Within-host environmental heterogeneity is associated with phenotypic but not genomic diversity in *Wolbachia* endosymbionts

Romain Pigeault[1*], Yann Dussert[1], Raphaël Jorge[2,3], Marie Panza[1], Maryline Raimond[1], Carine Delaunay[1], David Ogereau[4], Bouziane Moumen[1], Jean Peccoud[1], Richard Cordaux[4]

[1] Université de Poitiers, CNRS, EBI, UMR 7267, Poitiers, France

[2] Université Lyon 1, CNRS, VetAgroSup, Laboratoire de Biométrie et Biologie Evolutive, UMR 5558, Villeurbanne, France

[3] INSA Lyon, INRAE, Biologie fonctionnelle, insectes et interactions, UMR203, Villeurbanne, France

[4] Université Paris-Saclay, CNRS, IRD, UMR Évolution Génomes Comportement Écologie, 91190 Gif-sur-Yvette, France

ORCID :

Romain Pigeault: 0000-0002-8011-4600

Yann Dussert: 0000-0003-1900-7715

David Ogereau : 0000-0002-4496-3007

Bouziane Moumen: 0000-0002-1789-6160

Jean Peccoud: 0000-0002-3356-7869

Richard Cordaux: 0000-0003-4997-735X

*Corresponding author: **Romain Pigeault - romain.pigeault@univ-poitiers.fr**

Abstract

The mechanisms underlying within-host diversification in parasite populations are still poorly understood. Yet, diversification in parasite phenotype and genotype may have important clinical and epidemiological consequences. Our aim was to determine if the constraints associated with the infection of different host tissues lead to the emergence and co-existence of multiple parasite sub-populations with different phenotypes. We tested this hypothesis by injecting the most widespread bacterial endosymbiont, *Wolbachia*, isolated from three tissues of the common pill-bug *Armadillidium vulgare* into uninfected individuals, and by monitoring the growth rate and virulence of each bacterial sub-population in the new hosts. Our results highlighted that within-host environmental heterogeneity led to diverse phenotypes in *Wolbachia*. This variability did not appear to result from genomic variation, but from phenotypic plasticity. Indeed, we did not find recurrent tissue-specific genomic variation among infected individuals. Whole genome resequencing of the *Wolbachia* sub-populations and pedigree data from the animals used in this study demonstrated however the coexistence of two *Wolbachia* variants within our host lineage for almost 20 years. Our results further supported the stability of the *Wolbachia* genome with respect to small-scale genomic variation, in agreement with reports on other symbiotic systems.

Keywords: adaptation, acclimation, feminizing *Wolbachia*, infection dynamics, subpopulation, tissue tropism, variant calling, within-host evolution.

46 # 1. Introduction

47 Among the many factors that influence species richness and community structure,
48 environmental heterogeneity is generally considered to be a particularly important driver
49 [1,2]. Since heterogeneous environments offer a variety of ecological niches and refuges from
50 adverse conditions, species richness and persistence should increase with greater spatial
51 heterogeneity [1,3–5]. Environmental heterogeneity also promotes genetic and phenotypic
52 diversity within species through diversifying selection [6–9], which may underpin local
53 adaptations [10–13]. However, most studies that investigated relationships between
54 environmental heterogeneity and diversity have focused on free-living species (e.g.,[2,14]).
55 The relationship between environmental heterogeneity and both endoparasites species
56 diversity and phenotypic variability remains underexplored.

57 For endoparasites, environmental heterogeneity is found both within and between
58 hosts. Host population heterogeneity is known to shape parasite evolutionary trajectories
59 [15–17], leading to rapid parasite diversification [15,18]. However, a single host is treated by
60 most theoretical and empirical studies on endoparasite evolution as a homogeneous
61 environment – a single population of target cells without any structure – neglecting
62 heterogeneities in cell type, immune response, nutrient supply or microbiota composition.
63 Yet, several parasites exploit a heterogeneous environment composed of different biological
64 tissues and fluids which form a spatially structured fitness landscape for the parasite.
65 Successful colonization of different microenvironments, with their contrasted constraints,
66 may involve the expression or selection of different parasitic phenotypes leading to within-
67 host differentiation into several sub-populations [19]. This phenotypic variability can reflect

68 phenotypic plasticity (e.g., expression plasticity, [20]), selection of genetic variants initially
69 present in the inoculum [21] or of new variants arising from spontaneous mutations [22].
70 The co-occurrence of polymorphic sub-populations of parasites in different tissues or
71 cell types has been described in several biological models. RNA viruses (e.g., plum pox virus,
72 influenza, SARS-CoV-2, poliovirus) may display genetic diversity within single infected hosts,
73 and co-existing viral variants may evolve differently in distinct cell types [23–25]. The ability
74 of some RNA viruses to adapt to tissue-specific innate immune microenvironments may
75 actually be critical for establishing robust infections [25]. Gastric biopsies from multiple
76 stomach regions of *Helicobacter pylori*-infected hosts showed location-specific evolution of
77 the bacteria, suggesting the existence of structured niches with distinct selective pressures
78 within the stomach [22]. In other cases, such as *Pseudomonas aeruginosa* or *Mycobacterium*
79 *tuberculosis* chronic infections, the emergence of polymorphic sub-populations seems to be
80 driven by tissue-specific adaptation [26,27]. Some within-host parasite sub-population may
81 constitute a reservoir from which new virulent variants can emerge [28], and which can delay
82 clearance of parasite infection by the host immune system or drugs [29–31]. Understanding
83 the evolutionary forces and mechanisms at the origin of within-host parasite diversity is
84 essential, yet it is rarely studied. Descriptive studies (e.g., long-term ambulatory monitoring
85 of chronically infected patients) cannot trace the divergence of variants found within a host
86 and cannot identify the drivers of their diversification without controlling or assessing the
87 composition of the infectious inoculum and ensuring the absence of subsequent infections.
88 For this purpose, vertically transmitted parasites are particularly useful, as their transmission
89 mode ensures the common origin of all within-host parasite sub-populations.
90 Among the most widespread vertically transmitted endosymbionts in animals is
91 *Wolbachia pipientis*. This primarily maternally transmitted intracellular alpha-

92 proteobacterium is widely distributed among arthropods and has been extensively studied
93 with respect to its peculiar effects on its hosts [32,33]. *Wolbachia* not only manipulates host
94 reproduction, it may also negatively or positively affect host life history traits (e.g., immunity,
95 senescence) and the evolution of host genomes [34,35]. The diversity of effects of *Wolbachia*
96 on its hosts certainly relates to its wide distribution within infected individuals [35]. In addition
97 to colonizing the female germline, ensuring vertical transmission, *Wolbachia* has indeed been
98 observed in all major tissues (e.g., nerve chord, immune and digestive compartments) of many
99 arthropod species. The presence of vertically transmitted endosymbionts in somatic tissues
100 may seem paradoxical, especially considering that systemic infection can decrease host
101 fitness. Given the disadvantages of reducing the fitness of their host (i.e., fitness alignment),
102 these extensive somatic localizations must somehow benefit the parasite. Localization of
103 *Wolbachia* in somatic tissues could influence host biology in such a way that favours its vertical
104 and, possibly, horizontal transmission. The ability of the bacteria to colonize and establish in
105 different tissular microenvironments could involve the expression or selection of different
106 phenotypes.

107 To investigate this hypothesis, we assessed the relationship between environmental
108 heterogeneity and phenotype variability in the *w*VulC *Wolbachia* strain, which infects the
109 common pill-bug *Armadillidium vulgare* [36]. In this study, we focused on three naturally
110 infected tissues, which we selected for their physiological characteristics and the benefits that
111 *Wolbachia* could gain by colonizing them: (i) gonads, which are a prime target for vertical
112 transmission, (ii) haemocytes, as these circulating immune cells could contribute to spread
113 *Wolbachia* to different host tissues and may facilitate its horizontal transmission [37], and (iii)
114 the nerve chain, as the localization of *Wolbachia* in the brain and nerve chord could induce
115 adaptive modifications of host behaviour [38,39]. We hypothesized that the main phenotypic

116 variation between bacteria sub-populations colonizing these tissues pertained to their
117 replication rate. A high replication rate in vital tissues, such as the nerve chain, could disrupt
118 tissue structure and reduce host lifespan [40,41]. Thus, we expected *Wolbachia* populations
119 from the host's nerve system to show lower replication rates than those from the other
120 tissues. We tested this hypothesis by injecting *Wolbachia* extracted from the three tissue
121 types into uninfected *A. vulgare* individuals, and by monitoring their growth rates in their new
122 hosts. We then tested whether *Wolbachia* sub-populations were adapted and not merely
123 acclimated to their tissue types. In this case, we expected to find recurrent tissue-specific
124 genomic variation among infected individuals. We verified this prediction by sequencing the
125 whole genomes of focal bacterial sub-populations.

126 2. Materials and Methods

127 2.1 Biological models

128 All *A. vulgare* individuals used in this experiment were reared as previously described [42].
129 Given that the *Wolbachia* strain used in this study is feminizing, all the individuals used were
130 females. As *Wolbachia* is not cultivable, phenotyping sub-populations requires trans-infection
131 of an uninfected host with bacterial isolates from naturally infected animals. For this purpose,
132 we used a common pill-bug source line (WXw) naturally infected by the *w*VulC *Wolbachia*
133 strain (source individuals) and a naturally uninfected recipient line (WXa, recipient individuals)
134 into which we injected bacteria. The two laboratory lines came from the same original
135 population sampled in Helsingør (Denmark) in 1991.

136 2.2 Experimental infection

137 To characterize the phenotype of the different *Wolbachia* sub-populations associated with
138 three different tissues of their native host, five batches of trans-infection were carried out.

139 Three batches were used to monitor early infection dynamics within the recipient host (day 0
140 to day 60 post-infection), and two batches were used to estimate bacterial persistence in
141 tissues (day 200 post-infection). For each batch, one tissue suspension was prepared from
142 each of the following tissues: haemolymph, ovaries and nerve chain. Tissues were collected
143 from five source females from the WXw line (infection status previously validated by
144 quantitative PCR, see section 2.4). For each suspension, tissues were crushed in 1 mL of Ringer
145 solution (1.4 mM CaCl2, 2.4 mM HNaCO3, 2 mM KCl, 0.4 M NaCl) for haemolymph and 2 mL
146 for ovaries and nerve chain. The resulting suspension was filtered through a 1.2 µm pore
147 membrane to remove cell debris and host cell nuclei. As the number of bacteria differed
148 between the three focal tissues (see **Figure S1** in appendix 1), the solutions were further
149 diluted in Ringer solution to adjust the final concentration of bacteria at 2500 ± 300 cells/µL
150 for the trans-infection batches used to describe *Wolbachia* infection dynamics, and at 1282 ±
151 205 cells/µL for the two batches made to investigate bacterial persistence 200 days after
152 injection. One µL of each filtrate was injected using a thin glass needle into the general cavity
153 of each recipient individual through a small hole pierced at its posterior part. Twenty-one
154 recipient females of the WXa line were injected with each tissue filtrate from each of the three
155 trans-infection batches used to describe *Wolbachia* infection dynamics (i.e., 63 females per
156 tissue filtrate). Then, on days 20, 40 and 60 post-injection, 6-7 recipient females were
157 randomly selected to quantify *Wolbachia* in the focal tissues (see section 2.4). To study
158 bacterial persistence in recipient host tissues, 10 females were injected with each tissue
159 filtrate for each of the two trans-infection batches. Two hundred days post-injection, all
160 surviving females were dissected to quantify *Wolbachia* in the three focal tissues (see section
161 2.4). For each trans-infection batch, control groups were created using the same protocol, but
162 with female sources originating from the WXa line.

163 ## 2.3 Mortality monitoring

164 The effect of *Wolbachia* injection on pill-bug mortality was monitored every 10 days as part
165 of the study of the influence of *Wolbachia* tissue origin on the dynamics of early infection
166 within the recipient host, and every 15-25 days as part of the study of *Wolbachia* tissue origin
167 on the persistence of the bacterium in tissues. When individuals were harvested during the
168 experiment to quantify *Wolbachia*, they were censored for the survival analysis (see section
169 2.9).

170 ## 2.4 Quantification of *Wolbachia* in recipient host's tissues

171 Total DNA was extracted from the haemolymph, nerve chain and ovaries of recipient females
172 using standard protocols (Qiagen DNeasy 96 Blood & Tissue kit). For each sample, the purity
173 of the extracted DNA (OD 260/280 nm and 260/230 nm ratios) was measured using a
174 Nanodrop 1000 spectrophotometer (Thermofisher). To quantify *Wolbachia* load, we
175 developed a fluorescent probe-based quantitative PCR (qPCR) approach amplifying a
176 *Wolbachia* locus and a reference host nuclear locus in the same reaction. We used primers
177 wsp208f (5'- TGG-TGC-AGC-ATT-TAC-TCC-AG-3') and wsp413r (5'-TCG-CTT-GAT-AAG-CAA-
178 AAC-CA-3') targeting the *Wolbachia* protein surface gene (*wsp*, [40]) and we designed primers
179 amplifying a portion of the single-copy nuclear gene encoding the mitochondrial leucine-tRNA
180 ligase of *A. vulgare*: TLeuF (5'- TGT-ACA-CAT-CGA-GCA-GCA-AG-3') and TLeuR (5'- AAA-GAG-
181 GAG-CGG-AGA-GTT-TCA-G-3') [43]. The *Wolbachia wsp* double-dye probe (5'-TTG-CAG-ACA-
182 GTG-TGA-CAG-CGT-T-3') was labelled with Hexachlorofluorescein (HEX) as a reporter at the 5'
183 end and Black Hole Quencher 1 (BHQ-1) at the 3' end. *A. vulgare* TLeu probe (5'-ACG-AAG-
184 TTC-GCC-CTG-TTC-TGG-A-3') was labelled with 6-carboxy-fluorescein (FAM) as a reporter and
185 BHQ-1 as a quencher. The qPCR reactions were performed using Roche LIGHTCYCLER 480 with
186 the following program: 2 min at 50°C, 10 min at 95°C and 45 cycles, 15 s at 95°C, and 1 min at

187 60°C. For each sample, two independent technical replicates were carried out and the mean
188 cycle-threshold value was calculated. *Wolbachia* load was then calculated relative to the
189 reference gene (TLeu) using the delate CP method: $2^{-(CP(wsp) - CP(TLeu))}$.

190 ## 2.5 Genome assembly of *Wolbachia w*VulC

191 To study the within-host genetic diversity of *Wolbachia*, we first assembled the genome of the
192 *w*VulC *Wolbachia* strain from the WXw line. Three female individuals were used for Oxford
193 Nanopore Technologies (ONT) long-read sequencing and one for Illumina short-read
194 sequencing. Total DNA from haemolymph, nerve chain and ovaries of each female was
195 extracted using standard protocols (DNA used for ONT sequencing: Macherey-Nagel
196 NucleoBond HMW DNA kit, DNA used for Illumina sequencing: Qiagen DNeasy 96 Blood &
197 Tissue kit). The long-read sequencing library was prepared using the SQK-LSK109 Ligation
198 Sequencing Kit (ONT) and sequenced with an R9.4.1 flow cell on a MinION Mk1B (ONT). Reads
199 were basecalled with Guppy 6.3.8 (RRID:SCR_023196) using the SUP model. The short-read
200 library was prepared using the TruSeq Nano DNA Library Prep kit (Illumina) and sequenced on
201 a HiSeq X (Illumina, paired-end 2 × 150 pb).

202 ONT adapters were first trimmed with porechop 0.2.4
203 (https://github.com/rrwick/Porechop), and reads were filtered using NanoFilt 2.7.1 (minimum
204 average quality: 7, minimum length: 1000 bp) implemented in NanoPack [44]. Reads were
205 then assembled with Flye 2.8.3 [45] using default parameters. The entire assembly was first
206 polished with the ONT reads using medaka 1.7 (https://github.com/nanoporetech/medaka)
207 (r941_min_fast_g507 model, 2 rounds). For assembly polishing using short reads, Illumina
208 reads were mapped onto the assembly with bwa-mem2 2.2.1 [46] and used for further
209 polishing with Polypolish 0.4.3 [47]. This step was carried out twice. The largest assembled
210 contig was reported detected as circular by the Flye assembler and had the expected length

211 (~1.6 Mb) of the *w*VulC *Wolbachia* genome. We aligned this contig against a previous version
212 of the genome (GenBank accession: ALWU00000000.1) with nucmer in MUMmer 4.0.0rc1
213 [48], confirming that the contig corresponded to the *w*VulC strain. After mapping the short
214 reads for a third time onto this contig, short variants were called with Freebayes 1.3.1 [49].
215 Variants were then visually reviewed using the IGV genome browser (version 2.14.0) [50].
216 Eight of them, corresponding to errors, were kept. We corrected the genome sequence with
217 these eight variants using the consensus command in BCFtools 1.17 [51]. Following the
218 recommendations of Ioannidis [52], the genome sequence was rotated to start at the *hemE*
219 gene using the fixstart command of Circlator 1.5.5 [53] . The genome was annotated with the
220 NCBI Prokaryotic Genome Annotation Pipeline (PGAP) 2022-12-13.build6494 [54–56]. A
221 search of Insertion Sequence (IS) elements was carried out with digIS v1.2 [57]. Additionally,
222 bacteriophage-derived regions were detected using PHASTEST [58]. Results of these
223 annotations were plotted with Circos v0.69-9 [59]. The completeness of the genome and the
224 gene annotation was assessed using BUSCO 5.3.2 [60] using the rickettsiales_odb10 dataset.

225 ## 2.6 Whole-genome resequencing and variant calling

226 Three *A. vulgare* females from the WXw line were used to study the genetic diversity of the
227 different *Wolbachia* sub-populations associated with the nerve chain, ovaries and
228 haemolymph. Two of these females were sisters (names: F1_2015 & F2_2015). The matriline
229 of the third one (F_1999) diverged from the matriline of the sisters 18 years ago (18
230 generations). Females were dissected and their tissues isolated individually. Total DNA of each
231 biological sample was extracted using the Qiagen DNeasy 96 Blood & Tissue kit. DNA purity
232 and quantity were measured as described above. The haemolymph of two females could not
233 be sequenced because it did not yield enough DNA. DNA libraries were prepared using the
234 TruSeq Nano DNA Library Prep kit and sequenced on the Illumina HiSeq 2000 platform to

235 generate 2×150 pb reads, by Genoscreen (Lille, France). Read quality was assessed with
236 FastQC 0.11.9 (http://www.bioinformatics.babraham.ac.uk/projects/fastqc). Fastp 0.23.0
237 [61] was used to remove adapter contamination and systematic base calling errors. To
238 mitigate misalignments of reads originating from the host genome, notably from *Wolbachia*-
239 derived nuclear insertions, reads were aligned using bwa-mem2 [46] against a reference
240 including our *w*VulC genome assembly, the assembly of the partial mitochondrial genome of
241 *A. vulgare* (GenBank accession number: MF187614.1, [62]) and the genome assembly of *A.*
242 *vulgare* (accession number: GCA_004104545.1, [63]). PCR duplicates were marked with
243 samtools markdup (v1.14). Read coverage along the *w*VulC genome was computed for each
244 sample for adjacent 5 kb windows using mosdepth 0.2.9 [64]. For a given sample, normalized
245 coverage for each window was calculated by dividing the coverage value by the median of all
246 values in R (v. 4.2.1) [65]. SNPs and short indels were detected using Freebayes 1.3.1 [49], with
247 a filter on the minimum fraction of observations supporting a variant (F = 0.03), disabling
248 population priors (-k argument) and using an aggregate probability of observation balance
249 between alleles (-a argument). Variants were then filtered using a custom R script. This script
250 selected only variable positions for which the read total depth was over 200 and for which the
251 alternative allele was supported by at least 10 reads within a tissue (*Wolbachia* genome
252 coverage - mean ± se: ovaries = 723.9 reads ± 134.8, nerve chain = 769.8 ± 81.0, haemolymph
253 = 533.7).

254 ## 2.7 Variant validation

255 The *A. vulgare* reference genome was assembled from animals belonging to a different lineage
256 from the one used in our experiments [63]. To exclude spurious variants originating from
257 *Wolbachia* genome insertions into the genome of the *A. vulgare* line used in this study, we
258 implemented a four-step approach. (**1**) We aligned short reads from five individuals and five

259 pools of individuals from three different *A. vulgare* lineages uninfected by *Wolbachia* [66] on
260 the reference sequence. Variant calling was carried out with the same parameters as above,
261 but only at positions of detected candidate *Wolbachia* variants. All variants for which the
262 alternative allele was present in the reads sequenced from uninfected individuals were
263 eliminated. (**2**) We defined a 300-bp window around the position of each remaining variant
264 (starting 150 bp before the site) and excluded variants in windows containing at least two
265 reads from uninfected individuals. (**3**) For each remaining variable region, we then designed
266 PCR primers to amplify 120-250 bp targets encompassing the focus variant. These primers
267 were tested on two females from the recipient lineage (i.e., uninfected by *Wolbachia*), on two
268 females from the source lineage (i.e., infected by *Wolbachia*) and on two of their brothers (not
269 infected by *Wolbachia,* since infection induces feminization). If the primers amplified DNA
270 from at least one non-infected host, then the variant was eliminated. (**4**) Finally, read
271 alignments around all the variable regions that passed the previous filters were visualized
272 independently by two researchers using the IGV genome browser. Reads carrying the
273 alternative allele were checked to ensure that their insert size (distance between mates) was
274 consistent with that of the other reads, that they did not display unusual clipping (which may
275 indicate alignment errors) and that there was no significant increase in sequencing depth in
276 the region (which could indicate the presence of duplications rather than variants of the same
277 locus). To predict the effect of retained SNPs and indels on proteins coded by annotated genes,
278 we used SNPeff 5.2c [67]. Functional domains in these proteins were annotated using the
279 InterProScan online service (http://www.ebi.ac.uk/interpro/, database release 100.0 [68]).

280 ## 2.8 Statistical Analyses

281 Statistical analyses were carried out using the R software [65]. The influence of the tissular
282 origin of the injected *Wolbachia* on *A. vulgare* mortality was evaluated using the 'survreg'

283 function with either exponential or Weibull errors distribution ('survival' package, [69]). The
284 model included survival as a response variable, and infection status (i.e., injected with
285 *Wolbachia*-infected or uninfected tissue filtrate), tissue origin of the injected filtrate and
286 experimental block as explanatory variables. The influence of *Wolbachia* tissue origin on its
287 ability to colonize and infect different focal tissues of recipient hosts over time was assessed
288 using a generalized linear mixed model (GLMM). Log-transformed relative quantification (RQ)
289 values representing the per-host cell *Wolbachia* dose were fitted as a response variable,
290 assuming normal error distribution. The tissular origin of *Wolbachia* and the focal tissue of
291 recipient injected hosts and the day post injection were used as explanatory variables. The
292 quadratic term day post injection was added to assess whether it significantly improved the
293 model fit. Individual IDs nested in experimental block were specified as random-effect
294 variables in the model.

295 Maximal models including all higher-order interactions were simplified by sequentially
296 eliminating non-significant terms and interactions to establish a minimal model [70]. The
297 significance of the effects of explanatory variables was established using a likelihood ratio test
298 [70] whose statistics approximately follows a Chi-square distribution [71]. The significant Chi-
299 square values given in the Results section are for the minimal model, whereas non-significant
300 values correspond to those obtained before the deletion of the variables from the model. *A*
301 *posteriori* contrasts were carried out by aggregating factor levels together and by testing the
302 fit of the simplified model using a likelihood-ratio test [70,71].

303 # 3. Results

304 ## 3.1 *Wolbachia* infection does not significantly influence host survival

305 Under the assumption of both exponential errors and non-constant hazard, no effect of
306 *Wolbachia* tissue origin on the survival rate of *A. vulgare* during early stage of infection (from
307 day 0 to 60) was detected (exponential errors: LRT = 0.600, p = 0.740, Weibull errors: LRT =
308 0.602, p = 0.741, **Fig. 1A**). Overall, the injection of *Wolbachia*, irrespective of its tissue origin,
309 was not shown to influence the survival of individuals (exponential errors: LRT = 1.040, p =
310 0.308, Weibull errors: LRT = 1.039, p = 0.307). Sixty days after injection, the survival rate of
311 infected individuals was 0.87 (± 95% CI, ± 0.79-0.96), compared to 0.80 (± 0.71-0.91) for
312 control animals.

313 In the trans-infection batches used to investigate the influence of *Wolbachia* tissue
314 origin on its ability to persist in recipient host tissue, the mortality of animals was also not
315 significantly influenced by *Wolbachia* tissue origin (exponential errors: LRT = 2.65, p = 0.265,
316 Weibull errors: LRT = 2.88, p = 0.236) or even just by the fact of being infected by the bacteria
317 (whatever the origin of the tissue, exponential errors: LRT = 2.22, p = 0.136, Weibull errors:
318 LRT = 2.45, p = 0.117, **Fig. 1B**). The survival rate of infected individuals 200 days after the
319 experimental infection was 0.59 ± 0.48-0.74 and 0.47 ± 0.36-0.61 for infected and control
320 animals respectively.

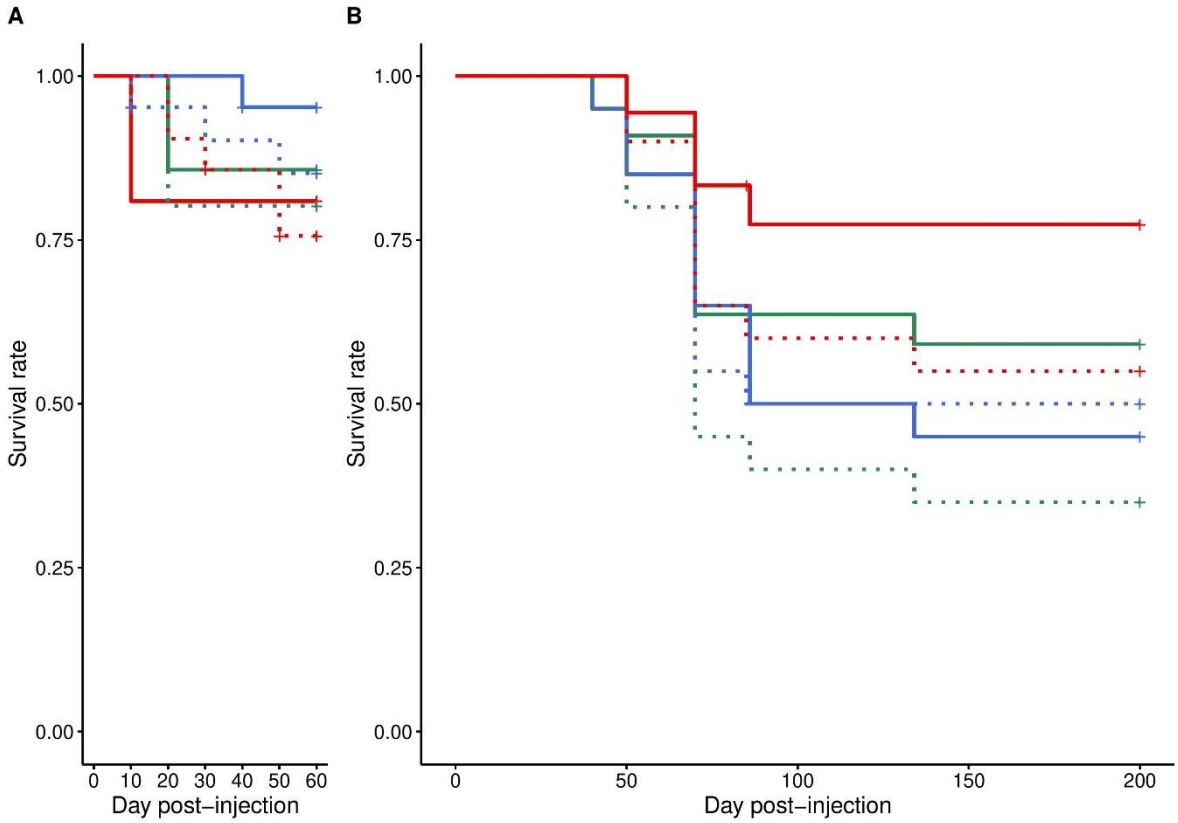


Figure 1. Survival rate of *Armadillidium vulgare* following *Wolbachia* (*w*VulC) trans-infection. Effect of tissue origin of injected *Wolbachia* on mortality of recipient hosts monitored (**A**) every 10 days from day 0 to day 60 post-injection and (**B**) every 15-20 days from day 0 to day 200 post-injection. Solid lines: recipient hosts injected with crushed filtered tissue from animals naturally infected by *Wolbachia w*Vulc. Dotted lines: control individuals injected with crushed filtered tissue *from A. vulgare* not infected by *Wolbachia.* The three colours correspond to the tissue of origin of Wolbachia. Green: nerve chain, red: ovaries, blue: haemolymph.

3.2 The tissue origin of *Wolbachia* influence its early infection dynamics in recipient host tissues

After injection, the presence of *Wolbachia* was detected by qPCR in the three tested tissues of all recipient hosts, confirming its ability to colonize a new host. *Wolbachia* density increased over time (post-injection day: LRT = 16.239, $p < 0.0001$) and fitting the quadratic term (day_post_injection^2) strongly improved model fit (LRT = 32.181, $p < 0.0001$), suggesting that *Wolbachia* burden was more accurately modelled by an accelerated polynomial function of day post-injection (**Fig. 2A**). The density of *Wolbachia* in tissues started very low on day 20

335 post-injection, then rose slightly between day 20 and day 40, to drastically increase between
336 day 40 and day 60 post-injection. Although the pattern of infection in the different recipient
337 tissues was overall similar, *Wolbachia* load was significantly influenced by the interaction
338 between the colonized recipient tissue and the day post-injection (LRT = 74.309, $p < 0.0001$).
339 The dynamics of infection within the nerve chain were significantly different from those
340 observed in the haemolymph and ovaries, leading to a higher bacterial load in nerve tissues
341 60 days post-infection (LRT = 57.228, $p < 0.0001$, LRT = 122.240, $p < 0.0001$, respectively, **Fig.**
342 **2A**). The dynamics of infection were also significantly different within haemolymph and
343 ovaries (LRT = 17.771, p = 0.0013, **Fig. 2A**), as bacterial load was lower in ovaries at the end of
344 monitoring (mean RQ ± se, nerve chain = 6.582 ± 1.596, haemolymph = 4.864 ± 1.612, ovaries
345 = 2.533 ± 0.594, **Fig. 2A**). *Wolbachia* infection dynamics were also influenced by the tissue
346 origin of the injected bacteria (interaction between days post-injection and tissue origin: LRT
347 = 24.205, $p < 0.0001$). Whether the bacteria came from the haemolymph or ovaries of
348 naturally infected animals did not significantly affect the colonization dynamics of *Wolbachia*
349 in the injected recipients (LRT = 0.647, p = 0.723, **Fig. 2B**). However, bacteria from the nerve
350 chain colonized recipient hosts less rapidly than bacteria from other tissues (nerve chain
351 versus haemolymph: LRT = 28.981, $p < 0.0001$, nerve chain versus ovaries: LRT = 33.64, $p <$
352 0.0001). At 60 days post-injection, the *Wolbachia* load was on average more than twice as
353 high when the bacteria came from the ovaries and haemolymph than when they came from
354 the nerve chain of infected source hosts (mean RQ ± se, ovaries = 8.191 ± 2.18, haemolymph
355 = 5.26 ± 1.077, nerve chain = 0.899 ± 0.197, **Fig. 2B**). However, there was no significant
356 interaction between the colonized recipient tissue and the *Wolbachia's* original tissue,
357 suggesting that bacteria originating from a tissue do not colonize that same tissue more
358 effectively in recipient hosts (LRT = 6.957 p = 0.138).

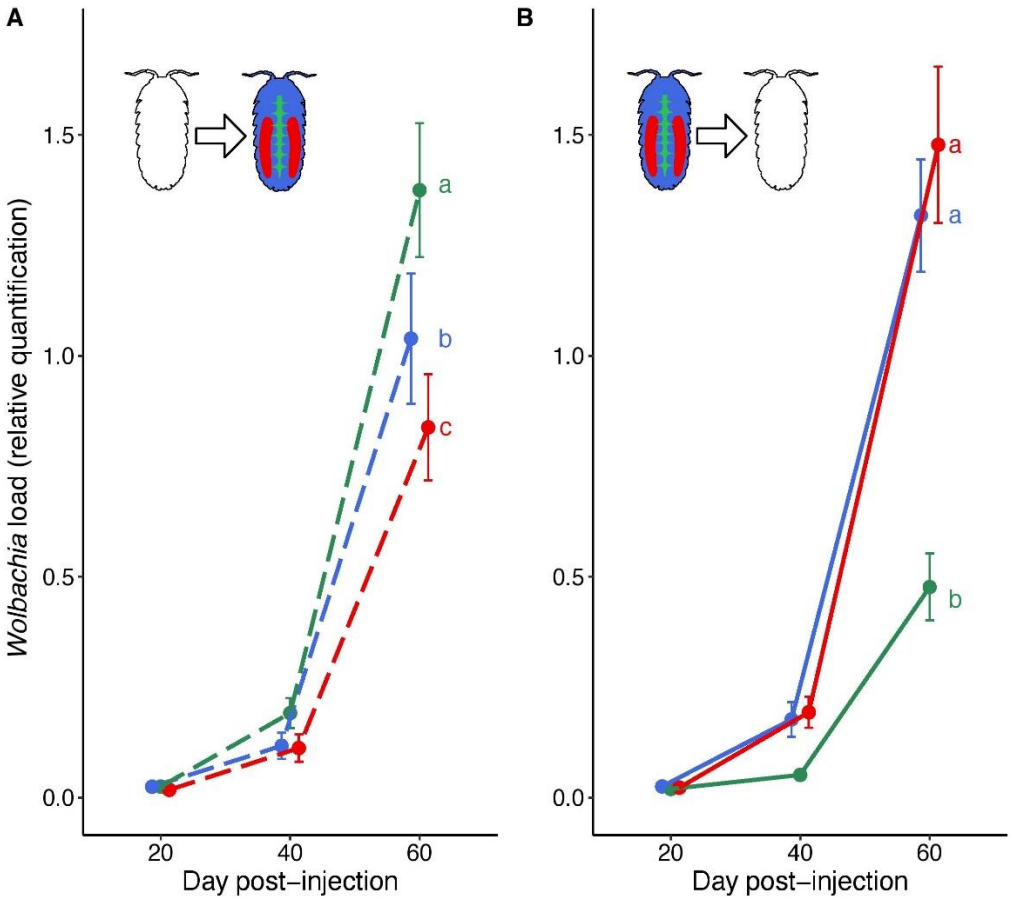
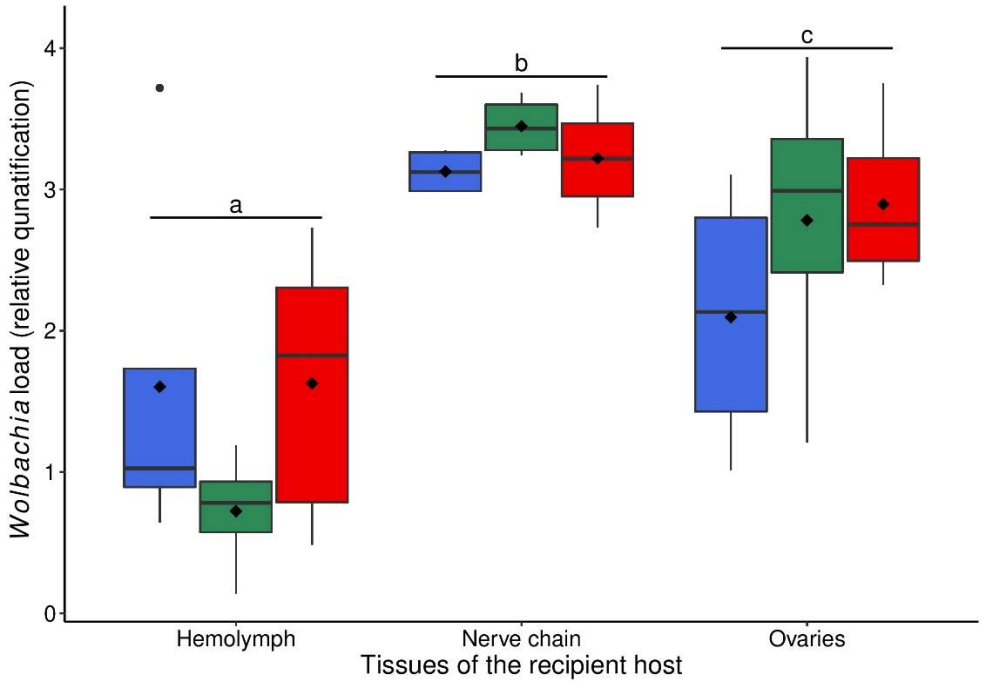


Figure 2. Infection dynamics of *Wolbachia* (*w*VulC) in *Armadillidium vulgare* from day 0 to day 60 post-injection. (A) *Wolbachia* loads in nerve chain (green), haemolymph (blue) and ovaries (red) of recipient hosts trans-infected by *Wolbachia*, all source tissue considered. (B) Influence of *Wolbachia* tissue origin on infection dynamics in tissue loads in recipient host, all recipient tissue considered. Green: nerve chain, red: ovaries, blue: haemolymph. Error bars correspond to the standard error. Infection dynamics not connected by the same letter are significantly different ($p < 0.05$).

3.3 The tissue origin of Wolbachia does not influence significantly its persistence in recipient host tissues

Wolbachia densities at 200 days after infection were significantly different between tissues from the recipient hosts (LRT = 39.224, $p < 0.0001$, **Fig. 3**). Bacterial load was the highest in the nerve chain and the lowest in the haemolymph (contrast analyses, haemolymph versus nerve chain: LRT = 38.133, $p < 0.0001$, haemolymph versus ovaries: LRT = 20.823, $p < 0.0001$,

372 nerve chain versus ovaries: LRT = 6.969, p = 0.008, **Fig. 3**). Contrary to what was observed
373 during the early phase of infection, there was no effect of the tissue origin of *Wolbachia* on
374 the bacterial load in recipient tissues at 200 days post-injection (LRT = 0.368, p = 0.832).
375 Similarly, *Wolbachia* density in recipient tissues was not influenced by the interaction between
376 the colonized tissue and the tissue origin of the injected bacteria (LRT = 9.539, p = 0.0589).



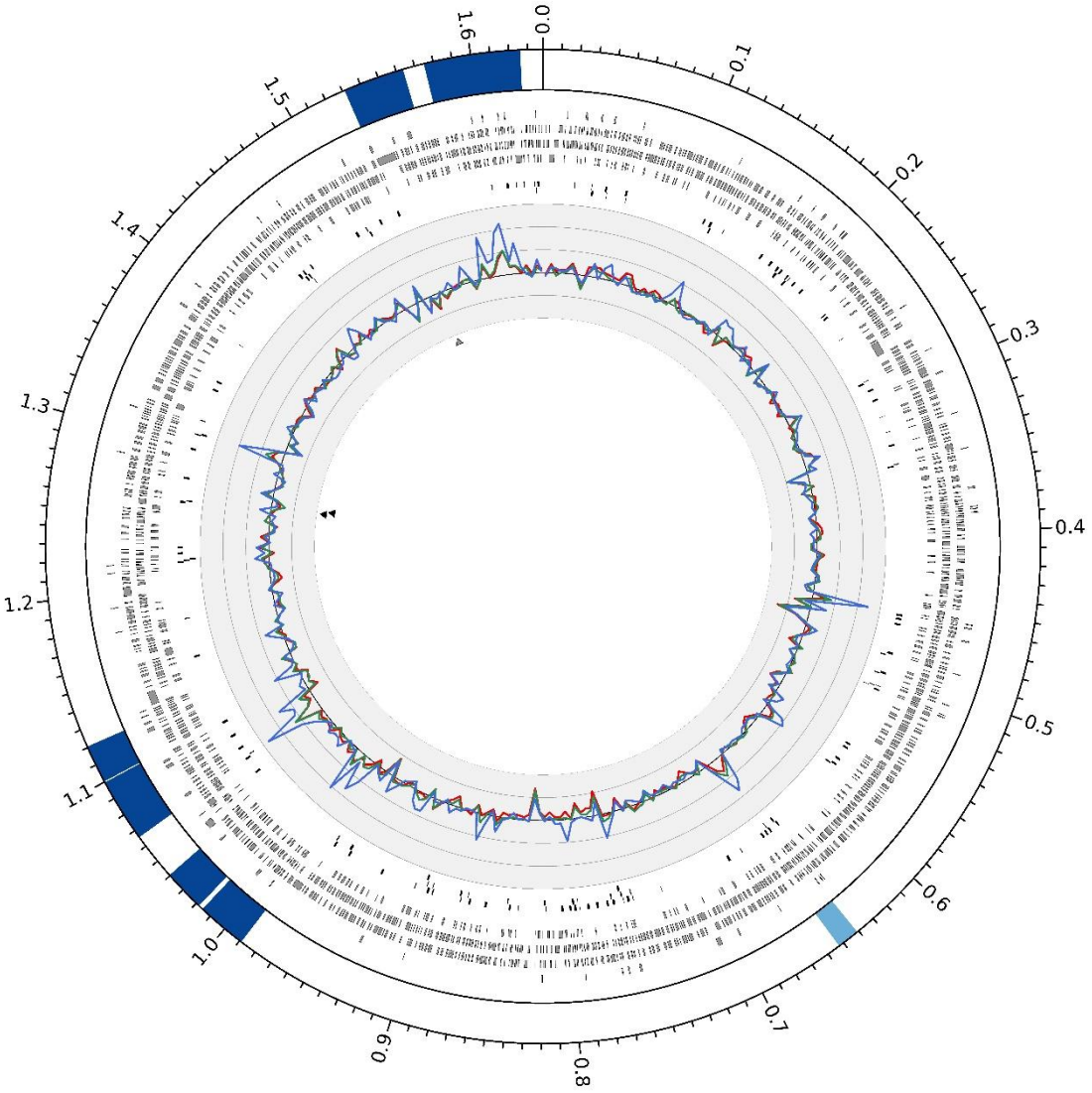
377

378 **Figure 3. *Wolbachia* (*w*VulC) load in *Armadillidium vulgare* tissues six months post-injection.** The colored
379 boxplots illustrate the tissue origin of Wolbachia (green: nerve chain, red: ovaries, blue: haemolymph). However,
380 the analysis did not reveal any significant effect of the tissue origin of the bacteria on the intensity of infection
381 in the recipient host tissues. Boxes above and below the medians (horizontal lines) show the first and third
382 quartiles, respectively. Black diamonds represent the means. Levels not connected by the same letter are
383 significantly different (p < 0.05).

384 ## 3.4 Complete genome assembly of the *w*VulC *Wolbachia* strain

385 Using ONT reads, a complete circular genome sequence (1,638,144 bp) of the *w*VulC
386 *Wolbachia* strain was assembled. The estimated completeness of the assembly was very high,
387 with 99.8% of the expected BUSCO genes found in the sequence (C: 99.8% [S: 99.5%, D: 0.3%],

388 F: 0.0%). The annotation of the genome resulted in 1,390 protein-coding sequences and 208
389 pseudogenes (**Fig. 4**). The *w*VulC genome contained 138 IS elements that are larger than 150
390 bp, representing 8.45% of the total length of the genome. This high proportion of repeated
391 sequences for a bacterium is similar to that of other *Wolbachia* strains [72]. We also identified
392 seven prophage regions (six potentially intact and one questionable region, following
393 PHASTEST scoring criteria), representing 12.1% of the *w*VulC genome.



394

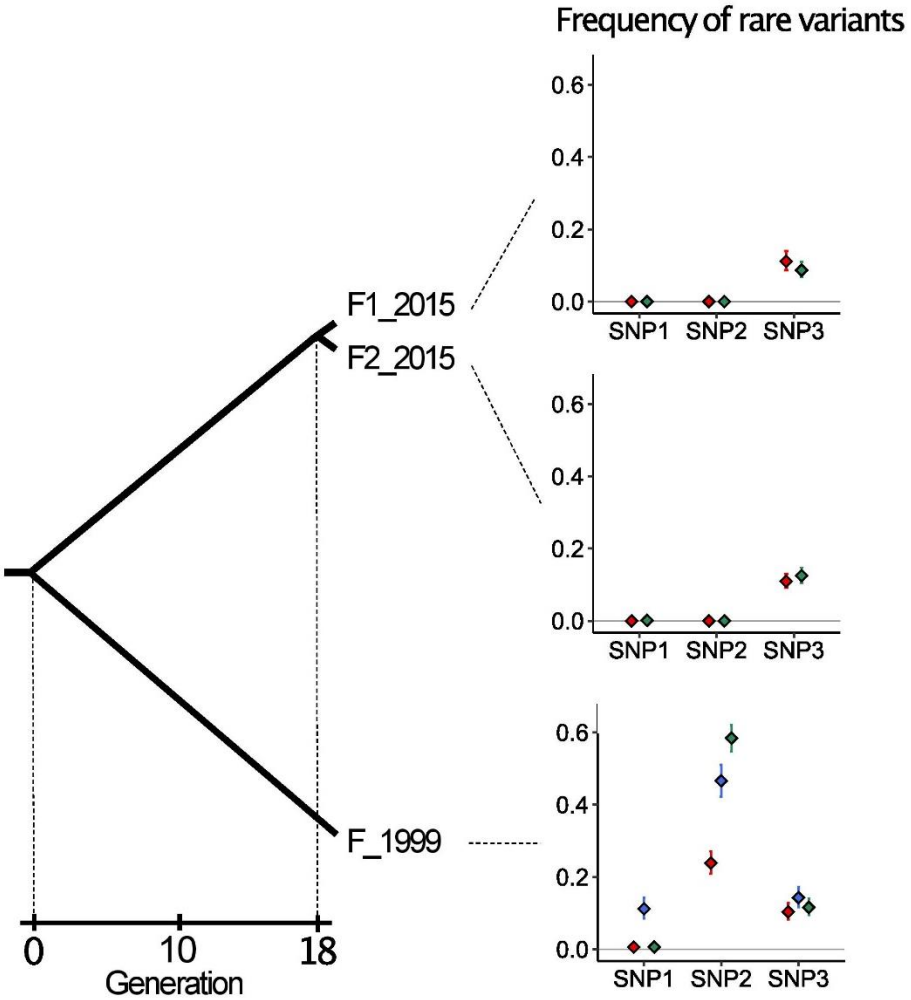
395 **Figure 4. Circos plot of the *w*VulC *Wolbachia* genome.** Prophage-derived regions are indicated inside the
396 genome ideogram (outer track), with their colour corresponding to the score attributed by PHASTEST (dark blue:
397 intact, light blue: questionable). Ticks on the ideogram are separated by 10 kb intervals, and indicated values are

398 in Mb. Predicted protein-coding genes are represented by grey rectangles (second track), and Insertion Sequence
399 elements by black rectangles (third track). In the fourth track, the normalized read coverage for 5 kb adjacent
400 windows (averaged over recipient individuals) is represented for the different resequenced tissues: ovaries (red),
401 nerve chain (green) and haemolymph (blue). The black line corresponds to a mean normalized coverage of one,
402 and grey lines are separated by 0.1 unit. Triangles indicates the position of detected variants after filtering (black
403 triangle: found in only one recipient individual, grey triangle: found in all recipient individuals).

404 3.5 Whole-genome resequencing and variant calling reveal *Wolbachia* genome stability

405 The variant calling performed using the whole-genome resequencing of tissues from three
406 females naturally infected by *Wolbachia* (i.e., source lineage) detected 641 SNPs and short
407 indels. However, 584 of these apparent variants were also detected in the resequenced
408 genome of several *A. vulgare* lineages not infected by *Wolbachia* (validation procedure, step
409 1), meaning that these variants probably amounted to *Wolbachia* sequences inserted into the
410 host nuclear genome. The presence of reads sequenced from uninfected hosts within a 300-
411 bp window around the 57 remaining polymorphic regions resulted in the exclusion of 25
412 variants (validation procedure step 2). Of the remaining 32 SNPs and small indels, 17 were
413 excluded because their regions were PCR-amplified in males not infected with *Wolbachia*
414 (infection status confirmed by qPCR, validation procedure step 3). Finally, during the last step
415 of manual validation of the variants, we excluded 18 variants due to detected anomalies, such
416 as inconsistent insert size, excessive soft-clipping, or substantial increases in sequencing depth
417 in their region. Overall, only 0.47% (two SNPs and one small indel) of the originally detected
418 variants were finally validated. Among the three variants, one SNP showed variation in all
419 tested tissues of all females (**Table 1**, **Fig. 5**). The frequency of the rarer allele was relatively
420 stable across tissues and varied between 8.7% and 14.3% (**Table 1**, **Fig. 5**). One SNP and one
421 indel of one nucleotide were observed in a single individual (F_1999, **Fig. 5**). The frequency of
422 these two variants varied slightly among the three tissues (**Fig. 5**). The three polymorphisms

423 affected three protein-coding genes, with two being non-synonymous substitutions and one
424 causing a frameshift (**Table 1**). Two of the putative proteins did not have any recognized
425 functional domain. The third one, whose gene is located in a prophage-derived region (**Fig. 4**)
426 contained ankyrin repeat domains and a latrotoxin C-terminal domain.



427

428 **Figure 5. Frequency of rare variants in the tissues of naturally infected hosts.** Pedigree of the three *Wolbachia*-
429 infected *Armadillidium vulgare* females associated to the frequency of the three variants identified. SNP1: variant
430 observed at the genome position 1266439, SNP2: variant observed at the genome position 1269492, SNP3:
431 variant observed at the genome position 1536425 (see Table 1). Colours correspond to host's tissues. Green:
432 nerve chain, red: ovaries, blue: haemolymph. Error bar: CI 95%.

433 ## 4. Discussion

434 Experimental trans-infections enabled us to study the influence of the tissue origin of
435 *Wolbachia* on the early infection dynamics of the endosymbiont in different recipient host
436 tissues, and the ability of each bacterial sub-population to establish a chronic infection. At the
437 same time, regular monitoring of injected individuals over a period of six months shed light
438 on the influence of *Wolbachia* on the survival rate of recipient hosts and did not reveal any
439 difference between *Wolbachia* sub-populations with respect to their virulence. In addition,
440 our results corroborate a previous report showing a similar survival rate between trans-
441 infected and un-infected animals [40]. As the source and recipient lines belonged to the same
442 field-collected *A. vulgare* population, they should be ecologically similar for *Wolbachia*. Low
443 virulence is therefore expected, as for most vertically transmitted endosymbionts transferred
444 between closely related hosts [73].

445 The quantification of *Wolbachia* in the main organs of recipient hosts showed that its
446 infection dynamics varies according to its tissue origin. Bacteria from the nerve chain
447 colonized recipient hosts much less rapidly than bacteria from ovaries and haemolymph. Two
448 months after infection with a similar dose of bacteria, *Wolbachia* load was more than three
449 times lower in hosts injected with bacteria from nerve chain, compared with hosts infected by
450 bacteria from the other two tissues. The growth rate of bacteria is known to largely vary
451 according to the type and amount of nutrients available in their environment [74,75]. The
452 lower growth rate of bacteria from nerve chain, compared with those from the other tissues,
453 could therefore be explained by the fact that in naturally infected hosts, this tissue is less
454 favourable to the development of *Wolbachia* (e.g., poorer in resources). If so, bacteria from
455 the nerve chain could be subject to growth-limiting stress at the time of their horizontal

456 transfer, which would explain their lower replication rate within the recipient host at the start
457 of infection. However, the recipient host tissue that was the most rapidly colonized by
458 *Wolbachia* after trans-infection, and the one showing the highest bacterial load at each
459 sampling point irrespective of the tissue origin of the bacteria, was the nerve chain. As well as
460 showing no tropism towards the bacteria's tissue of origin, this result suggests that nervous
461 tissue is favourable to the development of *Wolbachia.*

462 Alternatively, the lower growth rate from bacteria originating from the nerve chain
463 could be adaptive. Classic theory for the evolution of virulence is based on a trade-off between
464 parasite growth, transmission and host survival, which predicts that higher growth increases
465 not only transmission but also virulence [76,77]. In a tissue that is not directly involved in
466 parasite transmission, we predicted that virulence should constrain the parasite growth to low
467 levels. Our results corroborate this prediction, as *Wolbachia* from the nerve chain of *A. vulgare*
468 should not be transmissible – except maybe in rare cases of cannibalism and predation [73] –
469 and may prove detrimental to the host if they proliferate. Uncontrolled colonization of nerve
470 cells has indeed been shown to alter host behaviour in several *Wolbachia*-host pairs [38,40],
471 and to lead to severe tissue degeneration and premature host death [40,41,78,79]. In tissues
472 that allow parasite transmission, the balance between *Wolbachia* growth and host fitness
473 should favour higher growth rates. In terrestrial isopods, both ovaries and haemolymph play
474 a key role in both intra- and inter-host dissemination of *Wolbachia*. Ovaries are the route for
475 vertical transmission and haemolymph, in addition to participating in the dissemination of the
476 bacteria across host tissues [37,80,81], may also be involved in horizontal transfer via contact
477 between injured individuals [82].

478 Despite differences observed at two months following infection, the tissue origin of
479 the injected *Wolbachia* had no significant effect on the bacterial load in the three tissues of

480 the recipient hosts at six months after injection. The decrease of growth rate differences
481 within a few weeks suggests that the phenotypic variability observed at the start of the
482 infection corresponded to phenotypic plasticity. This assumption is supported by the lack of
483 recurring small genomic variation associated with tissue types in naturally infected individuals
484 from the source lineage. More generally, with respect to SNPs and small indels, we found very
485 little genomic variability within naturally infected hosts as only three mutations were detected
486 in sequence data from seven tissue samples. Interestingly however, a SNP showed variation
487 in all sequenced individuals. The two *A. vulgare* sisters (F1_2015 & F2_2015) have been
488 separated from individual F_1999 for 18 generations, suggesting that two *Wolbachia* lineages
489 varying at this SNP have been co-infecting the source lineage for at least 18 years (generation
490 time is one year in *A. vulgare*). The other two mutations were only observed in individual
491 F_1999, which suggests that they appeared after the divergence of its maternal lineage from
492 that of the two sisters. This very low number of mutations is consistent with reported low
493 mutation rates, as far as short genomic variation is concerned [83–86]. Long-read sequencing
494 of multiple individuals would allow studying structural genomic variation, which may play a
495 more important role than point mutations in *Wolbachia* evolution [87–89].

496 In conclusion, our study reveals that within-host environmental heterogeneity can lead
497 to diverse phenotypes in the most widespread bacterial endosymbiont in animals, *Wolbachia*.
498 This variability did not appear to result from genomic variation, but from phenotypic plasticity.
499 Knowing the pedigree of *Wolbachia*-infected animal lines proved pivotal in assessing the
500 stability of the *Wolbachia* genome over time. From a methodological point of view, we
501 showed that the detection of variants in endosymbiont populations requires considerable
502 caution, as our conservative approach led us to exclude more than 99.5% of the initially called
503 variants. We recommend the use of a rigorous step-by-step approach to eliminate spurious

genetic variants caused by the presence of endosymbiont genomic sequences inserted into the host genome.

Data accessibility statement

The R scripts and data supporting the conclusions of this article are available on the Figshare data repository (https://figshare.com/s/b1f84643b65b4f897c54). To ensure the stability and reliability of the command-lines used for genome assembly and SNP calling we have included all necessary information and options in the github repository (https://github.com/UMR-CNRS-7267/Wolbachia_endosymbiont_heterogeneity_paper). We have taken proactive measures to address potential volatility on GitHub by uploading a compressed archive of this GitHub methodology to the Figshare data repository. Raw sequencing data have been deposited in GenBank (*w*VulC genome assembly: BioProject PRJNA1116085; Wolbachia resequencing: PRJNA1117639).

Author's contributions

R.P. and R.C. conceived and designed the experiments. R.P., R.J., M.P. and M.R. performed the trans-infection experiments. R.P. and C.D prepared the samples for resequencing and carried out the PCRs. D.O. performed *Wolbachia* genome sequencing. Y.D. assembled the *Wolbachia* genome. R.P., Y.D., B.M. and J.P. analysed the data. R.P. wrote the first draft of the manuscript, and all authors contributed substantially to revision.

Competing interests

We declare we have no competing interests.

Funding

This work was funded by Agence Nationale de la Recherche Grants ANR-21-CE02-0004 (RESIST) to RC and ANR-20-CE02-004 (SymChroSex) to JP, and intramural funds from the CNRS and the University of Poitiers.

529 **Acknowledgements**

530 We thank Isabelle Giraud for the extraction of the DNA used to sequence and assemble the
531 genome of *Wolbachia w*VulC and Pierre Grève and Didier Bouchon for providing us with
532 preliminary data so that we could fine-tune the bioinformatics analyses. We would like to
533 thank Alexandra Lafitte for her invaluable work supervising the breeding of *A. vulgare*.

534 5. References

535 1. Ben-Hur E, Kadmon R. 2020 Heterogeneity–diversity relationships in sessile organisms: a
536 unified framework. *Ecol. Lett.* **23**, 193–207. (doi:10.1111/ele.13418)

537 2. Stein A, Kreft H. 2015 Terminology and quantification of environmental heterogeneity in
538 species-richness research. *Biol Rev Camb Philos Soc* **90**, 815–836.
539 (doi:10.1111/brv.12135)

540 3. Burnett MR, August PV, Brown JH, Killingbeck KT. 1998 The influence of
541 geomorphological heterogeneity on biodiversity: I. a patch-scale perspective. *Conserv.*
542 *Biol.y* **12**, 363–370.

543 4. Stein A, Gerstner K, Kreft H. 2014 Environmental heterogeneity as a universal driver of
544 species richness across taxa, biomes and spatial scales. *Ecol. Lett.* **17**, 866–880.
545 (doi:10.1111/ele.12277)

546 5. Yang Z, Liu X, Zhou M, Ai D, Wang G, Wang Y, Chu C, Lundholm JT. 2015 The effect of
547 environmental heterogeneity on species richness depends on community position along
548 the environmental gradient. *Sci. Rep.* **5**, 15723. (doi:10.1038/srep15723)

549 6. Houle C, Pelletier F, Bélisle M, Garant D. 2020 Impacts of environmental heterogeneity
550 on natural selection in a wild bird population. *Evolution* **74**, 1142–1154.
551 (doi:10.1111/evo.13939)

552 7. Huang B-H, Huang C-W, Huang C-L, Liao P-C. 2017 Continuation of the genetic
553 divergence of ecological speciation by spatial environmental heterogeneity in island
554 endemic plants. *Sci. Rep.* **7**, 5465. (doi:10.1038/s41598-017-05900-1)

555 8. Mcdonald JF, Ayala FJ. 1974 Genetic response to environmental heterogeneity. *Nature*
556 **250**, 572–574. (doi:10.1038/250572a0)

557 9. Rainey PB, Travisano M. 1998 Adaptive radiation in a heterogeneous environment.
558 *Nature* **394**, 69–72. (doi:10.1038/27900)

559 10. Huang Y, Tran I, Agrawal AF. 2016 Does genetic variation maintained by environmental
560 heterogeneity facilitate adaptation to novel selection? *Am. Nat.* **188**, 27–37.
561 (doi:10.1086/686889)

562 11. Qian H, Qian S, Zhang J, Kessler M. 2024 Effects of climate and environmental
563 heterogeneity on the phylogenetic structure of regional angiosperm floras worldwide.
564 *Nat. Comm.* **15**, 1079. (doi:10.1038/s41467-024-45155-9)

565 12. Trevail AM *et al.* 2021 Environmental heterogeneity promotes individual specialisation in
566 habitat selection in a widely distributed seabird. *J. Anim. Ecol.* **90**, 2875–2887.
567 (doi:10.1111/1365-2656.13588)

568 13. van Houte S, Padfield D, Gómez P, Luján AM, Brockhurst MA, Paterson S, Buckling A.
569 2021 Compost spatial heterogeneity promotes evolutionary diversification of a
570 bacterium. *J. Evol. Biol.* **34**, 246–255. (doi:10.1111/jeb.13722)

571 14. Langenheder S, Lindström ES. 2019 Factors influencing aquatic and terrestrial bacterial
572 community assembly. *Envir. Microb. Rep.* **11**, 306–315. (doi:10.1111/1758-2229.12731)

573 15. Johnson PTJ, Wood CL, Joseph MB, Preston DL, Haas SE, Springer YP. 2016 Habitat
574 heterogeneity drives the host-diversity-begets-parasite-diversity relationship: evidence
575 from experimental and field studies. *Ecol. Lett.* **19**, 752–761. (doi:10.1111/ele.12609)

576 16. Regoes RR, Nowak MA, Bonhoeffer S. 2000 Evolution of virulence in a heterogeneous
577 host population. *Evolution* **54**, 64–71. (doi:10.1111/j.0014-3820.2000.tb00008.x)

578 17. White PS, Choi A, Pandey R, Menezes A, Penley M, Gibson AK, de Roode J, Morran L.
579 2020 Host heterogeneity mitigates virulence evolution. *Biol. Lett.* **16**, 20190744.
580 (doi:10.1098/rsbl.2019.0744)

581 18. Kamiya T, O'Dwyer K, Nakagawa S, Poulin R. 2014 What determines species richness of
582 parasitic organisms? A meta-analysis across animal, plant and fungal hosts. *Biol. Rev.*
583 *Camb. Philos. Soc.* **89**, 123–134. (doi:10.1111/brv.12046)

584 19. Didelot X, Walker AS, Peto TE, Crook DW, Wilson DJ. 2016 Within-host evolution of
585 bacterial pathogens. *Nat. Rev. Microbiol.* **14**, 150–162. (doi:10.1038/nrmicro.2015.13)

586 20. Wernegreen JJ, Wheeler DE. 2009 Remaining flexible in old alliances: functional plasticity
587 in constrained mutualisms. *DNA Cell Biol.* **28**, 371–382. (doi:10.1089/dna.2009.0872)

588 21. Chrostek E, Teixeira L. 2018 Within host selection for faster replicating bacterial
589 symbionts. *PLoS One* **13**. (doi:10.1371/journal.pone.0191530)

590 22. Ailloud F, Didelot X, Woltemate S, Pfaffinger G, Overmann J, Bader RC, Schulz C,
591 Malfertheiner P, Suerbaum S. 2019 Within-host evolution of *Helicobacter pylori* shaped
592 by niche-specific adaptation, intragastric migrations and selective sweeps. *Nat. Comm.*
593 **10**, 2273. (doi:10.1038/s41467-019-10050-1)

594 23. Gupta K *et al.* 2022 Structural insights in cell-type specific evolution of intra-host
595 diversity by SARS-CoV-2. *Nat. Comm.* **13**, 222. (doi:10.1038/s41467-021-27881-6)

596 24. Jridi C, Martin J-F, Marie-Jeanne V, Labonne G, Blanc S. 2006 Distinct viral populations
597 differentiate and evolve independently in a single perennial host plant. *J. Virol.* **80**, 2349–
598 2357. (doi:10.1128/jvi.80.5.2349-2357.2006)

599 25. Xiao Y, Dolan PT, Goldstein EF, Li M, Farkov M, Brodsky L, Andino R. 2017 Poliovirus
600 intrahost evolution is required to overcome tissue-specific innate immune responses.
601 *Nat. Commun.* **8**, 375. (doi:10.1038/s41467-017-00354-5)

602 26. Jorth P *et al.* 2015 Regional Isolation Drives Bacterial Diversification within Cystic Fibrosis
603 Lungs. *Cell Host Micro.* **18**, 307–319. (doi:10.1016/j.chom.2015.07.006)

604 27. Lieberman TD, Wilson D, Misra R, Xiong LL, Moodley P, Cohen T, Kishony R. 2016
605 Genomic diversity in autopsy samples reveals within-host dissemination of HIV-
606 associated *Mycobacterium tuberculosis*. *Nat. Med.* **22**, 1470–1474.
607 (doi:10.1038/nm.4205)

608 28. Bessière P, Volmer R. 2021 From one to many: The within-host rise of viral variants. *PLOS*
609 *Path.* **17**, e1009811. (doi:10.1371/journal.ppat.1009811)

610 29. Avettand-Fenoel V, Hocqueloux L, Müller-Trutwin M, Prazuck T, Melard A, Chaix M-L,
611 Agoute E, Michau C, Rouzioux C. 2011 Greater diversity of HIV DNA variants in the
612 rectum compared to variants in the blood in patients without HAART. *J. Med. Virol.* **83**,
613 1499–1507. (doi:10.1002/jmv.22132)

614 30. Clement S, Vaudaux P, Francois P, Schrenzel J, Huggler E, Kampf S, Chaponnier C, Lew D,
615 Lacroix J-S. 2005 Evidence of an intracellular reservoir in the nasal mucosa of patients
616 with recurrent *Staphylococcus aureus* rhinosinusitis. *J. Infect. Disea.* **192**, 1023–1028.
617 (doi:10.1086/432735)

618 31. Obaldia N *et al.* 2018 Bone marrow is a major parasite reservoir in *Plasmodium vivax*
619 infection. *MBio* **9**. (doi:10.1128/mBio.00625-18)

620 32. Kaur R, Shropshire JD, Cross KL, Leigh B, Mansueto AJ, Stewart V, Bordenstein SR,
621 Bordenstein SR. 2021 Living in the endosymbiotic world of *Wolbachia*: A centennial
622 review. *Cell Host Microb.* **29**, 879–893. (doi:10.1016/j.chom.2021.03.006)

33. LePage D, Bordenstein SR. 2013 *Wolbachia*: Can we save lives with a great pandemic? *Trends Parasitol.* **29**, 385–393. (doi:10.1016/j.pt.2013.06.003)

34. Cordaux R, Gilbert C. 2017 Evolutionary significance of *Wolbachia*-to-animal horizontal gene transfer: Female sex determination and the f element in the Isopod *Armadillidium vulgare*. *Genes* **8**, 186. (doi:10.3390/genes8070186)

35. Sicard M, Dittmer J, Grève P, Bouchon D, Braquart-Varnier C. 2014 A host as an ecosystem: *Wolbachia* coping with environmental constraints. *Environ. Microbiol.* **16**, 3583–3607. (doi:10.1111/1462-2920.12573)

36. Cordaux R, Michel-Salzat A, Frelon-Raimond M, Rigaud T, Bouchon D. 2004 Evidence for a new feminizing *Wolbachia* strain in the isopod *Armadillidium vulgare*: evolutionary implications. *Heredity* **93**, 78–84. (doi:10.1038/sj.hdy.6800482)

37. Braquart-Varnier C, Raimond M, Mappa G, Chevalier FD, Le Clec'h W, Sicard M. 2015 The hematopoietic organ: A cornerstone for *Wolbachia* propagation between and within hosts. *Front. Microbiol.* **6**. (doi:10.3389/fmicb.2015.01424)

38. Bi J, Wang Y. 2020 The effect of the endosymbiont *Wolbachia* on the behavior of insect hosts. *Insect Sci* **27**, 846–858. (doi:10.1111/1744-7917.12731)

39. Templé N, Richard F-J. 2015 Intra-cellular bacterial infections affect learning and memory capacities of an invertebrate. *Front. Zool.* **12**, 36. (doi:10.1186/s12983-015-0129-6)

40. Le Clec'h WL, Braquart-Varnier C, Raimond M, Ferdy J-B, Bouchon D, Sicard M. 2012 High virulence of *Wolbachia* after host switching: When autophagy hurts. *PLOS Path.* **8**, e1002844. (doi:10.1371/journal.ppat.1002844)

41. Strunov A, Kiseleva E. 2016 *Drosophila melanogaster* brain invasion: pathogenic *Wolbachia* in central nervous system of the fly. *Insect Sci.* **23**, 253–264. (doi:10.1111/1744-7917.12187)

42. Sicard M, Chevalier F, De Vlechouver M, Bouchon D, Grève P, Braquart-Varnier C. 2010 Variations of immune parameters in terrestrial isopods: a matter of gender, aging and *Wolbachia*. *Naturwissenschaften* **97**, 819–826. (doi:10.1007/s00114-010-0699-2)

43. Durand S, Lheraud B, Giraud I, Bech N, Grandjean F, Rigaud T, Peccoud J, Cordaux R. 2023 Heterogeneous distribution of sex ratio distorters in natural populations of the isopod *Armadillidium vulgare*. *Biol. Lett.* **19**, 20220457. (doi:10.1098/rsbl.2022.0457)

654 44. De Coster W, D'Hert S, Schultz DT, Cruts M, Van Broeckhoven C. 2018 NanoPack:
655 visualizing and processing long-read sequencing data. *Bioinformatics* **34**, 2666–2669.
656 (doi:10.1093/bioinformatics/bty149)

657 45. Kolmogorov M, Yuan J, Lin Y, Pevzner PA. 2019 Assembly of long, error-prone reads using
658 repeat graphs. *Nat. Biotechnol.* **37**, 540–546. (doi:10.1038/s41587-019-0072-8)

659 46. Vasimuddin Md, Misra S, Li H, Aluru S. 2019 Efficient architecture-aware acceleration of
660 BWA-MEM for multicore systems. *IEEE International Parallel and Distributed Processing*
661 *Symposium (IPDPS)*, pp. 314–324. (doi:10.1109/IPDPS.2019.00041)

662 47. Wick RR, Holt KE. 2022 Polypolish: Short-read polishing of long-read bacterial genome
663 assemblies. *PLOS Comput. Biol.* **18**, e1009802. (doi:10.1371/journal.pcbi.1009802)

664 48. Marçais G, Delcher AL, Phillippy AM, Coston R, Salzberg SL, Zimin A. 2018 MUMmer4: A
665 fast and versatile genome alignment system. *PLOS Comput. Biol.* **14**, e1005944.
666 (doi:10.1371/journal.pcbi.1005944)

667 49. Garrison E, Marth G. 2012 Haplotype-based variant detection from short-read
668 sequencing. Arxiv:**1207**.3907 (doi:10.48550/arXiv.1207.3907)

669 50. Robinson JT, Thorvaldsdóttir H, Wenger AM, Zehir A, Mesirov JP. 2017 Variant Review
670 with the Integrative Genomics Viewer. *Cancer Res* **77**, e31–e34. (doi:10.1158/0008-
671 5472.CAN-17-0337)

672 51. Danecek P *et al.* 2021 Twelve years of SAMtools and BCFtools. *GigaScience* **10**, giab008.
673 (doi:10.1093/gigascience/giab008)

674 52. Ioannidis JPA. 2007 Non-replication and inconsistency in the genome-wide association
675 setting. *Hum. Hered*. **64**, 203–213. (doi:10.1159/000103512)

676 53. Hunt M, Silva ND, Otto TD, Parkhill J, Keane JA, Harris SR. 2015 Circlator: automated
677 circularization of genome assemblies using long sequencing reads. *Genom. Biol.* **16**, 294.
678 (doi:10.1186/s13059-015-0849-0)

679 54. Haft DH *et al.* 2024 RefSeq and the prokaryotic genome annotation pipeline in the age of
680 metagenomes. *Nucleic Acids Res.* **52**, D762–D769. (doi:10.1093/nar/gkad988)

681 55. Li W *et al.* 2021 RefSeq: expanding the prokaryotic genome annotation pipeline reach
682 with protein family model curation. *Nucleic Acids Res* **49**, D1020–D1028.
683 (doi:10.1093/nar/gkaa1105)

684 56. Tatusova T *et al.* 2016 NCBI prokaryotic genome annotation pipeline. *Nucleic Acids Res.*
685 **44**, 6614–6624. (doi:10.1093/nar/gkw569)

686 57. Puterová J, Martínek T. 2021 digIS: towards detecting distant and putative novel
687 insertion sequence elements in prokaryotic genomes. *BMC Bioinfo.* **22**, 258.
688 (doi:10.1186/s12859-021-04177-6)

689 58. Wishart DS, Han S, Saha S, Oler E, Peters H, Grant JR, Stothard P, Gautam V. 2023
690 PHASTEST: faster than PHASTER, better than PHAST. *Nucleic Acids Res.* **51**, W443–W450.
691 (doi:10.1093/nar/gkad382)

692 59. Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D, Jones SJ, Marra MA.
693 2009 Circos: an information aesthetic for comparative genomics. *Genome Res.* **19**, 1639–
694 1645. (doi:10.1101/gr.092759.109)

695 60. Manni M, Berkeley MR, Seppey M, Zdobnov EM. 2021 BUSCO: Assessing genomic data
696 quality and beyond. *Current Prot.* **1**, e323. (doi:10.1002/cpz1.323)

697 61. Chen S, Zhou Y, Chen Y, Gu J. 2018 fastp: an ultra-fast all-in-one FASTQ preprocessor.
698 *Bioinformatics* **34**, i884–i890. (doi:10.1093/bioinformatics/bty560)

699 62. Peccoud J, Chebbi MA, Cormier A, Moumen B, Gilbert C, Marcadé I, Chandler C, Cordaux
700 R. 2017 Untangling heteroplasmy, structure, and evolution of an atypical mitochondrial
701 genome by PacBio Sequencing. *Genetics* **207**, 269–280.
702 (doi:10.1534/genetics.117.203380)

703 63. Chebbi MA, Becking T, Moumen B, Giraud I, Gilbert C, Peccoud J, Cordaux R. 2019 The
704 genome of *Armadillidium vulgare* (Crustacea, Isopoda) provides insights into sex
705 chromosome evolution in the context of cytoplasmic sex determination. *Mol. Biol. Evol.*
706 **36**, 727–741. (doi:10.1093/molbev/msz010)

707 64. Pedersen BS, Quinlan AR. 2018 Mosdepth: quick coverage calculation for genomes and
708 exomes. *Bioinformatics* **34**, 867–868. (doi:10.1093/bioinformatics/btx699)

709 65. R Core Team. 2022 R: A language and environment for statistical computing.

710 66. Cordaux R, Chebbi MA, Giraud I, Pleydell DRJ, Peccoud J. 2021 Characterization of a sex-
711 determining region and its genomic context via statistical estimates of haplotype
712 frequencies in daughters and sons sequenced in pools. *Gen. Biol. Evol.* **13**, evab121.
713 (doi:10.1093/gbe/evab121)

714 67. Cingolani P, Platts A, Wang LL, Coon M, Nguyen T, Wang L, Land SJ, Lu X, Ruden DM.
715 2012 A program for annotating and predicting the effects of single nucleotide
716 polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118;
717 iso-2; iso-3. *Fly* **6**, 80–92. (doi:10.4161/fly.19695)

718 68. Paysan-Lafosse T *et al.* 2023 InterPro in 2022. *Nucleic Acids Res.* **51**, D418–D427.
719 (doi:10.1093/nar/gkac993)

720 69. Therneau TM, Lumley TL, Elizabeth A, Cynthia C. 2024 survival: Survival Analysis.

721 70. Crawley MJ. 2012 *The R Book*. Chichester: John Wiley & Sons,

722 71. Bolker BM. 2008 *Ecological Models and Data in R*. Princeton, NJ: Princeton University
723 Press.

724 72. Cerveau N, Leclercq S, Leroy E, Bouchon D, Cordaux R. 2011 Short- and long-term
725 evolutionary dynamics of bacterial insertion sequences: Insights from *Wolbachia*
726 endosymbionts. *Gen. Biol. Evol.* **3**, 1175–1186. (doi:10.1093/gbe/evr096)

727 73. Le Clec'h W, Raimond M, Guillot S, Bouchon D, Sicard M. 2013 Horizontal transfers of
728 feminizing versus non-feminizing *Wolbachia* strains: from harmless passengers to
729 pathogens. *Environ. Microbiol.* **15**, 2922–2936. (doi:10.1111/1462-2920.12172)

730 74. Keiblinger KM *et al.* 2010 The effect of resource quantity and resource stoichiometry on
731 microbial carbon-use-efficiency. *FEMS Microbiol. Ecol.* **73**, 430–440. (doi:10.1111/j.1574-
732 6941.2010.00912.x)

733 75. Liu M, Durfee T, Cabrera JE, Zhao K, Jin DJ, Blattner FR. 2005 Global transcriptional
734 programs reveal a carbon source foraging strategy by *Escherichia coli*. *J. Biol. Chem.* **280**,
735 15921–15927. (doi:10.1074/jbc.M414050200)

736 76. Alizon S, Hurford A, Mideo N, Baalen MV. 2009 Virulence evolution and the trade-off
737 hypothesis: history, current state of affairs and the future. *J. Evol. Biol.* **22**, 245–259.
738 (doi:10.1111/j.1420-9101.2008.01658.x)

739 77. Frank SA. 1996 Models of parasite virulence. *Q. Rev. Biol.* **71**, 37–78.
740 (doi:10.1086/419267)

741 78. Kosmidis S, Missirlis F, Botella JA, Schneuwly S, Rouault TA, Skoulakis EMC. 2014
742 Behavioral decline and premature lethality upon pan-neuronal ferritin overexpression in
743 *Drosophila* infected with a virulent form of *Wolbachia*. *Front. Pharmacol.* **5**, 66.
744 (doi:10.3389/fphar.2014.00066)

745 79. Min KT, Benzer S. 1997 *Wolbachia*, normally a symbiont of *Drosophila*, can be virulent,
746 causing degeneration and early death. *Proc. Natl. Acad. Sci.* **94**, 10792–10796.
747 (doi:10.1073/pnas.94.20.10792)

748 80. Chevalier F, Herbinière-Gaboreau J, Bertaux J, Raimond M, Morel F, Bouchon D, Grève P,
749 Braquart-Varnier C. 2011 The immune cellular effectors of terrestrial isopod

750 *Armadillidium vulgare*: Meeting with their invaders, *Wolbachia*. *PLoS ONE* **6**, e18531.
751 (doi:10.1371/journal.pone.0018531)

752 81. Sanaei E, Charlat S, Engelstädter J. 2021 *Wolbachia* host shifts: routes, mechanisms,
753 constraints and evolutionary consequences. *Biol. Rev.* **96**, 433–453.
754 (doi:10.1111/brv.12663)

755 82. Rigaud T, Juchault P. 1995 Success and failure of horizontal transfers of feminizing
756 *Wolbachia* endosymbionts in woodlice. *J. Evol. Biol.* **8**, 249–255. (doi:10.1046/j.1420-
757 9101.1995.8020249.x)

758 83. Dainty KR *et al.* 2021 *w*Mel *Wolbachia* genome remains stable after 7 years in Australian
759 *Aedes aegypti* field populations. *Microb. Genom.* **7**, 000641.
760 (doi:10.1099/mgen.0.000641)

761 84. Huang B, Yang Q, Hoffmann AA, Ritchie SA, Hurk AF van den, Warrilow D. 2020
762 *Wolbachia* genome stability and mtDNA variants in *Aedes aegypti* field populations eight
763 years after release. *iScience* **23**. (doi:10.1016/j.isci.2020.101572)

764 85. Ross PA *et al.* 2022 A decade of stability for *w*Mel *Wolbachia* in natural *Aedes aegypti*
765 populations. *PLOS Path.* **18**, e1010256. (doi:10.1371/journal.ppat.1010256)

766 86. Trouche B, Schrieke H, Duron O, Eren AM, Reveillaud J. 2024 *Wolbachia* populations
767 across organs of individual *Culex pipiens*: highly conserved intra-individual core
768 pangenome with inter-individual polymorphisms. *ISME Comm.* , ycae078.
769 (doi:10.1093/ismeco/ycae078)

770 87. Kaur R, Siozios S, Miller WJ, Rota-Stabelli O. 2017 Insertion sequence polymorphism and
771 genomic rearrangements uncover hidden *Wolbachia* diversity in *Drosophila suzukii* and
772 *D. subpulchrella*. *Sci. Rep.* **7**, 14815. (doi:10.1038/s41598-017-13808-z)

773 88. Vancaester E, Blaxter M. 2023 Phylogenomic analysis of *Wolbachia* genomes from the
774 Darwin Tree of Life biodiversity genomics project. *PLOS Biol.* **21**, e3001972.
775 (doi:10.1371/journal.pbio.3001972)

776 89. Vostokova EV, Dranenko NO, Gelfand MS, Bochkareva OO. 2023 Genome
777 rearrangements drive evolution of ANK genes in *Wolbachia. BioRxiv*. **10**.25.563763.
778 (doi:10.1101/2023.10.25.563763)

Table 1. Nucleotide differences between *Wolbachia* sub-populations from three tissues of naturally infected *Armadillidium vulgare*. HAEM: haemolymph, NC: nerve chain, OVA: ovaries. Nucl: Nucleotide, AA: amino acid; Ser: serine, Arg: arginine, Gln: glutamine, Ile: isoleucine, Thr: threonine, fs: Frameshift.

Genome position	Individual	Tissue	Frequency	Type of change	Nucl. change	AA change	ORF effect	Gene name	Functional domains
1266439	F1_2015	NC	0	Insertion	T -> TT	Ser419fs	Frameshift	ABOJ95_001272	No recognized functional domain
		OVA	0						
	F2_2015	NC	0						
		OVA	0						
	F_1999	HAEM	0.112						
		NC	0.006						
		OVA	0.006						
1269492	F1_2015	NC	0	Substitution	G -> A	Arg340Gln	Missense variant	ABOJ95_001275	No recognized functional domain
		OVA	0						
	F2_2015	NC	0						
		OVA	0						
	F_1999	HAEM	0.466						
		NC	0.584						
		OVA	0.239						
1536425	F1_2015	NC	0.087	Substitution	T -> C	Ile1264Thr	Missense variant	ABOJ95_001539	Domain of unknown function DUF1601 Ankyrin repeats Latrotoxin, C-terminal domain
		OVA	0.111						
	F2_2015	NC	0.125						
		OVA	0.110						
	F_1999	HAEM	0.143						

		NC	0.116						
		OVA	0.104						