szalafő	Szf						8/16/2002 (5) **8**	
Szomoróc	Szo						31/7/2003 (5) **10**	

Sampling site	**Abbr.**	**Region**	***M. alcon* cruciata type (ALct)**	***M. alcon* pneumonanthe type (ALpt)**	***M. arion* spring type (ARsp)**	***M. arion* summer type (ARsu)**	***M. nausithous* (NA)**	***M. teleius* (TE)**
Fülesd	Ful	Szatmár-Bereg (Hungary)		29/7/2006 (5) **24**				23/7/2002 (5) **32**
Hetefejércse	Hef			26/7/2013 (5) **19**				
Kaszonyi-hegy	Kah					23/7/2002 (1) **22** 21/7/2003 (5)		
Nagybajcs	Nab	Szigetköz (Hungary)					23/8/2002 (5) **7**	18/8/2002 (4) **22** 23/8/2002 (1)
Urhanya	Urh						15/8/2013 (1)	
Vérteskozma	Ver	Vértes (Hungary)	14/6/2003 (5) **25**			10/7/2002 (1) **20** 2-3/7/2003 (5)		
Baskó	Bas	Zemplén (Hungary)		29/7/2005 (5)				29/7/2005 (5)
Bodó-rét	Bor			28/7/2006 (5)				24/7/2013 (5) **27**
Drahos-rét	Dra			24/7/2013 (5) **23**		24/7/2013 (3)		1-2/8/2002 (5) **26**
Gyertyánkút	Gyk			9/8/2001 (5) **17**				9/8/2001 (5) **25**
Hochschwab	AHoch	Austria	11/7/2003 (3) **31** 4/7/2005 (3)					
Präbichl	APra		25/6/2012 (5)*					
Sankt Ilgen	ASIl		23/6/2012 (5)					
Zeiritzkampfel	AZek		24/6/2012 (5)*					
Hačava	Hac	Slovakia	13/7/2001 (5)					
Celje	SCel	Slovenia					11/7/2002 (3)	11/7/2002 (5) **19**
Kamnik	SKam		30/6/2003 (5) **16**			4/7/2003 (2) **3**		
Nanos plateau	SNan		1/7/2003 (5)** **27**					
Polovnik	SPol					12/7/2002 (2) **5** 4/7/2003 (1)		
Trnovski Gozd	STrG		3/7/2003 (5) **22**					

* *M. rebeli* from nominotypic localities which uses *Gentianella*.cf. *rhaetica* as initial food plant.
** This population uses an alternative food plant species, i.e. *Gentiana lutea*.

Sampling site	**Abbr.**	**Region**	***M. alcon* cruciata type (ALct)**	***M. alcon* pneumonanthe type (ALpt)**	***M. arion* spring type (ARsp)**	***M. arion* summer type (ARsu)**	***M. nausithous* (NA)**	***M. teleius* (TE)**
Lacul Roșu	TRos	Transylvania (Romania)				2/7/2002 (4) **14**		
Ghimeș-Făget	TGim					16/7/2011 (5)		
Șard	TSar			26/7/2006 (5)				26/7/2006 (5)
Senetea-1	T1Sen			10/7/2012 (5)				12/7/2012 (5)
Senetea-2	T2Sen			11/7/2012 (5)				
Fânațele Clujului	TFan						11/8/2003 (5) **21**	
Vârghiș	TVar		22/6/2003 (5) **21**					
Răscruci-1	T1Ras		15/6/2011 (5) **21**	3/8/2011 (5) **16**			31/7/2008 (5) **32**	
Răscruci-2	T2Ras						31/7/2008 (5) **24**	
Rădăuți	BRa	Bukovina (Romania)					23/7/2009 (5) **24**	
Runcu	BRu						24/7/2009 (5) **27**	
Suceava	BSu						23/7/2009 (5) **32**	
Total			140	55	46	60	94	111
Infected			140	55	46	60	34	16
Infected (%)			**100**	**100**	**100**	**100**	**36.2%**	**14.4%**

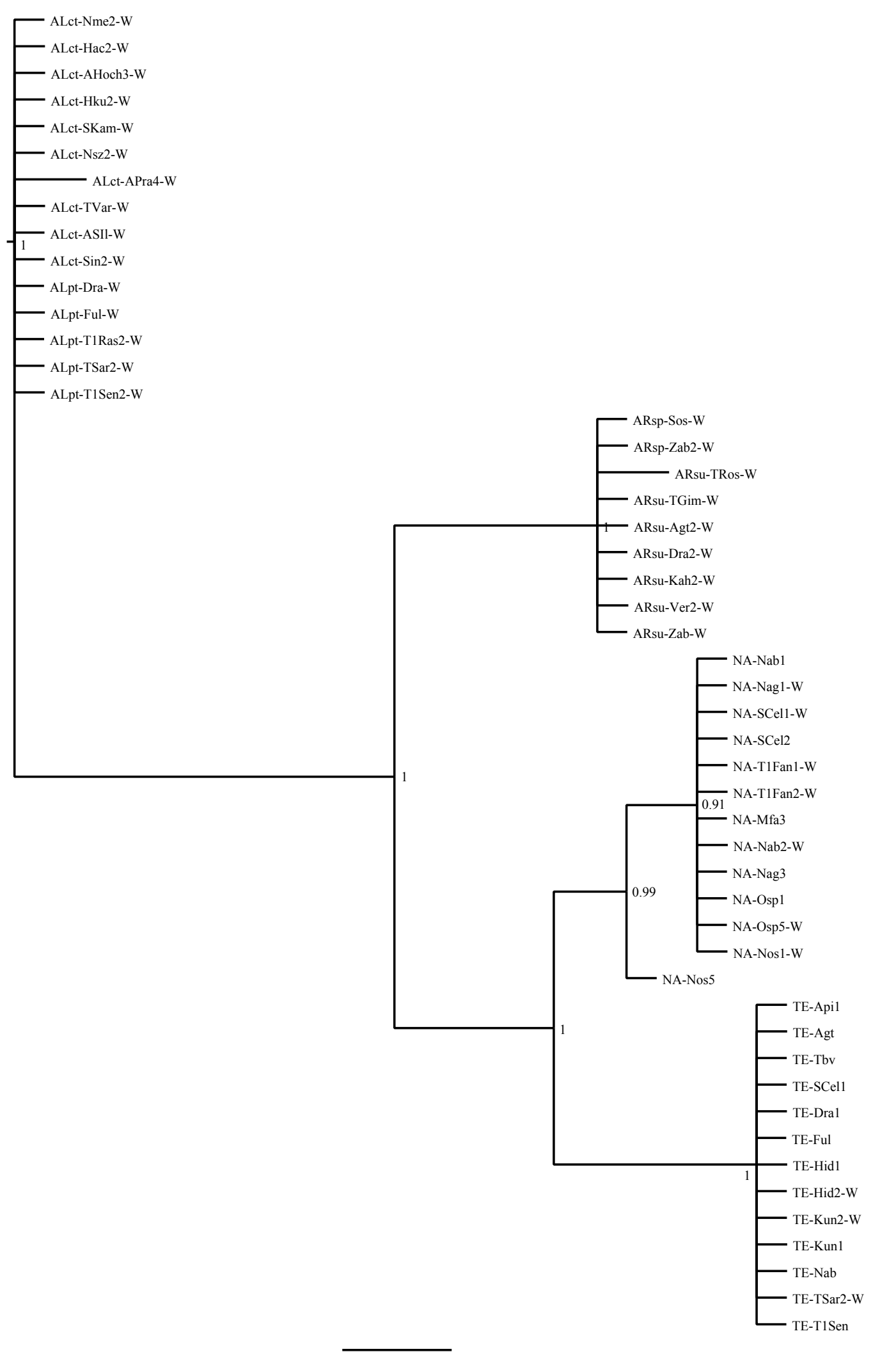
Species	ID	*Maculinea*					*Wolbachia*		
		COI part 1	COI part 2	GAPDH	MDH	wg	16S rRNA	WSP	MLST
	ALct-Hac1-W	KM517257					KM517522		+
	ALct-Hac2-W			KM517333	KM517384	KM517440			
	ALct-Sin1-W	KM517258					KM517533		+
	ALct-Sin2-W			KM517341	KM517394	KM517450			
	ALct-Toh-W	HG326645							
	ALct-Zab-W	KM517259					KM517531	+	
	ALct-Nme1-W	KM517260					KM517521		+
	ALct-Nme2-W			KM517332	KM517382	KM517438			
	ALct-Fes-W				KM517383	KM517439			
	ALct-Nsz1-W	KM517261					KM517526		+
	ALct-Nsz2-W			KM517337	KM517388	KM517444			
	ALct-Ver-W	KM517262					KM517530		+
	ALct-Hku1-W	KM517263					KM517524		+
	ALct-Hku2-W			KM517335	KM517386	KM517442			
M. alcon	ALct-AHoch1-W	HG326646							
cruciata type	ALct-AHoch2-W	KM517264					KM517523		+
	ALct-AHoch3-W			KM517334	KM517385	KM517441			
	ALct-ASIl-W	KM517265		KM517340	KM517393	KM517449	KM517532		+
	ALct-APra1-W	KM517266					KM517527		+
	ALct-APra2-W	KM517267							
	ALct-APra3-W	KM517268							
	ALct-APra4-W			KM517338	KM517389	KM517445			
	ALct-AZek1-W	KM517269					KM517534		+
	ALct-AZek2-W	KM517270							
	ALct-AZek3-W				KM517392	KM517448			
	ALct-SKam-W	KM517271		KM517336	KM517387	KM517443	KM517525		+
	ALct-STrG-W	KM517272							
	ALct-T1Ras1-W	KM517273					KM517528		+
	ALct-T1Ras2-W				KM517390	KM517446			
	ALct-TVar-W	KM517274		KM517339	KM517391	KM517447	KM517529		+

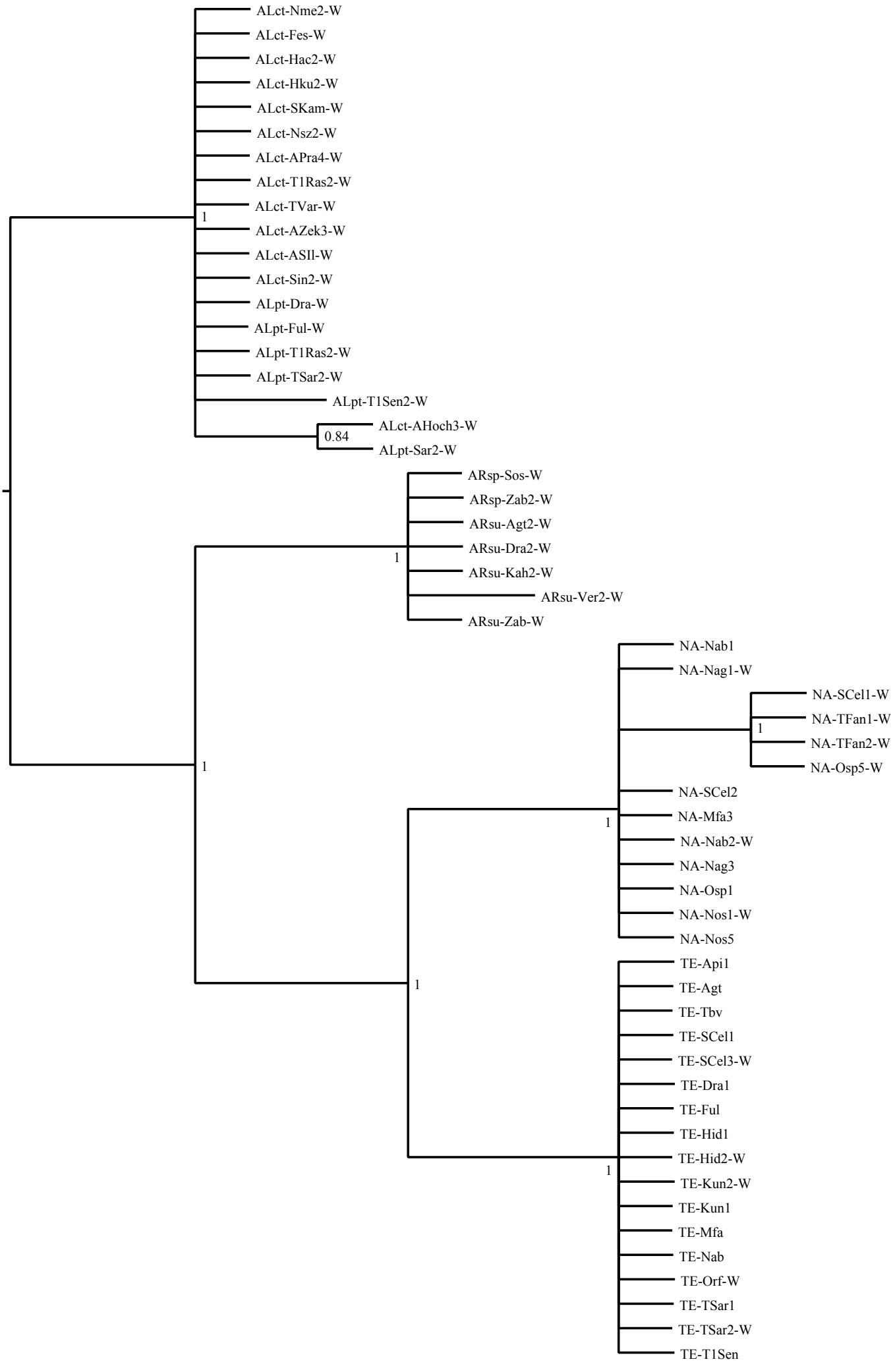
Species	ID	*Maculinea*					*Wolbachia*		
		COI part 1	COI part 2	GAPDH	MDH	wg	16S rRNA	WSP	MLST
M. alcon pneumonanthe type	ALpt-Dra-W	KM517250		KM517342	KM517395	KM517451	KM517536		+
	ALpt-Bas-W	KM517251					KM517535		+
	ALpt-Sar1-W	KM517252					KM517539		+
	ALpt-Sar2-W				KM517399	KM517455			
	ALpt-Ful-W	KM517253		KM517343	KM517396	KM517452	KM517537		+
	ALpt-T1Ras1-W	KM517254					KM517538		+
	ALpt-T1Ras2-W			KM517344	KM517397	KM517453			
	ALpt-T1Sen1-W	KM517255					KM517540		+
	ALpt-T1Sen2-W			KM517346	KM517400	KM517456			
	ALpt-TSar1-W	KM517256							
	ALpt-TSar2-W			KM517345	KM517398	KM517454			
M. arion spring type	ARsp-Har-W	HG326619							
	ARsp-Sin-W	HG326621					KM517506		+
	ARsp-Per-W	HG326623							
	ARsp-Alm-W	HG326620							
	ARsp-Zab1-W	HG326622					KM517508	+	+
	ARsp-Zab2-W			KM517348	KM517402	KM517458			
	ARsp-Bos-W	HG326624							
	ARsp-Kan-W	HG326626							
	ARsp-Kor-W	HG326625					KM517505		+
	ARsp-Sos-W	HG326627		KM517347	KM517401	KM517457	KM517507		+
M. arion summer type	ARsu-Sin-W	HG326630					KM517513		+
	ARsu-Szt-W	HG326629							
	ARsu-Per-W	HG326632							
	ARsu-Zab-W	HG326631		KM517355	KM517407	KM517465	KM517515		+
	ARsu-Agt1-W	HG326628							
	ARsu-Agt2-W			KM517351	KM517403	KM517461			
	ARsu-Kor-W	HG326633					KM517511		+
	ARsu-Dra1-W	KM517249					KM517510		+
	ARsu-Dra2-W			KM517352	KM517404	KM517462			

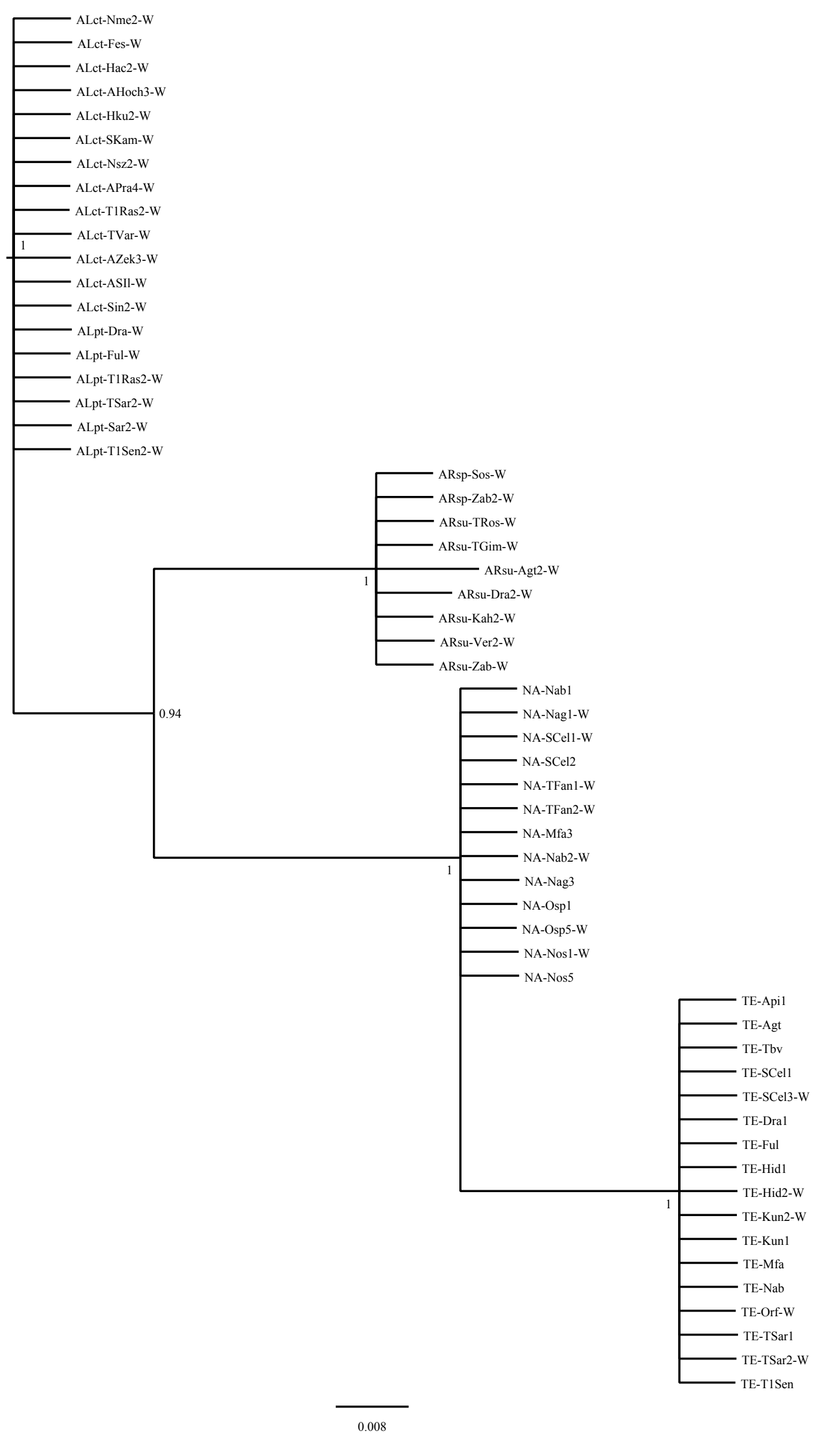
Species	ID	*Maculinea*					*Wolbachia*		
		COI part 1	**COI part 2**	**GAPDH**	**MDH**	**wg**	**16S rRNA**	**WSP**	**MLST**
M. arion summer type (cont.)	ARsu-Kah1-W	HG326634					KM517512		+
	ARsu-Kah2-W			KM517353	KM517405	KM517463			
	ARsu-Ver1-W	HG326635					KM517514		+
	ARsu-Ver2-W			KM517354	KM517406	KM517464			
	ARsu-SKam1-W	HG326638					KM517516		
	ARsu-SKam2-W	HG326639					KM517517		+
	ARsu-SPol1-W	HG326640					KM517519		
	ARsu-SPol2-W	HG326641					KM517520		
	ARsu-SPol3-W	HG326642					KM517518		+
	ARsu-TGim-W	HG326636		KM517350		KM517460	KM517509		+
	ARsu-TRos-W	HG326637		KM517349		KM517459			
M. nausithous	NA-Nab1	KM517298		KM517356	KM517408	KM517466			
	NA-Nab2-W	KM517299		KM517363	KM517415	KM517473	KM517547		
	NA-Urh	KM517300							
	NA-Api1	KM517301							
	NA-Api2-W	KM517302					KM517541		
	NA-Api3-W						KM517542		
	NA-Orf1-W	KM517303					KM517550		
	NA-Orf2-W						KM517551		
	NA-Orf3-W						KM517552		
	NA-Ket1	HG326644							
	NA-Ket2	KM517304							
	NA-Hha1	KM517305							
	NA-Hha2-W	KM517306					KM517544		
	NA-Szf1	KM517307							
	NA-Szf2-W						KM517561		
	NA-Osp1	KM517308		KM517365	KM517417	KM517475			
	NA-Osp2-W	KM517309					KM517553		
	NA-Osp3-W						KM517554		
	NA-Osp4-W						KM517555		

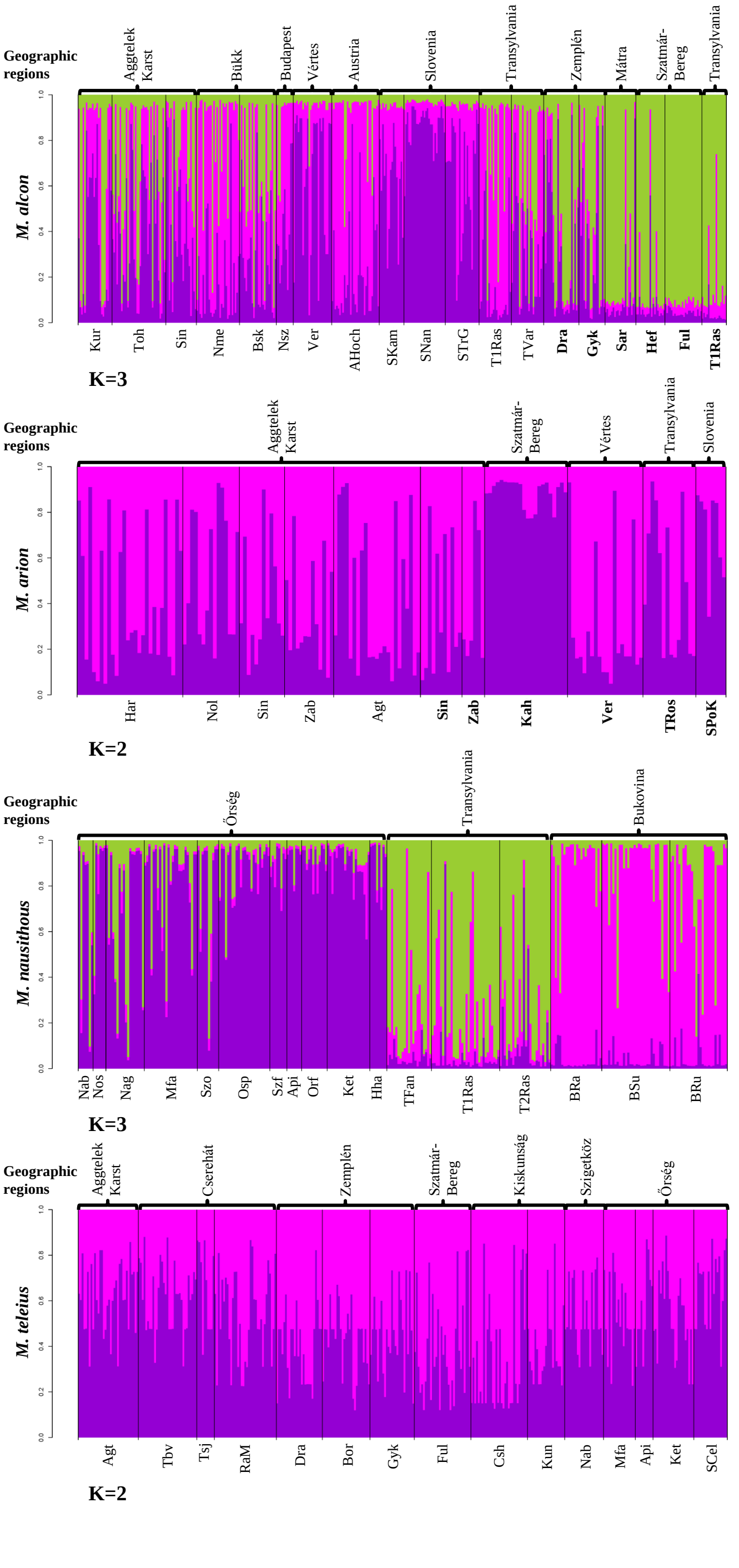
Species	ID	*Maculinea*					*Wolbachia*		
		COI part 1	**COI part 2**	**GAPDH**	**MDH**	**wg**	**16S rRNA**	**WSP**	**MLST**
M. nausithous (cont.)	NA-Osp5-W	KM517310		KM517366	KM517418	KM517476	KM517556		
	NA-Szo1-W	KM517311					KM517564		
	NA-Szo2-W	KM517312					KM517565		
	NA-Mfa1	KM517313							
	NA-Mfa2	KM517314							
	NA-Mfa3			KM517362	KM517414	KM517472			
	NA-Nos1-W	KM517315		KM517367	KM517419	KM517477			
	NA-Nos2-W						KM517557		
	NA-Nos3-W						KM517558		
	NA-Nos4-W	KM517316					KM517559		
	NA-Nos5			KM517368	KM517420	KM517478			
	NA-Nag1-W	KM517317		KM517357	KM517409	KM517467	KM517548		
	NA-Nag2-W						KM517549		
	NA-Nag3			KM517364	KM517416	KM517474			
	NA-Cst	KM517318							
	NA-SCel1-W	KM517319		KM517358	KM517410	KM517468	KM517543		
	NA-SCel2	KM517320		KM517359	KM517411	KM517469			
	NA-T1Ras	KM517321							
	NA-TFan1-W	KM517322		KM517360	KM517412	KM517470	KM517545		
	NA-TFan2-W	KM517323		KM517361	KM517413	KM517471	KM517546		
	NA-BRa1	KM517324							
	NA-BRa2	KM517325							
	NA-BSu1-W	KM517326							
	NA-BSu2-W						KM517562		
	NA-BSu3-W						KM517563		
	NA-BRu1	KM517327							
	NA-BRu2-W	KM517328					KM517560		

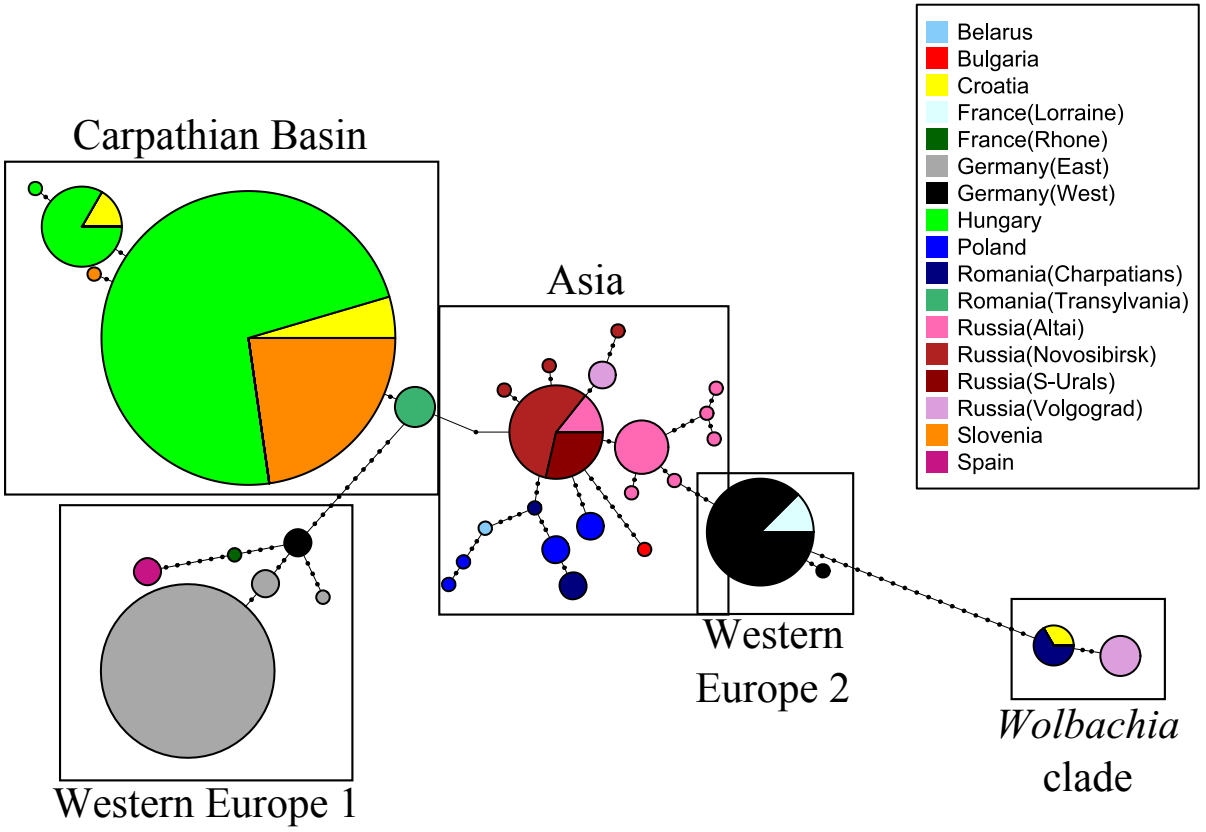
Species	ID	*Maculinea*					*Wolbachia*		
		COI part 1	**COI part 2**	**GAPDH**	**MDH**	**wg**	**16S rRNA**	**WSP**	**MLST**
M. teleius	TE-Agt	KM517275		KM517370	KM517422	KM517480			
	TE-Tsj	KM517276							
	TE-Tbv	KM517277		KM517371	KM517423	KM517481			
	TE-RaM-W	KM517278					KM517504		
	TE-Bor1	KM517281							
	TE-Bor2-W						KM517501		
	TE-Dra1	KM517279		KM517373	KM517426	KM517484			
	TE-Dra2-W	KM517280					KM517500		
	TE-Bas-W	KM517282					KM517502		
	TE-Ful	KM517283		KM517374	KM517427	KM517485			
	TE-Kun1	KM517284		KM517378	KM517431	KM517489			
	TE-Kun2-W	KM517285		KM517377	KM517430	KM517488	KM517498		
	TE-Nab	KM517286		KM517379	KM517433	KM517491			
	TE-Him	KM517329							
	TE-Hid1	KM517330		KM517375	KM517428	KM517486			
	TE-Hid2-W	KM517331		KM517376	KM517429	KM517487	KM517499		
	TE-Api1	KM517287		KM517369	KM517421	KM517479			
	TE-Api2	KM517288							
	TE-Orf-W	KM517289			KM517434	KM517492	KM517496		
	TE-Ket1	KM517290							
	TE-Ket2-W	KM517291					KM517497		
	TE-Mfa	KM517292			KM517432	KM517490			
	TE-SCel1	KM517293		KM517372	KM517424	KM517482			
	TE-SCel2	KM517294							
	TE-SCel3-W				KM517425	KM517483			
	TE-T1Sen	KM517295		KM517381	KM517437	KM517495			
	TE-TSar1	KM517296			KM517435	KM517493			
	TE-TSar2-W	KM517297		KM517380	KM517436	KM517494	KM517503		

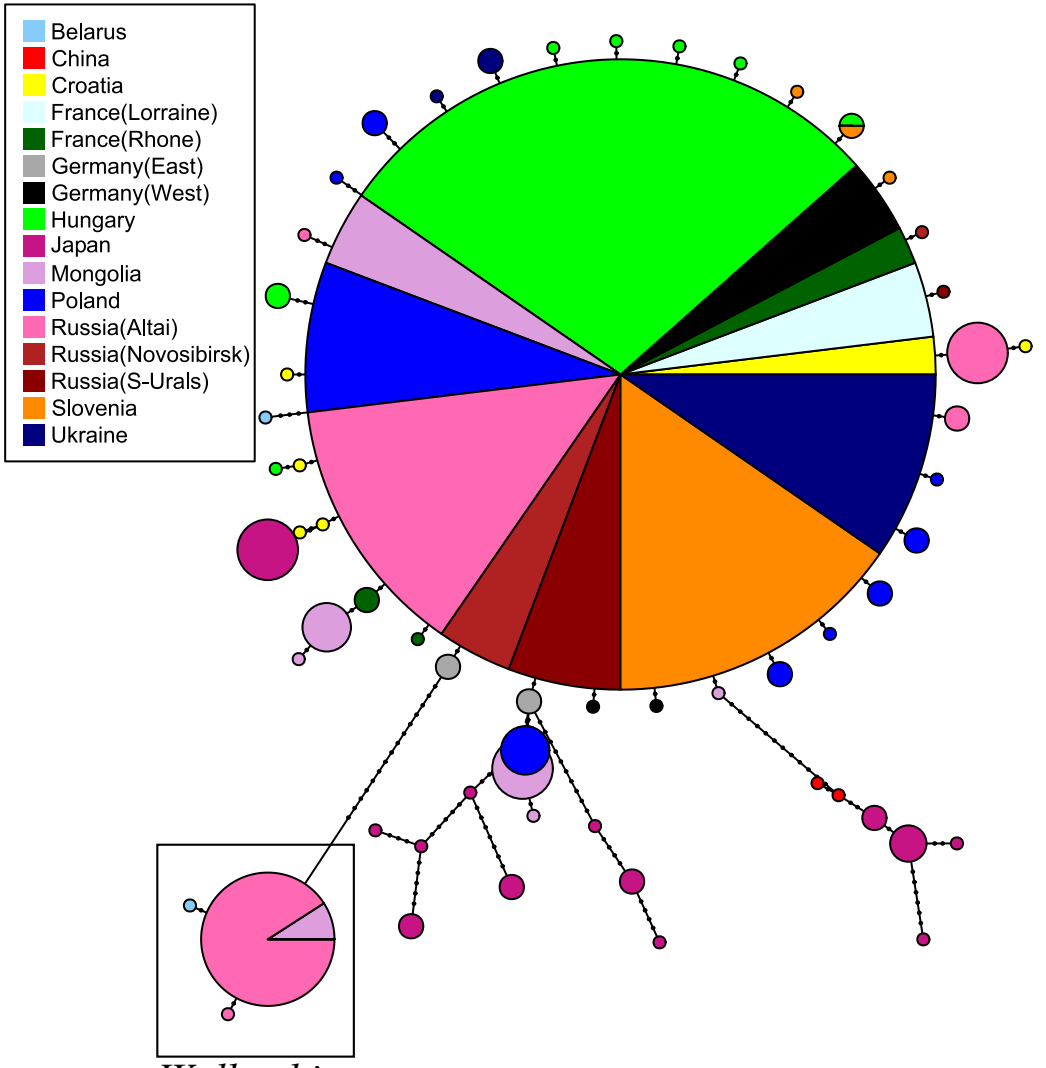

ALct-Nme2-W
ALct-Hac2-W
ALct-AHoch3-W
ALct-Hku2-W
ALct-SKam-W
ALct-Nsz2-W
ALct-APra4-W
ALct-TVar-W
ALct-ASIl-W
1
ALct-Sin2-W
ALpt-Dra-W
ALpt-Ful-W
ALpt-T1Ras2-W
ALpt-TSar2-W
ALpt-T1Sen2-W
ARsp-Sos-W
ARsp-Zab2-W
ARsu-TRos-W
ARsu-TGim-W
1
ARsu-Agt2-W
ARsu-Dra2-W
ARsu-Kah2-W
ARsu-Ver2-W
ARsu-Zab-W
NA-Nab1
NA-Nag1-W
NA-SCel1-W
NA-SCel2
NA-T1Fan1-W
NA-T1Fan2-W
0.91
NA-Mfa3
NA-Nab2-W
NA-Nag3
NA-Osp1
NA-Osp5-W
NA-Nos1-W
0.99
NA-Nos5
1
TE-Api1
TE-Agt
TE-Tbv
TE-SCel1
TE-Dra1
TE-Ful
TE-Hid1
1
TE-Hid2-W
TE-Kun2-W
TE-Kun1
TE-Nab
TE-TSar2-W
TE-T1Sen
1
0.02


ALct-Nme2-W
ALct-Fes-W
ALct-Hac2-W
ALct-Hku2-W
ALct-SKam-W
ALct-Nsz2-W
ALct-APra4-W
ALct-T1Ras2-W
ALct-TVar-W
1
ALct-AZek3-W
ALct-ASIl-W
ALct-Sin2-W
ALpt-Dra-W
ALpt-Ful-W
ALpt-T1Ras2-W
ALpt-TSar2-W
ALpt-T1Sen2-W
ALct-AHoch3-W
0.84
ALpt-Sar2-W
ARsp-Sos-W
ARsp-Zab2-W
ARsu-Agt2-W
1
ARsu-Dra2-W
ARsu-Kah2-W
ARsu-Ver2-W
ARsu-Zab-W
NA-Nab1
NA-Nag1-W
NA-SCel1-W
NA-TFan1-W
1
NA-TFan2-W
NA-Osp5-W
1
NA-SCel2
NA-Mfa3
NA-Nab2-W
NA-Nag3
NA-Osp1
NA-Nos1-W
NA-Nos5
TE-Api1
TE-Agt
TE-Tbv
TE-SCel1
TE-SCel3-W
TE-Dra1
TE-Ful
TE-Hid1
TE-Hid2-W
1
TE-Kun2-W
TE-Kun1
TE-Mfa
TE-Nab
TE-Orf-W
TE-TSar1
TE-TSar2-W
TE-T1Sen
1
1
0.02


ALct-Nme2-W
ALct-Fes-W
ALct-Hac2-W
ALct-AHoch3-W
ALct-Hku2-W
ALct-SKam-W
ALct-Nsz2-W
ALct-APra4-W
ALct-T1Ras2-W
ALct-TVar-W
1
ALct-AZek3-W
ALct-ASIl-W
ALct-Sin2-W
ALpt-Dra-W
ALpt-Ful-W
ALpt-T1Ras2-W
ALpt-TSar2-W
ALpt-Sar2-W
ALpt-T1Sen2-W
ARsp-Sos-W
ARsp-Zab2-W
ARsu-TRos-W
ARsu-TGim-W
ARsu-Agt2-W
1
ARsu-Dra2-W
ARsu-Kah2-W
ARsu-Ver2-W
ARsu-Zab-W
0.94
NA-Nab1
NA-Nag1-W
NA-SCel1-W
NA-SCel2
NA-TFan1-W
NA-TFan2-W
NA-Mfa3
NA-Nab2-W
1
NA-Nag3
NA-Osp1
NA-Osp5-W
NA-Nos1-W
NA-Nos5
TE-Api1
TE-Agt
TE-Tbv
TE-SCel1
TE-SCel3-W
TE-Dra1
TE-Ful
TE-Hid1
TE-Hid2-W
1
TE-Kun2-W
TE-Kun1
TE-Mfa
TE-Nab
TE-Orf-W
TE-TSar1
TE-TSar2-W
TE-T1Sen
0.008


Geographic regions
M. alcon
0.0
0.2
0.4
0.6
0.8
1.0
Aggtelek Karst
Bükk
Budapest
Vértes
Austria
Slovenia
Transylvania
Zemplén
Mátra
Szatmár-Bereg
Transylvania
Kur
Toh
Sin
Nme
Bsk
Nsz
Ver
AHoch
SKam
SNan
STrG
T1Ras
TVar
Dra
Gyk
Sar
Hef
Ful
T1Ras
K=3
Geographic regions
M. arion
Aggtelek Karst
Szatmár-Bereg
Vértes
Transylvania
Slovenia
Har
Nol
Sin
Zab
Agt
Sin
Zab
Kah
Ver
TRos
SPoK
K=2
Geographic regions
M. nausithous
Őrség
Transylvania
Bukovina
Nab
Nos
Nag
Mfa
Szo
Osp
Szf
Api
Orf
Ket
Hha
TFan
T1Ras
T2Ras
BRa
BSu
BRu
K=3
Geographic regions
M. teleius
Aggtelek Karst
Cserehát
Zemplén
Szatmár-Bereg
Kiskunság
Szigetköz
Őrség
Agt
Tbv
Tsj
RaM
Dra
Bor
Gyk
Ful
Csh
Kun
Nab
Mfa
Api
Ket
SCel
K=2


Belarus
Bulgaria
Croatia
France(Lorraine)
France(Rhone)
Germany(East)
Germany(West)
Hungary
Poland
Romania(Charpatians)
Romania(Transylvania)
Russia(Altai)
Russia(Novosibirsk)
Russia(S-Urals)
Russia(Volgograd)
Slovenia
Spain
Carpathian Basin
Asia
Western Europe 2
Wolbachia clade
Western Europe 1


Belarus
China
Croatia
France(Lorraine)
France(Rhone)
Germany(East)
Germany(West)
Hungary
Japan
Mongolia
Poland
Russia(Altai)
Russia(Novosibirsk)
Russia(S-Urals)
Slovenia
Ukraine
Wolbachia
clade