Diversification of Wolbachia Endosymbiont in the Culex pipiens Mosquito

Célestine Atyame Nten, Frédéric Delsuc, Nicole Pasteur, Mylène Weill, Olivier
Duron

1 **Diversification of *Wolbachia* endosymbiont in the *Culex pipiens***

2 **mosquito**

3

4 Célestine M. Atyame[*], Frédéric Delsuc[*], Nicole Pasteur[*], Mylène Weill[*] and Olivier Duron[*]

5

6 [*] Institut des Sciences de l'Evolution

7 Université Montpellier 2, CC 054

8 CNRS, UMR5554

9 Place Eugène Bataillon

10 34095 Montpellier cedex 05

11 France

12

13 Corresponding author: Olivier Duron; E-mail: olivier.duron@univ-montp2.fr

14

20

Abstract

The α-proteobacteria *Wolbachia* are among the most common intracellular bacteria and have recently emerged as important drivers of arthropod biology. *Wolbachia* commonly act as reproductive parasites in arthropods by inducing cytoplasmic incompatibility (CI), a type of conditional sterility between hosts harboring incompatible infections. In this study, we examined the evolutionary histories of *Wolbachia* infections, known as *w*Pip, in the common house mosquito *Culex pipiens*, which exhibits the greatest variation in CI crossing patterns observed in any insect. We first investigated a panel of twenty *w*Pip strains for their genetic diversity through a multi-locus scheme combining thirteen *Wolbachia* genes. Because *Wolbachia* depend primarily on maternal transmission for spreading within arthropod populations, we also studied the variability in the co-inherited *Cx. pipiens* mitochondria. In total, we identified fourteen *w*Pip haplotypes, which all share a monophyletic origin and clearly cluster into five distinct *w*Pip groups. The diversity of *Cx. pipiens* mitochondria was extremely reduced, which is likely a consequence of cytoplasmic hitchhiking driven by a unique and recent *Wolbachia* invasion. Phylogenetic evidence indicates that *w*Pip infections and mitochondrial DNA have co-diverged through stable co-transmission within the cytoplasm and shows that a rapid diversification of *w*Pip has occurred. The observed pattern demonstrates that a considerable degree of *Wolbachia* diversity can evolve within a single host species over short evolutionary periods. In addition, multiple signatures of recombination were found in most *w*Pip genomic regions, leading us to conclude that the mosaic nature of *w*Pip genomes may play a key role in their evolution.

Introduction

Arthropods are commonly infected by maternally transmitted endosymbionts. Although some endosymbionts confer direct benefits to their hosts by providing anabolic functions or resistance to pathogens (Haine 2008; Moran, McCutcheon, and Nakabachi 2008), others are associated with alterations of host reproduction (Werren, Baldo, and Clark 2008; Engelstadter and Hurst 2009). These symbionts represent reproductive parasites that include diverse unrelated bacteria, among which the alpha-proteobacteria *Wolbachia* are the most widespread (Duron et al. 2008; Hilgenboecker et al. 2008). In some host species, the successful spread of *Wolbachia* is achieved by biasing the host's sex ratio toward the production of females (the transmitting sex) through the induction of parthenogenesis, feminization or male-killing. More commonly, *Wolbachia* are able to induce a form of conditional sterility, termed cytoplasmic incompatibility (CI), between infected males and uninfected females or females infected by incompatible strains (Werren, Baldo, and Clark 2008; Engelstadter and Hurst 2009). Such manipulations enable *Wolbachia* to spread through arthropod populations and may drive arthropod evolution through their effects on host phenotypes (Moran, McCutcheon, and Nakabachi 2008; Werren, Baldo, and Clark 2008; Engelstadter and Hurst 2009).

The dynamics of *Wolbachia* infections within the common house mosquito, *Culex pipiens* complex, remain poorly understood. The most common members of the complex are the subspecies *Cx. p. quinquefasciatus* (Say) and *Cx. p. pipiens* (L.), representing the southern and northern mosquito populations, which are ubiquitous in tropical and temperate regions, respectively (Barr 1982). Members of the *Cx. pipiens* complex exhibit the greatest variation of CI crossing types observed in arthropods thus far (Laven 1967; O'Neill and Paterson 1992; Guillemaud, Pasteur, and Rousset 1997; Duron et al. 2006; Atyame et al. in press). However, an early genotyping approach using the *ftsZ* gene failed to reveal any polymorphism between

67 incompatible *Wolbachia* strains (Guillemaud, Pasteur, and Rousset 1997). Further analyses

68 used sequences of published complete *Wolbachia* genomes to characterize polymorphic

69 molecular markers. Genomes of *Wolbachia* strains infecting arthropod are scattered with

70 mobile genetic elements (MGE), such as prophages and transposable elements, which can

71 represent more than 20% of genome content (Wu et al. 2004; Klasson et al. 2008; Salzberg et

72 al. 2009; Klasson et al. 2009). Additionally, they contain an unusual high number of genes

73 encoding proteins with ankyrin (ANK) motif(s), which possibly mediate specific protein-

74 protein interactions (Sinkins et al. 2005; Duron et al. 2007; Walker et al. 2007). We further

75 developed genotyping approaches using ANK and MGE markers and used them to identify

76 more than 100 genetically distinct *Wolbachia* strains (referred to as *w*Pip strains) in natural

77 populations of *Cx. pipiens* (Duron et al. 2005; Duron et al. 2006; Duron et al. 2007; Atyame

78 et al. in press; Duron, Raymond, and Weill in press).

79

80 In this study, we characterized the evolutionary history of *Wolbachia* infections in the *Cx.*

81 *pipiens* complex by examining the association between the *w*Pip strains and *Cx. pipiens*

82 mitochondrial DNA variation. The predominant mode of *Wolbachia* transmission within a

83 species is vertical, *via* the egg cytoplasm (Werren, Baldo, and Clark 2008; Engelstadter and

84 Hurst 2009). Because *Wolbachia* and mitochondrial genomes are co-transmitted and,

85 therefore, in linkage disequilibrium, the spread of *Wolbachia* will strongly affects a host's

86 mtDNA diversity through indirect selection (review in Hurst and Jiggins 2005). However,

87 exceptions to strict vertical transmission have been found; in some cases, *Wolbachia* are also

88 transferred through horizontal transmission both within and among different host species,

89 although the mechanisms of transfer are not well understood (Ahrens and Shoemaker 2005;

90 Baldo et al. 2008; Raychoudhury et al. 2009). Consequently, the wide distribution of

91 *Wolbachia* among arthropods is generally assumed to result from complex interactions

92   between vertical and horizontal modes of transmission, modulated by their capacity to alter

93   host reproduction.

94

95   Here, we analyzed *w*Pip variability and the associated mtDNA diversity in twenty *Cx. pipiens*

96   lines encompassing uni- and bidirectionally incompatible strains that originated from different

97   geographic areas. Our results showed that the *w*Pip strains form a monophyletic clade of

98   closely related bacteria and that *Cx. pipiens* harbors a low level of mitochondrial variability,

99   which is a probable consequence of a recent *Wolbachia* invasion through cytoplasmic

100 hitchhiking. Investigation of *w*Pip sequences revealed extensive recombination between *w*Pip

101 strains, although multiple infections within a single mosquito were never detected using our

102 markers. However, a congruence between *w*Pip and mtDNA phylogenies was shown,

103 demonstrating that *Wolbachia* mainly use maternal inheritance to spread through *Cx. pipiens*

104 populations. The evolutionary implications of horizontal transfers and the question of whether

105 the *Cx. pipiens - Wolbachia* association is a unique case or a representative example is

106 discussed.

107

108

109 **Materials and Methods**

110 Mosquito collection

111 Twenty *Cx. pipiens* lines from a broad geographical range were examined (Table 1). This

112 collection encompassed the two main *Cx. pipiens* subspecies, *Cx. p. pipiens* and *Cx. p.*

113 *quinquefasciatus*, which are naturally infected with compatible and incompatible *w*Pip strains

114 (for more details, see Duron et al. 2006; Duron et al. 2007). The study also included the two

115 lines for which the *w*Pip genome has been sequenced, *w*Pip(Pel) (GenBank AM999887;

116 Klasson et al. 2008) and *w*Pip(JHB) (ABZA01000000; Salzberg et al. 2009).

117

Wolbachia markers

The *w*Pip strains were first genotyped for the five housekeeping genes developed for the

Wolbachia MultiLocus Strain Typing (MLST) methodology, *gatB*, *coxA*, *hcpA*, *ftsZ* and *fbpA*

(Baldo et al. 2006), and the *Wolbachia* surface protein gene *wsp* (Braig et al. 1998). The

MLST system is classically used to characterize the eight supergroups (A to I) currently

recognized within the *Wolbachia* genus (Lo et al. 2007; Ros et al. 2009). The polymorphism

of seven additional genes was also examined: the DNA mismatch repair protein gene *MutL*

(one copy in the *w*Pip(Pel) genome), three ANK genes, *ank2* (one copy), *pk1* (three identical

copies), and *pk2* (two identical copies), and three phage genes, the methylase gene *GP12* (four

identical copies), the putative secreted protein gene *GP15* (also known as *VrlC*; one copy) and

the regulatory protein gene *RepA* (one copy). None of these genes was amplified from

Wolbachia-free *Cx. pipiens* lines, which confirmed their *Wolbachia* origin. A total of 13

Wolbachia genes, encompassing 19 distinct loci with a wide distribution along the *w*Pip(Pel)

chromosome, were examined (Figure 1, Table S1).

132

Culex pipiens mitochondrial markers

The complete mitochondrial genome of the Pel line (15,587 bp) was obtained through BLAST

searches of the database of *w*Pip(Pel) contig DNA sequences from the Wellcome Trust-

Sanger Institute website (http://www.sanger.ac.uk/Projects/W_pipientis/) using the

mitochondrial genome of *Aedes albopictus* (GenBank AY072044) as a probe. A similar

approach to obtain the mitochondrial sequences of the JHB line from the VectorBase website

(http://www.vectorbase.org/) showed several divergent mitochondrial sequences, some of

which exhibited low read coverage, making the JHB sequences unreliable for further analysis.

141

142   A set of primers (Table S2) was designed from the Pel sequences and further used to obtain

143   the complete mitochondrial genomes (with the exception of the A+T-rich region) of

144   additional *Cx. pipiens* lines. Specific primers were also used to amplify a 613 bp fragment

145   from the NADH dehydrogenase subunit 2 (*ND2*) gene, a 1,132 bp fragment from the NADH

146   dehydrogenase subunit 5 (*ND5*) gene and an 852 bp fragment from the cytochrome b (*cytb*)

147   gene from all the investigated *Cx. pipiens* lines (Table S2).

148

149   PCR amplification and sequencing

150   DNA was extracted from individual mosquitoes using a CetylTrimethylAmmonium Bromide

151   (CTAB) protocol (Rogers and Bendich 1988). Amplification conditions were 3 min at 94°C,

152   followed by 30 cycles of 94°C for 30 s, 52°C for 30 s (58°C for *MutL*), and 72°C for 1 to 1.5

153   min depending on the fragment size. Amplified fragments were run in agarose gel (1.5%)

154   electrophoresis. The QIAquick gel extraction kit (QIAGEN, Valencia, CA) was then used to

155   purify the PCR products. Sequences were obtained directly for purified products using an ABI

156   Prism 3130 sequencer with the BigDye Terminator Kit (Applied Biosystems). The sequences

157   have been deposited in the GenBank database (accession numbers in Tables S1 and S2).

158

159   Sequence analyses

160   Sequence alignments were carried out using CLUSTALW (Thompson, Gibson, and Higgins

161   2002) and corrected using MEGA (Tamura et al. 2007). The GBLOCKS program (Castresana

162   2000) with default parameters was used to remove poorly aligned positions and to obtain

163   nonambiguous sequence alignments. The number of variable sites (VI), nucleotide diversity

164   (π), G+C content and the ratios of nonsynonymous *versus* synonymous substitutions (*Ka/Ks*)

165   were computed using DNASP (Librado and Rozas 2009). Nonrandom associations between

166   each pair of loci were estimated through the measure of allele linkage disequilibrium (LD)

167   using the D' statistic (Lewontin 1964; Hedrick 1987). The exact test procedure implemented

168   in GENEPOP (Raymond and Rousset 1995) was further used to test LD significance.

169   Statistical analyses for intragenic recombination were performed with the Sawyer's test

170   implemented in GENECONV (Sawyer 1999). A Bonferonni adjustment correction for

171   multiple testing was applied (Hochberg 1988).

172

173   Annotation of the *Cx. pipiens* mitochondrial genome was based on alignments with

174   mitochondrial sequences from *Ae. albopictus* (AY072044), *Ae. aegypti* (EU352212),

175   *Anopheles gambiae* (L20934), *An. funestus* (DQ146364) and *An. quadrimaculatus* (L04272).

176

177   Tree reconstruction

178   Phylogenetic relationships were evaluated for *Wolbachia* and *Cx. pipiens* mitochondrial

179   sequences. The best-fitting models of sequence evolution for each dataset were determined

180   using the Akaike information criterion with MODELTEST v3.7 (Posada and Crandall 1998).

181   The selected model was the General Time Reversible model, with Gamma distributed among-

182   site rate variation (GTR+G) for both *Wolbachia* and mitochondrial sequence datasets.

183   Bayesian inferences (BI) were used to reconstruct phylogenies using Mr. Bayes v 3.1.2

184   (Ronquist and Huelsenbeck 2003). Two independent replicates of four Metropolis-coupled

185   Monte Carlo Markov chains were run for 2,000,000 generations with Model parameters and

186   trees sampled every 200 generations. Bayesian posterior probabilities were obtained from the

187   50% majority-rule consensus of the sampled trees after discarding the initial burn-in period.

188   The resulting phylogenetic trees were visualized and edited in MEGA (Tamura et al. 2007).

189

190   *Wolbachia* genes were also analyzed within a phylogenetic network framework to account for

191   potentially conflicting signals due to recombination (Fitch 1997). A phylogenetic network

192   was constructed based on uncorrected *P*-distances using the Neighbor-net method (Bryant and

193   Moulton 2004) implemented in SPLITSTREE (Huson and Bryant 2006). Neighbor-net is a

194   distance-based method to construct a network as a generalization of all possible phylogenetic

195   trees that can be reconstructed from conflicting signals in the data.

196

197   Assessing the maximum age of mitochondrial sweep

198   We used the mitochondria data to infer the maximum age of the *Wolbachia* infection in *Cx.*

199   *pipiens* following the method of Rich et al. (1998). This method assumes that selection only

200   occurs at the protein level and that DNA polymorphism in degenerate sites is neutral. We

201   used fourfold and twofold synonymous sites from protein coding mtDNA sequences to assess

202   the age of the sweep. The number of twofold and fourfold synonymous sites in each coding

203   region was computed with MEGA (Tamura et al. 2007), and a conservative Jukes-Cantor

204   correction was applied for multiple hits.

205

206

207   **Results**

208   Monophyletic origin of the *w*Pip strains

209   The MLST and *wsp* genes did not exhibit sequence variation between the *w*Pip strains (eight

210   strains were examined here, i.e. *w*Pip(Sl), *w*Pip(Tn), *w*Pip(Ko), *w*Pip(Lv), *w*Pip(Is),

211   *w*Pip(Mc), *w*Pip(Pel) and *w*Pip(JHB)), establishing that these strains are very closely related.

212   The *w*Pip MLST sequence data were compared to sequences from 18 other strains belonging

213   to five distinct *Wolbachia* supergroups (A, B, D, F and H). The phylogenetic tree obtained

214   from the 2,079-bp concatenated MLST genes revealed that the *w*Pip strains form a robust

215   monophyletic clade within the B supergroup, which is closely related to the *w*Bol1 strain

216   present in the butterfly *Hypolimnas bolina* (identity > 99.9%; Figure S1).

217

High variability of *w*Pip genomes

219   Seven of the examined *Wolbachia* genes were polymorphic among the *w*Pip genomes: the

220   DNA mismatch protein gene *MutL*, three ANK genes *ank2*, *pk1* and *pk2*, and three phage

221   genes *GP12*, *GP15* and *RepA*. Analyses revealed considerable allelic variability among the 20

222   *w*Pip strains, with 2-8 alleles being found per gene (Table S3). This polymorphism was

223   mainly due to nucleotide substitutions, insertions or deletions (indels); note that an insertion

224   of the *Tr1* transposon (also known as ISWpi1; see Duron et al. 2005; Cordaux 2008) was

225   observed within the *RepA* sequence of three *w*Pip strains. An additional source of variability

226   arose from a *GP15* deletion in the *w*Pip(JHB) genome. A letter was attributed to each distinct

227   allele of the seven genes, the combination of which identified 14 *w*Pip haplotypes among the

228   twenty strains (Table 2).

229

230   Although the prophage related genes *pk1*, *pk2* and *GP12* were found to be present in several

231   copies in the *w*Pip(Pel) genome, divergent copies were never amplified from our *w*Pip strains:

232   direct sequences of PCR products were easily readable and showed no overlapping peaks.

233   This indicates that the different copies (if any) present in each *w*Pip strain examined here are

234   identical, as observed in *w*Pip(Pel). It further shows that only mono-*w*Pip-infections (or multi-

235   infections by closely related *w*Pip strains) are present within *Cx. pipiens* individuals.

236

High recombination in *w*Pip genomes

238   Recombination, both within and between *Wolbachia* genes, can blur molecular signals and

239   result in misleading observations related to strain relationships. For this reason, we checked

240   the possibility of recombination among the seven polymorphic markers obtained here (*MutL*,

241   *ank2*, *pk1*, *pk2*, *GP12*, *GP15* and *RepA*). Pairwise tests for intergenic recombination revealed

242   significant linkage disequilibrium (LD) for *MutL*, *ank2*, *pk1*, *GP12*, and *GP15* (Table S4).

243   Alleles at these five loci are not randomly associated and are stably co-transmitted within the

244   *w*Pip chromosome. However, nonsignificant LD was found between *pk2* and *RepA* and

245   between these two genes and the other five genes, showing that recombination has disrupted

246   genome clonality by shuffling the *RepA* and *pk2* alleles among *w*Pip strains. Intragenic

247   recombination was also detected for at least five genes by Sawyer's test (*MutL*, *pk1*, *pk2*,

248   *GP12* and *GP15*; Table S3). Intragenic recombination results in identical nucleotides or

249   amino acid motifs in *w*Pip strains divergent at other loci, which are readily apparent through

250   the examination of sequence alignments (Figure S2).

251

252   Gene rearrangements in *w*Pip genomes

253   Genome organization of *w*Pip strains was analyzed by comparing the locations of the thirteen

254   genes surveyed in this study in the *w*Pip(Pel) chromosome and in the five major *w*Pip(JHB)

255   contigs presently available (Figure S3). There are several rearrangements distinguishing these

256   genomes, in which diverse genes have been inverted (e.g. *gatB*, *coxA*), translocated (*MutL*,

257   *RepA*), duplicated or deleted (three and one *pk1* copies are found in *w*Pip(Pel) and *w*Pip(JHB),

258   respectively). Notably, rearrangements are not limited to phage regions, which are prone to

259   movements within and between genomes, but also affect housekeeping genes.

260

261   Inference of wPip strain relationships

262   Phylogenetic analyses of the 20 *w*Pip strains using the six *w*Pip genes, *MutL*, *ank2*, *pk1*, *pk2*,

263   *GP12* and *GP15,* revealed significant topological incongruence, as expected for a dataset

264   affected by recombination (Figure S4). For instance, the *w*Pip(Sl), *w*Pip(Bf-B) and *w*Pip(Mc)

265   strains are genetically similar for four markers (*ank2*, *pk1*, *GP12* and *GP15*), but appear

266   distantly related for two others markers (*MutL* and *pk2*).

267

268   To assess *w*Pip strain relationships, we performed phylogenetic analyses based on the

269   concatenated sequences of the seven genes. The concatenated tree deduced from Bayesian

270   inference splits the *w*Pip clade into five groups (designated *w*Pip-I to *w*Pip-V; Figure 2).

271   However, recombination can create artificial grouping of *w*Pip strains, and network analysis

272   was thus conducted to visualize recombination effects, which were illustrated by multiple

273   boxes (Figure 3). The evolutionary history of *w*Pip strains appears as a complex network with

274   multiple pathways interconnecting strains, emphasizing the mosaic nature of *w*Pip genomes.

275   Interestingly, despite recombination, the network analysis was congruent with the Bayesian

276   tree in recovering the same five *w*Pip groups with strong bootstrap values.

277

278   A spatial structuring of *w*Pip diversity emerged when the geographic distribution of *w*Pip

279   groups was examined, despite the limited number (19) of strains. The most common group,

280   *w*Pip-I, is distributed widely from Asia to Europe (Figure 4), and all *w*Pip strains recently

281   identified at La Réunion island (Indian Ocean) by Atyame et al. (in press) belong to that

282   group. The *w*Pip-V group is only found in East Asia, and the *w*Pip-II and *w*Pip-III groups

283   have an apparently discontinuous distribution, with strains being found in very distant

284   geographic areas (for example, the *w*Pip-II strains are from Australia and Europe).

285

286   Low mitochondrial diversity in *Cx. pipiens*

287   *Cx. pipiens* Pel mitochondrial DNA exhibits classical features found in other mosquito

288   species that have been analyzed. It contains tightly packed genes with high A+T content

289   (78.2%). There are twenty-two genes coding tRNAs, two coding ribosomal RNAs, thirteen

290   genes coding subunits of enzymes involved in oxidative phosphorylation and, finally, an A+T

291   rich noncoding region (Figure S5 and Table S5). The genes are arranged along the

292 chromosome in a manner similar to that of other mosquito species (Beard, Mills Hamm, and

293 Collins 1993; Mitchell, Cockburn, and Seawright 1993; Krzywinski, Grushko, and Besansky

294 1997).

295

296 The complete *Cx. pipiens* mitochondrial genome (14,856 bp without the A+T rich region) was

297 sequenced from the lines Ko, Tn, Sl and Is and compared to the Pel genome. Overall, the five

298 mtDNA sequences displayed a very low variability, with only 36 variable nucleotidic

299 positions being found (ca. 2‰), and two sequences were strictly identical (lines Ko and Tn).

300 Among the 13 protein-coding genes, five genes (*atp8, atp*6, *ND3*, *ND4L* and *ND6*) showed no

301 polymorphism, whereas *ND2*, *ND5* and *cytb* were the most polymorphic (Figure S5).

302

303 A likely explanation of the low mtDNA diversity in *Cx. pipiens* populations is that

304 cytoplasmic hitchhiking has occurred during *Wolbachia* invasion, as suggested earlier by

305 Guillemaud et al. (Guillemaud, Pasteur, and Rousset 1997) and Rasgon et al. (Rasgon,

306 Cornel, and Scott 2006). To confirm this hypothesis, we compared the nucleotide diversity

307 per site (π) at four mitochondrial loci in the eleven Culicidae species for which the presence

308 or absence of *Wolbachia* has been documented (Table 3 and supplementary materials). Only

309 two species, *Cx. pipiens* and *Ae. albopictus*, are known to be infected, whereas *Wolbachia*

310 infection was never found in the nine other species. These two *Wolbachia*-infected species

311 harbor significantly lower mtDNA diversity than the uninfected species (Wilcoxon test,

312 W=10, $P = 0.008$). For instance, the worldwide mtDNA diversity of *Cx. pipiens* is lower than

313 the diversity observed in the North American populations of *Cx. tarsalis*, an uninfected

314 species (Venkatesan et al. 2007). The low diversity of mtDNA observed in the *Cx. pipiens*

315 and *Ae. albopictus* populations led us to conclude that *Wolbachia* is most likely the causative

316 agent of mitochondrial sweeps in these taxa.

317

Recent mitochondrial sweep in *Cx. pipiens* complex

We then assessed the date of the mitochondrial sweep using the nucleotide divergence of 13

protein-coding mtDNA genes from the *Cx. pipiens* Is line and *Ae. albopictus* (GenBank

AY072044). We estimated the substitution rate for these genes at twofold and fourfold

degenerate sites with a conservative Jukes–Cantor correction. The genera *Culex* and *Aedes*

diverged approximately 172 to 226 My ago (Reidenbach et al. 2009). Using the most recent

estimate (172 My), the mtDNA substitution rates (substitution/site/year) were estimated at

5.1×10^{-8} and 19×10^{-8} for the twofold and fourfold degenerate sites, respectively, whereas,

when using the oldest estimate (226 My), the substitution rates were 39×10^{-9} and 15×10^{-8}.

Among the mitochondrial genomes of five *Cx. pipiens* lines (Is, Sl, Tn, Ko and Pel), we

observed 13 and 10 nucleotide differences among the twofold (n=2,938) and fourfold

(n=1,343) synonymous sites, respectively. Thus, this dates the *Cx. pipiens* mitochondrial

sweep between 12,000 and 16,000 years before present (95% confidence interval, if *Culex*

and *Aedes* diverged 172 My ago) or between 16,000 and 21,000 (95% confidence interval, if

the two genera diverged 226 My ago). It is possible that the date of the mitochondrial sweep

is even more recent as it was assumed here that the substitution rates are constant, an

assumption known to overestimate divergence times (Ho et al. 2005).

Clear co-divergence of *w*Pip and *Cx. pipiens* mitochondria

The co-divergence of mitochondria and *w*Pip was assessed by studying the sequences of the

three polymorphic *Cx. pipiens* mtDNA genes (*ND2*, *ND5* and *cytb*), encompassing 2,549 bp

(16.4% of the whole mitochondrial genome).

342   Analysis of the *Cx. pipiens* mtDNA sequences among the 19 lines indicated the presence of

343   14 haplotypes (named *pi*1 to *pi*14), which differed overall at 22 variable nucleotide sites

344   (Table S6). The mtDNA of the *Cx. pipiens* lines differed by only one to nine nucleotides,

345   confirming their very high homology (99.6-99.9%). Phylogenetic analyses revealed two main

346   mitochondrial lineages (*pi*1 to *pi*5 and *pi*6 to *pi*14) with strong branching support (Figure 5A).

347

348   The concatenated mtDNA phylogeny and the *w*Pip phylogeny were congruent (Figure 5A and

349   B). A significant association was found between mtDNA haplotypes and *w*Pip haplotypes

350   (Fisher's exact test, $P = 3 \times 10^{-5}$), as well as *w*Pip groups ($P = 8 \times 10^{-4}$). This demonstrates that

351   *w*Pip infections and mitochondrial DNA have co-diverged through stable co-transmission

352   within the cytoplasm in *Cx. pipiens* populations. Hence, the two main mitochondrial lineages

353   parallel the *w*Pip divergence pattern and strongly confirm the *w*Pip phylogeny. Additionally,

354   *Cx. pipiens* subspecies are not significantly associated with *w*Pip haplotypes ($P = 0.37$), *w*Pip

355   groups ($P = 0.26$) or mtDNA haplotypes ($P = 0.10$). Thus, *Cx. pipiens* nuclear genomes have

356   not co-diverged with mitochondria and *w*Pip infections and exhibit a different evolutionary

357   history.

358

359

360   **Discussion**

361   Here, we examined 20 isolates of *Wolbachia* and their associated mitochondria within the *Cx.*

362   *pipiens* complex. The combined use of *Wolbachia* and host mtDNA multi-locus sequencing

363   revealed the processes driving the evolution of *Wolbachia* infections in this mosquito and

364   raised the question of their likeliness to occur in other *Wolbachia*-arthropod associations.

365

366   *Wolbachia* and mitochondrial markers reveal a recent diversification of *w*Pip strains

367   The examined MLST and *wsp* markers showed that *w*Pip infections form a robust

368   monophyletic clade within the B group of *Wolbachia*, confirming the results of Baldo et al.

369   (Baldo et al. 2006). Although these markers are widely used to characterize the genetic

370   diversity of *Wolbachia*, even within a host species (e.g. Dedeine et al. 2004; Baldo et al.

371   2006; Baldo et al. 2008; Raychoudhury et al. 2009), they displayed no variation among the

372   *w*Pip infections in *Cx. pipiens*, which shows that *w*Pip strains have a unique and recent

373   evolutionary origin. As observed in other *Wolbachia* infected species (Hurst and Jiggins

374   2005), mitochondrial diversity was low in populations of *Cx. pipiens* (and highly significantly

375   lower than in non–*Wolbachia*-infected mosquitoes), suggesting that *Wolbachia* have affected

376   mitochondrial polymorphism in this species through cytoplasmic hitchhiking. The observed

377   polymorphism of mitochondrial protein coding genes indicates that the mitochondrial sweep

378   due to the spread of *Wolbachia* occurred within the last 21,000 years. This dating is within the

379   range of values classically estimated for other *Wolbachia* host species of ca. $< 100,000$ years,

380   (Jiggins 2003; Keller et al. 2004; Duplouy et al. 2010), and the evolutionary pathway of *w*Pip

381   in *Cx. pipiens* could be similar to other *Wolbachia/*arthropod associations.

382

383   Multi-locus typing using seven *w*Pip polymorphic markers, including domains of the MGE

384   and ANK genes, allowed the identification of 14 distinct *w*Pip haplotypes, which cluster into

385   five distinct *w*Pip groups. This typing approach also established that the two published *w*Pip

386   genomes, *w*Pip(Pel) and *w*Pip(JHB), are genetically very close to each other compared to

387   strains belonging to other *w*Pip groups, in spite of their genomic differences (Salzberg et al.

388   2009). The variability of the investigated mitochondrial markers corroborates the inferences

389   made from the *w*Pip markers; thus, in *Cx. pipiens,* different mitochondrial haplotypes may

390   indicate that *w*Pip infections are different. Overall, the observed genetic diversity indicates

391   that, after the spread of *Wolbachia*, diversification of *w*Pip and *Cx. pipiens* mitochondria

392   occurred.

393

394   The diversity found for *w*Pip exhibits geographic variations. A remarkable degree of diversity

395   was found in the Mediterranean area, where four of the five *w*Pip groups are found, whereas a

396   reduced diversity was observed in other regions. The most common group, *w*Pip-I, has a wide

397   distribution (Asia, Africa and Europe) and was also recently reported at La Réunion Island

398   (Indian Ocean) (Atyame et al. in press). In contrast, some *w*Pip groups have a discontinuous

399   distribution, as exemplified by the *w*Pip-II strains, which were found in Europe and in

400   Australia. Such a geographic pattern is likely to be a consequence of a recent worldwide

401   expansion due to human activity (Raymond et al. 1991; Fonseca et al. 2004; Fonseca et al.

402   2006) or/and to selective advantages, possibly including CI selection. However, the 20 *w*Pip

403   infections investigated in this study represent a restricted sampling, occasionally from old

404   mosquito colonies, and further investigations are required to improve our knowledge of the

405   spatial structure of the *w*Pip groups worldwide.

406

407   *w*Pip strains are independent of *Cx. pipiens* subspecies

408   Strict vertical transmission must have favored the co-divergence of *w*Pip and mtDNA within

409   shared cytoplasm. However, there was no clear association between *Cx. pipiens* subspecies

410   (nuclear diversity) and cytoplasmic diversity (i.e. *Wolbachia* and mtDNA): identical *w*Pip

411   strains and identical mitochondrial haplotypes were found in the two subspecies, *Cx. p.*

412   *pipiens* and *Cx. p. quinquefasciatus*. A likely explanation for this is that the transfer of

413   cytoplasm between *Cx. pipiens* subspecies occurred through hybridization events, as observed

414   in *Drosophila* species (Rousset and Solignac 1995; Ballard 2000) and in butterfly species

415   (Jiggins 2003; Narita et al. 2006; Charlat et al. 2009). In *Cx. pipiens*, this hypothesis is well

416 supported by the many reports of genetic introgression between the two subspecies in areas

417 where they come into contact (Cornel et al. 2003; Fonseca et al. 2004). Hence, we can predict

418 that DNA barcoding programs using mtDNA will fail to discriminate between *Cx. p. pipiens*

419 and *Cx. p. quinquefasciatus*. Overall, these observations support the call of Hurst and Jiggins

420 (Hurst and Jiggins 2005) to not use mtDNA alone as a reliable means of taxa resolution.

421

422 Intense recombination impacts the structure of *w*Pip genomes

423 The existence of extensive recombination among *w*Pip strains sheds light on the mechanisms

424 shaping the evolution of *w*Pip genomes since recombination can influence the adaptive

425 dynamics of *Wolbachia* by creating new alleles and thus allow the emergence of new

426 phenotypes. Recombination between distant *Wolbachia* genomes has been previously

427 documented (Jiggins et al. 2001; Bordenstein and Wernegreen 2004; Baldo, Lo, and Werren

428 2005; Gavotte et al. 2007), although in this study, we found recombination among very

429 closely related *Wolbachia* genomes. Evidence of recombination was found at almost all of the

430 examined *w*Pip loci, WO-phage genes, as well as nonrelated phage loci. This shows that a

431 high level of gene flow occurs among the *Wolbachia* genomes in *Cx. pipiens*. Hence, the *w*Pip

432 strains do not form a set of clones in which evolution is independent but, rather, represent a

433 large population of bacteria exchanging genetic information through lateral transfers.

434 Although no instances of multiple infection were detected using our markers, we must assume

435 that they occur, at least during a period long enough to allow recombination between strains.

436

437 Another consequence of recombination is that it can lead to misinterpretation of phylogenetic

438 relationships between strains. However, despite the extensive recombination observed, the

439 *w*Pip and mitochondrial phylogenies are congruent: recombinations have not disrupted our

440 grasp of the evolutionary history of *w*Pip strains, probably because the contribution of

441 recombinant regions in the phylogeny is weak compared to the diversity existing in

442   nonrecombinant DNA fragments. Therefore, as suggested by Baldo et al. (Baldo et al. 2006),

443   the use of a multi-locus approach, rather than single locus analysis, is required for a correct

444   understanding of the evolutionary history of *Wolbachia* infections.

445

446   The *Cx. pipiens-Wolbachia* association, a unique case?

447   The high number of *w*Pip strains, which is still certainly underestimated, makes the *Cx.*

448   *pipiens* system remarkable because lower diversity is usually reported in *Wolbachia* of other

449   host species (e.g. Vavre et al. 1999; Mercot and Charlat 2004; Charlat et al. 2006; Arthofer et

450   al. 2009). However, it is possible that genetic variations of *Wolbachia* in other host species

451   could have been missed due to the methodology generally used to characterize these bacteria,

452   as it is generally assumed that a single *Wolbachia* strain is present within a host species when

453   the MLST or *wsp* markers are not variable. In *D. melanogaster*, a single strain, *w*Mel, was

454   presumed to be present until Riegler et al. (Riegler et al. 2005) identified five distinct

455   genotypes by examining transposon insertion sites and chromosomal inversions. More recent

456   studies have reported different *Wolbachia* haplotypes solely on the basis of WO phage genes

457   in various host species, including crickets, beetles and butterflies (Charlat et al. 2009; Chafee

458   et al. 2010). Hence, the classical MLST system is well suited to characterize *Wolbachia*

459   belonging to distinct clades, but specific species typing systems based on markers with rapid

460   sequence evolution need to be developed to investigate the *Wolbachia* diversity that probably

461   exists in most associations.

462

463   Finally, the question remains of whether the *Cx. pipiens - Wolbachia* association is unique in

464   term of its extremely large CI diversity (e.g. Laven 1967; O'Neill and Paterson 1992;

465   Guillemaud, Pasteur, and Rousset 1997; Duron et al. 2006). We have clearly demonstrated

466   that the diversity of crossing types in this species is independent of nuclear backgrounds and

467 relies solely on *w*Pip variability (Duron et al. 2006; Atyame et al. in press). The reason that a

468 similar CI system has not been reported in other *Wolbachia*-infected species remains a matter

469 for speculation, but it is possible that the crossing studies conducted in *Cx. pipiens* have been

470 more exhaustive than in any other species because of the intensive investigations that have

471 been carried out for clarifying its systematics and studying the inheritance of morphological

472 characters since the 1930s (e.g. Marshall and Staley 1937; Roubaud 1941; Laven 1958;

473 Rozeboom 1958; Laven 1967; Barr 1975; Narang and Seawright 1982; Irving-Bell 1983). As

474 a result, the high variability of CI crossing types was investigated much earlier than the

475 causative agent was identified by Yen and Barr in 1971. Comparatively little work on the

476 variability of the effects of *Wolbachia* infection has been conducted in most arthropods,

477 except in *Drosophila* species, such as *D. simulans*, where five distinct crossing types

478 associated with distinct *Wolbachia* infections have thus far been identified (for review see

479 Mercot and Charlat 2004). Therefore, the possibility of the existence of variable reproductive

480 phenotypes in other host species remains to be examined.

481

482 In conclusion, the use of multi-locus typing combining *Wolbachia* and mitochondrial markers

483 highlights the processes underlying the evolutionary dynamics of *w*Pip infections. The

484 diversification inside the *w*Pip clade shows that a considerable amount of *Wolbachia* diversity

485 can be generated within a single host species in a short period of time. Further investigations

486 should examine the roles of recombination and MGE in the adaptive capacities of *Wolbachia*.

487 In particular, this could explain rapid changes of interactions between *Wolbachia* and their

488 hosts (Weeks et al. 2007; Echaubard et al. 2010) and play a key role in the evolution of

489 phenotypes induced by *Wolbachia*. Finally, an important question now is to determine

490 whether the *Cx. pipiens - Wolbachia* association is a unique case or, rather, a representative

491 example.

492

493

502

503

504   **Literature Cited**

505   Ahrens, M. E., and D. Shoemaker. 2005. Evolutionary history of *Wolbachia* infections in the

506   fire ant *Solenopsis invicta*. BMC Evol Biol **5**:35.

507   Arthofer, W., M. Riegler, D. Schneider, M. Krammer, W. J. Miller, and C. Stauffer. 2009.

508   Hidden *Wolbachia* diversity in field populations of the European cherry fruit fly,

509   *Rhagoletis cerasi* (Diptera, Tephritidae). Mol Ecol **18**:3816-3830.

510   Atyame, C., O. Duron, P. Tortosa, N. Pasteur, P. Fort, and M. Weill. 2011. Multiple

511   *Wolbachia* determinants control the evolution of cytoplasmic incompatibilities in

512   *Culex pipiens* mosquito populations. Mol Ecol. **20**: 286-298.

513   Baldo, L., N. A. Ayoub, C. Y. Hayashi, J. A. Russell, J. K. Stahlhut, and J. H. Werren. 2008.

514   Insight into the routes of *Wolbachia* invasion: high levels of horizontal transfer in the

515   spider genus *Agelenopsis* revealed by *Wolbachia* strain and mitochondrial DNA

516   diversity. Mol Ecol **17**:557-569.

517 Baldo, L., J. C. Dunning Hotopp, K. A. Jolley, S. R. Bordenstein, S. A. Biber, R. R.

518   Choudhury, C. Hayashi, M. C. Maiden, H. Tettelin, and J. H. Werren. 2006.

519   Multilocus sequence typing system for the endosymbiont *Wolbachia pipientis*. Appl

520   Environ Microbiol **72**:7098-7110.

521 Baldo, L., N. Lo, and J. H. Werren. 2005. Mosaic nature of the *Wolbachia* surface protein. J

522   Bacteriol **187**:5406-5418.

523 Ballard, J. W. 2000. When one is not enough: introgression of mitochondrial DNA in

524   Drosophila. Mol Biol Evol **17**:1126-1130.

525 Barr, A. R. 1975. Culex. Pp. 347-375 *in* R. C. King, ed. Handbook of Genetics. Plenum

526   Publishing Corporation, New York.

527 Barr, A. R. 1982. The *Culex pipiens* complex. Pp. 551-572 *in* W. W. M. Steiner, W. J.

528   Tabachnik, K. S. Rai, and S. Narang, eds. Recent developments in the genetics of

529   insect disease vectors. Stipes Publishing Company, Champaign.

530 Beard, C. B., D. Mills Hamm, and F. H. Collins. 1993. The mitochondrial genome of the

531   mosquito *Anopheles gambiae*: DNA sequence, genome organization, and comparisons

532   with mitochondrial sequences of other insects. Insect Mol Biol **2**:103-124.

533 Bordenstein, S. R., and J. J. Wernegreen. 2004. Bacteriophage flux in endosymbionts

534   (*Wolbachia*): infection frequency, lateral transfer, and recombination rates. Mol Biol

535   Evol **21**:1981-1991.

536 Braig, H. R., W. Zhou, S. L. Dobson, and S. L. O'Neill. 1998. Cloning and characterization of

537   a gene encoding the major surface protein of the bacterial endosymbiont *Wolbachia*

538   *pipientis*. J Bacteriol **180**:2373-2378.

539   Bryant, D., and V. Moulton. 2004. Neighbor-net: an agglomerative method for the

540   construction of phylogenetic networks. Mol Biol Evol **21**:255-265.

541   Castresana, J. 2000. Selection of conserved blocks from multiple alignments for their use in

542   phylogenetic analysis. Mol Biol Evol **17**:540-552.

543   Chafee, M. E., D. J. Funk, R. G. Harrison, and S. R. Bordenstein. 2010. Lateral phage transfer

544   in obligate intracellular bacteria (*Wolbachia*): verification from natural populations.

545   Mol Biol Evol **27**:501-505.

546   Charlat, S., A. Duplouy, E. A. Hornett, E. A. Dyson, N. Davies, G. K. Roderick, N. Wedell,

547   and G. D. Hurst. 2009. The joint evolutionary histories of *Wolbachia* and

548   mitochondria in *Hypolimnas bolina*. BMC Evol Biol **9**:64.

549   Charlat, S., J. Engelstadter, E. A. Dyson, E. A. Hornett, A. Duplouy, P. Tortosa, N. Davies, G.

550   K. Roderick, N. Wedell, and G. D. Hurst. 2006. Competing selfish genetic elements in

551   the butterfly *Hypolimnas bolina*. Curr Biol **16**:2453-2458.

552   Cordaux, R. 2008. ISWpi1 from *Wolbachia pipientis* defines a novel group of insertion

553   sequences within the IS5 family. Gene **409**:20-27.

554   Cornel, A. J., R. D. McAbee, J. Rasgon, M. A. Stanich, T. W. Scott, and M. Coetzee. 2003.

555   Differences in extent of genetic introgression between sympatric *Culex pipiens* and

556   *Culex quinquefasciatus* (Diptera: Culicidae) in California and South Africa. J Med

557   Entomol **40**:36-51.

558   Dedeine, F., F. Vavre, D. D. Shoemaker, and M. Bouletreau. 2004. Intra-individual

559   coexistence of a *Wolbachia* strain required for host oogenesis with two strains

560   inducing cytoplasmic incompatibility in the wasp *Asobara tabida*. Evolution **58**:2167-

561   2174.

562   Duplouy, A., G. D. Hurst, S. L. O'Neill, and S. Charlat. 2010. Rapid spread of male-killing

563   *Wolbachia* in the butterfly *Hypolimnas bolina*. J Evol Biol **23**:231-235.

564   Duron, O., C. Bernard, S. Unal, A. Berthomieu, C. Berticat, and M. Weill. 2006. Tracking

565   factors modulating cytoplasmic incompatibilities in the mosquito *Culex pipiens*. Mol

566   Ecol **15**:3061-3071.

567   Duron, O., D. Bouchon, S. Boutin, L. Bellamy, L. Zhou, J. Engelstadter, and G. D. Hurst.

568   2008. The diversity of reproductive parasites among arthropods: *Wolbachia* do not

569   walk alone. BMC Biology **6**:27.

570   Duron, O., A. Boureux, P. Echaubard, A. Berthomieu, C. Berticat, P. Fort, and M. Weill.

571   2007. Variability and expression of ankyrin domain genes in *Wolbachia* variants

572   infecting the mosquito *Culex pipiens*. J Bacteriol **189**:4442-4448.

573   Duron, O., P. Fort, and M. Weill. 2006. Hypervariable prophage WO sequences describe an

574   unexpected high number of *Wolbachia* variants in the mosquito *Culex pipiens*. Proc

575   Biol Sci **273**:495-502.

576   Duron, O., J. Lagnel, M. Raymond, K. Bourtzis, P. Fort, and M. Weill. 2005. Transposable

577   element polymorphism of *Wolbachia* in the mosquito *Culex pipiens*: evidence of

578   genetic diversity, superinfection and recombination. Mol Ecol **14**:1561-1573.

579   Duron, O., M. Raymond, and M. Weill. In press. Many compatible *Wolbachia* strains coexist

580   within natural populations of *Culex pipiens* mosquito. Heredity.

581   Echaubard, P., O. Duron, P. Agnew, C. Sidobre, V. Noel, M. Weill, and Y. Michalakis. 2010.

582   Rapid evolution of *Wolbachia* density in insecticide resistant *Culex pipiens*. Heredity

583   **104**:15-19.

584   Engelstadter, J., and G. D. D. Hurst. 2009. The ecology and evolution of microbes that

585   manipulate host reproduction. Annl Rev Ecol Evol Syst **40**:127-149.

586   Fitch, W. M. 1997. Networks and viral evolution. J Mol Evol **44 Suppl 1**:S65-75.

587   Fonseca, D. M., N. Keyghobadi, C. A. Malcolm, C. Mehmet, F. Schaffner, M. Mogi, R. C.

588   Fleischer, and R. C. Wilkerson. 2004. Emerging vectors in the *Culex pipiens* complex.

589   Science **303**:1535-1538.

590   Fonseca, D. M., J. L. Smith, R. C. Wilkerson, and R. C. Fleischer. 2006. Pathways of

591   expansion and multiple introductions illustrated by large genetic differentiation among

592   worldwide populations of the southern house mosquito. Am J Trop Med Hyg **74**:284-

593   289.

594   Gavotte, L., H. Henri, R. Stouthamer, D. Charif, S. Charlat, M. Bouletreau, and F. Vavre.

595   2007. A Survey of the bacteriophage WO in the endosymbiotic bacteria *Wolbachia*.

596   Mol Biol Evol **24**:427-435.

597   Guillemaud, T., N. Pasteur, and F. Rousset. 1997. Contrasting levels of variability between

598   cytoplasmic genomes and incompatibility types in the mosquito *Culex pipiens*. Proc

599   Biol Sci **264**:245-251.

600   Haine, E. R. 2008. Symbiont-mediated protection. Proc R Soc B: Biol Sci **275**:353-361.

601   Hedrick, P. W. 1987. Gametic disequilibrium measures: proceed with caution. Genetics

602   **117**:331–341.

603   Hertig, M., and S. B. Wolbach. 1924. Studies on *Rickettsia*-like micro-organisms in insects. J

604   Med Res **44**:329-374 327.

605   Hilgenboecker, K., P. Hammerstein, P. Schlattmann, A. Telschow, and J. H. Werren. 2008.

606   How many species are infected with *Wolbachia*?--A statistical analysis of current data.

607   FEMS Microbiol Lett **281**:215-220.

608   Ho, S.Y.W., M.J. Phillips, A. Cooper and A.J. Drummond. 2005. Time dependency of

609   molecular rate estimates and systematic overestimation of recent divergence times.

610   Mol Biol Evol**22**: 1561-1568.

611   Hochberg, Y. 1988. A sharper Bonferroni procedure for multiple tests of significance.

612   Biometrika **75**:800-802.

613   Hurst, G. D. D., and F. M. Jiggins. 2005. Problems with mitochondrial DNA as a marker in

614   population, phylogeographic and phylogenetic studies: the effects of inherited

615   symbionts. Proc Biol Sci **272**:1525-1534.

616   Huson, D. H., and D. Bryant. 2006. Application of phylogenetic networks in evolutionary

617   studies. Mol Biol Evol **23**:254-267.

618   Irving-Bell, R. J. 1983. Cytoplasmic incompatibility within and between *Culex molestus* and

619   *Cx. quinquefasciatus* (Diptera: Culicidae). J. Med. Entomol. **20**:44-48.

620   Jiggins, F. M. 2003. Male-killing *Wolbachia* and mitochondrial DNA: selective sweeps,

621   hybrid introgression and parasite population dynamics. Genetics **164**:5-12.

622   Jiggins, F. M., J. H. von Der Schulenburg, G. D. Hurst, and M. E. Majerus. 2001.

623   Recombination confounds interpretations of *Wolbachia* evolution. Proc Biol Sci

624   **268**:1423-1427.

625   Keller, G. P., D. M. Windsor, J. M. Saucedo, and J. H. Werren. 2004. Reproductive effects

626   and geographical distributions of two *Wolbachia* strains infecting the Neotropical

627   beetle, *Chelymorpha alternans* Boh. (Chrysomelidae, Cassidinae). Mol Ecol **13**:2405-

628   2420.

629   Kittayapong, P., K. J. Baisley, V. Baimai, and S. L. O'Neill. 2000. Distribution and diversity

630   of *Wolbachia* infections in Southeast Asian mosquitoes (Diptera: Culicidae). J Med

631   Entomol **37**:340-345.

632   Klasson, L., T. Walker, M. Sebaihia, M. J. Sanders, M. A. Quail, A. Lord, S. Sanders, J. Earl,

633   S. L. O'Neill, N. Thomson, S. P. Sinkins, and J. Parkhill. 2008. Genome evolution of

634   *Wolbachia* strain *w*Pip from the *Culex pipiens* group. Mol Biol Evol **25**:1877-1887.

635　Klasson, L., J. Westberg, P. Sapountzis, K. Naslund, Y. Lutnaes, A.C. Darby, Z. Veneti, L.

636　　　　Chen, H.R. Braig, R. Garrett, K. Bourtzis and S.G. Andersson. 2009. The mosaic

637　　　　genome structure of the *Wolbachia w*Ri strain infecting *Drosophila simulans*. Proc

638　　　　Natl Acad Sci USA **106**: 5725-5730.

639　Krzywinski, J., O. G. Grushko, and N. J. Besansky. 1997. Analysis of the complete

640　　　　mitochondrial DNA from *Anopheles funestus*: An improved dipteran mitochondrial

641　　　　genome annotation and a temporal dimension of mosquito evolution. Mol Phyl Evol

642　　　　**39**:417-423.

643　Laven, H. 1958. Speciation by cytoplasmic isolation in the *Culex pipiens* complex. Cold

644　　　　Spring Harbor Symposia in Quantitative Biology **24**:166-173.

645　Laven, H. 1967. Speciation and evolution in *Culex pipiens in* J. Wright, and R. Pal, eds.

646　　　　Genetics of Insect Vectors of Disease. Elsevier, Amsterdam.

647　Lewontin, R. C. 1964. The interaction of selection and linkage. I. General considerations,

648　　　　heterotic models. Genetics **49**: 49–67.

649　Librado, P., and J. Rozas. 2009. DnaSP v5: A software for comprehensive analysis of DNA

650　　　　polymorphism data. Bioinformatics **25**:1451-1452.

651　Lo, N., C. Paraskevopoulos, K. Bourtzis, S. L. O'Neill, J. H. Werren, S. R. Bordenstein, and

652　　　　C. Bandi. 2007. Taxonomic status of the intracellular bacterium *Wolbachia pipientis*.

653　　　　Int J Syst Evol Microbiol **57**:654-657.

654　Marshall, J. F., and J. Staley. 1937. Some notes regarding the morphological and biological

655　　　　differentiation of *Culex pipiens* Linnaeus and *Culex molestus* Forskal (Diptera,

656　　　　Culicidae). Entomol Soc London **12**:17-27.

657　Mercot, H., and S. Charlat. 2004. *Wolbachia* infections in *Drosophila melanogaster* and *D.*

658　　　　*simulans*: polymorphism and levels of cytoplasmic incompatibility. Genetica **120**:51-

659　　　　59.

660   Mitchell, S. E., A. F. Cockburn, and J. A. Seawright. 1993. The mitochondrial genome of

661   *Anopheles quadrimaculatus* species A: complete nucleotide sequence and gene

662   organization. Genome **36**:1058-1073.

663   Moran, N. A., J. P. McCutcheon, and A. Nakabachi. 2008. Genomics and evolution of

664   heritable bacterial symbionts. Ann Rev Genet **42**:165-190.

665   Narang, S., and J. A. Seawright. 1982. Linkage relationships and genetic mapping in *Culex*

666   and *Anopheles*. Pp. 231-272 *in* W. W. M. Steiner, W. J. Tabachnick, K. S. Rai, and S.

667   Narang, eds. Recent developments in the genetics of insect disease vectors. Stipes

668   Publishing Company, Champaign.

669   Narita, S., M. Nomura, Y. Kato, and T. Fukatsu. 2006. Genetic structure of sibling butterfly

670   species affected by *Wolbachia* infection sweep: evolutionary and biogeographical

671   implications. Mol Ecol **15**:1095-1108.

672   O'Neill, S. L., R. Giordano, A. M. Colbert, T. L. Karr, and H. M. Robertson. 1992. 16S rRNA

673   phylogenetic analysis of the bacterial endosymbionts associated with cytoplasmic

674   incompatibility in insects. Proc Natl Acad Sci USA **89**:2699-2702.

675   O'Neill, S. L., and H. E. Paterson. 1992. Crossing type variability associated with cytoplasmic

676   incompatibility in Australian populations of the mosquito *Culex quinquefasciatus* Say.

677   Med Vet Entomol **6**:209-216.

678   Posada, D., and K. A. Crandall. 1998. MODELTEST: testing the model of DNA substitution.

679   Bioinformatics **14**:817-818.

680   Rasgon, J. L., A. J. Cornel, and T. W. Scott. 2006. Evolutionary history of a mosquito

681   endosymbiont revealed through mitochondrial hitchhiking. Proc Biol Sci **273**:1603-

682   1611.

683   Rasgon, J. L., and T. W. Scott. 2004. An initial survey for *Wolbachia* (Rickettsiales:

684   Rickettsiaceae) infections in selected California mosquitoes (Diptera: Culicidae). J

685   Med Entomol **41**:255-257.

686   Raychoudhury, R., L. Baldo, D. C. Oliveira, and J. H. Werren. 2009. Modes of acquisition of

687   *Wolbachia*: horizontal transfer, hybrid introgression, and codivergence in the *Nasonia*

688   species complex. Evolution **63**:165-183.

689   Raymond, M., A. Callaghan, P. Fort, and N. Pasteur. 1991. Worldwide migration of amplified

690   insecticide resistance genes in mosquitoes. Nature **350**:151-153.

691   Raymond, M., and F. Rousset. 1995. Genepop (version 1.2), a population genetics software

692   for exact tests and ecumenicism. J Hered **86**:248-249.

693   Reidenbach, K. R., S. Cook, M. A. Bertone, R. E. Harbach, B. M. Wiegmann, and N. J.

694   Besansky. 2009. Phylogenetic analysis and temporal diversification of mosquitoes

695   (Diptera: Culicidae) based on nuclear genes and morphology. BMC Evol Biol **9**:298.

696   Ricci, I., G. Cancrini, S. Gabrielli, S. D'Amelio, and G. Favi. 2002. Searching for *Wolbachia*

697   (Rickettsiales: Rickettsiaceae) in mosquitoes (Diptera: Culicidae): large polymerase

698   chain reaction survey and new identifications. J Med Entomol **39**:562-567.

699   Rich, S. M., M. C. Licht, R. R. Hudson, and F. J. Ayala. 1998. Malaria's Eve: evidence of a

700   recent population bottleneck throughout the world populations of *Plasmodium*

701   *falciparum*. Proc Natl Acad Sci U S A **95**:4425-4430.

702   Riegler, M., M. Sidhu, W. J. Miller, and S. L. O'Neill. 2005. Evidence for a global *Wolbachia*

703   replacement in *Drosophila melanogaster*. Curr Biol **15**:1428-1433.

704   Rogers, S. O., and A. J. Bendich. 1988. Extraction of DNA from plant tissues. Pp. 1-10 *in* S.

705   B. Gelvin, and R. A. Schilperoort, eds. Plant Molecular Biology Manuel. Kluwer

706   Academic Publishers, Boston.

707   Ronquist, F., and J. P. Huelsenbeck. 2003. MrBayes 3: Bayesian phylogenetic inference under

708   mixed models. Bioinformatics **19**:1572-1574.

709   Ros, V. I., V. M. Fleming, E. J. Feil, and J. A. Breeuwer. 2009. How diverse is the genus

710   *Wolbachia*? Multiple-gene sequencing reveals a putatively new *Wolbachia* supergroup

711   recovered from spider mites (Acari: Tetranychidae). Appl Environ Microbiol **75**:1036-

712   1043.

713   Roubaud, E. 1941. Phénomènes d'amixie dans les intercroisements de Culicides du groupe

714   *pipiens*. C R Acad Sci **212**:257-259.

715   Rousset, F., and M. Solignac. 1995. Evolution of single and double *Wolbachia* symbioses

716   during speciation in the *Drosophila simulans* complex. Proc Natl Acad Sci U S A

717   **92**:6389-6393.

718   Rozeboom, L. E. 1958. Hybridization of *Culex pipiens fatigan*s Wiedemann from the

719   Philippine Islands with American strains of *Culex pipiens* group (Diptera: culicidae).

720   Amer J Trop Med Hyg **7**:526-530.

721   Salzberg, S. L., D. Puiu, D. D. Sommer, V. Nene, and N. H. Lee. 2009. The genome sequence

722   of *Wolbachia* endosymbiont of *Culex quinquefasciatus* JHB. J Bacteriol **191**:1725.

723   Sawyer, S. A. 1989. Statistical tests for detecting gene conversion. Mol Biol Evol **6**:526–538.

724   Sawyer, S. A. 1999. GENECONV: A computer package for the statistical detection of gene

725   conversion. Pp. Distributed by the author, Department of Mathematics, Washington

726   University in St. Louis, available at http://www.math.wustl.edu/~sawy.

727   Sinkins, S. P., T. Walker, A. R. Lynd, A. R. Steven, B. L. Makepeace, H. C. Godfray, and J.

728   Parkhill. 2005. *Wolbachia* variability and host effects on crossing type in *Culex*

729   mosquitoes. Nature **436**:257-260.

730   Tamura, K., J. Dudley, M. Nei, and S. Kumar. 2007. MEGA4: Molecular evolutionary

731   genetics analysis (MEGA) software version 4.0. Mol Biol Evol **24**:1596-1599.

732   Thompson, J. D., T. J. Gibson, and D. G. Higgins. 2002. Multiple sequence alignment using

733   ClustalW and ClustalX. Curr Protoc Bioinformatics **Chapter 2**:Unit 2 3.

734   Vavre, F., F. Fleury, D. Lepetit, P. Fouillet, and M. Bouletreau. 1999. Phylogenetic evidence

735   for horizontal transmission of *Wolbachia* in host-parasitoid associations. Mol Biol

736   Evol **16**:1711-1723.

737   Venkatesan, M., C.J. Westbrook, M.C. Hauer, and J.L. Rasgon. 2007. Evidence for a

738   population expansion in the West Nile Virus vector *Culex tarsalis*. Mol Biol Evol **24**:

739   1208-1218.

740   Walker, T., L. Klasson, M. Sebaihia, M. J. Sanders, N. R. Thomson, J. Parkhill, and S. P.

741   Sinkins. 2007. Ankyrin repeat domain-encoding genes in the *w*Pip strain of *Wolbachia*

742   from the *Culex pipiens* group. BMC Biol **5**:39.

743   Weeks, A. R., M. Turelli, W. R. Harcombe, K. T. Reynolds, and A. A. Hoffmann. 2007. From

744   parasite to mutualist: rapid evolution of *Wolbachia* in natural populations of

745   *Drosophila*. PLoS Biol **5**:e114.

746   Werren, J. H., L. Baldo, and M. E. Clark. 2008. *Wolbachia*: master manipulators of

747   invertebrate biology. Nature Rev Microbiol **6**:741-751.

748   Wu, M., L. V. Sun, J. Vamathevan, M. Riegler, R. Deboy, J. C. Brownlie, E. A. McGraw, W.

749   Martin, C. Esser, N. Ahmadinejad, C. Wiegand, R. Madupu, M. J. Beanan, L. M.

750   Brinkac, S. C. Daugherty, A. S. Durkin, J. F. Kolonay, W. C. Nelson, Y. Mohamoud,

751   P. Lee, K. Berry, M. B. Young, T. Utterback, J. Weidman, W. C. Nierman, I. T.

752   Paulsen, K. E. Nelson, H. Tettelin, S. L. O'Neill, and J. A. Eisen. 2004.

753   Phylogenomics of the reproductive parasite *Wolbachia pipientis w*Mel: a streamlined

754   genome overrun by mobile genetic elements. PLoS Biol **2**:E69.

755   Yen, J. H., and A. R. Barr. 1971. New hypothesis of the cause of cytoplasmic incompatibility

756   in *Culex pipiens* L. Nature **232**:657-658.

757

Mosquito line	Abbreviation	*Wolbachia* strain	*Culex pipiens* subspecies	Origin	Year of collection	Reference
Pel	Pel	*w*Pip(Pel)	*quinquefasciatus*	Sri Lanka	1984	(Klasson et al. 2008)
JHB	JHB	*w*Pip(JHB)	*quinquefasciatus*	South Africa	2001	(Salzberg et al. 2009)
Manille-A	Ma-A	*w*Pip(Ma-A)	*quinquefasciatus*	Philippines	2003	(Duron et al. 2006)
Manille-B	Ma-B	*w*Pip(Ma-B)	*quinquefasciatus*	Philippines	2003	(Duron et al. 2006)
Kara-C	Ka-C	*w*Pip(Ka-C)	*quinquefasciatus*	China	2003	(Duron, Fort, and Weill 2006)
MaClo	Mc	*w*Pip(Mc)	*quinquefasciatus*	California	1984	(Duron et al. 2005)
Slab	Sl	*w*Pip(Sl)	*quinquefasciatus*	California	1950	(Duron et al. 2005)
Cotonou-A	Cot-A	*w*Pip(Cot-A)	*quinquefasciatus*	Benin	2005	This study
Cotonou-B	Cot-B	*w*Pip(Cot-B)	*quinquefasciatus*	Benin	2005	This study
Australie	Au	*w*Pip(Au)	hybrid (*quinquefasciatus/pipiens*)	Australia	2004	(Duron, Fort, and Weill 2006)
El Palmar-A	Ep-A	*w*Pip(Ep-A)	*pipiens*	Spain	2005	(Duron et al. 2007)
El Palmar-B	Ep-B	*w*Pip(Ep-B)	*pipiens*	Spain	2005	(Duron et al. 2007)
LaVar	Lv	*w*Pip(Lv)	*pipiens*	France	2003	(Duron et al. 2005)
Bifa-A	Bf-A	*w*Pip(Bf-A)	*pipiens*	France	2002	(Duron et al. 2006)
Bifa-B	Bf-B	*w*Pip(Bf-B)	*pipiens*	France	2002	(Duron et al. 2006)
Kol	Ko	*w*Pip(Ko)	*pipiens*	Crete	2002	(Duron, Fort, and Weill 2006)
Keo-A	Ke-A	*w*Pip(Ke-A)	*pipiens*	Cyprus	2003	(Duron et al. 2006)
Keo-B	Ke-B	*w*Pip(Ke-B)	*pipiens*	Cyprus	2003	(Duron et al. 2006)
Tunis	Tn	*w*Pip(Tn)	*pipiens*	Tunisia	1992	(Duron et al. 2005)
Istanbul	Is	*w*Pip(Is)	*pipiens*	Turkey	2003	(Duron et al. 2005)

758

759   **Table 1.** Description of *Culex pipien*s lines and *w*Pip strains.

Strain	Gene							Haplotype
	MutL	*ank2*	*pk1*	*pk2*	*GP12*	*GP15*	*RepA*	
*w*Pip(Pel)	a	a	a	a	a	a	a	A
*w*Pip(Cot-A)	a	a	a	a	a	a	a	
*w*Pip(Cot-B)	a	a	a	a	a	a	a	
*w*Pip(Ko)	a	a	a	a	a	a	a	
*w*Pip(Tn)	a	a	a	a	a	a	a	
*w*Pip(Ma-B)	a	a	a	a	a	a	a	
*w*Pip(JHB)	a	a	a	a	a	-	a	B
*w*Pip(Ep-A)	a	a	a	d	a	a	a	C
*w*Pip(Ep-B)	a	a	a	d	a	a	a	
*w*Pip(Bf-A)	a	a	a	a	a	a	b	D
*w*Pip(Lv)	b	e	c	a	b	b	a	E
*w*Pip(Au)	d	e	c	a	b	f	a	F
*w*Pip(Ke-A)	c	e	c	a	d	e	a	G
*w*Pip(Ke-B)	c	e	c	a	e	e	a	H
*w*Pip(Sl)	e	b	b	b	b	c	b	I
*w*Pip(Bf-B)	e	b	b	c	b	c	b	J
*w*Pip(Mc)	b	b	b	b	b	c	b	K
*w*Pip(Is)	c	c	d	a	c	d	a	L
*w*Pip(Ka-C)	f	d	e	a	f	g	a	M
*w*Pip(Ma-A)	f	d	e	a	g	g	a	N

Table 2. Allelic profiles of the seven polymorphic *w*Pip genes in 20 *w*Pip strains. Letters A to N represent the 14 *w*Pip haplotypes. Dash indicates a gene deletion.

Taxon (subfamily, species)	π (no of sequences)				*Wolbachia* infection (reference)
	cytb	*ND4*	*COI*	*COII*	
Culicinae					
Culex pipiens	0.0021 (19)	0.0004 (14)	0.0000 (24)	0.0007 (8)	yes (Hertig and Wolbach 1924)
Cx. tarsalis	_	0.0116 (64)	_	_	no (Rasgon and Scott 2004)
Culex sp.	_	0.0090 (10)	_	_	no (Rasgon, Cornel, and Scott 2006)
Aedes aegypti	0.0094 (16)	0.0202 (46)	_	_	no (Kittayapong et al. 2000)
Ae. albopictus	0.0043 (14)	_	0.0039 (23)	_	yes (O'Neill et al. 1992)
Ae. caspius	_	_	0.0094 (7)	0.0063 (21)	no (Ricci et al. 2002)
Ae. vexans	_	_	0.0185 (7)	0.0084 (7)	no (Kittayapong et al. 2000; Ricci et al. 2002)
Anophelinae					
Anopheles aconitus	_	_	0.0053 (13)	0.0066 (35)	no (Kittayapong et al. 2000)
An. funestus	0.0066 (11)	_	_	_	no (Ricci et al. 2002)
An. gambiae	_	_	0.0053 (48)	_	no (Ricci et al. 2002)
An. maculipennis	_	_	0.0063 (62)	_	no (Ricci et al. 2002)

765

766 **Table 3.** Nucleotide diversity (π) of mitochondrial genes in eleven mosquito species (Culicidae).

Figure legends

Figure1. Map of the *w*Pip(Pel) genome showing the position of the thirteen genes examined. Number in brackets (1 to 4) indicates identical copies of a given gene located in different positions along the chromosome. Black boxes indicate the locations of prophage regions; the MLST and *wsp* genes are reported in grey.

Figure 2. Phylogenic tree of *w*Pip strains obtained from concatenated data set (*MutL*, *ank2*, *pk1*, *pk2*, *GP12*, *GP15* and *RepA* sequences) by Bayesian analysis. Posterior probabilities obtained are shown at major nodes. The scale bar is in units of substitutions/site.

Figure 3. Network analysis obtained from concatenated data set (*MutL*, *ank2*, *pk1*, *pk2*, *GP12*, *GP15* and *RepA* sequences) using the Neighbor-net method. Each edge (or a set of parallel edges) corresponds to a split in the data set and has length equal to the weight of the split. Incompatible splits, produced by recombination are represented by boxes in the network. Only bootstrap values for major grouping are indicated. The five *w*Pip groups (highlighted) are connected by multiple pathways resulting from recombination between *Wolbachia* genomes.
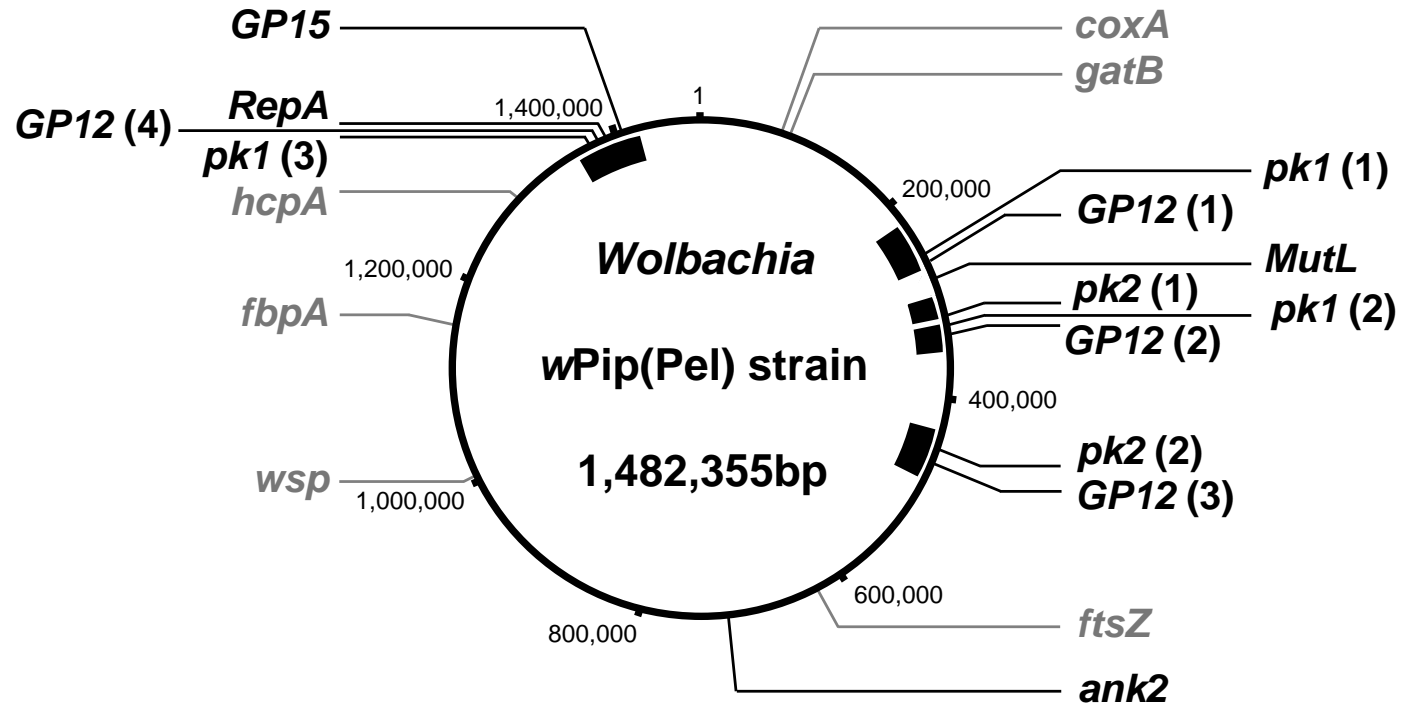
Figure 4. Distribution of *w*Pip haplotypes and *w*Pip groups in *Culex pipiens* populations. Letters and symbols represent the *w*Pip haplotypes and *w*Pip groups, respectively. * *w*Pip haplotypes recently described by Atyame et al. (Atyame et al. in press).

789 **Figure 5.** Comparisons between phylogeny of *Culex pipiens* mitochondria and phylogeny of

790 the *w*Pip strains. **A,** Mitochondrial phylogeny constructed using Bayesian inferences based on

791 *ND2*, *ND5* and *cytb* concatenated sequences. Names on branches indicate the mtDNA

792 haplotypes (*pi*I to *pi*XIV). **B,** *w*Pip phylogeny obtained from concatenated data set (*MutL*,

793 *ank2*, *pk1*, *pk2*, *GP12*, *GP15* and *RepA* sequences). The five *w*Pip groups are reported. The

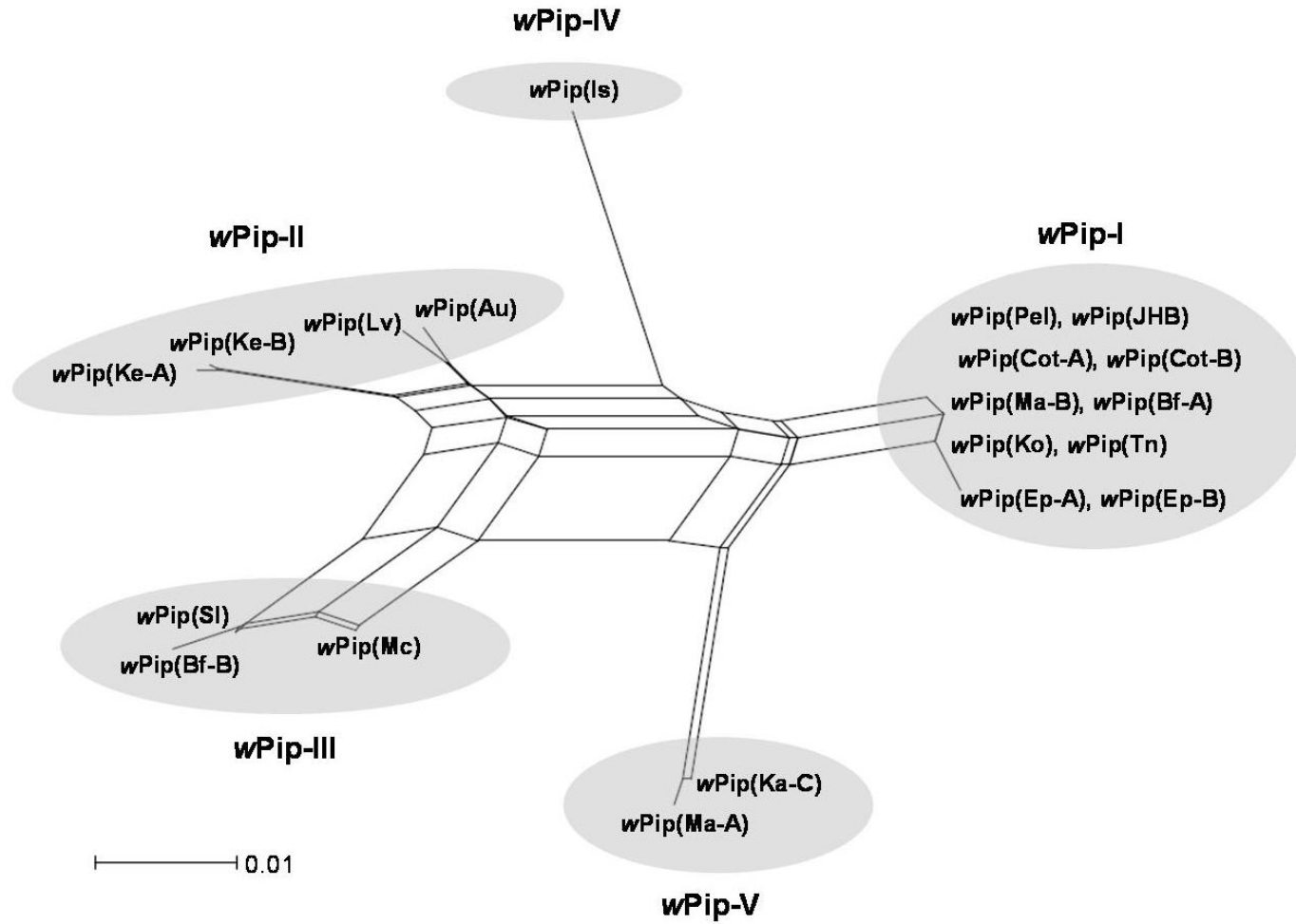794 scale bar is in units of substitutions/site.

Figure 1

799

800

801

802

Figure 2
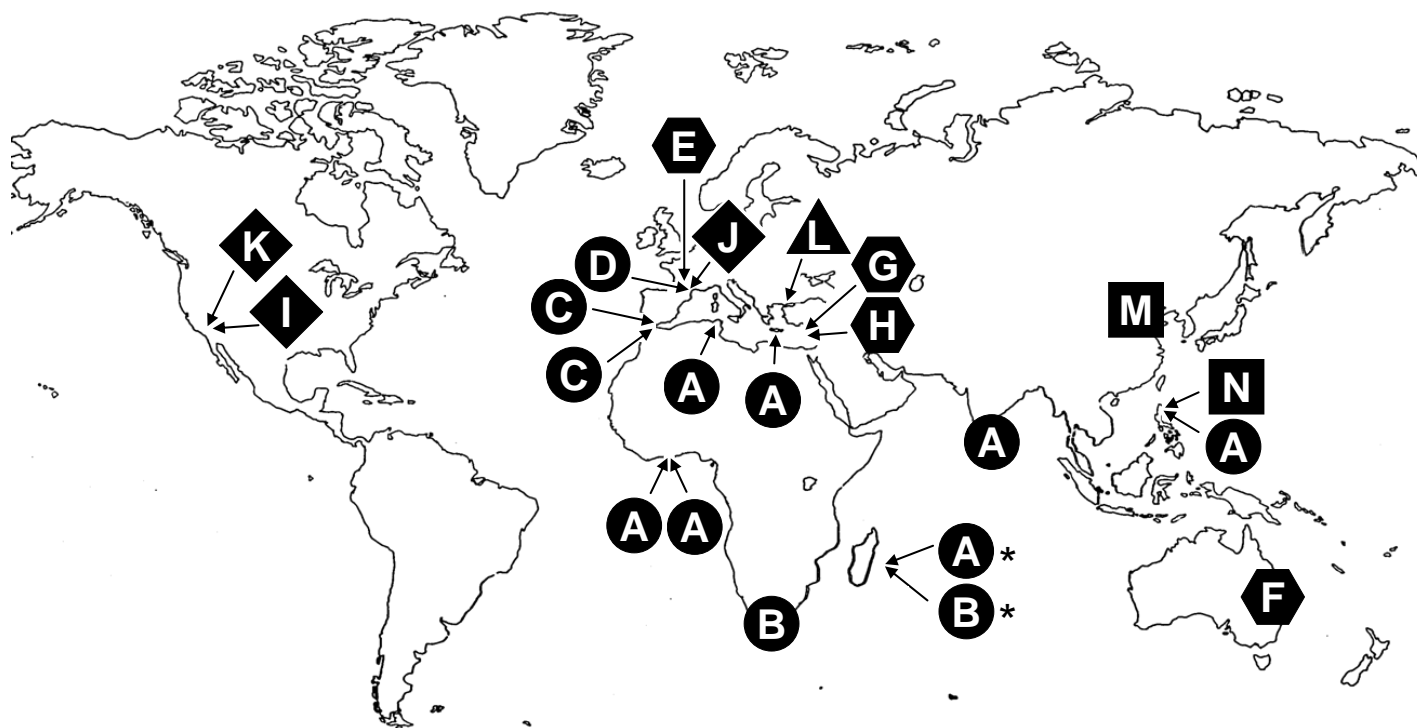
Figure 3

805

Figure 4

806

807

808

40

Figure 5

809

810

41