Multi-host *Bartonella* parasites display covert host-specificity even when transmitted by generalist vectors.

Running head: Covert-specificity of *Bartonella* in rodents

Susan M. Withenshaw[*a,b], Godefroy Devevey [c], Amy B. Pedersen [c] and Andy Fenton [a]

[a] Institute of Integrative Biology, University of Liverpool, Liverpool, England.

[b] NERC Centre for Ecology and Hydrology, Crowmarsh Gifford, England.

[c] School of Biology & Centre for Immunity, Infection and Evolution, University of Edinburgh, Edinburgh, Scotland.

[*] Corresponding author: suswit@liverpool.ac.uk

Summary

1. Many parasites infect multiple sympatric host species and there is a general assumption that parasite transmission between co-occurring host species is commonplace. Such between-species transmission could be key to parasite persistence within a disease reservoir and is consequently an emerging focus for disease control.
2. However, while a growing body of theory indicates the potential importance of between-species transmission for parasite persistence, conclusive empirical

evidence from natural communities is lacking, and the assumption that between-species transmission is inevitable may therefore be wrong.

3. We investigated the occurrence of between-species transmission in a well-studied multi-host parasite system. We identified the flea-borne *Bartonella* parasites infecting sympatric populations of *Apodemus sylvaticus* (Linneaus, 1978) (wood mice) and *Myodes glareolus* (Schreber, 1780) (bank voles) in the UK and confirmed that several *Bartonella* species infect both rodent species. However, counter to previous knowledge, genetic characterisation of these parasites revealed covert host-specificity, where each host species is associated with a distinct assemblage of genetic variants, indicating that between-species transmission is rare.
4. Limited between-species transmission could result from rare encounters between one host species and the parasites infecting another and/or host-parasite incompatibility. We investigated the occurrence of such encounter and compatibility barriers by identifying the flea species associated with each rodent host, and the *Bartonella* variants carried by individual fleas. We found that the majority of fleas were host-generalists but the assemblage of *Bartonella* variants in fleas tended to reflect the assemblage of *Bartonella* variants in the host species they were collected from, thus providing evidence of encounter barriers mediated by limited between-species flea transfer. However, we also found several fleas that were carrying variants never found in the host species from which they were collected, indicating some degree of host-pathogen incompatibility when barriers to encounter are overcome.
5. Overall, these findings challenge our default perceptions of multi-host parasite persistence, as they show that despite considerable overlaps in host species

ecology, separate populations of the same parasite species may circulate and persist independently in different sympatric host species. This questions our fundamental understanding of endemic transmission dynamics and the control of infection within natural reservoir communities.

Key-words *Apodemus sylvaticus*, *Bartonella*, fleas, host-generalist, host-specialist, *Myodes glareolus*, pathogen genotypes, rodents, sequencing, vector-borne diseases.

Introduction

Most parasites are able to infect multiple host species (Cleaveland *et al.* 2001; Woolhouse *et al.* 2001); a realisation that has fundamentally changed how we approach issues of disease control. This is because the endemic persistence of such "multi-host" parasites in wild host populations may rely on transmission between individuals of different host species (between-species transmission) as well as, or even instead of, transmission between conspecifics (within-species transmission) (Haydon *et al.* 2002; Holt *et al.* 2003; Dobson 2004; Fenton & Pedersen 2005; Streicker *et al.* 2013; Fenton *et al.* 2015). Consequently, successful control of infection in one host species may require interventions (e.g. vaccination or cullings) that target other species that dominate transmission in the host community (Laurenson *et al*. 2003; Donnelly *et al*. 2006; Serrano *et al*. 2011).

However, while a growing body of theory indicates the potential importance of between-species transmission for endemic multi-host parasite persistence (Holt &

Pickering 1985; Bowers & Begon 1991; Begon *et al*. 1992; Bowers & Turner 1997; Greenman & Hudson 1999; Greenman & Hudson 2000; Haydon *et al*. 2002; Holt *et al*. 2003; Dobson 2004; Fenton & Pedersen 2005; Begon 2008), conclusive empirical evidence from natural communities is often lacking. The occurrence of between-species transmission is often just assumed given that a parasite infects multiple sympatric host species (Dobson & Meagher 1996), or is concluded on the basis of indirect evidence such as correlations between parasite prevalence in one host species and population densities of another (Telfer *et al*. 2007a). However, such correlations may arise as a result of other processes not related to between-species transmission, and therefore the general importance of between-species transmission in endemic parasite persistence in nature remains largely unknown.

The study of parasite genetics in wild communities represents an important means to address this knowledge gap (Streicker *et al*. 2010; Forrester & Hall 2014). Fine-scale genetic characterisation of multi-host parasites may uncover structure within a parasite population that can provide direct evidence of the occurrence of between-species transmission. Intriguingly, of the relatively few studies that have employed such techniques, many have found that sympatric host species are infected with different genetic variants of the same parasite species (Sehgal *et al*. 2006; Whiteman *et al*. 2006; Martinez-Aquino *et al*. 2009). Such “covert host-specificity” indicates that discrete subsets of the same parasite species can circulate independently and persist within populations of sympatric host species with little or no between-species transmission. This fundamentally challenges our default perceptions of endemic multi-host parasite persistence, and it is therefore crucial to determine whether covert host-specificity is a widespread phenomenon.

A lack of transmission between co-occurring host species may result from limited between-species contact opportunities and/or physiological incompatibility between variants and host species ("encounter" and "compatibility" barriers respectively; Combes 2001). Encounter barriers may easily break down if contact rates increase but between-species transmission will remain inhibited if host-parasite incompatibility persists. Identifying the primary drivers of current covert host-specificity (i.e., whether it arises due to current limitations in contact or exposure, or due to current incompatibility between parasite and host) could therefore indicate how stable the host-specificity is, and enable predictions of how rapidly transmission dynamics are likely to change given future alterations to interactions within the host community.

Wild rodent communities are commonly used as model systems in which to study parasite infection and transmission dynamics within natural settings (Begon *et al*. 1999; Telfer *et al*. 2007a,b; Knowles *et al*. 2013; Turner *et al*. 2014), and they have been the focus of much multi-host parasite research (Begon *et al*. 1999; Carslake *et al*. 2006; Streicker *et al*. 2013; Fenton *et al*. 2015). In particular, several species of rodent *Bartonella* are considered model examples of endemic multi-host parasites, as these bacterial flea-borne haemoparasites are commonly found to infect several sympatric rodent species (Birtles *et al*. 2001; Telfer *et al.* 2007a; Paziewska *et al*. 2012). However, previous inferences of between-species *Bartonella* transmission within rodent populations have relied on observed differences in prevalence across different host community compositions (Telfer 2007a) and, importantly, the possibility of covert host-specificity (discrete populations of host-specific variants) has not been directly addressed. Where genetic variation in populations of rodent *Bartonella* has

been described (e.g. Birtles *et al*. 2001; Inoue *et al*. 2008; Berglund *et al.* 2010; Paziewska *et al*. 2011; Kosoy *et al*. 2012), it has largely been compared across broad geographic regions, or interpreted in relation to within-individual and within-species infection dynamics. In contrast, such variation has been rarely discussed in the context of between-species transmission and multi-host parasite persistence (although see Paziewska *et al*. 2012).

The vector-borne nature of rodent *Bartonella* transmission (Bown *et al*. 2004; Morick *et al*. 2010; Gutiérrez *et al*. 2015) allows an assessment of whether any covert host-specificity arises through current encounter barriers to between-species transmission (i.e., limited exposure of one rodent species to fleas from another species), or through host-*Bartonella* incompatibility. Although some rodent fleas are known to display differential host preferences (Khokhlova *et al*. 2012), close overlap between the flea communities of sympatric rodent species has also been demonstrated (Harris *et al*. 2009), and many flea species are documented as being able to infest several host species (Marshall 1981). Even so, host-generalist fleas may still present a barrier to between-species *Bartonella* transmission, as the rate of movement between different host species is likely to depend on the frequency and nature of between-host contacts (Krasnov & Khokhlova 2001) or rate of visitation to another host species' burrow, given that flea dispersal rates are generally low (Marshall 1981; Krasnov 2008). As such we do not currently know the extent to which flea biting behaviour acts as a barrier to between-species parasite transmission.

Through the genetic characterisation of *Bartonella* infections in wild sympatric populations of *Apodemus sylvaticus* Linneaus, 1978 (wood mice) and *Myodes*

glareolus Schreber, 1780 (bank voles) we provide conclusive evidence of covert host-specificity in this well-studied parasite system and therefore highlight that between-species transmission of multi-host parasites is potentially more rare than previously expected. Additionally, through characterising the communities of fleas associated with each host species, and identifying the genetic variants of *Bartonella* carried by individual fleas taken from the different host species, we show that while vectors of multi-host parasites may be generalists, ecological opportunities for vector transfer between different host species may be rare and therefore still represent a major impediment to between-species parasite transmission.

Materials and methods

Field sampling

Wood mice and bank voles were trapped using Sherman live-traps (Alana Ecology, UK; dimensions 8.9cm x 7.6cm x 22.9cm) and monitored longitudinally during 2011 and 2012 at three woodland sites in northwest England: Manor Wood (MW; N 53.3301°, E -3.0516°), Maresfield & Gordale woods (MFG; N 53.2729°, E -3.0615°) and Rode Hall (RH; N 53.1213°, E -2.2798°). When first captured, all rodents were given a sub-cutaneous electronic PIT-tag (AVID MicroChips, UK) enabling individual identification. A small blood sample (~25µL) was taken from the tail tip of each individual at each monthly capture to assess *Bartonella* infection. Blood samples were centrifuged at 12000rpm for 10 minutes to separate blood pellets (containing cells) from sera. Pellets were then frozen at -20°C until further processing (see below). Further details of field methods are given in Appendix S1.

Fleas were collected from rodents at MFG and RH in 2012 and during further field sampling at these sites in 2013 and 2014. Fleas were also collected from rodents at a fourth nearby site, Haddon Wood (HW; N 53.2709°, E -3.0268°; ~1.6 km from MFG and ~52 km from RH) during 2012. Fleas were removed from individuals by brushing the fur over a water bath, then stored individually in 90% ethanol and identified to species using a morphological key (Whitaker 2007). Some rodents were exposed to insecticide treatment as part of a concurrent experiment, but excluding these animals did not qualitatively affect the results obtained (compare Tables S2 and S3) and so data from all animals are presented throughout the main text.

Identification of *Bartonella* DNA in rodents and fleas

DNA was extracted from rodent blood pellets and individual fleas using standard protocols (Appendix S1). *Bartonella* DNA was detected by PCR targeting a partial region of the 16S-23S internal transcribed spacer (hereafter referred to as the pITS region) following standard methodology (Roux & Raoult 1995; Birtles *et al*. 2000; Houpikian & Raoult 2001; Telfer *et al*. 2005; Telfer *et al*. 2007a,b). As a non-coding region of DNA, the pITS region can withstand many point mutations and insertion/deletion events, and varies in length between different species of *Bartonella* (Roux & Raoult 1995; Birtles *et al*. 2000; Houpikian & Raoult 2001). We therefore assigned a *Bartonella* species identity to positive samples by first determining the size of the pITS amplicon(s) present when run on an agarose gel. This initial step also allowed identification of “coinfections” (where multiple species of *Bartonella* were present in the same sample), which were visible as multiple bands of different size on the gel.

198 Further to this species-level classification, we identified genetic variation within these
199 *Bartonella* species groups by sequencing a random subset of pITS amplicons of each
200 size from each host species and site (see Appendix S1 and Fig. S1 for methods and
201 assessment of sampling bias). We also sequenced amplicons from all *Bartonella*-
202 positive, non-coinfected fleas. Species classifications of variants were confirmed by
203 identifying the validated *Bartonella* species in Genbank with which each shared
204 highest percentage similarity. This process also allowed differentiation between pITS
205 sequences that are similar in length but somewhat divergent, and therefore likely to
206 represent different *Bartonella* species.
207
208 **Investigating covert host-specificity of *Bartonella* infecting wood mice and bank**
209 **voles**
210 We investigated whether wood mice and bank voles were associated with
211 significantly different assemblages of *Bartonella* parasites using linear discriminant
212 analyses (LDA) in the "*MASS*" package of R (v2.14.2). This analysis tests whether
213 individuals can be identified to host species based only on the identity of the
214 *Bartonella* DNA they were carrying (Venables & Ripley 2002). First, a random 75%
215 subset of the true host-*Bartonella* associations were used to train a host assignment
216 model, which was then used to predict the host identity of the remaining 25% of the
217 data. This was repeated 1000 times, each with a randomly selected set of training data,
218 and mean prediction success was calculated. We then determined the mean prediction
219 success of 1000 models trained using data that simulated random distributions of
220 parasites across host species. The prediction successes of these two sets of models
221 were compared using a χ^2 test to determine if host-parasite associations varied
222 significantly from random expectations.

This analysis was first conducted on assemblages of *Bartonella* DNA identified to species-level according to length of the pITS region. It was then repeated using the subset of *Bartonella* DNA that was sequenced and identified to pITS variant level to see if this afforded greater power to discriminate between host species (thus indicating covert host-specificity). The analyses used combined data from all woodland sites (results were consistent when data from each site was analysed separately; Table S5). *Bartonella* species or variants observed on <5 occasions were omitted, as inclusion of very rare species/variants introduced computational problems when performing model validation. Since host-specific *Bartonella* species comprising a single pITS variant have no potential for covert specificity, but may influence the power of parasite assemblages to discriminate between host species, we checked whether LDA results were affected by the inclusion of these species by re-running all species-level and variant-level LDAs using multi-host *Bartonella* infections only (i.e. infections with *B. grahamii*, *B. taylorii* and *B. birtlesii*). We also confirmed that none of the results were biased by any particular *Bartonella* species, or by repeat sampling of individual rodents (Table S6).

Comparison of flea communities associated with wood mice and bank voles

Opportunities for between-species *Bartonella* transmission may be limited by strong host preferences of different flea species. We investigated this possibility by using an LDA, as described above, to assess the similarity of flea assemblages infecting wood mice and bank voles. Host assignment models were trained on the associations between host and flea species, and we verified that sampling of multiple fleas from individual rodents did not affect the results (Table S7).

Investigating potential flea transfer between wood mice and bank voles

In the absence of strong host preferences, fleas may still limit opportunities for between-species *Bartonella* transmission if individual fleas rarely disperse between different host species. We therefore sought evidence of structure within the flea community that could indicate a general lack of movement/transfer between host species. We used an LDA, as described above, to determine whether the species identity of the host from which a flea was taken could be predicted based only on the *Bartonella* variant carried by a flea (results were not biased by any particular flea species, or by sampling of multiple flea specimens from individual rodents; Table S10). We also sought specific cases where fleas carried *Bartonella* variants never detected in the host species from which they were collected. Such occurrences would be evidence of host exposure to *Bartonella* variants from another host species but lack of infection, so suggesting the presence of a host-parasite compatibility barrier rather than a lack of ecological opportunity for infection. Since the host-specificity of *Bartonella* variants were determined from data collected in 2011 and 2012, whereas fleas were collected from hosts during 2012-2014 and at an additional site (HW), we checked for the consistency of these results using only data for which the characterisation of *Bartonella* DNA in rodents and fleas at the same sites and in the same sampling year were available (i.e. MFG and RH in 2012).

Results

Bartonella in rodents: overall prevalence

Blood samples were taken from 743 wood mice (1376 samples) and 751 bank voles (1224 samples). *Bartonella* DNA was detected in 816 (59.3%) wood mouse and 599

(48.9%) bank vole samples. *Bartonella* coinfections were detected in 23.2% of positive samples from wood mice and 15.2% of positive samples from bank voles.

Bartonella in rodents: species-level data

Amplicons of five broad size categories were obtained from the genus-specific *Bartonella* PCR. Sequencing analyses (see below) confirmed that seven distinct species groups were represented, according to similarity to validated species in GenBank. Patterns of host associations were consistent across woodland sites (Fig. S2, Table S2); we therefore describe the combined data here. Three species (*B. grahamii*, *B. taylorii* and *B. birtlesii*) were found in both wood mice and bank voles (Fig. 1, Fig. S2). Two species (*B. rochalimae*-like and *B. doshiae*) were found only in bank voles, and two species (BGA and *B. doshiae*-like) were found only in wood mice (Fig. 1, Table S2).

Bartonella in rodents: pITS variant-level data

Sequences were obtained for 439 *Bartonella* pITS amplicons from wood mice (43.5% of pITS amplicons) and 391 amplicons from bank voles (56.6% of amplicons) (Table S2). Twenty-six unique variants were identified (Table S2), including ten variants that were new to GenBank (see Table S4 for accession numbers). All variants shared at least 94% similarity (with the majority sharing 99-100% similarity) to their closest species match within GenBank, with their next closest species match sharing lower similarity (Table S11). We found no association between the proportion of pITS amplicons sequenced and the number of variants per *Bartonella* species found within each host species (Appendix S1.3; Figure S1). We therefore assume that the host-associations described below would not be affected by increased sequencing effort.

298 Samples that were not sequenced were classified to species according to amplicon
299 size only, and denoted as “unknown” variant within that species group.
300
301 Twenty-two of the variants identified constituted three different *Bartonella* species
302 groups and displayed varying degrees of host-specificity. Five variants, each ~315bp
303 in length, shared highest percentage similarity with *B. grahamii* in GenBank (Table
304 S11); three were bank vole-specific (grahamii-1, grahamii-2 and grahamii-3), and two
305 were found in both host species (“host-shared”; grahamii-4 and grahamii-5), and
306 while none were wood mouse-specific, the majority of wood mouse infections
307 comprised variants that were relatively rare in bank voles (Fig. 2a, Table S2). Ten
308 variants, each ~350bp in length, shared highest similarity with *B. taylorii* (Table S11);
309 five were wood mouse-specific (taylorii-6, taylorii-7, taylorii-8, taylorii-9 and
310 taylorii-10), and two were bank vole-specific (taylorii-1 and taylorii-2; Fig. 2b, Table
311 S2). The remaining three variants were host-shared, although one was more common
312 in bank voles (taylorii-3) and two more common in wood mice (taylorii-4 and taylorii-
313 5; Fig. 2b, Table S2). Finally, seven variants shared highest similarity with *B. birtlesii*
314 (Table S11). Each was 370bp in length, except for one, birtlesii-4, which was 351bp.
315 The majority were wood mouse-specific (birtlesii-2, birtlesii-3, birtlesii-4, birtlesii-5,
316 birtlesii-6 and birtlesii-7), while one was host-shared (birtlesii-1) but far more
317 common in bank voles (Fig. 2c, Table S2).
318
319 The four remaining variants each shared highest percentage similarity with a separate
320 *Bartonella* species in GenBank. There were two variants with a pITS length of
321 approximately 290bp. One matched most closely to *B. doshiae* (doshiae-1, 292bp)
322 whereas the other (doshiae-like-1) was identical to variant ‘wbs011’ found in previous

studies of rodent *Bartonella* in the UK (Table S11). This latter variant was classified as a *B. doshiae*-like species (Telfer *et al*., 2005), owing to its high similarity to *B. doshiae* at the citrate synthase marker but divergence at the ITS region, and we retain that nomenclature here. Finally, there were two variants with a pITS length of ~460bp. One (BGA-1, 466bp) was identical to a variant previously classified as a species called BGA (Telfer *et al*., 2007b), whereas the other (rochalimae-like-1, 461bp) was identical to a sequence from a non-isolated candidate species called *B. rudakovii* (Table S11). As this species is unconfirmed, we classify this variant as *B. rochalimae*-like here, as candidatus *B. rudakovii* has been found to group closely with the species *B. rochalimae* according to similarity at the ITS region and at other markers (e.g. Diniz *et al*., 2009). Each of these four species groups was host-specific: all amplicons of ~290bp sequenced from bank voles (2/2) were identified as *B. doshiae*, while all those sequenced from wood mice (58/161) were *B. doshiae*-like, and all amplicons of ~460bp sequenced from bank voles (66/152) were identified as *B. rochalimae*-like, while all of those sequenced from wood mice (35/55) were identified as BGA (Table S2).

Comparison of *Bartonella* parasites found in wood mice and bank voles

The assemblages of *Bartonella* detected in wood mice and bank voles were highly distinguishable according to the LDAs. Models trained on true host-parasite associations were consistently better at predicting host species than models trained on random associations (comparisons a-f Fig. 3A, Table S5). This was true whether *Bartonella* were identified to species-level (Fig. 3A comparison 'a' [77.1% versus 21.5%, χ^2=61.8, p<0.001] and comparison 'b' [66.7% versus 19.8%, χ^2=44.8, p<0.001]) or to variant-level (Fig. 3A comparison 'c' [97.8% versus 66.4%, χ^2=33.5,

$p<0.001$] and comparison 'd' [97.1% versus 66.9%, $\chi^2=30.9$, $p<0.001$]), and when considering associations of the variants within individual *Bartonella* species (Fig. 3A comparison 'e' [85.0% versus 44.9%, $\chi^2=33.8$, $p<0.001$] and comparison 'f' [95.5% versus 33.6%, $\chi^2=83.7$, $p<0.001$]). However, the success of models trained on species-level data was significantly reduced when the associations of the four host-specific, single-variant *Bartonella* species (*B. doshiae*, *B. doshiae*-like, *B. rudakovii* and BGA) were omitted (Fig. 3A comparison 'g' [77.1% versus 66.7%, $\chi^2=26.8$, $p<0.001$]). In contrast, models trained on variant-level data performed equally well whether incorporating all or just host-shared *Bartonella* species (Fig. 3A comparison 'h' [97.8% versus 97.1%, $\chi^2=0.99$, $p=0.32$]), and were always superior to models trained on species-level data (Fig. 3A comparison 'i' [97.8% versus 77.1%, $\chi^2=19.5$, $p<0.001$] and comparison 'j' [97.1% versus 66.7%, $\chi^2=31.2$, $p<0.001$]).

Rodent flea assemblages

Fleas were collected from 224 wood mice (WM; 325 fleas) and 357 bank voles (BV; 589 fleas). Seven species were identified: *Amalareus penicilliger mustelae* (from 91 BV and 23 WM), *Ctenophthalmus nobilis vulgaris* (231 BV, 188 WM), *Hystrichopsylla talpae talpae* (18 BV, 8 WM), *Megabothris turbidus* (88 BV, 22 WM), *Palaeopsylla sorcis* (1 BV, 2 WM), *Rhadinopsylla pentacantha* (27 BV, 12 WM) and *Typhlocerus poppei poppei* (0 BV, 4 WM). All species of flea except *T. p. poppei* were found on both rodent species (Fig. 4). The assemblages of flea species collected from wood mice and bank voles were not distinguishable according to the LDA. Models trained on true host-flea associations were no better at predicting host species than models trained on random associations (Fig. 3B [30.4% mean prediction success versus 30.7%, $\chi^2=0.212$, $p=0.88$], Table S7).

373

374 ***Bartonella* in rodent fleas**

375 DNA was extracted from 881 fleas. *Bartonella* DNA was detected in 460 (52%)
376 individual fleas, and in all flea species except *T. p. poppei*. pITS sequences were
377 obtained for 382 *Bartonella* pITS amplicons, each from a separate flea. The remaining
378 78 *Bartonella*-positive fleas were coinfected and pITS amplicons were not sequenced.
379 Thirty different variants were found (Table S8), representing eight *Bartonella* species
380 (Table S11). Twenty variants matched those identified in rodent blood samples in this
381 study; nine of which were wood mouse-specific (doshiae-like-1, BGA-1, taylorii-6,
382 taylorii-7, taylorii-8, taylorii-9, taylorii-10, birtlesii-5 and birtlesii-7), five were bank
383 vole-specific (doshiae-1, rudakovii-1, grahamii-1, grahamii-2 and taylorii-2) and six
384 were host-shared (grahamii-4, grahamii-5, taylorii-3, taylorii-4, taylorii-5 and
385 birtlesii-1) (Table S2). The remaining ten variants were novel to this study and to
386 GenBank (they have now been added; Table S9). There were three *B. grahamii*
387 variants (grahamii-6, grahamii-7 and grahamii-8), two *B. taylorii* variants (taylorii-11
388 and taylorii-12), two *B. birtlesii* variants (birtlesii-8 and birtlesii-9) and one *B. doshiae*
389 variant (doshiae-2) (Table S11). One variant (tribocorum-1) was most similar to *B.*
390 *tribocorum*; a species previously found to infect rats (Heller *et al*. 1998), and never
391 recorded from wood mice or bank voles in this study. One further variant (unknown-
392 1) did not closely match any known *Bartonella* species in GenBank (Table S11).

393

394 **Comparing *Bartonella* in fleas collected from wood mice and bank voles**

395 A range of *Bartonella* pITS variants, including wood mouse-specific, bank vole-
396 specific and host-shared, were found in all flea species in which *Bartonella* DNA was
397 detected (except *P. sorcis*, for which only a single *Bartonella* pITS amplicon was

characterised; Table S8). However, the LDA showed that the species of rodent from which a *Bartonella*-positive flea was collected was highly predictable based on the variant of *Bartonella* it was carrying (Fig. 3C [85.3% mean prediction success for models trained on true associations between flea *Bartonella* variants and rodent species versus 49.8% for models trained on random associations, χ^2=28.7, p<0.001], Table S10). In other words, the assemblage of *Bartonella* variants found within fleas tended to reflect the assemblage of *Bartonella* variants found within the host species they were collected from. This pattern is unlikely to simply reflect recent acquisition of infections by fleas feeding on their current host, as the variants carried by fleas often did not match the variants carried by the rodent host from which they were collected (Table S13).

Host-specific pITS variants were occasionally found in fleas collected from the alternative rodent species (Fig. 5). Wood mouse-specific variants were found in *C. n. vulgaris* (doshiae-like-1, taylorii-6, taylorii-7, taylorii-8, BGA-1; Fig. 5a and Table S8) collected from bank voles, and bank vole-specific variants were found in *C. n. vulgaris* (doshiae-1, grahamii-1, taylorii-2; Fig. 5a and Table S8), *M. turbidus* (grahamii-1, grahamii-2, rudakovii-1; Fig. 5b and Table S8), *A. p. mustelae* (grahamii-1, grahamii-2, rudakovii-1; Fig. 5c and Table S8) and *H. t. talpae* (grahamii-2; Fig. 5d and Table S8) collected from wood mice. No such pattern was found in *R. pentacantha* (Fig. 5e and Table S8) even though a similar number of pITS amplicons were sequenced for this flea species (n=6) as for *H. t. talpae* (n=7) for which evidence of between-host species flea transfer was present. There was also no evidence of flea transfer for *P. sorcis,* but only a single specimen of this flea species was positive for *Bartonella* DNA. Examples of host-specific variants in fleas collected from the

alternative rodent species were also evident when considering only data from 2012 at MFG and RH, the site-year combinations for which *Bartonella* sequences from both hosts and fleas were available (Fig. S3).

Discussion

An ever-expanding body of evidence clearly demonstrates that most parasite species infect multiple host species (Cleaveland *et al*. 2001; Taylor *et al*. 2001; Pedersen *et al*. 2005; Streicker *et al*. 2013). Where the same parasite endemically infects sympatric host species, between-species transmission is assumed to be commonplace (i.e. a "true multi-host parasite"; Fenton & Pedersen 2005), meaning a parasite reservoir potentially comprises an entire multi-host community (Haydon *et al*. 2002) with transmission occurring somewhat freely between species. For medically important parasites, such a scenario would require potentially complex disease management across all host species (Fenton *et al*. 2015). In contrast to this conventional wisdom, however, we have shown that even with considerable overlaps in host species ecology, and despite the presence of host-generalist vectors, the transmission of multi-host parasites between endemically infected sympatric host species in the wild is surprisingly infrequent.

Overall we found seven *Bartonella* species circulating within a host community of two sympatric rodent species, and three of these (*B. grahamii*, *B. taylorii* and *B. birtlesii*) infected both wood mice and bank voles. This is consistent with a previous study that used one of the same field sites (Manor Wood) (Telfer *et al*. 2007a). However, our genetic characterisation of these parasites revealed considerable diversity within a partial ITS region of these three *Bartonella* species. Crucially, we

found that each host species was associated with highly distinguishable assemblages of variants, and many of these variants were host specific. Furthermore, while some variants were shared across host species, these shared variants were always far more common in one host species than the other. Together, these results provide strong evidence for 'covert host-specificity' among these variants, implying a general lack of parasite transmission between these two common sympatric rodent species, despite such transmission having previously been suggested from observed relationships between parasite prevalence and host densities (Telfer *et al*. 2007a).

We found clear evidence that most flea species are host-generalists; in fact, all flea species except *T. p. poppei* were found on both wood mice and bank voles, and overall the assemblages of fleas associated with each host species were indistinguishable according to our linear discriminant analyses. However, the dispersal of these generalist vectors between host species appeared to be limited, which may restrict opportunities for between-species *Bartonella* transmission. We identified the genetic variants of *Bartonella* being carried by fleas and found that overall, the identity of the host species from which a flea was taken could be determined by looking only at the *Bartonella* variant carried by that flea. The assemblage of *Bartonella* variants found within the flea community therefore has clear structure, which is strongly correlated with the rodent host species that fleas were collected from. This suggests that separate communities of the same flea species may circulate largely independently within each host species population, and that transfer of individual fleas between these discrete pools is rare. This seems reasonable, as the flea species found at our study sites are mostly nest-dwellers that feed opportunistically on hosts entering their nests (Marshall 1981; Krasnov 2008).

Flea movement between species is therefore likely to require close mouse-vole contact, or use of the same habitat space by different host individuals for a sufficient period of time (Krasnov & Khokhlova 2001), which may be infrequent due to differences in activity patterns and microhabitat usage by wood mice and bank voles (Watts 1968; Crawley 1969; Greenwood 1978; Canova 1993). Indeed, wood mice and bank voles were only occasionally captured at the same trap location during a given monthly session across our study sites (median proportion of multi-species trap locations per session was 0.2 across all sessions and sampling sites; Table S12), indicating some differentiation in microhabitat use within the same broad woodland area.

As a consequence of limited between-species vector dispersal, opportunities for between-species parasite transmission may be rare (i.e. an encounter barrier), even when host species are infected by the same vector species. This potentially counters the complex view of parasite persistence and control within multi-species reservoirs (Haydon *et al*. 2002). However, if host-specific variants are physiologically capable of infecting a wider range of host species given the opportunity, between-species transmission may occur if barriers to encounter break down, for example due to anthropogenic shifts in community structure (i.e. a 'potential multi-host parasite' becoming a 'true multi-host parasite'; Fenton & Pedersen 2005). Here, however, we found evidence that at least some host-specific *Bartonella* variants were unable to infect the other species, possibly due to physiological incompatibility, as some fleas were found carrying these host-specific variants on the other, uninfected host species. In fact, as we did not sequence *Bartonella* DNA from any coinfected fleas, it is possible that we underestimate the occurrence of between-species flea transfer here,

as coinfected fleas may arise as a result of feeding sequentially on multiple host individuals, and possibly different host species, infected with different pathogens. Such compatibility barriers have been found in Irish rodent communities, where wood mice were endemically infected with *Bartonella* but sympatric bank voles were not, despite harbouring *Bartonella*-positive fleas (Telfer *et al*. 2005). Laboratory inoculation experiments have also shown that *Bartonella* infections often only establish in species of wild rodents when challenged with a variant originally obtained from that same species (Kosoy *et al*. 2000). It therefore seems likely that should ecological barriers to between-host vector transfer break down in the future (e.g. due to environmentally-driven changes to host or vector movement), initial incompatibility barriers may prevent or slow the emergence of regular between-species *Bartonella* transmission, until new variants able to infect multiple host species evolve and increase in frequency (Antia *et al*. 2003; Lloyd-Smith *et al*. 2009).

Interestingly, we found six shared variants of *Bartonella*, all of which were far more common in one host species than the other. Given the occasional occurrence of fleas carrying variants never found in the host species from which they were collected, it seems that between-species flea transfer does occur, at a rate which is sufficient for those few shared variants to maintain a relatively constant, but low, degree of host-generalism (indicative of spillover dynamics; Fenton & Pedersen 2005). Alternatively, it may be that host generalism is a more dynamic phenomenon, and that our data represent a snapshot in evolutionary time such that we are witnessing the evolution of these variants from host-specialists to host-generalists (or vice versa). It would therefore be fascinating to conduct a longer-term study of this system to see the extent

to which variants change in frequency in the two host species over time, and therefore whether between-species transmission is becoming more or less common.

Our findings provide compelling evidence that the ecology of host-generalist vectors could inhibit between-species parasite transmission. We acknowledge that our conclusions about between-species vector movement are drawn from proxy evidence of associations between individual fleas and host species, and an investigation of the genetic structure of the flea populations may help to assess the frequency with which individual fleas transfer between sympatric species and promote between-species transmission. The generality of our findings will also depend on parasite transmission mode and the off-host dispersal capabilities of other vector types (Randolph 1998). For example, vectors that engage in frequent host-independent dispersal (e.g. dipterans such as mosquitoes) have the opportunity to feed sequentially on different host species more often and thus are less likely to represent a barrier to the between-species transmission of multi-host parasites. Furthermore, parasites transmitted by direct contact may have fewer opportunities to cross between host species. For example, it was previously shown that risk of infection with the directly transmitted cowpox virus is not influenced by between-species transmission for sympatric populations of wood mice and bank voles (Begon *et al*. 1999; Carslake *et al*. 2006), presumably due to infrequent appropriate inter-species encounters. In contrast, opportunities for between-species exposure for parasites with environmental transmission stages (e.g. intestinal helminths) may be more frequent, with different vector-borne parasites lying at different points along a continuum between these two extremes. Identifying general trends in the occurrence of between-species

546 transmission based on broad host and parasite ecology would improve our
547 understanding of disease transmission within complex ecological communities.
548
549 In conclusion, our results show that the transmission of multi-host parasites between
550 sympatric host species is not inevitable, and cannot necessarily be predicted based on
551 shared host ecologies alone, nor on the presence of host-generalist vectors. We
552 emphasise that, in fact, between-species transmission may be a lot more rare than
553 previously assumed. Thus, separate populations of the same parasite species may
554 often circulate and persist independently in different sympatric host species
555 populations. This challenges conventional wisdom surrounding the control of multi-
556 host parasites and, if a general phenomenon, suggests that control interventions would
557 likely need to be multi-pronged, aiming to reduce infection independently in multiple
558 host species.
559

560 ## Data Accessibility

561 All data associated with this study have been deposited in the Dryad Digital
562 Repository (http://dx.doi.org/10.5061/dryad.gm061). As data analyses are ongoing,
563 release of data has been embargoed for 1 year from the date of publication. GenBank
564 accession numbers of all sequences included in this paper, including those identified
565 for the first time here, are shown in Tables S4, S9 and S11.
566

567 ## Acknowledgements

568 We would like to thank the field technicians Rebecca Barber, Łukasz Łukomski,
569 Stephen Price, Trevor Jones and Niamh Quinn, along with numerous undergraduate
570 assistants, for collecting much of the data used in this paper, and the various

landowners for allowing us to work on their land. Thank you to Ami Patrick and Evelyn Rynkiewicz for help with identification and sequencing of fleas, Mike Begon and Roman Biek for helpful suggestions about data analysis and interpretation, and two anonymous reviewers for useful comments on an earlier draft of this manuscript. The work was funded by a Natural Environment Research Council (NERC) studentship to SW (NE/H528442/1), NERC grants (NE/G007349/1 and NE/G006830/1) awarded to AF, ABP and Owen L. Petchey (University of Zurich), the Wellcome Trust Strategic Grant for the Centre for Immunity Infection and Evolution (095831) Advanced Fellowship to ABP, and the University of Edinburgh Chancellors Fellowship to ABP. Animal work was licensed under Home Office Project Licence PPL40/3235.

References

Antia, R., Regoes, R. R., Koella, J. C. & Bergstrom, C. T. (2003). The role of evolution in the emergence of infectious diseases. *Nature*, **426,** 658–661.

Begon, M., Bowers, R. G., Kadiankis, N. and Hodgkinson D. E. (1992) Disease and community structure: the importance of host self-regulation in a host-host-pathogen model. *The American Naturalist*, **139,** 1131–1150.

Begon, M., Hazel, S. M., Baxby, D., Bown, K., Cavanagh, R., Chantrey, J,, Jones, T. & Bennett, M. (1999). Transmission dynamics of a zoonotic pathogen within and between wildlife host species. *Proceedings of the Royal Society of London B: Biological Sciences*, **266,** 1939-1945.

Berglund, E.C., Ellegaard, K., Granberg, F., Xie, Z., Maruyama, S., Kosoy, M., Birtles, R. J. & Andersson, S. G. E. (2010). Rapid diversification by recombination in *Bartonella grahamii* from wild rodents in Asia contrasts with low levels of

genomic divergence in Northern Europe and America. *Molecular Ecology*, **19,** 2241–2255.

Birtles, R., Hazel, S., Bown, K., Raoult, D., Begon, M. & Bennett, M. (2000). Subtyping of uncultured bartonellae using sequence comparison of 16 S/23 S rRNA intergenic spacer regions amplified directly from infected blood. *Molecular and Cellular Probes*, **14,** 79-87.

Birtles, R., Hazel, S. M., Bennett, M., Bown, K., Raoult, D. & Begon, M. (2001). Longitudinal monitoring of the dynamics of infections due to *Bartonella* species in UK woodland rodents. *Epidemiology and Infection*, **126,** 323-329.

Bowers, R. G. & Turner, J. (1997) Community structure and the interplay between interspecific infection and competition. *Journal of Theoretical Biology,* **187,** 95–109.

Bowers, R. & Begon, M. (1991). A host-host-pathogen model with free-living infective stages, applicable to microbial pest control. *Journal of Theoretical Biology*, **148,** 305–329.

Bown, K., Bennet, M. & Begon, M. (2004). Flea-borne *Bartonella grahamii* and *Bartonella taylorii* in bank voles. *Emerging Infectious Diseases*, **10,** 684-687.

Canova, L. (1993). Resource partitioning between the bank vole *Clethrionomys glareolus* and the wood mouse *Apodemus sylvaticus* in woodland habitats. *Bollettino di Zoologia*, **60,** 193-198.

Carslake, D., Bennett, M., Hazel, S., Telfer, S. & Begon, M. (2006). Inference of cowpox virus transmission rates between wild rodent host classes using space-time interaction. *Proceedings of the Royal Society of London B: Biological Sciences*, **273,** 775-782.

620 Cleaveland, S., Laurenson, M. & Taylor, L., 2001. Diseases of humans and their
621 domestic mammals: pathogen characteristics, host range and the risk of emergence.
622 *Philosophical Transactions of the Royal Society of London B: Biological Sciences*,
623 **356,** 991-999.
624 Combes, C. (2001). Parasitism: *The ecology and evolution of intimate interactions*.
625 University of Chicago Press, Chicago, Illinois, USA.
626 Crawley, M. C. (1969). Movements and home-ranges of *Clethrionomys glareolus*
627 Schreber and *Apodemus sylvaticus* L. in North-East England. *Oikos,* **20,** 310–319.
628 Diniz, P. P. V. P., Billeter, S. A., Otranto, D., De Caprariis, D., Petanides, T.,
629 Mylonakis, M. E., Koutinas, A. F. & Breitschwerdt, E. B. (2009). Molecular
630 documentation of *Bartonella* infection in dogs in Greece and Italy. *Journal of*
631 *Clinical Microbiology*, **47,** 1567–1567.
632 Dobson, A. (2004). Population dynamics of pathogens with multiple host species. *The*
633 *American Naturalist*, **164,** S64-S78.
634 Dobson, A. & Meagher, M. (1996). The population dynamics of Brucellosis in the
635 Yellowstone National Park. *Ecology*, **77,** 1026-1036.
636 Donnelly, C.A., Woodroffe, R., Cox, D. R., Bourne, F. J., Cheeseman, C. L., Clifton-
637 Hadley, R. S., Wei, G., Gettinby, G., Gilks, P., Jenkins, H., Johnston, W. T., Le
638 Fevre, A. M., McInerney, J. P. & Morrison, W. I. (2006). Positive and negative
639 effects of widespread badger culling on tuberculosis in cattle. *Nature*, **439,** 843-
640 846.
641 Fenton, A. & Pedersen, A. (2005). Community epidemiology framework for
642 classifying disease threats. *Emerging Infectious Diseases*, **11,** 1815-1821.

643 Fenton, A., Streicker, D. G., Petchey, O. L. & Pedersen, A. B. (2015). Are all hosts
644 created equal? Partitioning host species contributions to parasite persistence in
645 multihost communities. *The American Naturalist*, **186,** 610-622.
646 Forrester, S. J. & Hall, N. (2014). The revolution of whole-genome sequencing to
647 study parasites. *Molecular and Biochemical Parasitology*, **195,** 77-81.
648 Greenman, J. V. & Hudson, P. J. (1999) Host exclusion and coexistence in apparent
649 and direct competition: an application of bifurcation theory. *Theoretical*
650 *Population Biology*, **56,** 48–64.
651 Greenman, J. V. & Hudson, P. J. (2000) Parasite-mediated and direct competition in a
652 two-host shared macroparasite system. *Theoretical Population Biology*, **57,** 13–
653 34.
654 Greenwood, P. J. (1978). Timing of activity of the bank vole *Clethrionomys glareolus*
655 and the wood mouse *Apodemus sylvaticus* in a deciduous woodland. *Oikos*, **31,**
656 123-127.
657 Gutiérrez, R., Krasnov, B., Morick, D., Gottlieb, Y. Khokhlova, I. S. & Harrus, S.
658 (2015). *Bartonella* infection in rodents and their flea ectoparaistes: an overview.
659 *Vector Borne and Zoonotic Diseases*, **15,** 27-39.
660 Harris, P. D., Paziewska, A., Zwolinska, L. & Sinski, E. (2009). Seasonality of the
661 ectoparasite community of woodland rodents in a Mazurian forest, Poland. *Wiad*
662 *Parazytol*, **55,** 377-388.
663 Haydon, D., Cleaveland, S., Taylor, L. H. & Laurenson, M. K. (2002). Identifying
664 reservoirs of infection: A conceptual and practical challenge. *Emerging Infectious*
665 *Diseases*, **8,** 1468-1473.
666 Heller, R., Riegel, P., Hansmann, Y., Delacour, G., Bermond, D., Dehio, C.,
667 Lamarque, F., Monteil, H., Chomel, B. & Piémont, Y. (1998) *Bartonella*

668 *tribocurum* sp. nov., a new *Bartonella* species isolated from the blood of wild rats.
669 *International Journal of Systemic Bacteriology*, **48,** 1333-1339.
670 Holt, R. D. & Pickering, J. (1985). Infectious disease and species coexistence: a
671 model of Lotka-Volterra form. *The American Naturalist*, **126,** 196–211.
672 Holt, R., Dobson, A. P., Begon, M., Bowers, R. G. & Schauber, E. M. (2003). Parasite
673 establishment in host communities. *Ecology Letters*, **6,** 837–842.
674 Houpikian, P. & Raoult, D. (2001). 16S/23S rRNA intergenic spacer regions for
675 phylogenetic analysis, identification, and subtyping of *Bartonella* species. *Journal*
676 *of Clinical Microbiology*, **39,** 2768–2778.
677 Inoue, K., Maruyama, S., Kabeya, H., Yamada, N., Ohashi, N., Sato, Y., Yukawa, M.,
678 Masuzawa, T., Kawamori, F., Kadosaka, T., Takada, N., Fujita, H. and Kawabata,
679 H. (2008). Prevalence and genetic diversity of *Bartonella* species isolated from
680 wild rodents in Japan. *Applied and Environmental Microbiology*, **74,** 5086-5092.
681 Khokhlova, I.S., Fielden, L. J., Gegen, A. A. & Krasnov, B. R. (2012). Feeding
682 performance of fleas on different host species: is phylogenetic distance between
683 hosts important? *Parasitology*, **139,** 60–68.
684 Kilpatrick, A.M., Daszak, P., Jones, M. J., Marra, P. P. & Kramer, L. D. (2006). Host
685 heterogeneity dominates West Nile virus transmission. *Proceedings of the Royal*
686 *Society of London B: Biological Sciences*, **273,** 2327–2333.
687 Knowles, S. C. L., Fenton, A., Petchey, O. L., Jones, T. R., Barber, R. & Pedersen, A.
688 B. (2013). Stability of within-host parasite communities in a wild mammal system.
689 *Proceedings of the Royal Society of London B: Biological Sciences*, **280,**
690 20130598
691 Kosoy, M., Saito, E. K., Green, D., Marston, E. L., Jones, D. C. & Childs, J. E.
692 (2000). Experimental evidence of host specificity of *Bartonella* infection in

693 rodents. *Comparative Immunology, Microbiology and Infectious Diseases*, **23,**
694 221–238.
695 Kosoy, M., Hayman, D.T.S. & Chan, K.-S. (2012). *Bartonella* bacteria in nature:
696 Where does population variability end and a species start? *Infection, Genetics and*
697 *Evolution*, **12,** 894–904.
698 Krasnov, B. (2008). *Functional and evolutionary ecology of fleas: A model for*
699 *ecological parasitology*. Cambridge University Press, Cambridge, UK.
700 Krasnov, B.R. & Khokhlova, I.S. (2001). The effect of behavioural interactions on the
701 transfer of fleas (Siphonaptera) between two rodent species. *Journal of Vector*
702 *Ecol*ogy, **26,** 181–190.
703 La Scola, B., Zeaiter, Z., Khamis, A. & Roult, D. (2003). Gene-sequence-based
704 criteria for species definition in bacteriology: the *Bartonella* paradigm. *Trends in*
705 *Microbiology*, **11,** 318–321.
706 Laurenson, M.K., Norman, R. A., Gilbert, L., Reid, H. W. & Hudson, P. J. (2003).
707 Identifying disease reservoirs in complex systems: mountain hares as reservoirs of
708 ticks and louping□ill virus, pathogens of red grouse. *Journal of Animal Ecology*,
709 **72,** 177–185.
710 Lloyd-Smith, J.O., George, D., Pepin, K. M., Pitzer, V. E., Pulliam, J. R. C., Dobson,
711 A. P., Hudson, P. J. & Grenfell, B. T. (2009). Epidemic Dynamics at the Human-
712 Animal Interface. *Science*, **326,** 1362–1367.
713 Marshall, A. G. (1981). *The ecology of ectoparasitic insects*. Academic Press Ltd.,
714 London, UK.
715 Martínez-Aquino, A., Reyna-Fabián, M., Rosas-Valdez, R., Razo-Mandivil, U.,
716 Pérez-Ponce de León, G. & García-Varela, M. (2009). Detecting a complex of
717 cryptic species within *Neoechinorhynchus golvani* (Acanthocephala:

718 Neoechinorhynchidae) inferred from ITSs and LSU rDNA gene sequences.
719 *Journal of Parasitology*, **95,** 1040–1047.
720 Morick, D., Krasnov, B., Khokhlova, I. S., Shrenbrot, G. I., Kosoy, M. Y. & Harrus,
721 S. (2010). *Bartonella* Genotypes in Fleas (Insecta: Siphonaptera) Collected from
722 Rodents in the Negev Desert, Israel. *Applied and Environmental Microbiology*, **76,**
723 6864–6869.
724 Paziewska, A., Harris, P. D., Zwolinska, Bajer, A. & Sinski, E. (2011).
725 Recombination within and between species of the alpha proteobacterium
726 *Bartonella* infecting rodents. *Microbial Ecology*, **61,** 134–145.
727 Paziewska, A., Harris, P. D., Zwolinska, L, Bajer, A. & Sinski, E. (2012). Differences
728 in the ecology of *Bartonella* infections of *Apodemus flavicollis* and *Myodes*
729 *glareolus* in a boreal forest. *Parasitology*, **139,** 881-893.
730 Pedersen, A. B., Altizer, S., Poss, M., Cunningham, A. A. & Nunn, C. L. (2005).
731 Patterns of host specificity and transmission among parasites of wild primates.
732 *International Journal of Parasitology*, **35,** 647-657.
733 Randolph, S.E. (1998). Ticks are not insects: Consequences of contrasting vector
734 biology for transmission potential. *Parasitology Today*, **14,** 186–192.
735 Roux, V. & Raoult, D. (1995). Inter-species and intra-species identification of
736 *Bartonella* (*Rochalimaea*) species. *Journal of Clinical Microbiology*, **33,** 1573–
737 1579.
738 Sehgal, R.N.M., Hull, A. C., Anderson, N. L., Valkiunas, G., Markovets, M. J.,
739 Kawamura, S. & Tell, L. A. (2006). Evidence for cryptic speciation of
740 *Leucocytozoon* spp. (Haemosporida, Leucocytozoidae) in diurnal raptors. *Journal*
741 *of Parasitology*, **92,** 375–379.

742 Serrano, E., Cross, P. C., Beneria, M., Ficapal, A., Curia, J., Marco, X., Lavin, S. &
743 Marco, I. (2011). Decreasing prevalence of brucellosis in red deer through efforts
744 to control disease in livestock. *Epidemiology and Infection*, **139,** 1626-1630.
745 Streicker, D.G., Turmelle, A. S., Vonhof, M. J., Kuzmin, I. V., McCracken, G. F. &
746 Rupprecht, C. E. (2010). Host phylogeny constrains cross-species emergence and
747 establishment of Rabies virus in bats. *Science*, **329,** 676–679.
748 Streicker, D. G., Fenton, A., & Pedersen, A. B. (2013). Differential sources of host
749 species heterogeneity influence the transmission and control of multihost parasites.
750 *Ecology Letters*, **16,** 975-984.
751 Taylor, L., Latham, S. & Woolhouse, M. (2001). Risk factors for human disease
752 emergence. *Philosophical Transactions of the Royal Society of London Series B*,
753 **356,** 983–989.
754 Telfer, S., Bown, K., Sekules, R. & Begon, M. (2005). Disruption of a host-parasite
755 system following the introduction of an exotic host species. *Parasitology*, **130,**
756 661-668.
757 Telfer, S., Clough, H.E., Birtles, R. J., Bennett, M., Carslake, D., Helyar, S. & Begon,
758 M. (2007a). Ecological differences and coexistence in a guild of microparasites:
759 *Bartonella* in wild rodents. *Ecology*, **88,** 1841-1849.
760 Telfer, S., Begon, M., Bennett, M., Bown, K. J., Burthe, S., Lambin, X., Telford, G. &
761 Birtles, R. (2007b). Contrasting dynamics of *Bartonella* spp. in cyclic field vole
762 populations: the impact of vector and host dynamics. *Parasitology*, **134,** 413-425.
763 Turner, A. K., Beldomenico, P. M., Bown, K., Burthe, S. J., Jackson, J. A., Lambin,
764 X. & Begon, M. (2014). Host-parasite biology in the real world: the field voles of
765 Kielder. *Parasitology*, **141,** 997-1017.

Venables, W.N. & Ripley, B.D. (2002) *Modern Applied Statistics with S,* 4th Edition. Springer, New York, USA.

Watts, C. (1968). Foods eaten by wood mice (*Apodemus sylvaticus*) and bank voles (*Clethrionomys glareolus*) in Wytham Woods, Berkshire. *Journal of Animal Ecology*, **37,** 25–41.

Whitaker, A. P. (2007). *Handbooks for the identification of British insects: Fleas (Siphonaptera)*. 2nd Edition. Royal Entomological Society, London, UK.

Whiteman, N.K., Sanchez, P., Merkel, J., Klompen, H. & Parker, P. G. (2006). Cryptic host specificity of an avian skin mite (Epidermoptidae) vectored by louseflies (Hippoboscidae) associated with two endemic Galapagos bird species. *Journal of Parasitology*, **92,** 1218–1228.

Withenshaw, S. M., Devevey, G., Pedersen, A. B. & Fenton, A. (2016). Data from: Multi-host *Bartonella* parasites display covert host-specificity even when transmitted by generalist vectors. Dryad Digital Repository. doi:10.5061/dryad.gm061

Woolhouse, M., Taylor, L. & Haydon, D. (2001). Population biology of multihost pathogens. *Science,* **292,** 1109-1112.

Supporting Information

The following Supporting Information is available for this article online:

Appendix S1. Additional methodological details.

Figure S1. Relationship between the proportions of positive samples per *Bartonella* species that were sequenced and the number of *Bartonella* variants detected.

Figure S2. Proportion of blood samples testing positive for infection with each *Bartonella* species at each field site.

Figure S3. Number of *Bartonella*-positive fleas of each species taken from wood mice and bank voles at MFG and RH during 2012.

Table S1. Number of individual wood mice and bank voles captured and number of blood samples collected from each rodent species at each field site.

Table S2. The twenty-six *Bartonella* partial 16S-23S ITS sequence variants detected in this study and where they were found.

Table S3. As Table S2 but only samples from animals not exposed to treatment are presented.

Table S4. GenBank accession numbers of the ten novel Bartonella partial 16S-23S ITS sequence variants detected in rodent blood samples in this study.

Table S5. Results of linear discriminant analyses that modelled host species identity based on either the species-level (S) or variant-level (V) identification of *Bartonella* parasites with which they were infected.

Table S6. As Table S5, but using a reduced data set that includes only a single record of a particular *Bartonella* species or pITS variant for each individual.

Table S7. Results of linear discriminant analyses that modelled host species identity based on the morphological identification of the flea species collected from them.

Table S8. The species identity and *Bartonella* infection status of fleas collected from rodents.

Table S9. GenBank accession numbers of the ten *Bartonella* partial 16S-23S ITS sequence variants detected in fleas only.

Table S10. Results of linear discriminant analyses that modelled host species identity based on the variant of *Bartonella* carried by fleas collected from them.

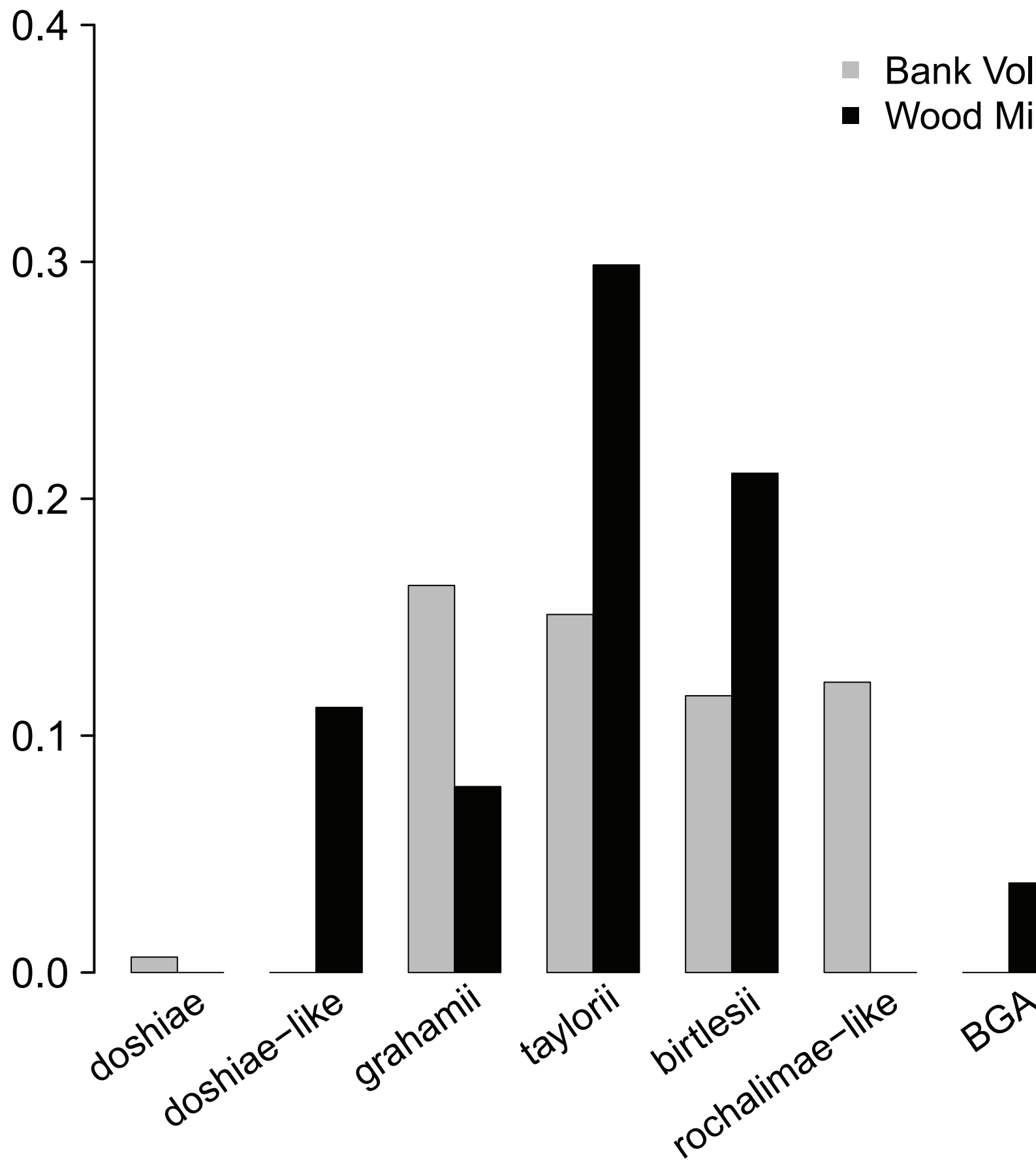
Table S11. *Bartonella* species submissions in GenBank with which each pITS variant in this study shares highest and second highest similarity.

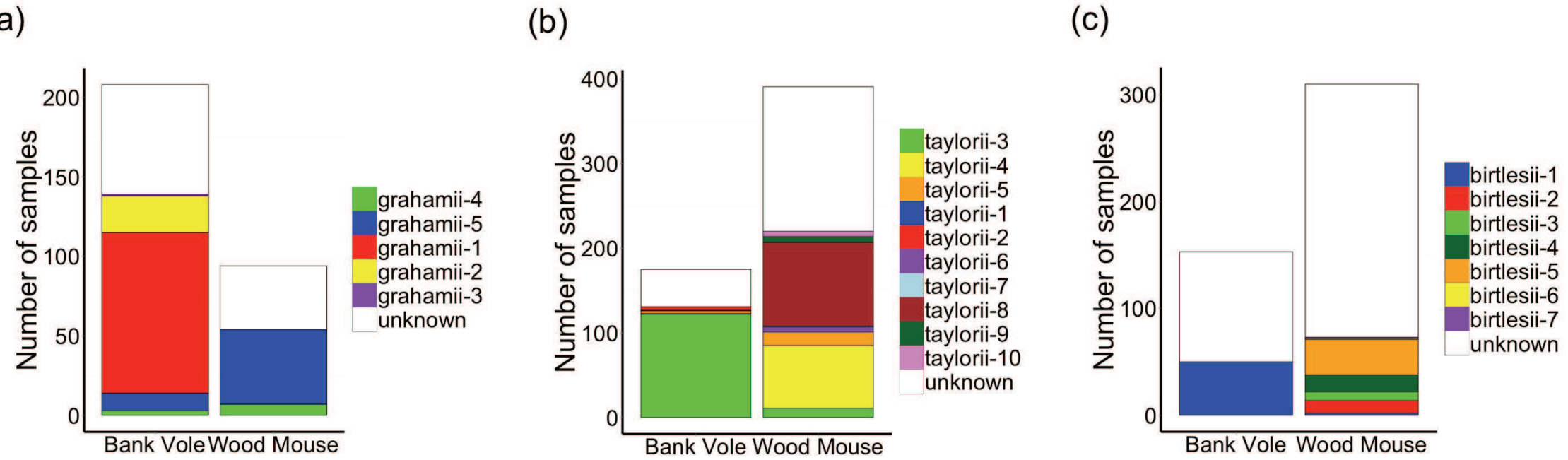
Table S12. Proportions of trap locations at which both wood mice and bank voles were captured during each trapping session at each site.

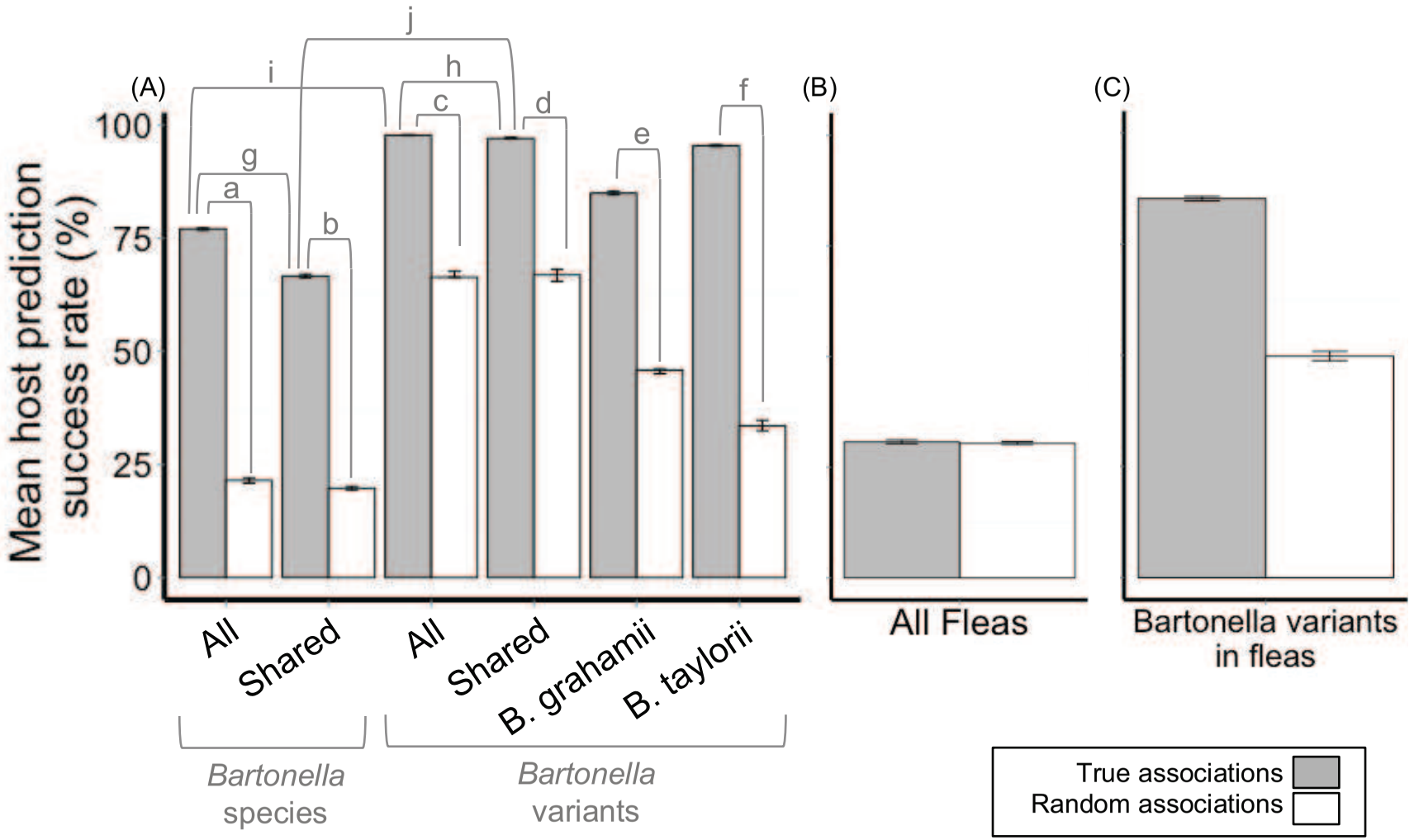
Table 13. Comparison of the *Bartonella* pITS variants found in individual fleas and the variants found in the rodent hosts from which each flea was collected.

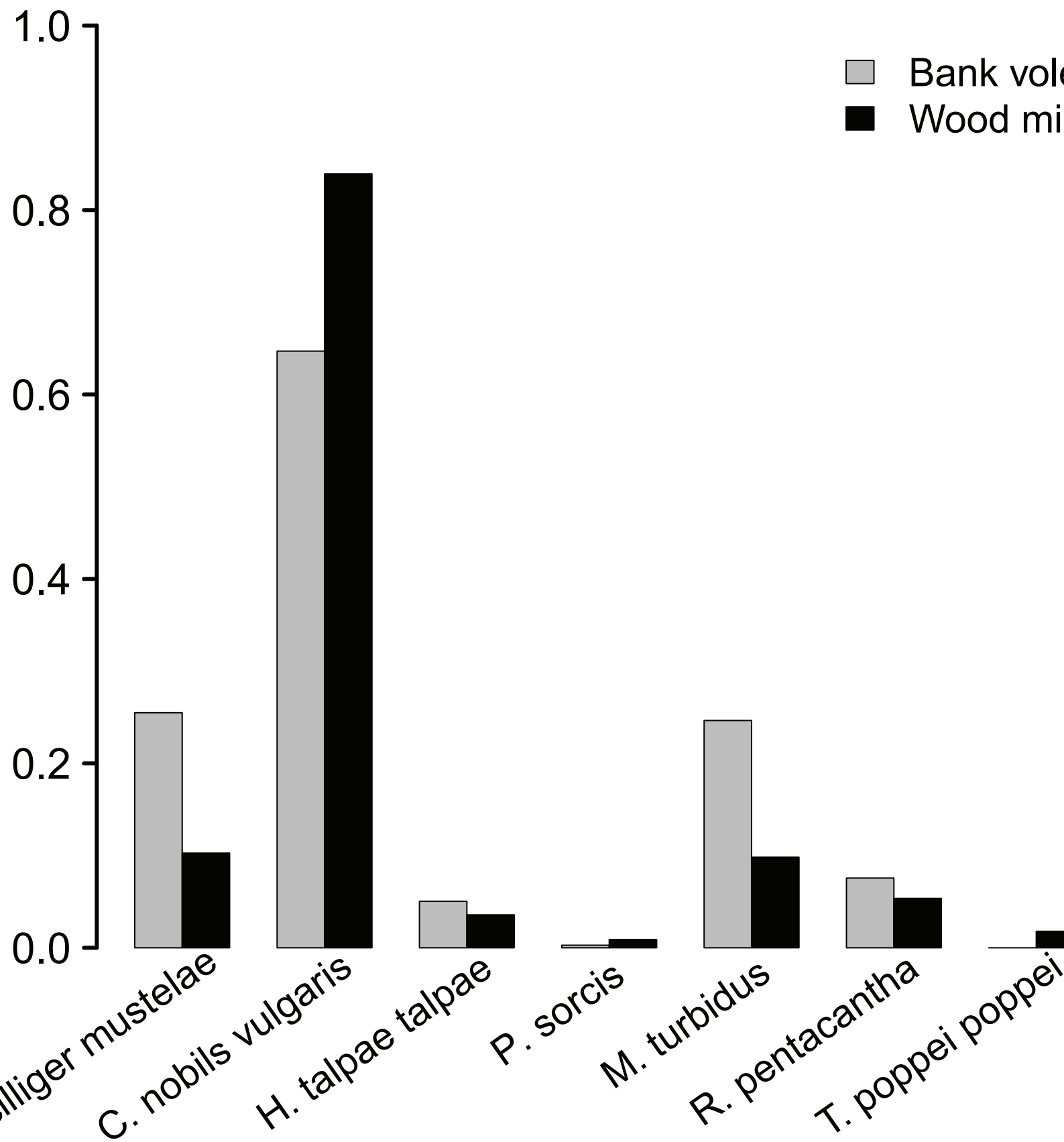
820 **Figure 1** The proportion of blood samples that tested positive for infection with each
821 *Bartonella* species in bank voles and wood mice across all sites. Infections were
822 identified to species according to sequencing of the pITS region where possible, and
823 according to the length of the pITS region in all other cases.
824
825 **Figure 2** The number of each (a) *B. grahamii* (b) *B. taylorii* and (c) *B. birtlesii* variant
826 detected within wood mice and bank voles across all sites. Colour-coding represents
827 different variants within each *Bartonella* species group. Infections that were not
828 sequenced are classed as "unknown" variants (white). Classification of "unknown"
829 variants into their respective *Bartonella* species groups is based on pITS length.
830
831 **Figure 3** Mean percentage of individuals correctly identified to host species according
832 to linear discriminant analyses where models were trained on **(A)** *Bartonella*
833 infections of the hosts, **(B)** flea infestations of the hosts (χ^2=0.02, p=0.88) and **(C)**
834 *Bartonella* infections of the fleas infesting the hosts (χ^2=28.7, p<0.001), using data
835 from all three woodland sites combined. In each case, models were trained on random
836 selections of 75% of host-parasite associations and used to predict the host identity of
837 the remaining 25% of the data. This was done 1000 times in each case. Grey bars
838 represent models trained on true host-parasite associations while white bars represent
839 models trained on random host-parasite associations. Differences between the
840 predictive capabilities of each model were assessed using χ^2 analyses. In **(A)**, models
841 were trained on host *Bartonella* infections identified either to species-level
842 ("*Bartonella* species") or to pITS variant-level ("*Bartonella* variants"), and ten
843 comparisons were made, represented by the letters a-j. **a:** χ^2=61.8, p<0.001, **b:**
844 χ^2=44.8, p<0.001, **c:** χ^2=33.5, p<0.001, **d:** χ^2=30.9, p<0.001, **e:** χ^2= 33.8, p<0.001, **f:**

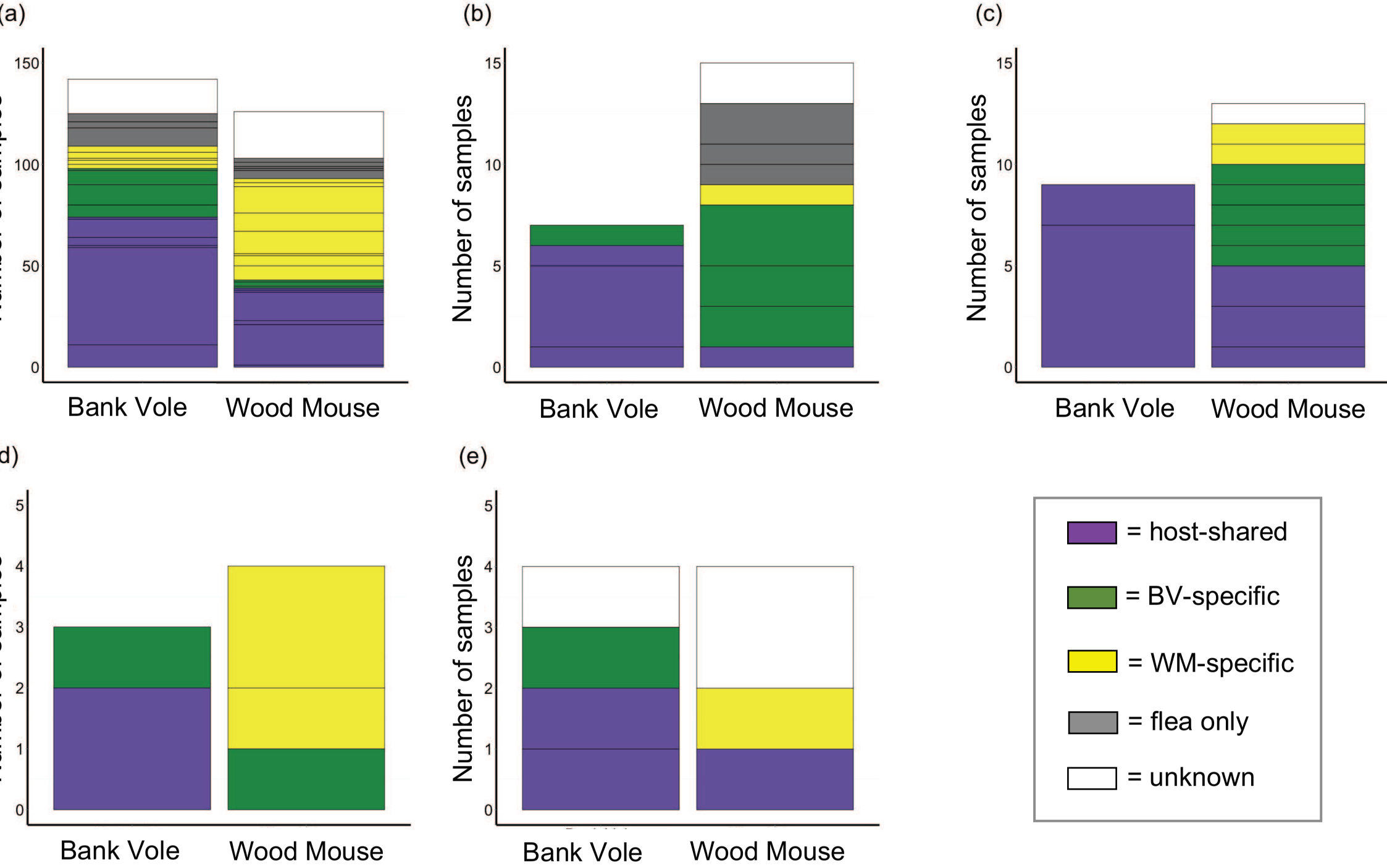
845 χ^2=83.7, p<0.001, **g:** χ^2=26.8, p<0.001, **h:** χ^2=0.99, p=0.32, **i:** χ^2=19.5, p<0.001, **j:**
846 χ^2=31.2, p<0.001. LDA models could not be computed for *B. birtlesii* variants alone
847 as the distribution of the one variant shared between host species was highly skewed
848 (birtlesii-1, found only twice in wood mice but 50 times in bank voles; Table S2).
849
850 **Figure 4** The proportion of flea-infested wood mice and bank voles that were infested
851 with at least one specimen of each species of flea detected in this study.
852
853 **Figure 5** The number of (a) *C. n. vulgaris* (b) *M. turbidus* (c) *A. p. mustelae* (d) *H. t.*
854 *talpae* and (e) *R. pentacantha* taken from wood mice and bank voles that tested
855 positive for *Bartonella* infection. Colour-coding represents the host associations
856 (according to this study) of the *Bartonella* pITS variants found within the fleas: purple
857 = found in wood mice and bank voles, green = found only in bank voles, yellow =
858 found only in wood mice, grey = found only in fleas. White represents infections in
859 fleas that were not sequenced. Horizontal divisions within colour blocks represent
860 multiple pITS variants within a host-association category. The specific identities of
861 variants identified in each flea species collected from each host species are shown in
862 Table S8.


0.4
0.3
0.2
0.1
0.0
Bank Vol
Wood Mi
doshiae
doshiae−like
grahamii
taylorii
birtlesii
rochalimae−like
BGA


a)
Number of samples
200
150
100
50
0
Bank Vole
Wood Mouse
grahamii-4
grahamii-5
grahamii-1
grahamii-2
grahamii-3
unknown
(b)
Number of samples
400
300
200
100
0
Bank Vole
Wood Mouse
taylorii-3
taylorii-4
taylorii-5
taylorii-1
taylorii-2
taylorii-6
taylorii-7
taylorii-8
taylorii-9
taylorii-10
unknown
(c)
Number of samples
300
200
100
0
Bank Vole
Wood Mouse
birtlesii-1
birtlesii-2
birtlesii-3
birtlesii-4
birtlesii-5
birtlesii-6
birtlesii-7
unknown


(A)
(B)
(C)
Mean host prediction success rate (%)
100
75
50
25
0
a
b
c
d
e
f
g
h
i
j
All
Shared
All
Shared
B. grahamii
B. taylorii
All Fleas
Bartonella variants in fleas
Bartonella species
Bartonella variants
True associations
Random associations


1.0
0.8
0.6
0.4
0.2
0.0
Bank vol
Wood mi
lliger mustelae
C. nobils vulgaris
H. talpae talpae
P. sorcis
M. turbidus
R. pentacantha
T. poppei poppei


(a)
150
100
50
0
Bank Vole
Wood Mouse
(b)
Number of samples
15
10
5
0
Bank Vole
Wood Mouse
(c)
Number of samples
15
10
5
0
Bank Vole
Wood Mouse
(d)
5
4
3
2
1
0
Bank Vole
Wood Mouse
(e)
Number of samples
5
4
3
2
1
0
Bank Vole
Wood Mouse
= host-shared
= BV-specific
= WM-specific
= flea only
= unknown