

Small-scale spatial variation in infection risk shapes the evolution of a *Borrelia* resistance gene in wild rodents

Luca Cornetti^{1,2}, Daniela Hilfiker¹, Mélissa Lemoine¹ & Barbara Tschirren^{3*}

¹ Department of Evolutionary Biology and Environmental Studies, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland

² University of Basel, Zoological Institute, Vesalgasse 1, 4051 Basel, Switzerland

³ Centre for Ecology and Conservation, University of Exeter, Penryn TR10 9FE, UK

Keywords: Toll-like receptor 2 (*TLR2*), local adaptation, *Borrelia afzelii*, spatially heterogeneous natural selection, elevational gradients, *Myodes glareolus*, host-parasite interactions, wildlife disease ecology, parasite-mediated selection, immune system evolution

*Correspondence:

Barbara Tschirren, Email: b.tschirren@exeter.ac.uk

Running title: Microgeographic *TLR2* evolution

Abstract

Spatial variation in pathogen-mediated selection is predicted to influence the evolutionary trajectory of host populations and lead to spatial variation in their immunogenetic composition. However, to date few studies have been able to directly link small-scale spatial variation in infection risk to host immune gene evolution in natural, non-human populations. Here we use a natural rodent-*Borrelia* system to test for associations between landscape-level spatial variation in *Borrelia* infection risk along replicated elevational gradients in the Swiss Alps and *Toll-like receptor 2* (*TLR2*) evolution, a candidate gene for *Borrelia* resistance, across bank vole (*Myodes glareolus*) populations. We found that *Borrelia* infection risk (i.e. the product of *Borrelia* prevalence in questing ticks and the average tick load of voles at a sampling site) was spatially variable and significantly negatively associated with elevation. Across sampling sites, *Borrelia* prevalence in bank voles was significantly positively associated with *Borrelia* infection risk along the elevational clines. We observed a significant association between naturally occurring *TLR2* polymorphisms in hosts and their *Borrelia* infection status. The *TLR2* variant associated with a reduced likelihood of *Borrelia* infection was most common in rodent populations at lower elevations that face a high *Borrelia* infection risk, and its frequency changed in accordance with the change in *Borrelia* infection risk along the elevational clines. These results suggest that small-scale spatial variation in parasite-mediated selection affects the immunogenetic composition of natural host populations, providing a striking example that the microbial environment shapes the evolution of the host's immune system in the wild.

Introduction

The immune system plays a key role in host defence and pathogens are therefore predicted to be major drivers of immune gene evolution (Murphy *et al.* 2008; Barreiro & Quintana-Murci 2010; Paterson & Piertney 2011). Indeed, molecular patterns of positive or balancing selection have been detected in a range of immune genes across taxa (e.g. Piertney & Oliver 2006; Ferrer-Admetlla *et al.* 2008; Tschirren *et al.* 2011a).

The strength of pathogen-mediated selection acting on host populations is, however, often spatially heterogeneous and influenced by local differences in infection risk. Such spatial variation in pathogen-mediated selection is predicted to influence the evolutionary trajectories of host populations and ultimately lead to spatial variation in the (immuno)genetic composition of natural populations (i.e. local adaptation; Barrett & Schluter 2008; Richardson *et al.* 2014). To date, however, few studies have been able to link small-scale spatial variation in pathogen infection risk to host immune gene evolution in natural, non-human populations (but see Gonzalez-Quevedo *et al.* 2016; see also Kwiatkowski 2005; Ferwerda *et al.* 2007; Garamszegi 2014 for examples of adaptive immune gene evolution in humans). This is likely due to a lack of information on spatial variation in pathogen abundance and host exposure, but also because of a lack of candidate immune genes for most wildlife diseases (Spurgin & Richardson 2010).

Borrelia afzelii is a common tick-transmitted pathogen in rodents and one of the causative agents of human Lyme borreliosis, the most common vector-borne disease on the Northern hemisphere (Steere *et al.* 2004; Mead

2015). A recent experimental study in one of the main rodent hosts of *B. afzelii* in Europe, the bank vole (*Myodes glareolus*) (Kurtenbach *et al.* 1995; Hellgren *et al.* 2011), showed that an infection with this spirochete has negative fitness consequences for its natural rodent host, demonstrating that *Borrelia* is an agent of selection for host resistance or tolerance mechanisms in this species (Cayol 2017).

In Europe, the main vector of *Borrelia sp.* is the sheep tick *Ixodes ricinus*, a hard-shelled tick that can infest a wide range of vertebrates (Mannelli *et al.* 2012). The distribution and abundance of *Ixodes* ticks is strongly influenced by climatic factors, in particular temperature and humidity (Burri *et al.* 2007). Average temperature decreases with increasing elevation and accordingly *I. ricinus* abundance decreases with increasing elevation (Jouda *et al.* 2004; Materna *et al.* 2005, Lemoine *et al.* 2018). Furthermore, *Borrelia* prevalence in *I. ricinus* nymphs significantly decreases with increasing elevation (Lemoine *et al.* 2018). The combination of these two factors, a lower vector abundance as well as a lower *Borrelia* prevalence in these vectors, suggests that the risk of *Borrelia* infection is spatially heterogeneous and decreases along elevational clines. Consequently, *Borrelia*-mediated selection acting on host defence mechanisms is predicted to be strongest at low elevations and to decrease with increasing elevation. Elevational clines thus provide a unique opportunity to quantify how small-scale variation in pathogen pressure affects the evolution of resistance genes in natural host populations.

Toll-like receptors (*TLRs*) are innate immune receptors that play a key role in pathogen recognition and in the initiation of innate and acquired

immune responses (Akira & Takeda 2004). Polymorphisms at *TLRs* are associated with resistance or susceptibility to many infectious diseases (Schröder & Schumann 2005; Netea *et al.* 2012). Most mammals have 10-12 different *TLRs* (Roach *et al.* 2005), each recognising different pathogen ligands (Akira & Takeda 2004). *TLR2*, which recognizes lipoproteins from bacterial cell walls (Ma & Weis 1993; Garantziotis *et al.* 2008), has been identified as a candidate gene for *Borrelia* resistance in laboratory mice (Hirschfeld *et al.* 1999; Roper *et al.* 2001; Alexopoulou *et al.* 2002; Wooten *et al.* 2002; Dennis *et al.* 2009). Transferring this knowledge from the laboratory to the field we have recently shown that naturally occurring variants of *TLR2* are associated with a reduced likelihood of *Borrelia* infection also in natural, wild-living host populations (Tschirren *et al.* 2013). Furthermore, evidence for an association between *Borrelia* infection risk and the frequency of the protective *TLR2* variant has been found across rodent populations on a large spatial scale (i.e. across Europe; Tschirren 2015).

Here we use a natural rodent-*Borrelia* system to test for associations between small-scale spatial variation in *Borrelia* infection risk along elevational gradients in the Swiss Alps and patterns of *TLR2* evolution across natural host populations. We predict that *Borrelia* infection risk is highest at low elevations and decreases with increasing elevation. Furthermore, we predict that selection acting on host *TLR2* will change accordingly, resulting in spatial variation in the immunogenetic composition of natural host populations along elevational clines.

Material and Methods

Sample collection

The bank vole (*Myodes glareolus*) is one of the main hosts of *B. afzelii* in Europe (Kurtenbach *et al.* 1995; Hellgren *et al.* 2011). We caught bank voles at four locations (i.e. valleys) in the Kanton Graubünden, Switzerland using live-traps (Longworth Mammal Traps, Anglian Lepidopterist Supplies, Norfolk, UK). At each location three sampling sites, one at low (622-702 metres above sea level (masl)), one at middle (1096-1217 masl), and one at high (1471-1737 masl) elevation, were identified (12 sampling sites in total; Fig. 1, Table 1). Trapping sessions were performed at each of these 12 sites for three to five nights from June to September 2014. We sexed and weighed (± 0.1 g) the caught bank voles (N = 266) and determined their tick load by counting the number of ticks on their ears (Tschirren *et al.* 2013). We then took a small ear biopsy for DNA extraction and the detection of *Borrelia* infection, and stored it in 95% ethanol. After these procedures, the animals were released at their capture site. Age classes were defined as adults (≥ 20 g), subadults (15-20 g), and juveniles (≤ 15 g) following Gliwicz (1988). Animal catching, handling and sampling were performed under the permit 2012_17, issued by the Department of Food Safety and Animal Health, Kanton Graubünden, Switzerland.

Spatial variation in Borrelia infection risk

Previous studies have found that both the abundance of the tick vector *Ixodes ricinus* (Jouda *et al.* 2004; Materna *et al.* 2005; Lemoine *et al.* 2018) as well as *Borrelia* prevalence in questing *I. ricinus* ticks (Lemoine *et al.* 2018) are

highest at low elevations and decrease with increasing elevation. We therefore used elevation (masl) as a first proxy of spatially variable infection risk. Elevation is an indirect measure of spatially variable infection risk, but more robust against short-term fluctuations in pathogen abundance across years and sampling errors associated with more direct measures of infection risk (see below).

Second, to test more directly if small-scale spatial variation in *Borrelia* infection risk shapes the evolution of host *TLR2*, we determined local *Borrelia* infection risk by calculating $\log_{10}(\text{average tick load on bank voles} \times \textit{Borrelia} \text{ prevalence in questing ticks} + 1)$ for each sampling site (hereafter referred to as '*Borrelia* infection risk'). Data on *Borrelia* prevalence in questing *I. ricinus* at our sampling sites were obtained from Lemoine *et al.* (2018). This second measure is a more direct quantification of spatial variation in *Borrelia* infection risk, but it is associated with a higher measurement error because voles and ticks were sampled in only one season and sampled individuals represent only a subsample of the entire host population at a site. Importantly, both measures of spatial variation in infection risk are independent of host resistance or resilience. Statistical analyses were performed with both measures of spatially variable infection risk (see below) to confirm the robustness of the results.

Borrelia infection in bank voles

Genomic DNA of bank voles was extracted from the ear biopsies using Qiagen Biosprint 96 DNA Blood Kit (Qiagen, Venlo, the Netherlands). We determined the *B. afzelii* infection status of bank voles (N = 266, 13-37

individuals per site) using a *flaB* -specific quantitative PCR (qPCR) as described in Råberg (2012). We focused on *B. afzelii* because a pilot study using reverse line blot (RLB, Herrmann *et al.* 2013) had revealed that *B. afzelii* is the only *Borrelia* genospecies present in bank voles at our study sites (unpublished data).

The amplification was carried out in a total volume of 20 µl, including 10 µl SYBR® Select Master Mix (2x, Applied Biosystems), 0.8 µl of each primer (10 µM) and 4 µl extracted genomic DNA on a StepOnePlus® Real-Time PCR machine (Applied Biosystems, Foster City, CA, USA). The following primers were used: Fla5F: 5'-CACCAGCATCACTTTCAGGA-3' and Fla6R: 5'-CTCCCTCACCAGCAAAAAGA-3' (Råberg 2012). The qPCR protocol consisted of two initial holding steps first at 50 °C and then at 95 °C for 2 min. These steps were followed by 42 cycles of 95 °C for 15 sec, 59 °C for 30 sec, and 72 °C for 30 sec (Råberg 2012). Eight negative controls and eight serially diluted positive controls were included on each plate. Samples with a melting temperature between 76.4 °C and 77.8 °C were considered to be *B. afzelii*-positive (Råberg 2012). All samples were analyzed in duplicate on two different plates. For 92% of the samples, the results were identical on the two plates (i.e. either twice negative or twice positive). The remaining 8% of samples were re-run a third time. All of these samples turned out to be *Borrelia*-negative upon reanalysis.

Bank vole TLR2 sequencing

We sequenced a 1173 bp long fragment of *TLR2* (from bp 691 to 1863), which contains all sites involved in ligand-binding (Gautam *et al.* 2006; Jin *et*

al. 2007). Amplification was carried out in a total volume of 10 μ l, including 0.2 μ l JumpStart Taq DNA Polymerase (2.5 U/ μ l; Sigma-Aldrich, St. Louis, MO, USA), 1 μ l dNTP mix (2 mM), 0.3 μ l of each primer (10 μ M) and 1 μ l of genomic DNA (approximately 35 ng). The following primers were used: Mg1TLR2F: 5'-CATCCATCACCTGACCCTTC-3' and Mg1TLR2R: 5'-CCAGTAGGAATCCTGCTCG-3' (Tschirren *et al.* 2012). The PCR protocol started with an initial denaturation step at 94 °C for 1 min, followed by 35 cycles at 94 °C for 30 sec, 57 °C for 30 sec and 72 °C for 90 sec, followed by a final extension step at 72 °C for 10 min (Tschirren *et al.* 2011b). Before sequencing, PCR products were purified by adding 0.5 μ l rAPid Alkaline phosphatase (1 U/ μ l; Roche Diagnostics, Mannheim, Germany), 0.25 μ l Exonuclease I (20 U/ μ l, Roche Diagnostics, Mannheim, Germany) and 9.25 μ l RNase-free water (Qiagen, Hilden, Germany) and incubated at 37 °C for 45 min. Subsequently, the added enzymes were inactivated at 80 °C for 15 min. PCR products were sequenced in both directions using Big Dye terminator version 3.1 (Applied Biosystems) on an ABI Prism 3730 capillary sequencer (Applied Biosystems). Raw sequences were edited, assembled and aligned in GENEIOUS version 5.6.7 (Kearse *et al.* 2012). Polymorphisms were identified automatically by the program and confirmed by visual inspection of the sequences.

TLR2 haplotypes were reconstructed using the program PHASE 2.1, integrated in DnaSP 5.10.1 (Librado & Rozas 2009), using default settings. For haplotypes that occurred only once, the PCR and sequencing reactions were repeated in order to confirm the sequences. A median-joining haplotype network was generated using Popart, (Leigh & Bryant 2015), including all

bank vole *TLR2* sequences deposited in GenBank (Fig. S1, Table S1). *TLR2* clusters were defined as groups of haplotypes that are separated from other groups of haplotypes by at least seven mutations (*sensu* Morger *et al.* 2015). The total number of *TLR2* haplotypes (h) within each population, the number of haplotypes considering only nonsynonymous substitutions (h_{nons}), haplotype diversity (H), and nucleotide diversity (π) were determined using DnaSP 5.10.1 (Librado & Rozas 2009). Newly described haplotypes were submitted to NCBI GenBank (see Table S1 for accession numbers).

Statistical analyses

Spatial variation in Borrelia infection risk

To test for spatial variation in *Borrelia* infection risk along elevational clines we used a general linear mixed model that included elevation (masl) of the sampling site and elevation² as explanatory variables and location as a random effect. Elevation was standardised to ensure independence between the main effect and its quadratic term.

Borrelia prevalence in bank vole populations

To test if *Borrelia* prevalence in vole populations varies across elevations we ran a generalised linear mixed model with a binomial error structure that included elevation (masl) of the sampling site and elevation² as explanatory variables and location as a random effect. Individual ID was included as an additional random effect to account for overdispersion in this model.

Second, to test more directly if spatial variation in *Borrelia* infection risk explains variation in *Borrelia* prevalence across vole populations, we ran the

same model as described above, but replaced elevation (masl) with *Borrelia* infection risk. No evidence for overdispersion was found in this model.

TLR2 genotype and Borrelia infection in bank voles

To test for a relationship between the *Borrelia* infection status of a vole (*Borrelia*-infected / non-infected) and its *TLR2* genotype we ran generalised linear mixed models with a binomial error structure that included sex of the vole (male / female), elevation of the sampling site, *TLR2* genotype, and the number of *TLR2* clusters observed at a location (two or three; see Results) as explanatory variables and location as a random effect. *TLR2* genotype was defined in two different ways: first, as the presence or absence (coded as 0 or 1) of the different *TLR2* haplotype clusters ($TLR2_{c1}$, $TLR2_{c3}$ and $TLR2_{c4}$, see Results), and second, as the cumulative number (coded as 0, 1 or 2) of the different *TLR2* haplotype clusters ($TLR2_{c1}$, $TLR2_{c3}$ and $TLR2_{c4}$, see Results). We also tested for a possible heterozygote advantage of individuals carrying haplotypes belonging to one (i.e. 'homozygous'; coded as 0) or two (i.e. 'heterozygous'; coded as 1) *TLR2* clusters. All *TLR2* clusters were initially included in the model but removed from the final model if $P > 0.10$. All two-way interactions were initially also included, but dropped from the final models because $P > 0.10$.

Borrelia prevalence in young animals is substantially lower than in adults (Tschirren *et al.* 2013; Supplementary analyses S1), which is likely due to a lower exposure. Only adults (N = 217) were therefore used to test for an association between *TLR2* variants and *Borrelia* infection (Tschirren *et al.* 2013). There was no significant difference in *Borrelia* infection prevalence

across the season and no significant elevation x season interaction (Supplementary analyses S2).

Spatial variation in host immunogenetic composition

To test if small-scale spatial variation in infection risk is associated with variation in the frequency of the *TLR2*_{c4} variant across host populations, we ran generalised linear mixed models with a binomial error structure that included either elevation (masl) and its quadratic term or *Borrelia* infection risk at a sampling site as fixed effect and location as a random effect. No evidence for overdispersion was found in these models. Additional analyses of molecular footprints of selection acting on host *TLR2* are presented in the Supplements (Supplementary analyses S2).

In all analyses, significance of fixed effects was determined by comparing nested models with and without the factor of interest using likelihood ratio tests. All statistical analyses were performed in R 3.3.3 (R Core Team 2016) using the package lme4 (Bates *et al* 2015).

Results

Spatial variation in *Borrelia* infection risk

Borrelia infection risk was highest at low and middle elevation sites and decreased significantly with increasing elevation ($b = -0.474$, $\chi^2 = 4.963$, $P = 0.026$; Fig. 2). Including the quadratic term of elevation in the model furthermore revealed that the relationship was non-linear ($\chi^2 = 12.261$, $P <$

0.001) with a steeper decrease in *Borrelia* infection risk between mid and high elevation sites than between low and mid elevation sites ($\chi^2 = 4.380$, $P = 0.036$; Fig. 2).

Borrelia prevalence across bank vole populations

28% (range: 0 - 55%) of bank voles were *B. afzelii* infected across sampling sites (Table 1). *Borrelia* prevalence in voles was highest at low elevation sites and decreased significantly with increasing elevation ($\chi^2 = 10.618$, $P = 0.001$; Fig. 3A). Including the quadratic term of elevation in the model furthermore revealed that the relationship was non-linear ($\chi^2 = 12.261$, $P < 0.001$) with a steeper decrease in *Borrelia* prevalence in voles between mid and high elevation sites than between low and mid elevation sites (Fig. 3A). The highest elevation at which a *Borrelia*-infected vole was caught was 1680 masl (in PAH).

Similarly, we observed a significant positive relationship between the *Borrelia* infection risk at a site and *Borrelia* prevalence in the local bank vole population ($\chi^2 = 26.649$, $P < 0.001$; Fig. 3B). There was no indication that this relationship is non-linear (*Borrelia* infection risk²: $\chi^2 = 0.575$, $P = 0.448$; Fig. 3B). Full model outputs are presented in Table S2.

The same patterns were observed when restricting the analyses to adult bank voles (Supplementary analyses S1; Fig. S2, Fig. S3).

TLR2 genotype and *Borrelia* infection in bank voles

Thirty-four unique *TLR2* haplotypes were found in the 12 analyzed bank vole populations (N = 266 animals). Of those, 16 haplotypes differed at the amino

acid level. Overall, genetic variation was high within and homogeneous among populations (Table 2; Supplementary analyses S1).

All *TLR2* haplotypes grouped into the distinct haplotype clusters that have previously been described for the European bank vole (Morger *et al.* 2015; Tschirren 2015; Fig. S1). Haplotypes belonging to clusters *TLR2*_{c1} and *TLR2*_{c3} were present in all populations (Table 2). Surprisingly, cluster *TLR2*_{c2}, which is common in bank vole populations across Europe (Morger *et al.* 2015; Tschirren 2015) and has previously been found to confer partial qualitative resistance against *Borrelia* (Tschirren *et al.* 2013) was absent in the study populations. Instead, cluster *TLR2*_{c4}, which has previously been described only from Italian bank vole populations (Morger *et al.* 2015) was found in Flims and Feldis valleys (Table 2).

We observed a significant association between the *TLR2* genotype of a bank vole and its *Borrelia* infection status. Adult bank voles that carried *TLR2*_{c4} were significantly less likely to be *Borrelia* infected (30.4% infected, N = 23) than animals without *TLR2*_{c4} (33.6% infected, N = 194; $\chi^2 = 4.331$, $P = 0.037$; no animal was homozygote for *TLR2*_{c4}). No association between *TLR2*_{c1} or *TLR2*_{c3} and *Borrelia* infection (all $P > 0.320$) and no association between zygosity state and *Borrelia* infection ($\chi^2 = 0.115$, $P = 0.734$) was observed. Full model outputs are presented in Table S2.

Spatial variation in host immunogenetic composition

In line with the hypothesis that small-scale spatial variation in pathogen-mediated selection shapes the immunogenetic composition of host populations, we found that *TLR2*_{c4} was most common in bank vole

populations at low elevations, and its frequency significantly decreased with increasing elevation (elevation: $\chi^2 = 13.348$, $P < 0.001$; elevation²: $\chi^2 = 0.857$, $P = 0.355$; Table 2).

Similarly, when replacing elevation with *Borrelia* infection risk in the model we found that the frequency of *TLR2_{c4}* in host populations significantly increased with increasing *Borrelia* infection risk at a site ($\chi^2 = 5.038$, $P = 0.025$; Table 2). Full model outputs are presented in Table S2. These patterns were also observed when the analyses were restricted to the valleys where *TLR2_{c4}* did occur (elevation: $\chi^2 = 13.348$, $P < 0.001$; *Borrelia* infection risk: $\chi^2 = 5.038$, $P = 0.025$).

Discussion

Borrelia infection risk was highest at low and middle elevation sites and significantly decreased with increasing elevation. This spatial variation in *Borrelia* infection risk translated into spatial variation in *Borrelia* infection prevalence in natural rodent hosts, which was highest at low and middle elevation sites where *Borrelia* infection risk is high and significantly decreased with increasing elevation. Although infection prevalence was low at high elevation sites, a few infected bank voles were found up to 1680 masl, which is 500 - 600 m higher than the highest previously reported *Borrelia* infection in natural hosts in the Alps (Italy, 1200 masl, Pintore *et al.* 2015; Austria, 1090 masl, Sonnleitner *et al.* 2015). This finding is consistent with the hypothesis that *Borrelia* is currently undergoing a range expansion (Jaenson *et al.* 2012; Medlock *et al.* 2013) and that host populations at higher latitudes and

elevations that were previously *Borrelia*-free are newly becoming exposed. However, additional work is required to exclude alternative explanations such as differences in sampling effort among studies.

In line with previous findings (Wooten *et al.* 2002; Tschirren *et al.* 2013), we observed a significant association between naturally occurring *TLR2* variants and *Borrelia* infection, highlighting that this innate immune receptor is a target of *Borrelia*-mediated selection in natural hosts. Surprisingly however, whereas it was *TLR2_{c2}* that was associated with *Borrelia* resistance in a Swedish bank vole population (Tschirren *et al.* 2013), this variant, which is common across whole Europe (Morger *et al.* 2015), was absent at our study sites. Instead, a variant that has previously only been reported from the Italian bank vole lineage (*TLR2_{c4}*; Morger *et al.* 2015) was associated with a lower likelihood of *Borrelia* infection in our study populations. This is remarkable because bank voles at our study sites belong to the Western lineage (Cornetti *et al.* 2016), which has been separated from the Italian lineage > 0.27 Ma (Deffontaine *et al.* 2005). It suggests that either *TLR2_{c4}* is evolutionarily old (i.e. its origin predates the split of the Western and Italian bank vole lineages) but has been lost from most other European populations (Morger *et al.* 2015), that *TLR2_{c4}* has been introduced into our populations through immigrants from the Italian clade, or that pathogen-mediated selection has favoured the formation and spread of a new protective variant (*TLR2_{c4}*) after *TLR2_{c2}* has been lost through neutral processes, such as bottlenecks and drift (i.e. independent convergent evolution in populations of the Western and Italian clades). Irrespective of the exact mechanism, our findings suggest that *TLR2_{c4}* is functionally similar to *TLR2_{c2}* (i.e. both are associated with a reduced

likelihood of *Borrelia* infection (this study, Tschirren *et al.* 2013), although the effect size of $TLR2_{c2}$ is substantially larger than of $TLR2_{c4}$, suggesting that $TLR2_{c4}$ is less effective in conferring resistance against *Borrelia* than $TLR2_{c2}$), and that the maintenance, spread or introgression of the $TLR2_{c4}$ variant is favoured by natural selection when $TLR2_{c2}$ is lost. Indeed, $TLR2_{c4}$ has only been observed in populations where $TLR2_{c2}$ is absent (this study, Morger *et al.* 2015). *In vitro* analyses will be required to investigate these functional similarities (as well as the differences in effectiveness) between $TLR2_{c4}$ and $TLR2_{c2}$ further. Of particular interest for such analyses will be the non-synonymous substitution A988G, resulting in an Asn - Asp change at *Mus musculus* TLR2 amino acid position 330, which is shared by the $TLR2_{c4}$ and $TLR2_{c2}$ variants. Unfortunately, the currently known crystal structure of the *Mus musculus* TLR2 complex (PDB 5D3I) does not expand to this position, making any inference about the biological consequences of this shared substitution difficult at this point.

Although $TLR2_{c4}$ is associated with a reduced likelihood of *Borrelia* infection, the variant was comparably rare within and across host populations. This may be due to several, non-mutually exclusive processes. First, $TLR2_{c4}$ may be comparably rare because *Borrelia*-mediated selection acting on host resistance may be weak. Indeed, until recently little was known about the fitness consequences of *B. afzelii*-infection in natural rodent hosts. In a recent field study, however, Cayol (2017) documented a significant reduction in reproductive success in bank voles experimentally infected with *B. afzelii*, especially at low population densities, demonstrating that selection imposed

by *B. afzelii* on host traits that prevent or control infection is substantial. We can thus exclude this first explanation.

Alternatively, the formation or introgression of *TLR2_{c4}* could be recent and local (see also above) and given the relatively low dispersal capability of voles (Le Galliard *et al.* 2012), it might take a long time for the variant to spread to other populations. Indeed, small but significant population differentiation at neutral genetic markers has been observed among all our study populations, even among populations within the same valley (Cornetti *et al.* 2016), providing support that gene flow among populations is limited. Whereas reduced gene flow may explain why the variant is present in some valleys but not in others, importantly within valleys gene flow is higher (rather than lower) at high elevation sites (Cornetti *et al.* 2016). The low frequency of *TLR2_{c4}* at higher elevations thus cannot be explained by reduced gene flow at these sites.

Finally, *TLR2_{c4}* frequency might be low because the variant reduces *Borrelia* susceptibility, but may at the same time increase susceptibility to other pathogens and / or induce immunopathological costs. Indeed, *TLRs* activate inflammatory responses, which can induce self-harm when too strong (Netea *et al.* 2012). Such deleterious effects of exaggerated inflammatory responses may balance the benefits of reduced *Borrelia* susceptibility and lead to comparably low frequencies of variants that confer *Borrelia*-resistance (Netea *et al.* 2012). Further studies are required to explore the relative importance of these different processes in our system.

Borrelia infection risk in the environment and *Borrelia* prevalence in voles were both highest at low and middle elevation sites and decreased

significantly with increasing elevation. Given the negative fitness consequences of *Borrelia*-infection in voles (Cayol 2017), it indicates that the strength of *Borrelia*-mediated selection acting on host defence mechanisms is spatially heterogeneous and decreases with increasing elevation. In line with the hypothesis that pathogen-mediated selection is a main driver of host immune gene evolution, we found evidence that the frequency of the protective *TLR2_{c4}* variant in host populations changed in accordance with this spatially variable pathogen pressure: the frequency of the *TLR2_{c4}* variant was highest in populations at low elevation sites where *Borrelia* infection risk is high, and the change in *TLR2_{c4}* frequency across populations mirrored the change in *Borrelia* infection risk along the elevational cline (see also Tschirren 2015 for an association between *TLR2_{c2}* frequency and *Borrelia* infection risk on a large spatial scale). This finding is in line with previous studies that documented associations between small-scale spatial variation in infection risk and the frequency of resistance or susceptibility alleles in humans (e.g. Kwiatkowski 2005; Ferwerda et al. 2007; Garamszegi 2014) and birds (Gonzalez-Quevedo et al. 2016). It highlights that spatial variation in pathogen pressure can shape the immunogenetic composition of natural host populations even on a small spatial scale, leading to microgeographic adaptation.

As most immunogenetics studies in natural vertebrate populations, our study has a number of limitations. First, the study was performed during only one season and four valleys. Especially the strength of the relationship between *TLR2* genotype and *Borrelia* infection may vary across years (e.g. because of variation in food abundance). Confirming the patterns we

observed in additional valleys and across multiple years would thus be an interesting but logistically challenging next step. Second, our study is correlational and we cannot exclude the possibility that other pathogens (or other abiotic or biotic factors) that vary along elevational clines might contribute to patterns of *TLR2* evolution observed in our study. Third, we cannot exclude the possibility that variation at a gene closely linked to *TLR2* (rather than at *TLR2* itself) underlies the observed differences in resistance. However, there are several lines of evidence that support our conclusion that *Borrelia* is a key driver of host *TLR2* evolution in our system. First, the results of the statistical analyses were very similar when including elevation (which is an indirect measure of spatially variable *Borrelia* infection risk that also incorporates other potential selective agents that vary along elevational clines) or *Borrelia* infection risk (i.e. the product of *Borrelia* prevalence in questing ticks and average tick load on hosts at a site, which is a direct measure of small-scale spatial variation in host *Borrelia* exposure) as explanatory variable. It highlights that other pathogens would have to be closely linked to *Borrelia* (i.e. co-occur within ticks) to explain this pattern. No such patterns of pathogen co-occurrence were observed using 16S sequencing of ticks at our study sites (T. Aivelo *et al*, unpublished data). Second, our conclusion is supported by extensive molecular and immunological research that has identified the mechanisms by which *TLR2* recognises *Borrelia* and initiates innate and acquired immune responses against the spirochete (Hirschfeld *et al.* 1999; Wooten *et al.* 2002; Jin *et al.* 2007; Dennis *et al.* 2009), as well as QTL and knock-out studies in laboratory mice, which have identified *TLR2* as a candidate gene for *Borrelia* resistance

(Roper *et al.* 2001; Alexopoulou *et al.* 2002). Finally, negative fitness consequences of *Borrelia* infection in bank voles have been experimentally demonstrated (Cayol 2017) and significant associations between *TLR2* genotype and *Borrelia* infection status have been observed in other bank vole populations (Tschirren *et al.* 2013).

Although our conclusions are thus parsimonious and supported by molecular, immunological, biomedical as well as ecological and evolutionary research, an experimental evolution approach would be required to establish causal relationships and fully exclude alternative explanations. However, such an approach remains a major practical challenge in natural vertebrate populations in general, and wildlife disease systems in particular.

In conclusion, our study provides correlative evidence that small-scale spatial variation in pathogen-mediated selection shapes host immune gene evolution, leading to small-scale spatial variation in the immunogenetic composition of host populations. Ultimately, such differences in the strength of selection acting on immune genes will affect the evolutionary trajectories of host populations, leading to spatial heterogeneity in resistance, tolerance and life history.

Acknowledgements

We thank Kevin Reeh and numerous field assistants for help with data collection, Jennifer Morger for help in the lab, Michail Isupov for discussion of functional consequences of *TLR2* substitutions and Lars Råberg and five anonymous reviewers for comments on the manuscript. The study was

financially supported by the Swiss National Science Foundation (PP00P3_128386 and PP00P3_157455), the University of Zurich Research Priority Program 'Evolution in Action: from Genomes to Ecosystems', the Faculty of Science of the University of Zurich, the Baugarten Stiftung, the Stiftung für wissenschaftliche Forschung an der Universität Zürich (STWF 17_027) and the Georges und Antoine Claraz-Schenkung. We thank the local authorities for their permission to trap bank voles in their municipal areas.

References

- Akira S, Takeda K (2004) Toll-like receptor signalling. *Nature Reviews. Immunology*, **4**, 499–511.
- Alexopoulou L, Thomas V, Schnare M, *et al.* (2002) Hyporesponsiveness to vaccination with *Borrelia burgdorferi* OspA in humans and in TLR1- and TLR2-deficient mice. *Nature Medicine*, **8**, 878-884.
- Barreiro LB, Quintana-Murci L (2010) From evolutionary genetics to human immunology: how selection shapes host defence genes. *Nature Reviews. Genetics*, **11**, 17–30.
- Barrett RDH, Schluter D (2008) Adaptation from standing genetic variation. *Trends in Ecology and Evolution*, **23**, 38–44.
- Bates, D, Maechler M, Bolker, B, Walker, S (2015) Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, **67**, 1-48.
- Burri C, Moran Cadenas F, Douet V, Moret J, Gern L (2007) *Ixodes ricinus* density and infection prevalence of *Borrelia burgdorferi* sensu lato along a North-facing altitudinal gradient in the Rhône Valley (Switzerland). *Vector Borne and Zoonotic Diseases*, **7**, 50–8.
- Cayol, C (2017). *Eco-Epidemiology of Tick- and Rodent-Borne Pathogens in Boreal Forests* (PhD thesis). University of Jyväskylä, Jyväskylä, Finland. pp.118-157.
- Cornetti L, Lemoine M, Hilfiker D *et al.* (2016) Higher genetic diversity on mountain tops: The role of historical and contemporary processes in shaping genetic variation in the bank vole. *Biological Journal of the Linnean Society*, **118**, 233–244.
- Deffontaine V, Libois R, Kotlík P *et al.* (2005) Beyond the Mediterranean

peninsulas: Evidence of central European glacial refugia for a temperate forest mammal species, the bank vole (*Clethrionomys glareolus*).

Molecular Ecology, **14**, 1727–1739.

Dennis VA, Dixit S, O'Brien SM, *et al.* (2009) Live *Borrelia burgdorferi* spirochetes elicit inflammatory mediators from human monocytes via the Toll-like receptor signaling pathway. *Infection and Immunity*, **77**, 1238-1245.

Fay JC, Wu C-I (2000) Hitchhiking under positive Darwinian selection. *Genetics*, **155**, 1405–1413.

Ferrer-Admetlla A, Bosch E, Sikora M *et al.* (2008) Balancing selection is the main force shaping the evolution of innate immunity genes. *The Journal of Immunology*, **181**, 1315–1322.

Ferwerda B, McCall MBB, Alonso S *et al.* (2007) TLR4 polymorphisms, infectious diseases, and evolutionary pressure during migration of modern humans. *Proceedings of the National Academy of Sciences of the United States of America*, **104**, 16645–16650.

Le Galliard JF, Rémy A, Ims RA, Lambin X (2012) Patterns and processes of dispersal behaviour in arvicoline rodents. *Molecular Ecology*, **21**, 505–523.

Garamszegi LZ (2014) Global distribution of malaria-resistant MHC-HLA alleles: the number and frequencies of alleles and malaria risk. *Malaria Journal*, **13**, 349.

Garantziotis S, Hollingsworth JW, Zaas AK, Schwartz DA (2008) The effect of toll-like receptors and toll-like receptor genetics in human disease. *Annual Review of Medicine*, **59**, 343–59.

- Gautam JK, Ashish, Comeau LD, Krueger JK, Smith MF (2006) Structural and functional evidence for the role of the TLR2 DD loop in TLR1/TLR2 heterodimerization and signaling. *Journal of Biological Chemistry*, **281**, 30132–30142.
- Gliwicz J (1988) Seasonal dispersal in non-cyclic populations of *Clethrionomys glareolus* and *Apodemus flavicollis*. *Acta Theriologica*, **33**, 263–272.
- Gonzalez-Quevedo, C, Davies RG, Phillips KP, Spurgin LG & Richardson DS (2016) Landscape-scale variation in an anthropogenic factor shapes immune gene variation within a wild population. *Molecular Ecology*, **25** (17), 4234-4246.
- Hedrick PW (2005) *Genetics of Populations*. Jones & Bartlett Publishers, Sudbury, MA.
- Hellgren O, Andersson M, Råberg L (2011) The genetic structure of *Borrelia afzelii* varies with geographic but not ecological sampling scale. *Journal of Evolutionary Biology*, **24**, 159-167.
- Herrmann C, Voordouw MJ, Gern L (2013) *Ixodes ricinus* ticks infected with the causative agent of Lyme disease, *Borrelia burgdorferi* sensu lato, have higher energy reserves. *International Journal for Parasitology*, **43**, 477–483.
- Hirschfeld M, Kirschning CJ, Schwandner R, *et al.* (1999) Cutting edge: Inflammatory signaling by *Borrelia burgdorferi* lipoproteins is mediated by Toll-like receptor 2. *Journal of Immunology*, **163**, 2382-2386.
- Jin MS, Kim SE, Heo JY *et al.* (2007) Crystal structure of the TLR1-TLR2 heterodimer induced by binding of a tri-acylated lipopeptide. *Cell*, **130**,

1071–1082.

Jouda F, Perret J-L, Gern L (2004) *Ixodes ricinus* density, and distribution and prevalence of *Borrelia burgdorferi* sensu lato infection along an altitudinal gradient. *Journal of Medical Entomology*, **41**, 162–9.

Kearse M, Moir R, Wilson A *et al.* (2012) Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*, **28**, 1647–1649.

Kurtenbach K, Kampen H, Dizij A, *et al.* (1995) Infestation of rodents with larval *Ixodes ricinus* (Acari, Ixodidae) is an important factor in the transmission cycle of *Borrelia burgdorferi* s.l. in German woodlands. *Journal of Medical Entomology*, **32**, 807-817.

Kwiatkowski DPP (2005) How malaria has affected the human genome and what human genetics can teach us about malaria. *American Journal of Human Genetics*, **77**, 171–190.

Leigh JW, Bryant D (2015) Popart: Full-feature software for haplotype network construction. *Methods in Ecology and Evolution*, **6**, 1110–1116.

Lemoine M, Cornetti L, Tschirren B (2018) Does *Borrelia burgdorferi* sensu lato facilitate the colonisation of marginal habitats by *Ixodes ricinus*? A correlative study in the Swiss Alps. *bioRxiv*. doi.org/10.1101/273490

Li H (2011) A new test for detecting recent positive selection that is free from the confounding impacts of demography. *Molecular Biology and Evolution*, **28**, 365–75.

Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics (Oxford, England)*, **25**, 1451–2.

Ma Y, Weis JJ (1993) *Borrelia burgdorferi* outer surface lipoproteins OspA

- and OspB possess B-cell mitogenic and cytokine-stimulatory properties. *Infection and immunity*, **61**, 3843–53.
- Mannelli A, Bertolotti L, Gern L, Gray J (2012) Ecology of *Borrelia burgdorferi* sensu lato in Europe: transmission dynamics in multi-host systems, influence of molecular processes and effects of climate change. *FEMS Microbiology Reviews*, **36**, 837–61.
- Materna J, Daniel M, Danielová V (2005) Altitudinal distribution limit of the tick *Ixodes ricinus* shifted considerably towards higher altitudes in central Europe: results of three years monitoring in the Krkonose Mts. (Czech Republic). *Central European Journal of Public Health*, **13**, 24–8.
- Mead PS (2015) Epidemiology of Lyme Disease. *Infectious Disease Clinics of North America*, **29**, 187–210.
- Medlock JM, Hansford KM, Bormane A *et al.* (2013) Driving forces for changes in geographical distribution of *Ixodes ricinus* ticks in Europe. *Parasites & Vectors*, **6**, 1.
- Morger J, Råberg L, Hille SM *et al.* (2015) Distinct haplotype structure at the innate immune receptor Toll-like receptor 2 across bank vole populations and lineages in Europe. *Biological Journal of the Linnean Society*, **116**, 124–133.
- Murphy K, Travers P, Walport M (2008) *Janeway's immunobiology*. Garland Science, New York.
- Netea MG, Wijmenga C, O'Neill LAJ (2012) Genetic variation in Toll-like receptors and disease susceptibility. *Nature Immunology*, **13**, 535–42.
- Paterson S, Piertney SB (2011) Frontiers in host-parasite ecology and evolution. *Molecular Ecology*, **20**, 869–71.

- Piertney SB, Oliver MK (2006) The evolutionary ecology of the major histocompatibility complex. *Heredity*, **96**, 7–21.
- Pintore MD, Ceballos L, Iulini B *et al.* (2015) Detection of invasive *Borrelia burgdorferi* strains in north-eastern Piedmont, Italy. *Zoonoses and Public Health*, **62**, 365–74.
- R Core Team. 2016 R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria
- Råberg L (2012) Infection intensity and infectivity of the tick-borne pathogen *Borrelia afzelii*. *Journal of Evolutionary Biology*, **25**, 1448–53.
- Richardson, JL, Urban MC, Bolnick DI & Skelly DK (2014) Microgeographic adaptation and the spatial scale of evolution. *Trends in Ecology & Evolution* **29** (3), 165-176.
- Roach JC, Glusman G, Rowen L *et al.* (2005) The evolution of vertebrate Toll-like receptors. *Proceedings of the National Academy of Sciences of the United States of America*, **102**, 9577–82.
- Roper RJ, Weis JJ, McCracken BA, *et al.* (2001) Genetic control of susceptibility to experimental Lyme arthritis is polygenic and exhibits consistent linkage to multiple loci on chromosome 5 in four independent mouse crosses. *Genes and Immunity*, **2**, 388-397.
- Schröder NWJ, Schumann RR (2005) Single nucleotide polymorphisms of Toll-like receptors and susceptibility to infectious disease. *The Lancet. Infectious diseases*, **5**, 156–64.
- Sonnleitner ST, Margos G, Wex F *et al.* (2015) Human seroprevalence against *Borrelia burgdorferi* sensu lato in two comparable regions of the eastern Alps is not correlated to vector infection rates. *Ticks and Tick-*

- borne Diseases*, **6**, 221–227.
- Spurgin LG, Richardson DS (2010) How pathogens drive genetic diversity: MHC, mechanisms and misunderstandings. *Proceedings of the Royal Society B*, **277**, 979–88.
- Steere AC, Coburn J, Glickstein L (2004) The emergence of Lyme disease. *Journal of Clinical Investigation*, **113**, 1093–1101.
- Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, **123**, 585–95.
- Tschirren B (2015) *Borrelia burgdorferi* sensu lato infection pressure shapes innate immune gene evolution in natural rodent populations across Europe. *Biology Letters*, **11**, 20150263.
- Tschirren B, Andersson M, Scherman K et al. (2013) Polymorphisms at the innate immune receptor TLR2 are associated with *Borrelia infection* in a wild rodent population. *Proceedings of the Royal Society B*, **280**, 20130364.
- Tschirren B, Råberg L, Westerdahl H (2011a) Signatures of selection acting on the innate immunity gene Toll-like receptor 2 (TLR2) during the evolutionary history of rodents. *Journal of Evolutionary Biology*, **24**, 1232–1240.
- Tschirren B, Andersson M, Scherman K, Westerdahl H, Raberg L (2011b) Contrasting patterns of diversity and population differentiation at the innate immunity gene Toll-like receptor 2 (TLR2) in two sympatric rodent species. *Evolution*, **66**, 720–731.
- Wooten RM, Ma Y, Yoder RA et al. (2002) Toll-like receptor 2 is required for innate, but not acquired, host defense to *Borrelia burgdorferi*. *Journal of*

Immunology, **168**, 348–55.

Data accessibility

DNA sequences: GenBank accession numbers are listed in Table S1.

Additional data used for the analyses: Dryad doi:10.5061/dryad.7sr405h

Author contributions

BT designed the study, DH and ML collected samples and performed lab analyses, LC, DH, ML and BT analysed the data, LC and BT wrote the manuscript, all authors commented on the manuscript.

Figure 1. Location of sampling sites in the Swiss Alps.

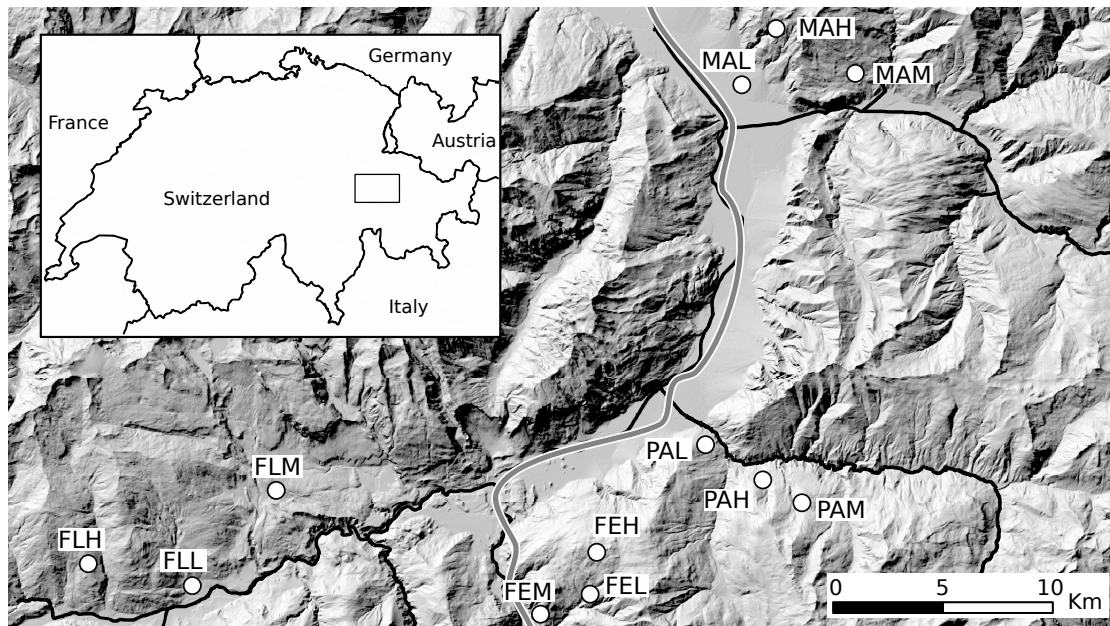


Table 1. Geographical location, elevation, coordinates, *Borrelia* infection risk, number of caught bank voles and *Borrelia* prevalence in bank voles at the different sampling sites. See Methods for the quantification of local *Borrelia* infection risk.

Location	Label	masl	GPS coordinates		<i>Borrelia</i> infection risk	Bank voles (N)	<i>Borrelia</i> infected (%)	Adult bank voles (N)	<i>Borrelia</i> infected adults (%)
			North	East					
Malans	MAL	560	46.992	9.558	0.16	33	42.4	29	44.8
Malans	MAM	1106	46.996	9.637	0.27	13	23.1	9	33.3
Malans	MAH	1774	47.020	9.583	0	18	0	16	0
Feldis	FEL	630	46.761	9.426	0.57	21	23.8	16	31.2
Feldis	FEM	1144	46.772	9.454	0.69	20	55.0	19	57.9
Feldis	FEH	1673	46.789	9.453	0	25	0	21	0
Passugg	PAL	732	46.841	9.538	0.47	23	43.5	13	61.5
Passugg	PAM	1094	46.826	9.570	0.46	18	33.3	18	33.3
Passugg	PAH	1582	46.818	9.590	0	22	4.5	16	6.2
Flims	FLL	693	46.783	9.233	0.29	37	43.2	31	48.4
Flims	FLM	1138	46.827	9.281	0.15	16	37.5	11	54.5
Flims	FLH	1454	46.795	9.169	0.03	20	25.0	18	22.2

Table 2. *TLR2* diversity within bank vole populations (all age classes). h : number of haplotypes, h_{nons} : number of haplotypes considering only nonsynonymous substitutions, H : haplotype diversity, π : nucleotide diversity. The frequency of the different *TLR2* clusters within each population is also given.

Label	Number of alleles	Frequency <i>TLR2</i> Cluster (%)						
		<i>h</i>	<i>h_{nons}</i>	<i>H</i>	π	<i>TLR2_{c1}</i>	<i>TLR2_{c3}</i>	<i>TLR2_{c4}</i>
MAL	66	9	4	0.785	0.00353	86.3	13.7	0
MAM	26	9	5	0.846	0.00586	75.0	25.0	0
MAH	36	8	5	0.805	0.00601	66.7	33.3	0
FEL	42	10	7	0.872	0.00745	50.0	33.3	16.7
FEM	40	8	5	0.819	0.00652	33.3	61.5	5.2
FEH	50	9	6	0.846	0.00689	52.0	44.0	4.0
PAL	46	8	4	0.739	0.00462	78.2	21.8	0
PAM	36	11	4	0.892	0.00652	69.4	30.6	0
PAH	44	10	5	0.903	0.00637	68.2	31.8	0
FLL	74	8	4	0.832	0.00517	75.6	9.5	14.9
FLM	32	9	5	0.865	0.00510	78.1	18.8	3.1
FLH	40	7	2	0.822	0.00498	80.0	20.0	0

Figure 2. *Borrelia* infection risk in the environment decreases with increasing elevation. Labels identify different populations (see Table 1 and Figure 1 for details).

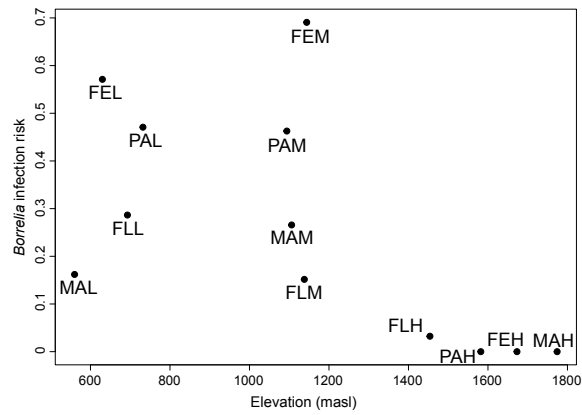
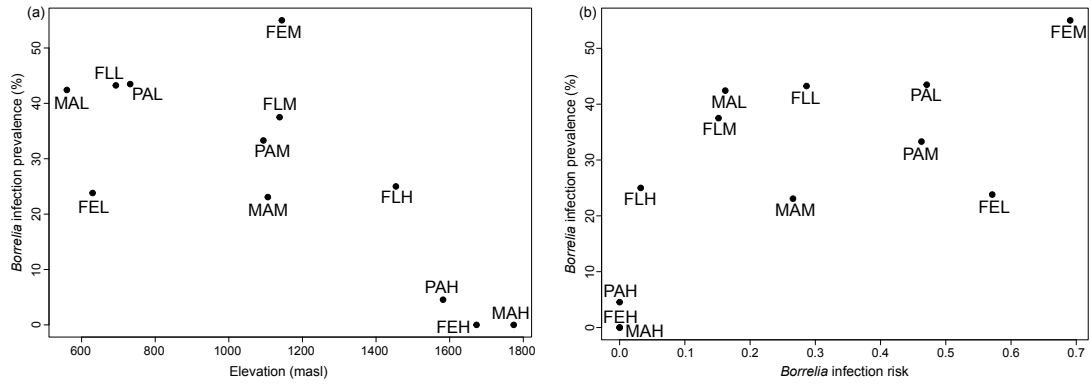


Figure 3. *Borrelia* prevalence in bank voles decreases with increasing elevation (masl) (A) and increases with increasing *Borrelia* infection risk in the environment (B). Labels identify different populations (see Table 1 and Figure 1 for details).



Supplemental Information for:

Small-scale spatial variation in infection risk shapes the evolution of a

***Borrelia* resistance gene in wild rodents**

Luca Cornetti, Daniela Hilfiker, Mélissa Lemoine & Barbara Tschirren

Figure S1. Network of bank vole *TLR2* haplotypes observed across Europe

Haplotypes belonging to different clusters are represented with different colours (Cluster 1: yellow; Cluster 2: green; Cluster 3: purple; Cluster 4: orange, following Morger et al 2015). Haplotypes that differ at the amino acid level are labelled with different numbers (e.g. H1, H2, and H3). Haplotypes that do not differ at the amino acid level are labelled with different letters (e.g. H1a, H1b, and H1c).

The circle size is relative to the number of alleles observed across our study populations. The number of mutations between two haplotypes is indicated by the number of marks on the branches. Haplotypes observed in this study are underlined and haplotypes that have been newly described in this study are in italics. See Table S1 for GenBank accession numbers.

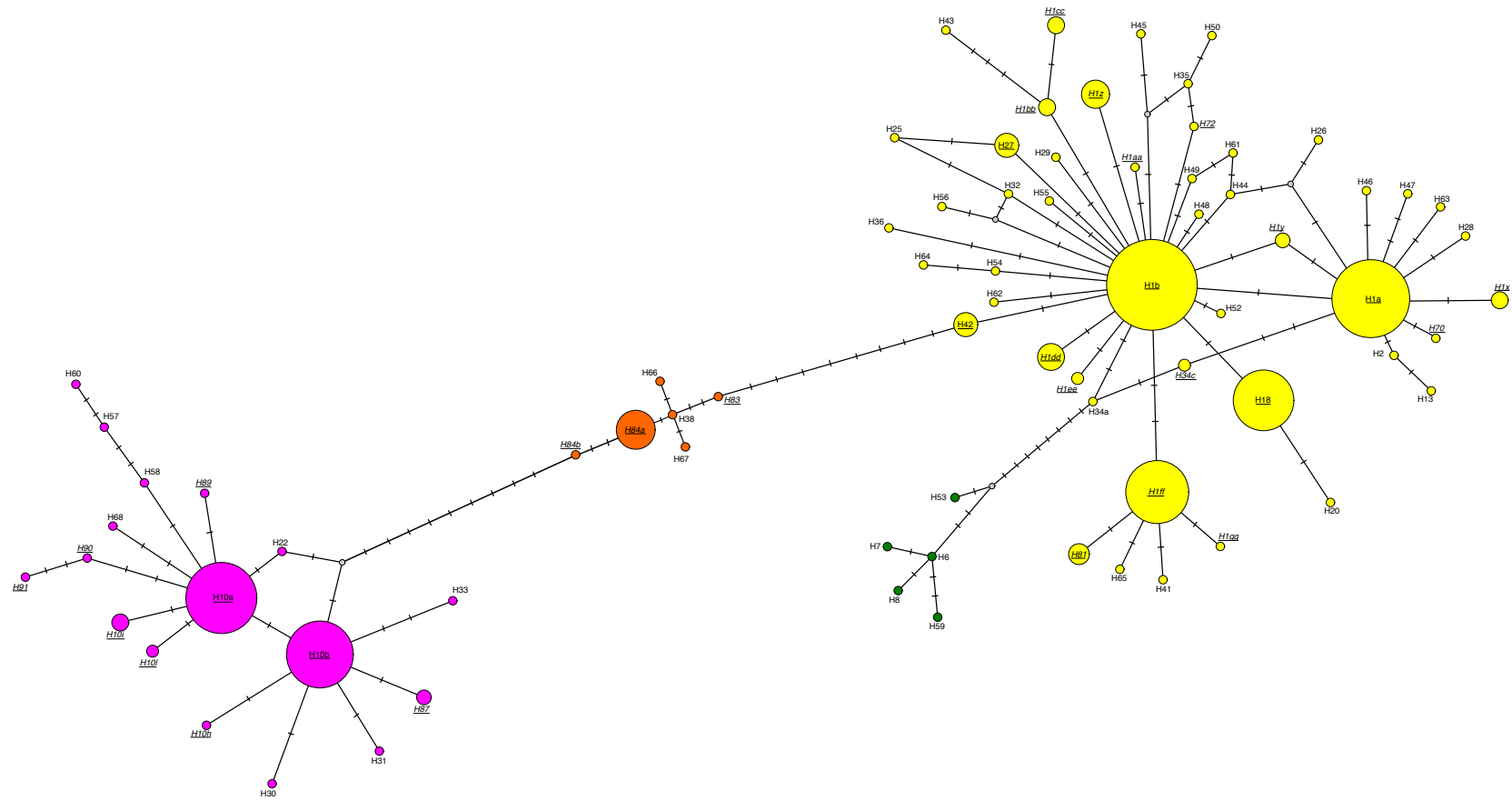


Table S1. GenBank accession numbers

GenBank accession numbers of bank vole *TLR2* sequences used in the network (Figure S1). Haplotypes that have been newly described in this study are underlined.

Haplotype	GenBank Accession	Haplotype	GenBank Accession
Hap1	JN674535	Hap60	KM979278
Hap2	JN674536	Hap61	KM979279
Hap6	JN674538	Hap62	KM979280
Hap7	JN674539	Hap63	KM979281
Hap8	JN674540	Hap64	KM979282
Hap10	JN674541	Hap65	KM979283
Hap13	JN674544	Hap66	KM979284
Hap18	KJ612421	Hap67	KM979285
Hap20	KJ612423	Hap68	KM979286
Hap21	KJ612424	<u>Hap69</u>	<u>KY660324</u>
Hap22	KJ612425	<u>Hap70</u>	<u>KY660325</u>
Hap25	KJ612428	<u>Hap71</u>	<u>KY660326</u>
Hap26	KJ612429	<u>Hap72</u>	<u>KY660327</u>
Hap27	KJ612430	<u>Hap73</u>	<u>KY660328</u>
Hap28	KJ612431	<u>Hap74</u>	<u>KY660329</u>
Hap29	KJ612432	<u>Hap75</u>	<u>KY660330</u>
Hap30	KJ612433	<u>Hap76</u>	<u>KY660331</u>
Hap31	KJ612434	<u>Hap77</u>	<u>KY660332</u>
Hap32	KJ612435	<u>Hap78</u>	<u>KY660333</u>
Hap33	KJ612436	<u>Hap79</u>	<u>KY660334</u>
Hap34	KJ612437	<u>Hap80</u>	<u>KY660335</u>
Hap35	KJ612438	<u>Hap81</u>	<u>KY660336</u>
Hap36	KJ612439	<u>Hap82</u>	<u>KY660337</u>
Hap37	KJ612440	<u>Hap83</u>	<u>KY660338</u>
Hap38	KJ612441	<u>Hap84</u>	<u>KY660339</u>
Hap41	KJ612444	<u>Hap85</u>	<u>KY660340</u>
Hap42	KJ612445	<u>Hap86</u>	<u>KY660341</u>
Hap43	KJ612446	<u>Hap87</u>	<u>KY660342</u>
Hap44	KJ612447	<u>Hap88</u>	<u>KY660343</u>
Hap45	KJ612448	<u>Hap89</u>	<u>KY660344</u>
Hap46	KJ612449	<u>Hap90</u>	<u>KY660345</u>
Hap47	KJ612450	<u>Hap91</u>	<u>KY660346</u>
Hap48	KJ612451	<u>Hap92</u>	<u>KY660347</u>
Hap49	KJ612452	<u>Hap93</u>	<u>KY660348</u>
Hap50	KJ612453		
Hap52	KJ612455		
Hap53	KJ612456		
Hap55	KJ612458		
Hap56	KJ612459		
Hap58	KJ612461		
Hap59	KJ612462		

Table S2. Full model outputs of the statistical analyses presented in the main text

a) Does a) elevation (masl) or b) *Borrelia* infection risk explain variation in *Borrelia* prevalence across vole populations?

Elevation and *Borrelia* infection risk were standardised before analysis to ensure independence of main effect and quadratic term. See main text (*Statistical analyses*) for a detailed description of the models and statistical approach.

Factors	Estimate (log odds)	SE	χ^2	<i>P</i>
<i>a.</i>				
Elevation	-1.107	0.249	10.618	0.001
Elevation ²	-0.943	0.266	12.261	<0.001
<i>b.</i>				
<i>Borrelia</i> infection risk	1.156	0.290	26.649	<0.001
<i>Borrelia</i> infection risk ²	-0.321	0.432	0.575	0.448

b) Is a) vole *TLR2* genotype (measured as: cluster absent / present), b) vole *TLR2* genotype (measured as cumulative number of alleles: 0, 1, 2) or c) vole *TLR2* heterozygosity (measured as: cluster homozygous / cluster heterozygous) associated with *Borrelia* infection status of voles? See main text (*Statistical analyses*) for a detailed description of the models and statistical approach. Results from the final models are presented for the significant terms.

Factors	Estimate (log odds)	SE	χ^2	P
<i>a.</i>				
<i>TLR2</i> _{c4}	-1.104	0.547	4.331	0.037
<i>TLR2</i> _{c1}	0.593	0.601	1.023	0.312
<i>TLR2</i> _{c3}	0.098	0.349	0.079	0.779
Cluster diversity (3)	0.577	0.341	2.839	0.092
Elevation	-0.928	0.186	30.193	<0.001
Sex (male)	-0.228	0.322	0.503	0.478

<i>b.</i>				
<i>TLR2</i> _{c4}	-1.104	0.547	4.331	0.037
<i>TLR2</i> _{c3}	-0.153	0.259	0.349	0.555
<i>TLR2</i> _{c1}	0.153	0.259	0.349	0.555
Cluster diversity (3)	0.577	0.341	2.839	0.092
Elevation	-0.928	0.186	30.193	<0.001
Sex (male)	-0.228	0.322	0.503	0.478
<i>c.</i>				
<i>TLR2</i> heterozygosity (heterozygous)	-0.106	0.312	0.115	0.734
Cluster diversity (3)	0.327	0.316	1.083	0.298
Elevation	-0.832	0.175	26.514	<0.001
Sex (male)	-0.143	0.316	0.205	0.651

c) Does a) elevation or b) *Borrelia* infection risk explain variation in the frequency of the *TLR2_{c4}* variant across vole populations? See main text (*Statistical analyses*) for a detailed description of the models and statistical approach.

Factors	Estimate (log odds)	SE	χ^2	<i>P</i>
a.				
Elevation	1.009	0.319	13.348	<0.001
Elevation ²	0.479	0.530	0.857	0.355
b.				
<i>Borrelia</i> infection risk	0.6175	0.313	5.038	0.025

Supplementary analyses S1

Borrelia prevalence within and across bank vole populations

Borrelia infection prevalence differed markedly across host age classes ($\chi^2 = 14.597$, $P < 0.001$), with 33.2% of adult bank voles, but only 11.1% of subadults and 7.7% of juveniles being *Borrelia* infected. This lower infection prevalence is likely due to a lower exposure of young animals (Tschirren *et al.* 2013).

The relationship between *Borrelia* prevalence in voles and elevation ($\chi^2 = 16.075$, $P < 0.001$; Fig. S2) or *Borrelia* infection risk ($\chi^2 = 11.764$, $P < 0.001$; Fig. S3) was also observed when restricting the analysis to adult bank voles (see main text for analysis across all age classes). There was no differences in the proportion of adult bank voles across elevations ($\chi^2 = 0.513$, $P = 0.474$) or across sites with different *Borrelia* infection risk ($\chi^2 = 0.004$, $P = 0.949$).

No significant difference in *Borrelia* prevalence across locations (i.e. valleys; considering all age classes: $\chi^2 = 3.223$, $P = 0.359$; considering only adults: $\chi^2 = 3.489$, $P = 0.322$) was observed. Furthermore, there was no difference in *Borrelia* infection prevalence across the season (all bank voles: month: $\chi^2 = 1.833$, $P = 0.608$, interaction month x elevation: $\chi^2 = 0.510$, $P = 0.917$; adults only: month: $\chi^2 = 1.824$, $P = 0.610$, interaction month x elevation: $\chi^2 = 0.392$, $P = 0.942$).

TLR2 diversity across bank vole populations

There was no significant relationship between the number of sampled TLR2 alleles (i.e. number of caught bank voles; $\chi^2 = 2.636$, $P = 0.105$), the number of TLR2 haplotypes ($\chi^2 = 0.020$, $P = 0.887$), the number of TLR2 amino acid variants ($\chi^2 = 0.001$, $P = 0.995$), TLR2 haplotype diversity ($\chi^2 = 1.001$, $P = 0.317$) or TLR2 nucleotide diversity ($\chi^2 = 2.477$, $P = 0.116$) and elevation. Similarly, there was no significant relationship between the number of sampled TLR2 alleles (i.e. number of caught bank voles; $\chi^2 = 0.047$, $P = 0.829$), the number of TLR2 haplotypes ($\chi^2 = 0.044$, $P = 0.834$), the number of TLR2 amino acid variants ($\chi^2 = 0.159$, $P = 0.690$), TLR2 haplotype diversity ($\chi^2 = 0.119$, $P = 0.731$) or TLR2 nucleotide diversity ($\chi^2 = 0.101$, $P = 0.750$) and *Borrelia* infection risk.

Finally, there was no significant association between the number of sampled alleles (i.e. voles caught) and the number of TLR2 haplotypes ($\chi^2 = 0.048$, $P = 0.827$), the number of TLR2 amino acid variants ($\chi^2 = 0.100$, $P = 0.752$), TLR2 haplotype diversity ($\chi^2 = 1.096$, $P = 0.295$) or TLR2 nucleotide diversity ($\chi^2 = 2.857$, $P = 0.092$).

Figure S2. *Borrelia* prevalence in adult bank voles decreased with increasing elevation (masl). Labels identify different populations (see Table 1 and Figure 1 for details).

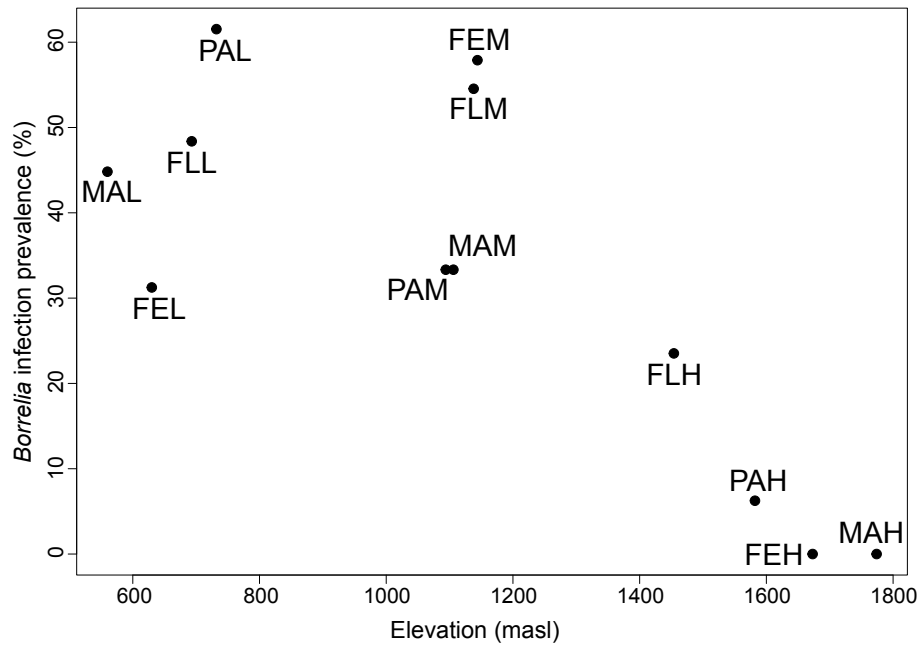
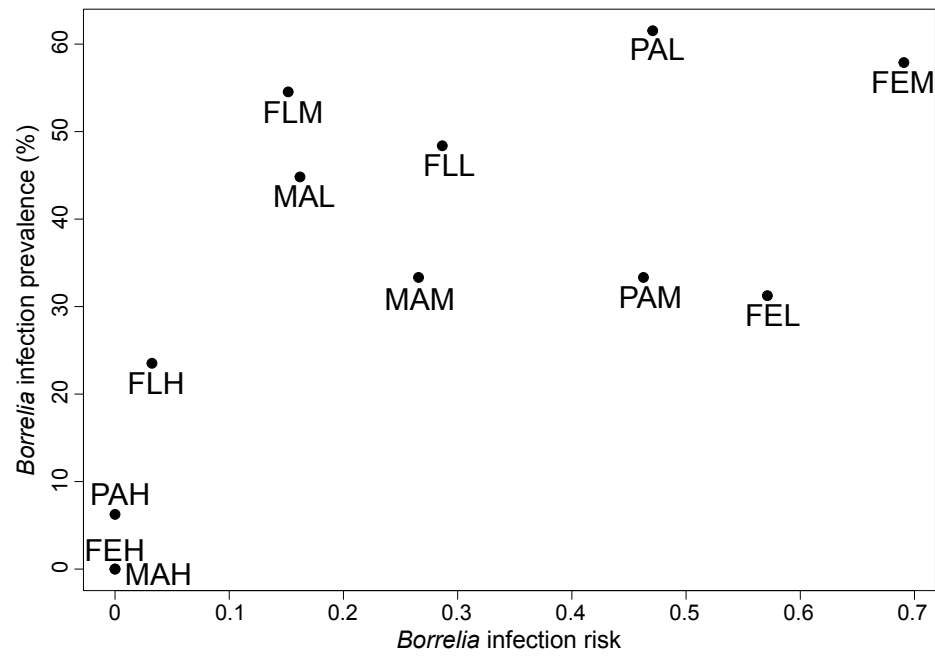


Figure S3. *Borrelia* prevalence in adult bank voles increases with increasing *Borrelia* infection risk in the environment.

Labels identify different populations (see Table 1 and Figure 1 for details).



Supplementary analyses S2

Molecular footprints of selection at TLR2

We used Fay & Wu's neutrality test (Fay & Wu 2000) to quantify if *TLR2* haplotype frequencies differ from neutral expectations across bank vole populations. Fay & Wu's H was calculated for each population in DnaSP 5.10.1 (Librado & Rozas 2009) using *Apodemus flavicollis* (JN674549.1) as an outgroup, which allows to determine the ancestral state of alleles. A positive Fay and Wu's H value suggests a lack of high-frequency derived haplotypes, which is a signal of purifying selection, while a negative Fay and Wu's H value suggests an excess of high-frequency derived haplotypes, which is a signal of positive selection (Fay & Wu 2000; Li 2011). Significance of Fay and Wu's H was determined using neutral coalescent simulations, based on the observed number of segregating sites and 10,000 replicates in DnaSP 5.10.1 (Librado & Rozas 2009).

We found that Fay and Wu's H was negative for all bank vole populations and reached statistical significance in half of them (Fig. S4).

In a next step, we used linear mixed models to test if variation in the strength of selection across populations (measured as Fay and Wu's H) is associated with elevation. Location was included as a random effect in this model.

We found a significant association between the strength of selection acting on host *TLR2* (as measured by Fay and Wu's H) and elevation (elevation: $\chi^2 = 4.707$, $P = 0.030$; elevation²: $\chi^2 = 1.644$, $P = 0.200$), with the most negative Fay and Wu's H values (i.e. the strongest selection) found at low elevation sites (Fig. S4).

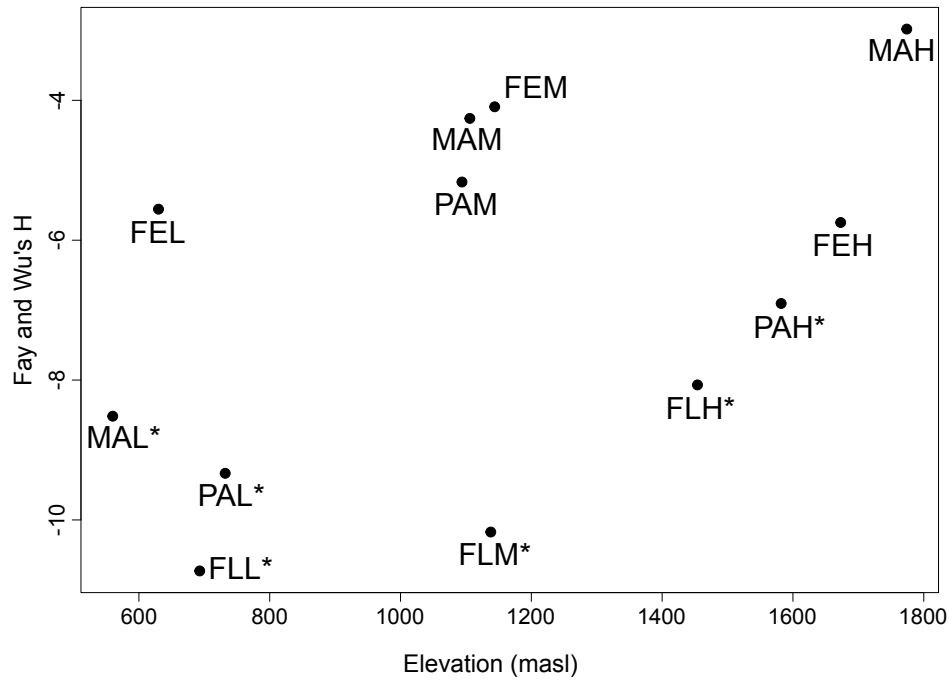
In addition, we also performed a second neutrality test, Tajima's D (Tajima 1989), for each population using DnaSP 5.10.1 (Librado & Rozas 2009). A Tajima's D value > 0 indicates a decline in population size and/or evidence of balancing selection, whilst a value < 0 is indicative of an increase in population size and/or purifying selection (Tajima 1989). In contrast to Fay & Wu's neutrality test, Tajima's D did not reach statistical significance in any population.

The discrepancy between Fay & Wu's neutrality test and Tajima's neutrality test is likely due to demographic processes (such as population size changes), which affect Tajima's D more strongly than Fay and Wu's H (Fay & Wu 2000; Hedrick 2005). Fay and Wu's neutrality test is thus more powerful to detect molecular footprints of selection in natural populations (Fay & Wu 2000; Hedrick 2005).

MFDM test is an alternative neutrality test that is 'free of confounding effects of demography' (Li 2011). However, unfortunately, it could not be performed in our study because for the majority of populations the number of sampled individuals was below the minimal sample size required for the test.

Across our study sites, significant population differentiation at neutral genetic markers was observed in all pairwise comparisons, also among populations from the same valley (Cornetti et al 2016). F_{ST} values were however comparably small, indicating that gene flow is limited but populations are not completely isolated. Genetic population clustering and / or non-complete isolation of populations could potentially have an influence on the outcome of neutrality tests. Furthermore, *TLR2* variants that confer resistance against *Borrelia* might at the same time increase susceptibility to other pathogens and / or induce immunopathological costs, which could balance selective pressures acting on *TLR2* and influence outcomes of neutrality tests (see also Discussion in main text). Caution should thus be used when interpreting the results of these neutrality tests.

Figure S4. Relationship between Fay and Wu's H and elevation (masl). Asterisk indicates Fay and Wu's H $P < 0.05$. Labels identify different populations (see Table 1 and Figure 1 for details).



Supplementary References

Fay JC, Wu CI (2000) Hitchhiking under positive Darwinian selection. *Genetics*, **155**, 1405-1413.

Hedrick PW (2005) *Genetics of Populations* Jones & Bartlett Publishers, Sudbury, MA.

Librado P, Rozas J (2009) DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, **25**, 1451-1452.

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

Tschirren B, Andersson A, Scherman K, *et al.* (2013) Polymorphisms at the innate immune receptor *TLR2* are associated with *Borrelia* infection in a wild rodent population *Proceedings of the Royal Society B*, **280**, 10.1098/rspb.2013.0364.