Wolbachia, *Cardinium* and climate: an analysis of global data

15 July 2019

Charlesworth, J.[1,3*], Weinert, L. A.[2], Araujo-Jnr., E. V.[1], Welch, J. J.[1]

1. Department of Genetics. University of Cambridge. Downing Street, Cambridge, CB23EH, UK.
2. Department of Veterinary Medicine. University of Cambridge. Madingley Road, Cambridge CB30ES, UK.
3. Current address: Warwick Medical School, Warwick, University of Warwick, Gibbet Hill Road Coventry CV47AL, UK

* Author for correspondence: janepipistrelle@googlemail.com**.**

2486 words

Abstract

Bacterial endosymbionts are very common in terrestrial arthropods, but infection levels vary widely among populations. Experiments and within-species comparisons suggest that environmental temperature might be important in explaining this variation. To investigate the importance of temperature, at broad geographical and taxonomic scales, we extended a global database of terrestrial arthropods screened for *Wolbachia* and *Cardinium*. Our final dataset contained data from >117,000 arthropods (>2,500 species) screened for *Wolbachia* and >18,000 arthropods (>800 species) screened for *Cardinium*, including samples from 137 different countries, with mean temperatures varying from -6.5 to 29.2°C. In insects and relatives, *Cardinium* infection showed a clear and consistent tendency to increase with temperature. For *Wolbachia*, a tendency to increase with temperature in temperate climates, is counteracted by reduced prevalence in the tropics, resulting in a weak negative trend overall. We discuss the implications of these results for natural and introduced symbionts, in regions affected by climate change.

Keywords

endosymbiosis, biogeography, Köppen climates, Beta-binomial modeling

Introduction

Bacterial endosymbionts can exert profound effects on their hosts, for example by manipulating host reproductive biology to maximise their own vertical transmission (1,2,3). Infection with such symbionts is very widespread among arthropods, with the genera *Wolbachia* and *Cardinium,* having been estimated to infect, respectively, around 50% and 12.5% of all terrestrial arthropod species (3). Nevertheless, symbiont prevalence (i.e., the proportion of infected individuals within a population) varies widely, and the reasons for this remain unknown. Several factors may shape variation in prevalence, including costs of reproductive parasitism (4), benefits of protection against viral pathogens (5), or host dispersal patterns (6). Another putative influence on endosymbiont prevalence is environmental temperature (7). Laboratory studies suggest that endosymbionts are more susceptible to thermal stress than their hosts (8,9,10), and the physiological costs or benefits of endosymbiont carriage or transmission might be temperature-sensitive (7). Positive temperature clines have been observed for *Cardinium* infection in *Culicoides* midges in Israel (11), for the global prevalence of *Wolbachia* infection in *Lepidoptera*, and across tropical/temperate gradients in Australian *Diptera* (12,13,14). However, these are isolated results, and the prevalence of *Wolbachia* in insect populations, for example, remains remarkably constant across broad continental scales (15).

Weinert *et al.* (3) collated a database of published PCR screens for *Cardinium* and *Wolbachia* in wild populations of terrestrial athropods and developed a maximum-likelihood modelling framework that they used to test hypotheses about endosymbiont incidence in different arthropod groups (3). Here, we greatly extended their database, and modified their methods to test whether temperature clines for prevalence of *Wolbachia* and *Cardinium* obtain more generally. *Wolbachia* is of special interest because of its use as a bio-control agent for mosquito disease vectors (5,6). *Cardinium* provides an interesting contrast because it adopts many of the same transmission strategies as *Wolbachia*, and is also known to infect both pests (16) and disease vectors (11,17) but is present at a much lower incidence.

Results

We extended the database of Weinert *et al.* (3) to yield 135,876 arthropods screened for *Wolbachia* and *Cardinium,* drawn from a total of 320 publications. When sampling location was specified, we obtained an estimate of mean temperature between 1970-2000 (18), and assigned each sample to a Köppen climate zone, which summarizes multiple ecologically relevant variables (19,20). The data are diverse taxonomically (with hosts from >40 arthropod orders), and geographically (with 27/31 Köppen climates, including all higher-level zones: tropical, arid, temperate, continental, and polar). The database is available as Supplementary Table S1, and is summarized in Figures 1 and S1-S3.

We first consider *Cardinium* (Figure 1a-e). This symbiont has very different prevalences in the two major arthropod groups, namely Chelicerata: mites, ticks, spiders and relatives, and Mandibulata: mainly insects, but also, wingless hexapods, crustaceans and myriapods (21,3), so we consider these two groups separately.

For *Cardinium* infection in Mandibulata (Figure 1b-c; Table 1a), there is a clear trend for increasing infection with temperature (Figure 1c). This is confirmed by a regression analysis

71   (Figure 1b; Table 1a). Temperature remained a significant predictor of prevalence when we
72   allowed for an effect of climatic zone, by estimating a typical prevalence level for each of the
73   five higher-level Köppen climates ("K5" in Table 1a), and when we allowed for systematic
74   differences between all finer-grained Köppen climates ("K31" in Table 1a). Furthermore, the
75   effects of temperature increased when we removed the subset of populations with imprecisely
76   specified locations, where the temperature estimates are least certain (see Supplementary
77   Information: Appendix 1 and Table S4a). The improvement in fit due to the Köppen
78   categories, suggest that elements of climate, other than temperature, might predict
79   prevalence. This was confirmed by a permutation test, which showed that much smaller
80   improvements were seen in equally large models, but with the Köppen climate labels
81   permuted randomly among the sampled populations (permutation p-value $< 10^{-4}$; see
82   Supplementary Appendix 1).

84   A major caveat to these results is the highly unrepresentative taxonomy in our database (see
85   Supplementary Figure S1 and Supplementary Table S2). For example, over half of the
86   individuals sampled (6127/11, 755), and three quarters of those infected (1596/2070) are
87   Hemiptera (true bugs). Nevertheless, the effect of temperature remained when we removed
88   all Hemiptera from the dataset, and when we considered Hemiptera alone (Supplementary
89   Table S2). Furthermore, the same trend (albeit non-significant) is evident in the two best-
90   sampled hemipteran groups, namely sternorrhynca (aphids, whiteflies and relatives), and
91   fulgoromorpha (planthoppers). Together, then, our results suggest that both temperature, and
92   other features of climate have predictable effects on levels of *Cardinium* infection in the
93   Mandibulata.

95   For *Wolbachia* in Mandibulata (Figure 1h-g; Table 1b), the results are quite different. Here,
96   there is a significant tendency for colder climates to harbor higher prevalence infections
97   (Table 1b). However, this effect is very small (an increase in temperature from 0° to 10° only
98   decreases expected mean prevalence from 42% to 40%), and not robust to removing
99   populations with imprecisely reported locations (Supplementary Table S4). Nevertheless,
100   similar effects appear in multiple taxonomic groups. As shown in Supplementary Table S2,
101   and Supplementary Figure S4, a negative effect of temperature is found in 4/6 of highly-
102   sampled insect orders: Coleoptera, Hymenoptera, Hemiptera and Orthoptera, and in pooled
103   data from the remaining, sparsely-sampled groups. The same trend was also seen in
104   mosquitoes (Diptera: Culicidae), and in the remainder of the Diptera, though not in the well-
105   sampled Lepidoptera.

107   As shown in Figure 2, this effect is driven by differences between the major climatic zones,
108   with several groups showing less infection in tropical regions, and more infection in cold,
109   continental regions (see also Figure S9). If we model this effect, assigning a typical
110   prevalence level to each of the five climates ("K5" in Table 1b), then the effect of
111   temperature reverses sign, and it becomes a significantly positive predictor of prevalence.
112   Finer-grained analyses show that this is driven by a strong effect of temperature within the
113   best-sampled "temperate" zones, with no consistent pattern elsewhere (Supplementary Tables
114   S3, S4, and Figure S5). The effect also disappears if we allow for systematic differences
115   between the finer-grained Köppen climates ("K31" in Table 1b); in this case, model fit
116   improves substantially, and more so than when climatic labels are randomly permuted
117   (permutation $p < 10^{-4}$), but including temperature as an explanatory variable adds little
118   predictive power, and fit is very similar for the two largest models.

120   Chelicerates are much less well sampled than insects, but this host group showed little
121   evidence of a consistent effect of temperature (Figure 1d-e; and h-i; Table 1c-d). Indeed, for
122   *Wolbachia* there is a notable lack of variation in the mean prevalence estimates. We find
123   meaningful variation between typical prevalences in the finer-grained climatic zones
124   (permutation $p<10^{-4}$), but the best-fit slope for temperature is close to zero (Table 1d;
125   Supplementary Table S4d), and confidence intervals overlap for most pairs of higher-level
126   climates (see Supplementary Figure S6). This lack of a clear trend holds for the well-sampled
127   Acari (mites, ticks and relatives), which includes many disease vectors, and for the remainder
128   of the chelicerate data (Supplementary Table S2).
129
130
131
132   **Discussion**
133
134   Climate and temperature might be associated with symbioses in multiple ways, including
135   direct effects of temperature on the symbiont (7,8,9) or effects on host density or diversity
136   that might affect symbiont transmission (10,32,33). A step towards understanding these
137   processes is to determine the patterns of variation that appear in nature. To this end, we have
138   compiled a global database of arthropod populations screened for the *Cardinium* and
139   *Wolbachia*, and tested for an effect of climate, and long-term environmental temperature, on
140   infection prevalence. Our data were extensive, but haphazardly sampled, and so we have
141   focused on patterns we found consistently in different subsets of the data.
142
143   A clear and consistent positive relationship between temperature and prevalence was found
144   for *Cardinium* infection in Mandibulata hosts, particularly in insects. This suggests that the
145   pattern reported for *Culicoides* midges from Israel (11), obtains much more generally. While
146   the overall incidence of *Cardinium* in insects is relatively low (3), infection is common in
147   Hemiptera (15), including the rice pest *Sogatella furcifera*, where *Cardinium* appears to
148   increase host fitness (33). Our results may therefore be of relevance to pest control efforts.
149
150   For infection with *Wolbachia* – where incidence levels are generally higher – we observed
151   more complex patterns. In Chelicerate hosts, no consistent trends were found, and infection
152   levels are surprisingly constant across very different geographical regions, and host groups.
153   In insects, by contrast, we found a tendency for infection to increase with temperature, but
154   only within temperate climates; and tropical climates, tended to have lower infection levels
155   overall. This last finding contrasts with previous studies of *Wolbachia* in single species or
156   families of Diptera (13,14), but we observed the trend in multiple host groups, including
157   mosquitoes, the major dipteran disease vectors.
158
159   The results for *Wolbachia* have a special interest, because of efforts to use this symbiont as a
160   biocontrol agent, particularly for mosquito-borne human pathogens. For example, in *Aedes*
161   *aegypti*, *Wolbachia* infection inhibits the replication of dengue, chikungunya and zika
162   viruses, as well as malarial parasites (5,22,6). The use of *Wolbachia* for biocontrol has
163   focused on tropical regions (in Australia, Brazil, Indonesia and Vietnam), where the problems
164   of mosquito-borne disease are most acute. But the trends we find within temperate zones
165   suggest that changes in temperature can have consistent effects on symbioses in regions
166   where, because of climate change, such diseases are re-emerging (22,23). Experimental
167   studies have shown that increases in temperature can perturb mosquito-*Wolbachia* symbioses
168   in the laboratory (24,25), and our results suggest these lab results could have a strong
169   ecological relevance.

170
171
172
173 **Methods**
174
175 We corrected and extended an existing database of arthropod screens (3), adding data on
176 sampling location. Full details are given in Supplementary Methods 1.
177
178 The main statistical analyses used standard Betabinomial modeling (3,26). Here, the number
179 of infected individuals in a given population sample was treated a binomially-distributed
180 random variable, parameterised with the true prevalence for that population (i.e., the
181 proportion of the population infected). We then assumed that the true prevalence for each
182 population was drawn from a Beta distribution, parameterized with an overall mean
183 prevalence, and a correlation parameter, describing how the variation in infection is
184 distributed within versus between populations (3). To test for effects of climatic zone and
185 temperature, the logit transformed mean prevalence was constrained to the linear model. For
186 example, we used

187 $$E\left(\log(\mu_i/(1-\mu_i)) \mid t_i, K_i\right) = \beta_0 + \beta_{K_i} + \beta_t t_i$$

188 where μ_i, is the mean prevalence of population i, t_i is the temperature estimate for its
189 sampling location, and K_i is a categorical variable representing its Köppen climate. We held
190 the correlation parameter of the Beta distribution constant across all populations, although
191 relaxing this assumption had little qualitative effect on results (Supplementary Table S2 and
192 Supplementary Figures S8-S11). This agrees with claims that Beta-binomial models yield
193 robust estimates of the mean prevalence, even if the shape of the distribution is misspecified
194 (3). All models were fit using the *vglm* function of the R package VGAM v1.0-3 (27). Best-
195 fit slopes for temperature, β_t, are reported in Table 1 and, the best-fit coefficients for the
196 climatic zones, β_{K_i}, in Supplementary Figures S8-S11. Symbiont infection is a highly
197 dynamic process, and so no set of predictors is likely to explain much of the variation (3).
198 Nevertheless, we also calculated a pseudo-r^2 ($1 - \ln\hat{L}/\ln\hat{L}_{null}$) (30), and the Akaike
199 Information Criterion ($-2\ln\hat{L} + 2n$), which is lower for preferred models (31). Our major aim
200 was to test for an effect of temperature, and for this purpose, we used two-tailed z-tests,
201 defining significance as $p<0.005$ (28). For the permutation tests, we recalculated the
202 maximized log likelihood after randomly permutating the Köppen classifications among
203 populations (following permutation procedures described in (3)). The p-value is the
204 proportion of 10,000 random permutations where the maximized log-likelihood is at least as
205 high as with the true classifications. Low p-values suggest that the climatic zones are adding
206 predictive power, beyond what would be expected from adding additional parameters to the
207 model.
208
209 For the illustrative plots (Figure 1c-e and h-i; Supplementary Figure S4), we binned
210 populations by mean temperature, centring the bins on the mean temperatures recorded in our
211 data set for the five higher-level Köppen climates (Supplementary Figure S3). A distinct
212 Beta-binomial model was then fit within each bin, with confidence intervals defined as mean
213 prevalence values that reduce the maximized log likelihood by two units (29). The same
214 approach was taken for estimating prevalence in each higher-level climatic zone (see Figures
215 3 and Supplementary S6).
216
217
218

References

1. Zchori-Fein, E. and K. Bourtzis (eds). 2011. Manipulative tenants - Bacteria associated with arthropods (CRC press).
2. Duron O, Hurst GD 2013. Arthropods and inherited bacteria: from counting the symbionts to understanding how symbionts count. *BMC Biology* 11:45. doi:10.1186/1741-7007-11-45.
3. Weinert LA, Araujo-Jnr EV, Ahmed MZ, Welch JJ (2015). The incidence of bacterial endosymbionts in terrestrial arthropods. Proc Biol Sci. 282.
4. Hurst GD, Frost CL. Reproductive parasitism: maternally inherited symbionts in a biparental world. Cold Spring Harb Perspect Biol. 2015;7(5):a017699.
5. Moreira LA, Iturbe-Ormaetxe I, Jeffery JA, Lu G, Pyke AT, Hedges LM, Rocha BC, Hall-Mendelin S, Day A, Riegler M, Hugo LE, Johnson KN, Kay BH, McGraw EA, van den Hurk AF, Ryan PA, O'Neill SL. (2009). A *Wolbachia* symbiont in *Aedes aegypti* limits infection with dengue, Chikungunya, and Plasmodium. Cell. 139:1268-78.
6. Jiggins FM 2017. The spread of Wolbachia through mosquito populations. PLoS Biol 15(6): e2002780.
7. Corbin C, Heyworth ER, Ferrari J, Hurst GD (2017). Heritable symbionts in a world of varying temperature. Heredity 118:10-20.
8. Wernegreen JJ. Mutualism meltdown in insects: Bacteria constrain thermal adaptation. Current opinion in microbiology. 15(3):255-262. doi:10.1016/j.mib.2012.02.001.
9. Kikuchi Y, Tada A, Musolin DL, Hari N, Hosokawa T, Fujisaki K, Fukatsu T. 2016. Collapse of Insect Gut Symbiosis under Simulated Climate Change. mBio. 7:e01578-16. doi:10.1128/mBio.01578-16.
10. Hussain, M., Akutse, K. S., Ravindran, K., Lin, Y., Bamisile, B. S., Qasim, M., Dash, C. K. and Wang, L. 2017, Effects of different temperature regimes on survival of *Diaphorina citri* and its endosymbiotic bacterial communities. Environ Microbiol, 19: 3439–3449.
11. Morag N, Klement E, Saroya Y, Lensky I, Gottlieb Y. Prevalence of the symbiont *Cardinium* in *Culicoides* (Diptera: Ceratopogonidae) vector species is associated with land surface temperature. FASEB J. 2012 26:4025-34.
12. Ahmed MZ, Araujo-Jnr EV, Welch JJ, Kawahara AY. 2015. Wolbachia in butterflies and moths: geographic structure in infection frequency. Front Zool. 12:16.
13. Morrow JL, Frommer M, Royer JE, Shearman DC, Riegler M. 2015. Wolbachia pseudogenes and low prevalence infections in tropical but not temperate Australian tephritid fruit flies: manifestations of lateral gene transfer and endosymbiont spillover? BMC Evolutionary Biology 15: 202.
14. Kriesner P, Conner WR, Weeks AR, Turelli M, Hoffmann AA. 2016. Persistence of a Wolbachia infection frequency cline in *Drosophila melanogaster* and the possible role of reproductive dormancy. Evolution 70(5):979-97. doi:10.1111/evo.12923.
15. Werren JH, Windsor DM. *Wolbachia* infection frequencies in insects: evidence of a global equilibrium? Proceedings of the Royal Society B: Biological Sciences. 2000;267(1450):1277-1285.
16. Nakamura Y, Kawai S, Yukuhiro F, Ito S, Gotoh T, Kisimoto R, Yanase T,Matsumoto Y, Kageyama D, Noda H. Prevalence of Cardinium bacteria in planthoppersand spider mites and taxonomic revision of "Candidatus Cardinium hertigii" based on detection of a new Cardinium group from biting midges. Appl Environ Microbiol.2009 Nov;75(21):6757-63.

278 17. Mee PT, Weeks AR, Walker PJ, Hoffmann AA, Duchemin JB. Detection of Low-
279   Level Cardinium and Wolbachia Infections in Culicoides. Appl Environ Microbiol.
280   2015 Sep;81(18):6177-88.
281 18. Fick SE and Hijmans RJ. 2017. WorldClim 2: new 1-km spatial resolution climate
282   surfaces for global land areas. Int. J. Climatol, 37: 4302–4315. doi:10.1002/joc.5086
283 19. Peel, M. C., Finlayson, B. L., and McMahon, T. A. 2007. Updated world map of the
284   Köppen-Geiger climate classification. Hydrol. Earth Syst. Sci., 11, 1633-1644, doi:
285   10.5194/hess-11-1633-2007.
286 20. Chen, D. and H. W. Chen, 2013: Using the Köppen classification to quantify climate
287   variation and change: An example for 1901–2010. Environmental Development, 6,
288   69-79, 10.1016/j.envdev.2013.03.007.
289 21. Martin, O. Y., Goodacre, S. L. (2009) Widespread infections by the bacterial
290   endosymbiont Cardinium in arachnids. J. Arachnol. 37:106–108. doi:10.1636/SH08-
291   05.1
292 22. Jeffries, C. L., and T. Walker 2016. Wolbachia Biocontrol Strategies for Arboviral
293   Diseases and the Potential Influence of Resident *Wolbachia* Strains in Mosquitoes.
294   Curr Trop Med Rep. 2016; 3: 20–25. doi:10.1007/s40475-016-0066-2
295 23. Amon JJ 2016. The impact of climate change and population mobility on neglected
296   tropical disease elimination. International Journal of Infectious Diseases 53:12.
297 24. Yixin H. Ye, Alison M. Carrasco, Yi Dong, Carla M. Sgrò, and McGraw E. A. 2016.
298   The Effect of Temperature on Wolbachia-Mediated Dengue Virus Blocking in Aedes
299   aegypti. Am J Trop Med Hyg. 94(4): 812–819. doi:10.4269/ajtmh.15-0801.
300 25. Ross, P. A., Wiwatanaratanabutr, I., Axford, J. K., White, V. L., Endersby-Harshman,
301   N. M., & Hoffmann, A. A. (2017). Wolbachia Infections in *Aedes aegypti* Differ
302   Markedly in Their Response to Cyclical Heat Stress. PLoS Pathogens, 13(1),
303   e1006006. doi:10.1371/journal.ppat.1006006
304 26. Hilgenboecker K, Hammerstein P, Schlattmann P, Telschow A, Werren JH 2008.
305   How many species are infected with Wolbachia? - a statistical analysis of current
306   data. FEMS Microbiol Lett. 281(2):215–20.
307 27. Yee TW 2015. Vector Generalized Linear and Additive Models: With an
308   implementation in R. New York, USA: Springer.
309 28. Benjamin, D. J., J. Berger, M. Johannesson, B. A Nosek, E.-J. Wagenmakers, R.
310   Berk, K Bollen, et al. 2017. "Redefine Statistical Significance". PsyArXiv. July 22.
311   psyarxiv.com/mky9j.
312 29. Edwards, AWF. 1992. Likelihood: Expanded edition. The Johns Hopkins University
313   Press, Oxford.
314 30. McFadden, D. (1974) "Conditional logit analysis of qualitative choice behavior." Pp.
315   105-142 in P. Zarembka (ed.), Frontiers in Econometrics. Academic Press.
316 31. Akaike, H 1974. A new look at the statistical model identification. IEEE Transactions
317   on Automatic Control, 19: 716–723, doi:10.1109/TAC.1974.1100705
318 32. Dunn RR, Davies TJ, Harris NC, Gavin MC. Global drivers of human pathogen
319   richness and prevalence. *Proc Biol Sci*. 2010;277(1694):2587–2595.
320   doi:10.1098/rspb.2010.0340
321 33. Kajtoch Ł, Kolasa M, Kubisz D, Gutowski JM, Ścibior R, Mazur MA, Holecová M.
322   Using host species traits to understand the Wolbachia infection distribution across
323   terrestrial beetles. Sci Rep. 2019 Jan 29;9(1):847.
324 34. Zhang XF, Zhao DX, Hong XY. *Cardinium*—the Leading Factor of Cytoplasmic
325   Incompatibility in the Planthopper *Sogatella furcifera* Doubly Infected
326   With *Wolbachia* and *Cardinium*, *Environmental Entomology*, 41: 833–840.
327
328

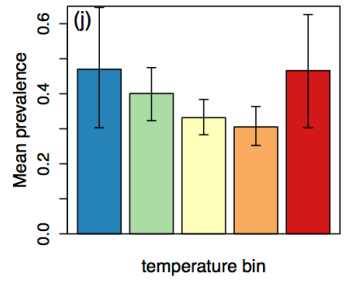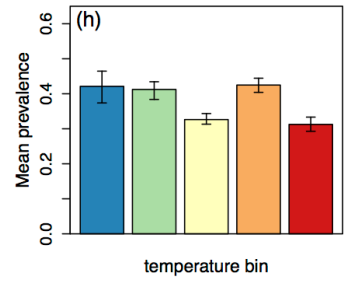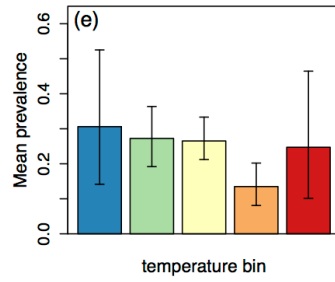329 **Table 1**: The effects of mean temperature and climatic zone on symbiont prevalence
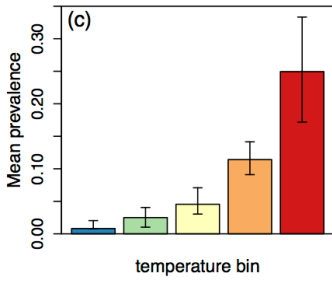
Data set:			KC	n	AIC	pseu-r^2	temp: slope (p val)	
(a) *Cardinium* Mandibulata			-	2	1436.37			
# Populations:	1,374		-	3	1385.92	0.037	0.128	$(7.04 \times 10^{-13}*)$
# Individuals (infected):	11,755 (2070)		K5	6	1434.74	0.007		
# Species min-max:	623-814		K5	7	1359.13	0.061	0.218	$(2.34 \times 10^{-17}*)$
			K31	20	1371.96	0.070		
			K31	21	**1339.39**	0.094	0.198	$(3.26 \times 10^{-9}*)$
(b) *Wolbachia* Mandibulata			-	2	22033.92			
# Populations:	7,986		-	3	22023.29	0.001	-0.011	$(0.000376*)$
# Individuals (infected):	102,267 (47,734)		K5	6	21971.81	0.003		
# Species min-max:	2,451-3,686		K5	7	21953.93	0.004	0.024	$(7.79 \times 10^{-6}*)$
			K31	27	**21741.97**	0.016		
			K31	28	21743.68	0.016	0.004	(0.590)
(c) *Cardinium* Chelicerata			-	2	979.57			
# Populations:	345		-	3	**978.14**	0.004	-0.038	(0.064)
# Individuals (infected):	4,718 (1,212)		K5	6	981.93	0.006		
# Species min-max:	138-157		K5	7	980.95	0.009	-0.042	(0.083)
			K31	21	980.75	0.038		
			K31	22	982.63	0.038	-0.012	(0.720)
(d) *Wolbachia* Chelicerata			-	2	2005.79			
# Populations:	638		-	3	2007.68	0.000	-0.005	(0.746)
# Individuals (infected):	8,413 (2,147)		K5	6	1996.68	0.009		
# Species min-max:	323-350		K5	7	1998.65	0.009	0.003	(0.843)
			K31	20	1949.50	0.046		
			K31	21	**1949.49**	0.047	0.035	(0.147)

330

331 *Note*: KC: climatic zones included in the model as categorical predictors, either K5 (five
332 higher-level Köppen classifications) or K31 (up to 31 finer-grained climates); n: the total
333 number of parameters in the model fit; pseu-r^2: McFadden's pseudo r-squared (30); AIC:
334 Akaike Information Criterion (31), with the preferred model shown in bold; temp: the best-fit
335 slope and p-value associated with the mean temperature, when this was included in the
336 model. * $p < 0.005$. The number of arthropod species in each subset of the data is given as a
337 minimum (counting named species only), and a maximum (including each partially identified
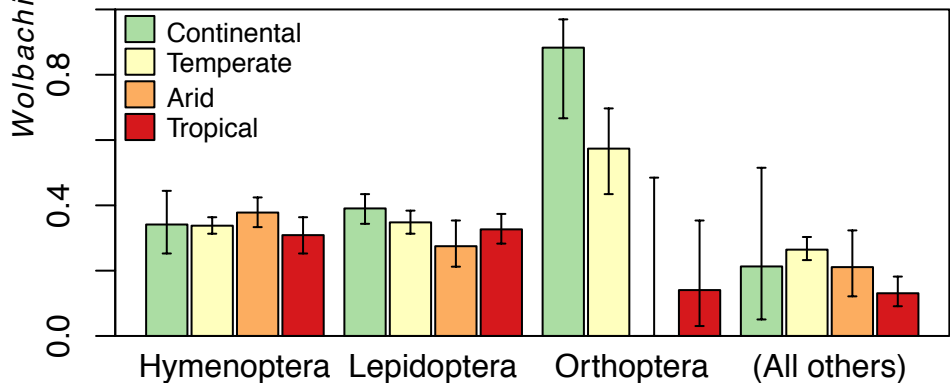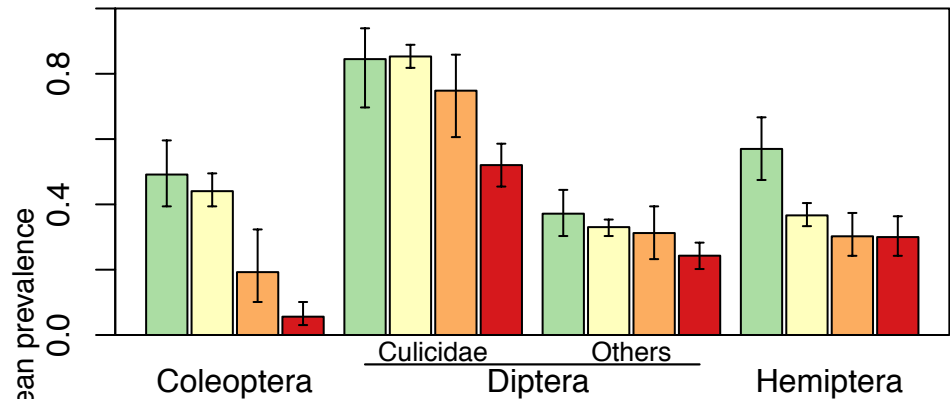338 taxon as a unique species).

339
340

340

341

342

343

344 **Figure 1**: Climatic variation in the prevalence of *Cardinium* (a)–(e) and *Wolbachia* (f)–(i)
345 infection in terrestrial arthropods. Top panel: world maps, with colours corresponding to
346 mean temperature over the period 1970-2000. Points indicate population screens, with point
347 size indicating the number of individuals sampled. Middle: regression of mean prevalence
348 (logit transformed) on the mean temperature of sampling location; the best-fit lines
349 correspond to the three-parameter models reported in Table 1. Bottom: illustrative plots,
350 showing the estimated mean prevalence for populations falling within temperature bins,
351 coloured to match the middle panels, and centred on the mean temperature for the five major
352 Köppen zones: Polar (blue); Continental (green); Temperate (yellow), Arid (orange) and
353 Tropical (red). Confidence intervals correspond to two units of log likelihood. Separate
354 results are given for the two major clades of arthropods, namely Mandibulata (mostly insects,
355 but also wingless hexapods, crustaceans and myriapods), and Chelicerata (ticks, mites,
356 spiders and relatives).

357

(a) Cardinium

(b) Mandibulata

(c)

(d) Chelicerata

(e)

(f) Wolbachia

(g) Mandibulata

(h)

(i) Chelicerata

(j)
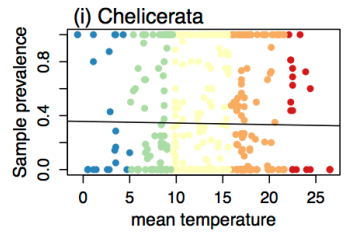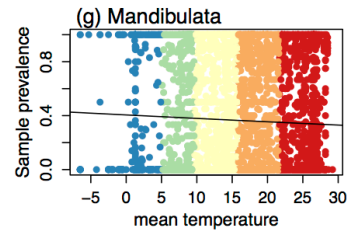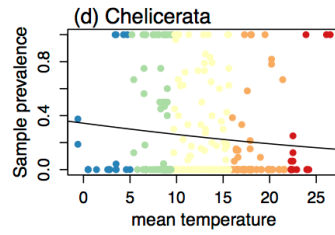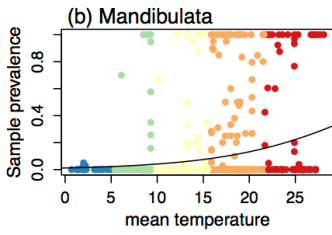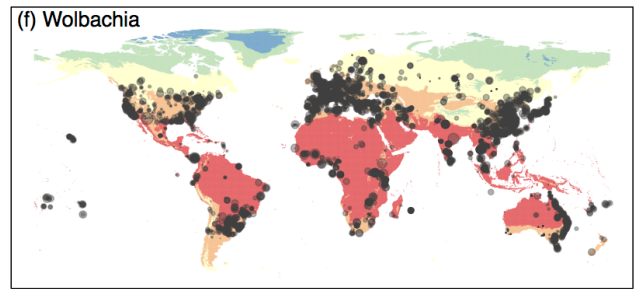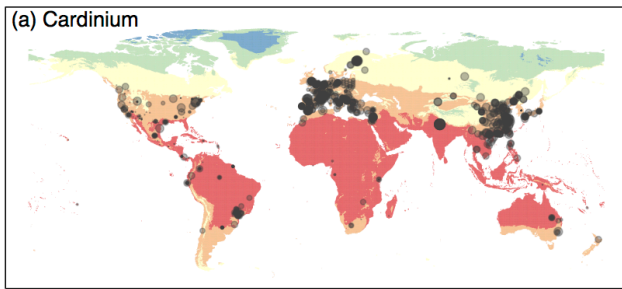
357
358
359   **Figure 2**: Estimated mean prevalence for *Wolbachia* infection in mosquitoes (Culicidae:
360   Diptera), the best-sampled insect orders, and the remainder of the mandibulate arthropods. In
361   each case, separate estimates are shown for populations from the four best-sampled climatic
362   zones, according to the Köppen system. The main database also contains samples from polar
363   climates, but only a few for each taxonomic group. Confidence intervals correspond to two
364   units of log likelihood.
365

Supplementary Information

Appendix 1:
In this Appendix we detail the sampling strategy for the database of terrestrial arthropod screens. Reference numbers refer to references in the main text.

Our starting point was the collation of PCR screens of terrestrial arthropods, for *Cardinium, Rickettsia* and *Wolbachia* from 364 publications prior to 2014, published by Weinert et al (3). From this database, we retained only screens for *Wolbachia* and *Cardinium*, from wild populations. Weinert et al. (3) removed screens from areas where host populations were known *a priori* to harbour infection. Given our different goals, we retained these screens, but removed studies where arthropod individuals were pooled before screening. We re-checked entries for all studies, and reentered data in accordance with the goals of the present study (for example, in several instances, multiple geographically proximate populations that were treated as a single large sample in the previous study, were separated here). This led to several changes in previous entries (~0.09% of the total). We then searched the literature for more recent studies, stopping at March 2018. This yielded data from an additional 125 published papers, increasing the number of *Wolbachia* screens by about a third. To associate each record with climate data, we first obtained the decimal latitude and longitude of each sampling location, either from the source publications, or by contacting the authors (see acknowledgements), or by finding named locations on Google Maps (2017). Where location information was very coarse-grained (e.g. "Texas" or "Scotland") we used two strategies to calculate approximate coordinates. If other studies had sampled from the same location, we used the median coordinates of the sampled data for our unspecified location. If no other studies sampled from a given location, we used coordinates for the centroid of that location. Together, these approximate methods were used for ~15.6% of records, as noted in the database. Our results are robust to removing this subset of the data as shown in Supplementary Table S4.

We then used the latitude and longitude information to obtain climatic data. First, we obtained mean temperature estimates over a 2.5-minute grid (~4.5km at the equator) from the WorldClim database (16). We associated each of our sampling locations with its closest grid entry, by Euclidean distance. We also excluded samples from some island or coastal regions, where WorldClim lacks data. Each of our temperature estimates is the mean of weekly maximum and minimum temperatures, averaged over the years 1970-2000. Where sampling dates of the arthropod populations were reported, these were also recorded in our database. But despite the volume of data, there were insufficient records to test directly for temporal change. For similar reasons, we were unable to include other potentially interesting predictors, such as *Wolbachia* supergroups or co-infection with both *Cardinium* and *Wolbachia* in our models. Any discrepancy between sampling dates and temperature estimates should add noise, but not bias to our results. Estimates of temperature variability, and of the maximum and minimum temperatures during 1970-2000, correlate very strongly with the mean temperatures (Supplementary Figure S7), increasing our confidence that our chosen predictor is ecologically meaningful. For the same reason, we used only the mean temperature in our analyses, and did not add additional, highly correlated predictors. We also classified each sampling location according to the Köppen system, again using Euclidean distance, to associate each sampled point with a grid reference (17, data downloaded from 18). These classifications correlate with the mean temperature estimates, and this increases the stringency of our tests, but as shown in Supplementary Figure S3, most climate types

include a wide range of temperatures, and there is overlap between the temperature ranges of different climatic zones.
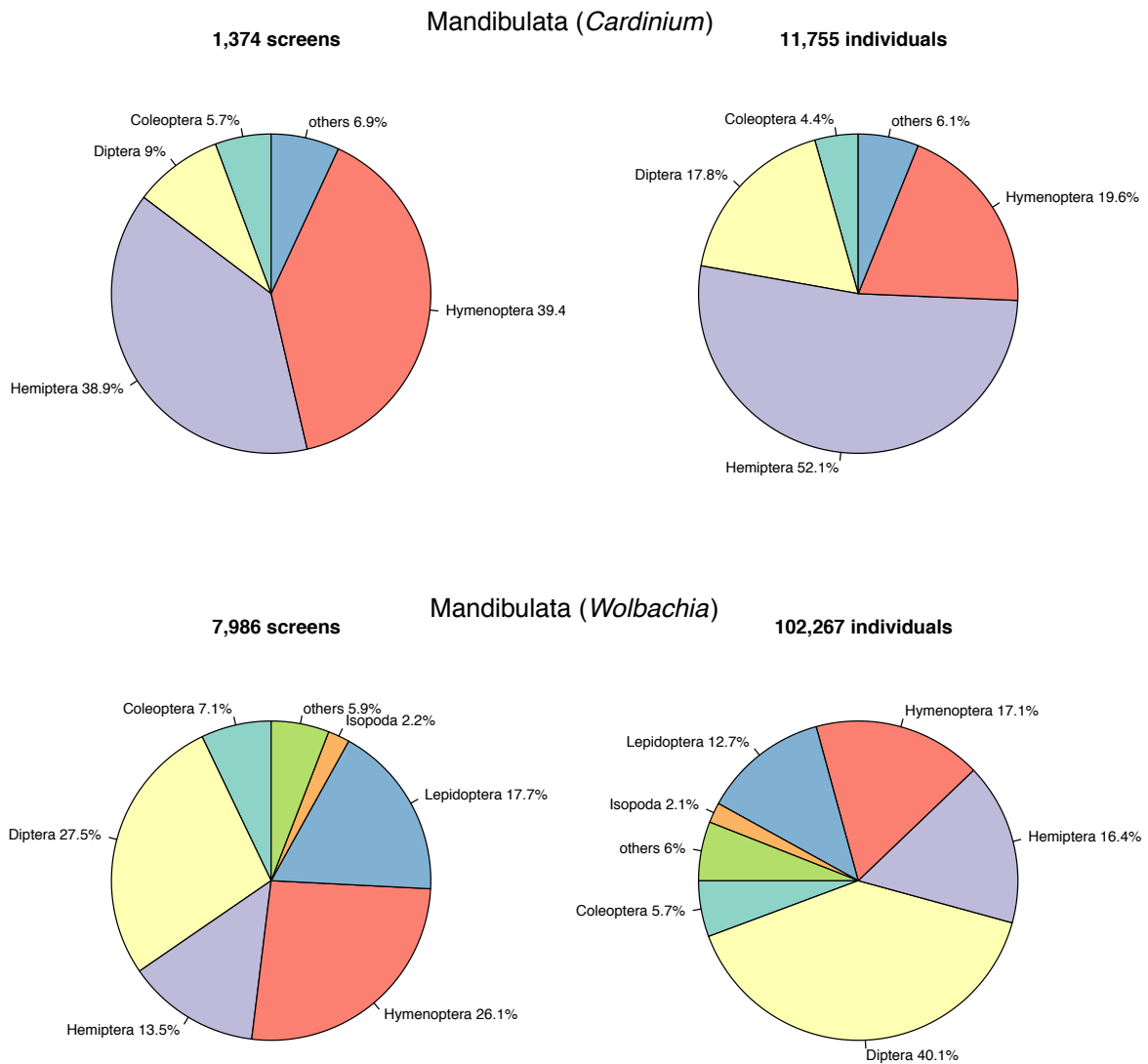
Figure S1:
Taxonomic breakdown of the screen database for arthropod groups classified as Mandibulata, i.e., Hexapoda (including insects), Crustacea and Myriapoda. Plots are shown for the number of screens (i.e., the number of populations sampled), and the total number of individuals screened, excluding pooled screens and samples for which temperature and climate data were not available. Ordinal or superordinal taxa are indicated if they comprise >2% of the database, on either measure. The full database is available as Supplementary Table S1.
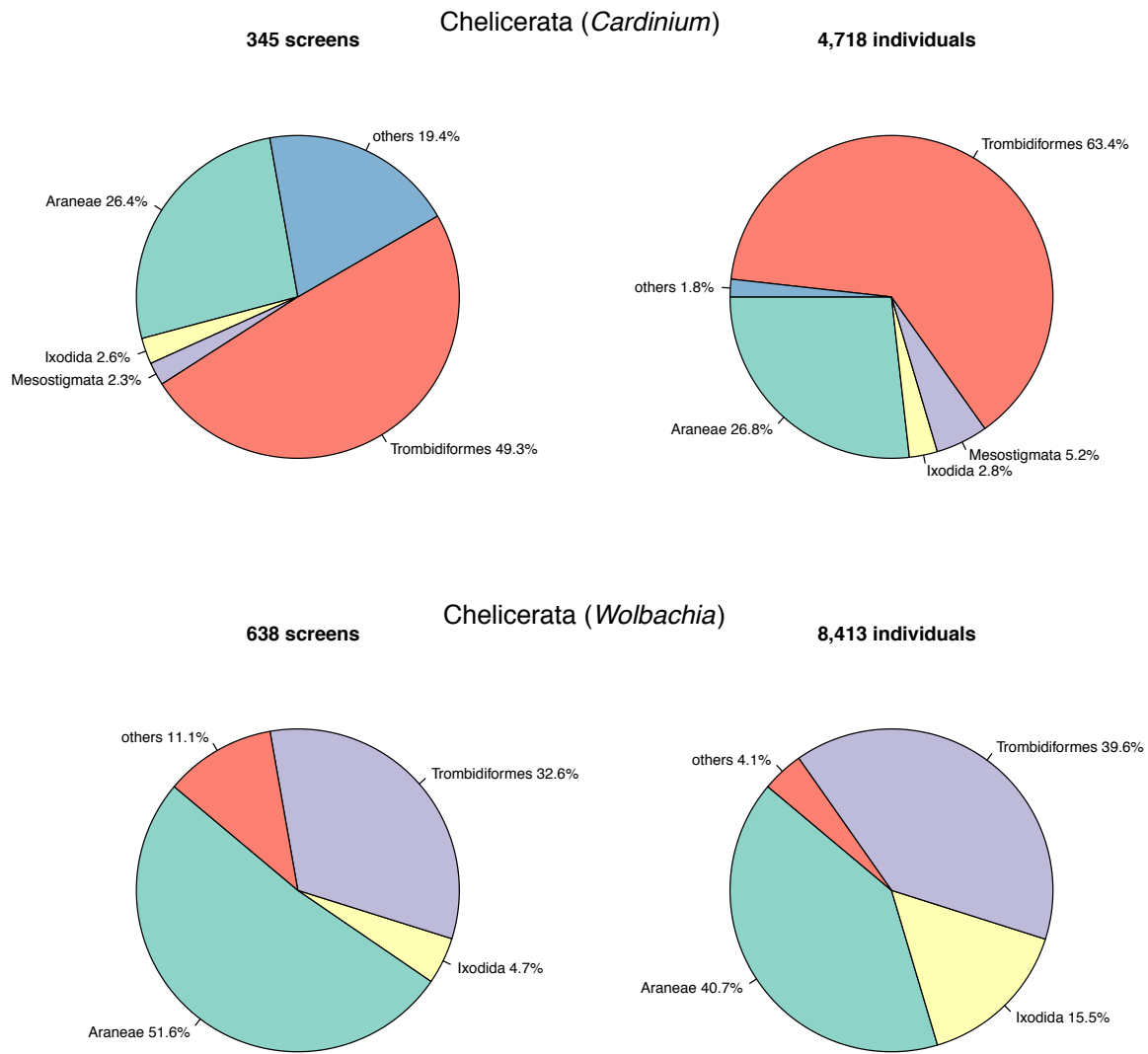
Figure S2:
Taxonomic breakdown of the screen database for arthropod groups classified as Chelicerata, i.e., Acari (ticks and mites), Araneae (spiders), and relatives. All other details match Figure S1.
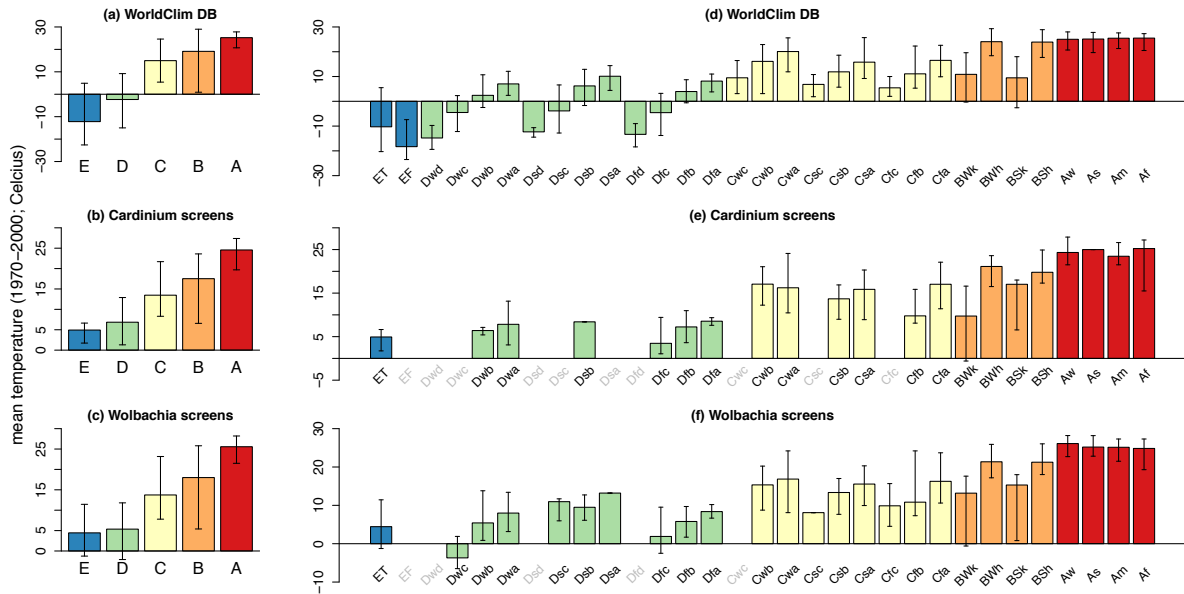
Figure S3:

The range of mean monthly temperatures for the period 1970-2010 associated with each of the Köppen climate classifications. These classifications are fit as a categorical predictor, along with mean temperature itself, in our analyses of symbiont prevalence (see Table 1; Supplementary Table S2). Bars show means, and confidence intervals show 95% quantiles. The upper row (panels (a) and (d)) shows all records in the WorldClim 2 database (Fick and Hijmans 2017). The middle row (panels (b) and (e)) shows the equivalent results for locations in our database of *Cardinium* screens, and the lower row (panels (c) and (f)) shows results for our database of *Wolbachia* screens. The left-hand column (panels (a)-(c)), shows results for the five higher-level climate classifications, namely A (tropical), B (arid), C (temperate), D (continental), and E (polar). The right-hand column (panels (d)-(f)): shows results for the 31 finer-grained climates, namely: Af (tropical rainforest), Am (tropical monsoon), Aw (tropical savanna, wet), As (tropical savanna, dry); BWh (hot desert), BWk (cold desert), BSh (hot steppe), BSk (cold steppe), Csa (hot-summer Mediterranean), Csb (warm-summer Mediterranean), Csc (cool-summer Mediterranean), Cwa (Monsoon-influenced humid subtropical), Cwb (subtropical highland or temperate oceanic climate with dry winters), Cwc (cold subtropical highland climate or subpolar oceanic climate with dry winters), Cfa (humid sub-tropical), Cfb (temperate oceanic), Cfc (subpolar oceanic); Dfa (hot-summer humid continental), Dfb (warm-summer humid continental), Dfc (subarctic), Dfd (extremely cold subarctic), Dwa (Monsoon-influenced hot-summer humid continental), Dwb (Monsoon-influenced warm-summer humid continental climate), Dwc (Monsoon-influenced subarctic climate), Dwd (Monsoon-influenced extremely cold subarctic climate); Dsa (hot, dry-summer continental), Dsb (warm, dry-summer continental), Dsc (sub-arctic, dry-summer continental), Dsd (extremely cold sub-arctic, dry-summer continental); ET (polar tundra), EF (eternal winter/ice cap). Comparing the rows shows that the some of the finer-grained climates were missing from our population sample (indicated by the gray labels), while others were represented by a biased sample with respect to temperature. For example, there were no population samples from EF (eternal winter/ice cap), while samples from ET (polar tundra) come from the warmer parts of this climatic zone.
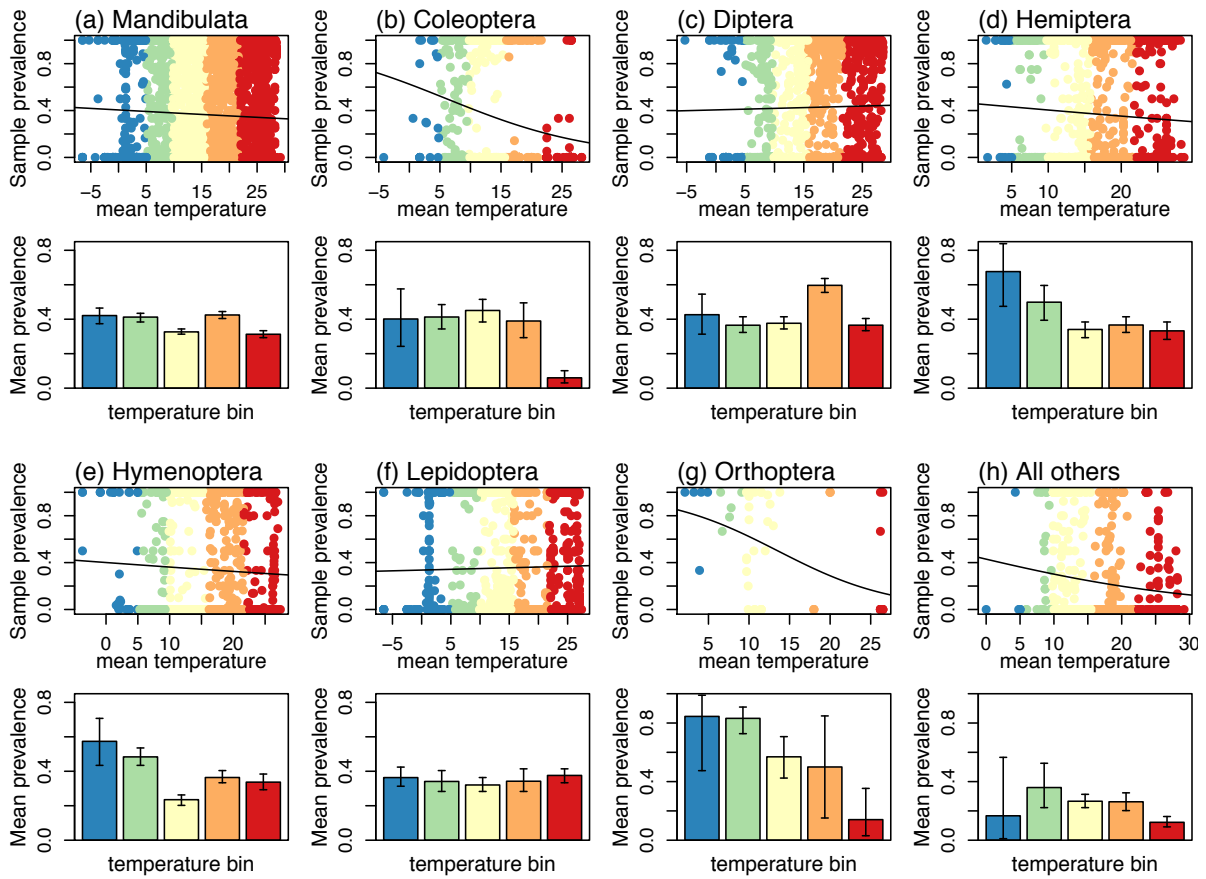
Figure S4:
The effects of temperature on *Wolbachia* prevalence in the major insect orders (panels (b)-(g)), and in the remainder of the sampled Mandibulata (panel (h)). The slopes correspond to results for three-parameter models given in Supplementary Table S2. All other details match Figure 2b-c, which is replicated for ease of comparison in panel (a).
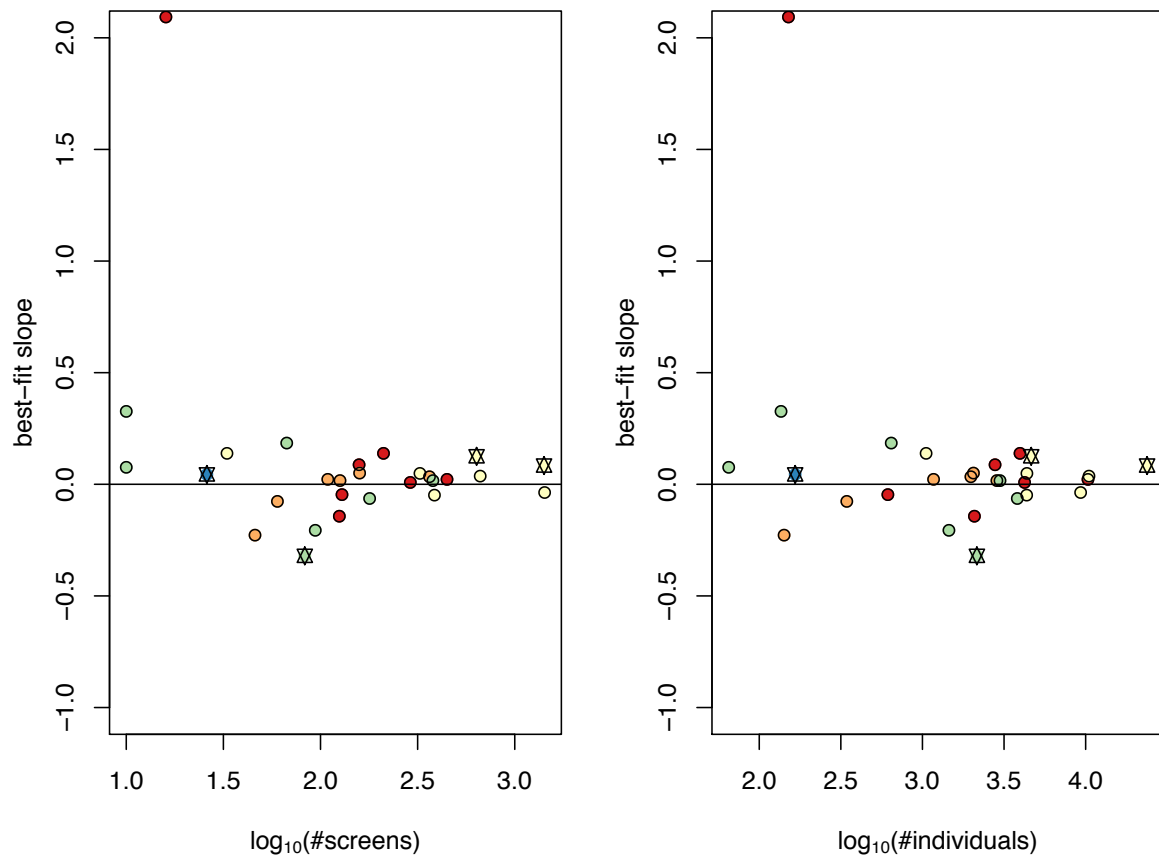
Figure S5:

Best-fit slopes for the regression of mean *Wolbachia* prevalence onto temperature. Each point represents samples from a particular group of mandibulate hosts (one of the major insect orders, or the remainder of the data), and a single higher-level Köppen zone (polar: blue; continental: green, temperate: yellow, arid: orange; and tropical: red). Regressions that were individually significant are indicated as asterisks, and a full description is found in Supplementary Table S3. As expected, data show that the estimates of the slopes become more erratic as sample size decreases, but there is little evidence for a consistent trend towards negative or positive slopes. The apparent trend for *Wolbachia* infection to increase with temperature is driven by the well-sampled Diptera and Lepidoptera in temperate zones, indicated by the yellow asterisks towards the right-hand side of the plots.
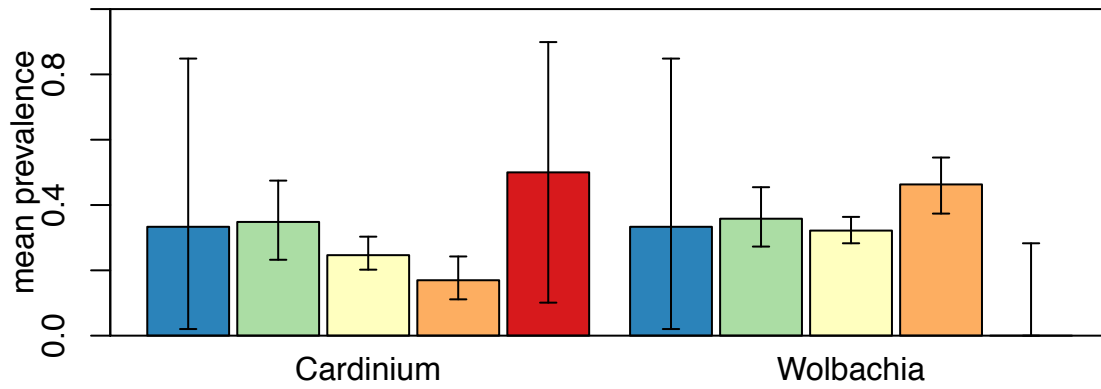
Figure S6:
Best-fit mean prevalence estimates in Chelicerata hosts, for the symbionts *Cardinium* (left-hand bars) and *Wolbachia* (right-hand bars). In each case, estimates are shown for the five higher-level Köppen climates: from left-to-right, these are polar (blue), continental (green), temperate (yellow), arid (orange) and tropical (red).

Correlations between WorldClim bioclimatic temperature variables
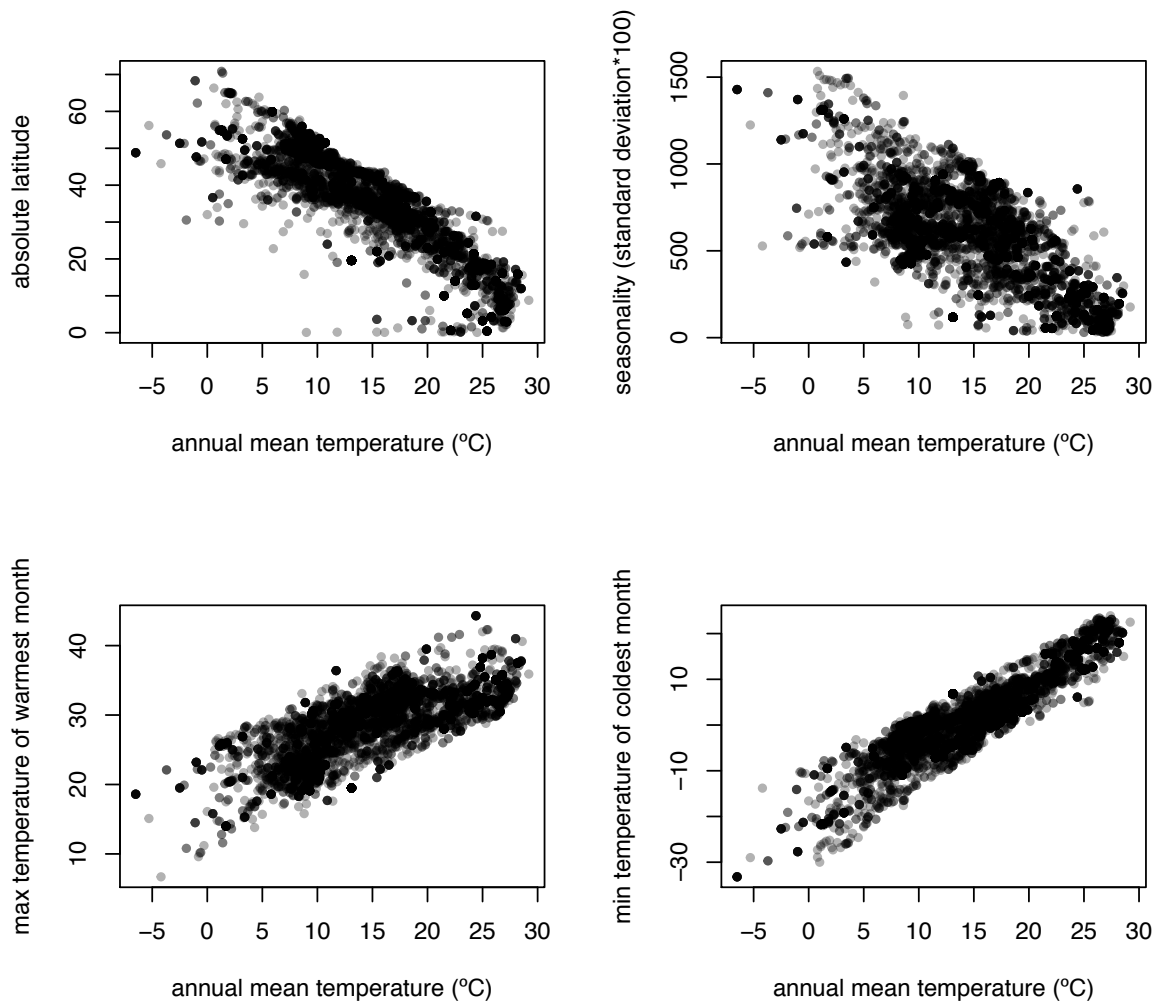
Figure S7:
Estimates of mean temperature show strong correlations with other potential predictors of endosymbiont prevalence. Results are shown for (a) absolute latitude, (b) "seasonality" (defined as the standard deviation in the recorded temperatures over the period 1970-2000), (c) the maximum recorded temperature over this period, and (d) the minimum recorded temperature. All data come from the WorldClim database (Fick and Hijmans 2017).

Figure S8:

Best-fit mean prevalence values for *Cardinium* prevalence in Mandibulata hosts, in each of the Köppen climate zones. The upper row of panels ((a), (c) and (e)) shows estimates when only the climatic zones were included in the model, while the lower row (panels (b), (d) and (f)), show estimates after the effects of mean temperature have been removed. The strong difference between these rows suggests that temperature is a strong predictor of mean prevalence in these data. Results are compared for the five higher-level classifications (panels (a)-(d)), and for the 31 finer-grained classifications (panels (e)-(f)), not all of which were represented in our sample of screens (shown by the gray labels). Results are also compared when the model was fit only to mean prevalence (panels (a)-(b)), and when it was also fit the correlation parameter, allowing the shape of the distribution of prevalences to vary between climate zones (panels (c)-(d)). The definitions of all climatic zones, and the range of temperatures that they encompass, are shown in Supplementary Figure S3.

Figure S9

Best-fit mean prevalence values for *Wolbachia* prevalence in Mandibulata hosts, in each of the Köppen climate zones. The broad similarity between the upper and lower row of panels suggests that mean temperature is not a strong predictor of prevalence for these data. All other details match Supplementary Figure S8.

Figure S10:

Best-fit mean prevalence values for *Cardinium* prevalence in Chelicerata hosts, in each of the Köppen climate zones. The broad similarity between the upper and lower row of panels suggests that mean temperature is not a strong predictor of prevalence for these data. All other details match Supplementary Figure S8.
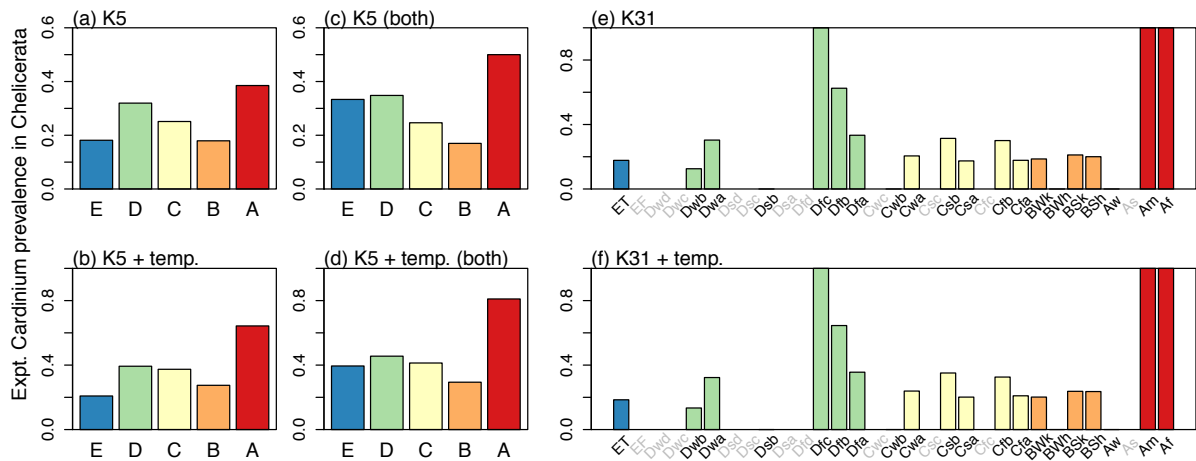
Figure S11:

Best-fit mean prevalence values for *Wolbachia* prevalence in Chelicerata hosts, in each of the Köppen climate zones. The broad similarity between the upper and lower row of panels suggests that mean temperature is not a strong predictor of prevalence for these data. All other details match Supplementary Figure S8.

Description of Supplementary Tables:
Supplementary Table S1: The database of screens of wild arthropod populations either *Cardinium* or *Wolbachia*.

Supplementary Table S2: Results of Beta-binomial model fitting to symbiont screen data.

Supplementary Table S3: Tests for an effect of mean temperature of *Wolbachia* prevalence in Mandibulata hosts, applied separately to populations from each of the five higher-level Köppen climates.

Supplementary Table S4: The effects of mean temperature and climatic zone on symbiont prevalence, excluding imprecisely reported locations (as described in Appendix 1).

Data set:		KC	*np*	AIC	pseu-r^2	temp: slope (*p* val)	
(a) *Cardinium* Mandibulata		-	2	1136.43			
# Populations:	921	-	3	1028.50	0.097	0.237	$(5.47\times10^{-22}*)$
# Individuals (infected):	10,451 (1,985)	K5	5	1111.23	0.028		
		K5	6	1002.71	0.125	0.306	$(1.34\times10^{-22}*)$
		K31	17	1014.18	0.134		
		K31	18	**983.31**	0.163	0.232	$(5.26\times10^{-9}*)$
(b) *Wolbachia* Mandibulata		-	2	19107.00			
# Populations:	6,749	-	3	19109.00	0.000	0.000	0.987
# Individuals (infected):	89,023 (43,890)	K5	6	19085.01	0.002		
		K5	7	19056.85	0.003	0.031	$(3.83\times10^{-8}*)$
		K31	26	18865.14	0.015		
		K31	27	18863.52	0.015	0.014	0.055
(c) *Cardinium* Chelicerata		-	2	929.29			
# Populations:	327	-	3	**924.82**	0.007	-0.054	0.0119
# Individuals (infected):	4,067 (1,119)	K5	6	931.88	0.006		
		K5	7	928.57	0.012	-0.058	0.0223
		K31	21	931.60	0.039		
		K31	22	932.29	0.040	-0.041	0.0240
(d) *Wolbachia* Chelicerata		-	2	1586.08			
# Populations:	459	-	3	1587.49	0.000	-0.012	0.439
# Individuals (infected):	6,378 (1,875)	K5	6	1582.68	0.007		
		K5	7	1584.63	0.007	0.004	0.825
		K31	20	**1553.69**	0.043		
		K31	21	1553.87	0.044	0.035	0.172

Note: KC: climatic zones included in the model as categorical predictors, either K5 (five higher-level Köppen classifications) or K31 (up to 31 finer-grained climates); *np*: the total number of parameters in the model fit; pseu-r^2: McFadden's pseudo *r*-squared (30); AIC: Akaike Information Criterion (31), with the preferred model shown in bold; temp: the best-fit slope and *p*-value associated with the mean temperature, when this was included in the model. * $p < 0.005$. The number of arthropod species in each subset of the data is given as a minimum (counting named species only), and a maximum (including each partially identified taxon as a unique species).