Uniwersytet im. Adama Mikiewicza w Poznaniu

Wydział Biologii

Pracownia Biologii Ewolucyjnej

Związek genów układu odpornościowego nornicy rudej (*Clethrionomys glareolus*) z genami białek powierzchniowych krętków *Borrelia afzelii*

Joanna Różańska-Wróbel

Rozprawa doktorska

Wykonana pod kierunkiem

prof. dr hab. Jacka Radwana

Poznań, 2025

Adam Mickiewicz University

Faculty of Biology

Evolutionary Biology Group

Associations between immune system genes of the bank vole (*Clethrionomys glareolus*) and outer surface protein genes of *Borrelia afzelii* spirochetes

Joanna Różańska-Wróbel

Doctoral thesis

Supervised by

prof. dr hab. Jacek Radwan

Poznań, 2025

Acknowledgments / Podziękowania

Dziękuję mojemu Promotorowi, prof. dr hab. Jackowi Radwanowi, za poświęcony czas, otwartość, życzliwość oraz wiele cennych wskazówek, które miały istotny wpływ na kształt tej rozprawy. Dziękuję również za stworzenie świetnego zespołu badawczego, w którym mogłam zdobywać doświadczenie i rozwijać się.

Z całego serca dziękuję moim Rodzicom – Mamie Dorocie i śp. Tacie Adamowi – za to, że zaszczepili we mnie ciekawość świata, wspierali mnie w trudnych chwilach i nieustannie motywowali do rozwoju i poszukiwania własnej drogi. To dzięki Wam odważyłam się sięgać dalej. Mamie dziękuję również za długie wieczorne rozmowy i cierpliwe wysłuchiwanie opowieści o genetyce i ewolucji.

Mojemu Mężowi Ireneuszowi dziękuję za nieocenione wsparcie, cierpliwość i zrozumienie, które pozwoliły mi z pełnym zaangażowaniem realizować się naukowo. Dziękuję Ci za to, że zawsze byłeś przy mnie i za to, że razem możemy dzielić tak wiele pasji.

Dziękuję również wszystkim współautorom oraz członkom Laboratorium Technik Biologii Molekularnej UAM – za inspirujące rozmowy, porady i wspólną pracę, która była dla mnie cennym doświadczeniem i przyjemnością. W szczególności dziękuję prof. ICHB PAN dr hab. Annie Urbanowicz za udostępnienie swojego laboratorium i pomoc w przygotowaniu analiz ELISA, prof. UAM dr hab. Mirosławie Dabert za sekwencjonowanie amplikonów CFH, dr Witoldowi Nowakowi za pomoc i cenne wskazówki przy optymalizowaniu reakcji qPCR, a także Sylwii Jedut za izolację RNA. Dziękuję także dr Katarzynie Dudek z Uniwersytetu Jagiellońskiego w Krakowie za sekwencjonowanie amplikonów OspC i OspE.

I would also like to thank the entire Evolutionary Biology Group. Your openness, commitment, and team spirit made this journey not only scientifically enriching but also personally meaningful.

Szczególnie dziękuję Agnieszce Szubert-Kruszyńskiej za pomoc w załatwianiu spraw administracyjnych i organizacyjne wsparcie oraz Karolinie Przesmyckiej za wspólne rozmowy i towarzystwo podczas zbierania próbek w trakcie wyjazdów terenowych.

Contents

Funding

This research was supported by the National Science Centre grant 2019/34/A/NZ8/00231 awarded to Jacek Radwan.

Streszczenie

Interakcje między żywicielami i pasożytami prowadzą do ich współzależnej ewolucji. Wynikająca z tych oddziaływań koewolucyjna dynamika wpływa na kształtowanie zmienności genetycznej, struktury populacji oraz rozprzestrzeniania się chorób. Może ona sprzyjać utrzymywaniu się wysokiego poziomu polimorfizmu, często przekraczającego założenia wynikające z neutralnej ewolucji. Teoria koewolucji zakłada specyficzne oddziaływania pomiędzy genami żywiciela i pasożyta, jednak niewiele jest bezpośrednich dowodów na takie interakcje, zwłaszcza wśród dzikich populacji kręgowców. Niniejsza rozprawa doktorska ma na celu wypełnienie tej luki poprzez badanie nornicy rudej (*Clethrionomys glareolus*) oraz infekującej ją bakterii *Borrelia afzelii*. Bakteria ta przenoszona jest przez kleszcze i u ludzi wywołuje chorobę zwaną boreliozą. Jednak człowiek jest przypadkowym żywicielem dla krętków *Borrelia*, a ich głównym nosicielem są małe ssaki i ptaki. Badania zawarte w rozprawie bazują na fakcie, że nornice rude są jednym z głównych, naturalnych żywicieli *B. afzelii* w Europie. W celu zgłębienia genetycznych podstaw interakcji żywiciel-patogen, praca skupia się na poszukiwaniu związków pomiędzy różnymi wariantami genów nornicy i krętków *Borrelia*, które kodują oddziałujące ze sobą białka, a także bada zmiany ekspresji genów nornic w odpowiedzi na infekcję bakterią *Borrelia*, wykorzystując sekwencjonowanie całych transkryptomów. Łącznie, oba te podejścia dążą do lepszego zrozumienia genetycznych oddziaływań gospodarz-pasożyt ukształtowanych przez koewolucję, a także zbadania szerszej perspektywy odpowiedzi immunologicznej nornic, która może leżeć u podstaw tolerancji na infekcję u tych gryzoni.

Pierwszy rozdział rozprawy skupił się na parze genów gospodarza i pasożyta, których produkty prawdopodobnie wchodzą ze sobą w bezpośrednią interakcję, napędzając wzajemną presję selekcyjną. Zbadane w nim zostały związki pomiędzy wariantami genów głównego kompleksu zgodności tkankowej (ang. *major histocompatibility complex*, MHC) nornicy rudej, który bierze udział w swoistej odpowiedzi odpornościowej organizmu, a wariantami genu ospC (ang. *outer surface protein C*) krętków *B. afzelii*, który koduje immunogenne białko znajdujące się na powierzchni tych bakterii, służące *Borrelii* do manipulacji układem odpornościowym gospodarza. OspC występował w badanej populacji nornic w 12 wariantach. Wyniki testu

Tajimy potwierdziły, że OspC ewoluuje pod wpływem doboru równoważącego. Analiza redundancji (ang. *redundancy analysis*, RDA) wykazała, że ani allele MHC klasy II, ani supertypy (grupy funkcjonalne) nie przewidywały statusu ani intensywności zakażenia, mierzonej metodą ilościowego PCR w czasie rzeczywistym (ang. *quantitative real-time PCR*, qPCR), szczepami *B. afzelii* charakteryzującymi się występowaniem różnych wariantów OspC. Niemniej jednak różnice w poziomach przeciwciał IgG przeciwko dwóm częstym w badanych populacjach wariantom OspC, wykazały statystycznie istotny związek z konkretnymi allelami MHC klasy II wśród seropozytywnych nornic. Wyniki te wskazują, że różne allele MHC mogą kształtować specyficzną dla wariantu odpowiedź immunologiczną na białko OspC krętków *Borrelia*, co jest zgodne z założeniami koewolucji gospodarz-pasożyt.

W drugim rozdziale rozprawy zbadano kolejną parę genów zaangażowanych w interakcję gospodarz-patogen: geny kodujące czynnik H (ang. *complement factor H*, CFH) nornicy rudej, który bierze udział w nieswoistej odpowiedzi immunologicznej i koduje białko regulatorowe alternatywnej ścieżki dopełniacza, oraz geny ospE (ang. *outer surface protein E*) kodujące białko powierzchniowe *B. afzelii*, które wiąże CFH gospodarza w celu uniknięcia wykrycia i zniszczenia przez jego układ odpornościowy. Wyniki testów PAML i MEME potwierdziły, że gen ospE ewoluuje pod wpływem działania doboru pozytywnego. Analiza redundancji nie wykazała jednak żadnych zależności pomiędzy wariantami genów CFH i należących do tej samej rodziny genów związanych z CFH (ang. *factor H-related*, FH-R), a wariantami OspE. Ponadto sieć haplotypów, skonstruowana w oparciu o sekwencje OspE, nie wykazała geograficznego zróżnicowania częstości wariantów OspE, które we wcześniejszych badaniach zaobserwowano dla CFH, a analiza ko-korespondencji (ang. *co-correspondence analysis*, CoCA) nie wykryła żadnej wspólnej struktury genetycznej między CFH i OspE. Wcześniejsze badania wykazały, że CFH charakteryzuje się silniejszą strukturą niż średnia genomowa u nornicy rudej, prawdopodobnie z powodu lokalnej adaptacji. Wyniki przedstawione w rozprawie wskazują natomiast, że OspE nie odzwierciedla tej struktury, co sugeruje, że ewolucja tych dwóch genów może być kształtowana przez odrębne presje selekcyjne.

W rozdziale trzecim wykorzystano sekwencjonowanie RNA w celu zbadania zmian ekspresji genów nornicy rudej w odpowiedzi na zakażenie *B. afzelii*. Pozwoliło to zidentyfikować geny zaangażowane w obronę żywiciela przed infekcją, a także zgłębić

w jaki sposób odpowiedź odpornościowa tego naturalnego gospodarza różni się od tej zgłaszanej dla przypadkowych żywicieli, takich jak człowiek. Analiza ekspresji genów, WGCNA (ang. *weigthed gene co-expression network analysis*) oraz *functional enrichment* ujawniły aktywację genów związanych z nabytą odpowiedzią odpornościową, w szczególności związanych z funkcjonowaniem limfocytów B i produkcją przeciwciał. Odpowiedź prozapalna była natomiast obniżona w porównaniu z tą opisywaną u ludzi lub myszy laboratoryjnych. Ponadto zaobserwowano zmiany w ekspresji genów związanych z macierzą pozakomórkową, których białkowe produkty są wykorzystywane przez krętki *Borrelia* w celu ułatwienia adhezji i przetrwania w tkankach gospodarza. Zmiany te mogą odzwierciedlać ewolucyjną strategię tolerancji, która ogranicza uszkodzenia tkanek i zmniejsza ryzyko przewlekłych, ogólnoustrojowych objawów choroby u nornic. Co ciekawe, wykryto również kilka biologicznych ścieżek związanych z chorobami neurodegeneracyjnymi. Obejmowały one geny zaangażowane w metabolizm, produkcję energii, reakcję na stres oksydacyjny, a także degradację białek. Sugeruje to, że infekcja *B. afzelii* może mieć szersze konsekwencje dla swojego naturalnego żywiciela, niż dotąd przypuszczano.

Podsumowując, niniejsza rozprawa doktorska dostarcza informacji na temat genetycznych mechanizmów leżących u podstaw interakcji między żywicielem a patogenem w oparciu o naturalny system nornica ruda-*B. afzelii*. Praca ukazuje w jaki sposób koewolucja kształtuje zarówno specyficzne zależności na poziomie genów i różnych ich wariantów, takie jak związek między MHC a OspC, jak i szerszą perspektywę modulacji ekspresji genów w odpowiedzi na infekcję, podkreślając ewolucyjnie ukształtowaną tolerancję naturalnych żywicieli na zakażenie.

Słowa kluczowe: oddziaływania gospodarz-pasożyt, nornica ruda, *Borrelia afzelii*, borelioza

Abstract

Interactions between hosts and their parasites drive a coevolutionary dynamic, shaping genetic diversity, population structure, and disease spread. This dynamic can promote the maintenance of high levels of polymorphism, often exceeding expectations under neutral evolution. While coevolutionary theory assumes specific genetic associations between hosts and parasites, direct evidence for such relationships, especially in wild vertebrate populations, remains scarce. This thesis aims to contribute to addressing this gap by studying the bank vole (*Clethrionomys glareolus*) and its pathogen *Borrelia afzelii*, a bacterium that causes Lyme disease in humans. The research builds on the fact that bank voles are natural reservoir hosts of *B. afzelii* in Europe. To explore the genetic basis of host–pathogen interactions, the thesis investigates specific gene pairs in both host and pathogen, as well as the bank vole's transcriptomic response to *B. afzelii* infection. Together, these two approaches aim to uncover both specific genetic associations shaped by coevolution and broader patterns of immune activation that may underlie infection tolerance in this natural reservoir host.

The first chapter focused on a pair of host and pathogen candidate genes whose products are likely to interact directly and drive reciprocal selective pressures. Specifically, it investigated associations between the major histocompatibility complex (MHC) of the bank vole, a component of the adaptive immune system, and the ospC gene of *B. afzelii*, which encodes an immunogenic outer surface protein used to manipulate the host's immune response. Twelve variants of OspC were detected in the studied populations of bank voles. The results of Tajima's test suggested that OspC is evolving under balancing selection. Redundancy analysis revealed that neither MHC class II alleles nor supertypes (functional groupings) predicted infection status or infection intensity, as measured by qPCR, with *B. afzelii* strains carrying distinct OspC variants. Nevertheless, variation in IgG antibody levels against two OspC variants common in the examined populations was significantly associated with specific MHC class II alleles among seropositive individuals. These findings imply that MHC alleles may shape variant-specific immune responses to *Borrelia* OspC, consistent with host–pathogen coevolution.

The second chapter examined another pair of genes involved in host–pathogen interactions: the bank vole complement factor H (CFH), an essential component of the innate immune system and a central regulator of the alternative complement pathway, and the ospE gene of *B. afzelii*, which encodes a surface protein that binds CFH to facilitate immune evasion. The results from PAML and MEME analyses provided evidence that the ospE gene is subject to positive selection. However, redundancy analysis revealed no associations between factor H family genes and OspE variants. Additionally, a haplotype network based on OspE sequences showed no geographical differentiation in the frequency of OspE variants, in contrast to the patterns previously observed for CFH, and co-correspondence analysis did not detect any spatial co-structure between CFH and OspE. Previous research has shown that CFH exhibits stronger population structure than the genomic average in bank voles, likely due to local adaptation. In contrast, findings presented in the thesis indicate that OspE does not reflect this structure, suggesting that the evolution of these genes may be shaped by distinct and possibly unlinked selective forces.

The third chapter utilized RNA sequencing to investigate the transcriptomic response of bank voles to *B. afzelii* infection, with the aim of identifying additional genes involved in host defense and examining how the response of this natural reservoir host differs from that reported in accidental hosts, such as humans. Differential gene expression analysis and weighted gene co-expression network analysis (WGCNA), followed by functional enrichment, revealed activation of genes involved in adaptive immune processes, particularly those related to B-cell function and antibody production. In contrast, the inflammatory response was subdued compared to that reported in humans or mice. Additionally, shifts were observed in the expression of genes related to extracellular matrix, whose protein products are exploited by *Borrelia* to facilitate adhesion and persistence within host tissues. These changes may reflect an evolved tolerance strategy that limits tissue damage and reduces the risk of chronic, systemic manifestations of the disease. Notably, several enriched pathways were associated with neurodegenerative diseases and included genes involved in energy metabolism, oxidative stress responses, and protein turnover. These findings suggest that *B. afzelii* infection may have broader physiological effects in its natural reservoir host than previously recognized.

In summary, this thesis provides insights into genetic mechanisms underlying host-pathogen interactions in natural bank vole-*B. afzelii* system. It explores how coevolution shapes both specific gene-level associations, such as between MHC and OspC, and broader transcriptomic responses to infection, highlighting immune modulation consistent with evolved infection tolerance. Together, these findings advance our understanding of host–parasite coevolution and the ways in which natural reservoir hosts manage persistent infections.

Keywords: host-pathogen interactions, bank vole, *Borrelia afzelii*, Lyme disease

General introduction

The association between hosts and their parasites, including pathogenic microorganisms, is crucial not only for the fields of parasitology and epidemiology but also plays a critical role in shaping evolutionary and ecological dynamics (Buckingham & Ashby, 2022; Juarez-Estrada et al., 2023). Host-parasite interactions drive coevolutionary processes that can influence genetic diversity and the structure of ecological communities (Karvonen & Seehausen, 2012). They have been involved in the evolution of the immune system (Mayer et al., 2016), sociality (Prado et al., 2009), sex and sexual selection (Ashby, 2020; Lively, 2010). Interactions between hosts and parasites often involve a continual evolutionary arms race, in which each party evolves in response to the selective pressures imposed by the other. Such dynamics can lead to different coevolutionary outcomes: they may drive a series of selective sweeps, known as an arms race scenario, in which novel adaptations and counter-adaptations replace previous genotypes. Alternatively, coevolution may result in the maintenance of high genetic diversity through balancing selection (Woolhouse et al., 2002). Mechanisms underlying balancing selection include heterozygote advantage, which could be observed in the major histocompatibility complex (MHC), in humans known as human leukocyte antigen (HLA) system (Doherty & Zinkernagel, 1975) and the sickle cell hemoglobin polymorphism, found in malaria endemic regions (Hedrick, 2011). Another mechanism is negative frequency-dependent selection (NFDS) which favors rare variants and has also been implicated in MHC/HLA diversity system (Bodmer, 1972; Ejsmond & Radwan, 2015; Radwan et al., 2020). A third mechanism involves spatially and temporally fluctuating selection, where the fitness of alleles varies across environments or over time (Hedrick, 2002).

Demonstrating coevolution in natural systems poses a significant challenge, as it requires knowledge of interacting host and parasite genes, as well as evidence of reciprocal adaptation over time. This can be achieved by studying host immunity genes, the products of which can directly recognize pathogenic molecules, and parasite infectivity genes, which are crucial for establishing infection. Tracking reciprocal evolution can be done in systems where hosts and parasites have coevolved over long

timescales, particularly in secondary contact zones where independent evolutionary histories intersect.

A promising system to study host-parasite coevolution is the interaction between *Borrelia afzelii*, a Gram-negative spirochete bacterium which is a causative agent of Lyme disease in Europe, and its rodent host, the bank vole (*Clethrionomys glareolus*). Bank voles are small rodents widespread in forests throughout Europe (Kotlík et al., 2022). Two bank vole lineages (Eastern and Carpathian) persisted in separate glacial refugia and recolonized Poland after the last glaciation, creating zones of secondary contact with detectable variability in mitochondrial DNA and microsatellite markers (Kotlík et al., 2006; Tarnowska et al., 2016, 2019). This evolutionary background, combined with the fact that bank voles are one of the main reservoir hosts for *B. afzelii* (Hanincová et al., 2003; van Duijvendijk et al., 2015; Zhong et al., 2019), makes this system particularly well-suited for studying coevolution.

B. afzelii, a member of the *B. burgdorferi* sensu lato (s.l.) complex, exhibits high antigenic diversity, which likely drives selection on host immune genes (Råberg et al., 2017). This diversity of *B. burgdorferi* s.l. is largely attributed to variation in outer surface proteins, which play key roles in infection establishment and host immune evasion, and serve as important targets of host immune responses (Bhattacharjee et al., 2013; Brisson & Dykhuizen, 2004; Caine et al., 2017). In Europe, genospecies within the *B. burgdorferi* s.l. complex are host-specialized, for example, *B. afzelii* is associated with rodents and *B. garinii* with birds (Kurtenbach et al., 2002). In North America, Lyme disease is primarily caused by *B. burgdorferi* sensu stricto, which does not exhibit host specialization (Stanek et al., 2012; Steere et al., 2016). *B. burgdorferi* s.l. spirochetes are transferred to vertebrate host by *Ixodes* tick vectors and can establish chronic infection in both natural hosts and humans (Hofmeister et al., 1999; Steere et al., 2016). In humans, infection can trigger a strong immune response and lead to Lyme disease, which may involve symptoms such as skin manifestations, arthritis, carditis, fatigue, and neurological disorders (Stanek et al., 2012). In contrast, natural reservoir hosts such as bank voles or white-footed mice (*Peromyscus leucopus*) often develop milder symptoms of infection with no effect on their survival (Schwanz et al., 2011; Sipari et al., 2022; Voordouw et al., 2015; Zhong et al., 2019). However, it has been shown that *B. afzelii* infection can reduce

reproductive success of bank voles, especially among males in low-density environments, where greater exploratory behavior is necessary (Cayol et al., 2018).

To investigate the molecular basis of coevolution between bank voles and *B. afzelii*, I focused on candidate genes from both the host and the pathogen that are likely to interact and exert reciprocal selective pressures. The bank vole major histocompatibility complex (MHC) genes, which in humans are known as human leucocyte antigens (HLA), are part of vertebrate adaptive immunity. They encode transmembrane proteins that present pathogen-derived antigens to T cells, triggering either cytotoxic or antibody-mediated immune responses (Kaufman, 2018). The classical MHC genes can be divided into two classes. MHC class I molecules present antigens from intracellular pathogens to $CD8^+$ cytotoxic T cells, while MHC class II molecules present extracellular antigens to $CD4^+$ helper T cells, which initiate antibody production (Kaufman, 2018). MHC genes are characterized by exceptional polymorphism in vertebrate genomes, and this diversity is thought to be maintained by balancing selection driven by host–parasite coevolution (Ejsmond & Radwan, 2015; Radwan et al., 2020). Pathogens may evolve to evade recognition by most common MHC variants, favoring rare alleles (Migalska et al., 2022; Phillips et al., 2018), or take the advantage of heterozygosity to present a broader range of antigens, potentially improving resistance (Doherty & Zinkernagel, 1975; Penn et al., 2002).

One of the *B. afzelii* antigens that may be recognized by host MHC class II molecules is the outer surface protein C (OspC). The ospC gene is highly polymorphic and exist as a single copy in the *B. afzelii* genome, located on the cp26 plasmid (Barbour & Travinsky, 2010; Schwartz et al., 2022). Its expression is initiated during tick feeding, and is downregulated during mammalian infection, becoming undetectable approximately two weeks post-infection (Schwan et al., 1995; Tilly et al., 2006). OspC is required for establishing infection in vertebrates (Grimm et al., 2004) and contributes to the migration of spirochetes from the tick midgut to the salivary glands (Pal et al., 2004; Schwan & Piesman, 2000). Importantly, bank voles produce specific antibodies against OspC, and these antibodies show little cross-reactivity between different OspC variants (Baum et al., 2012; Jacquet et al., 2015). This confirms OspC's immunogenicity and supports its potential role as a target of adaptive immune responses. Together with the preceding

overview of MHC, these insights form the basis for Chapter I, which investigates associations between bank vole MHC class II and *B. afzelii* OspC.

Another bank vole gene of interest in the context of *B. afzelii* infection is complement factor H (CFH), a key regulator of the alternative complement pathway. The complement system is a part of innate immunity, composed of numerous plasma and membrane molecules, which can identify pathogens, mark them for destruction and induce inflammation (Meri, 2016). It can be activated through three pathways: the classical pathway, triggered by the C1q protein binding to pathogen surfaces or antigen-antibody complexes; the lectin pathway, initiated by the mannan-binding lectin, a serum protein which binds bacterial of viral carbohydrates; and the alternative pathway, which is activated directly on pathogen surfaces (Meri, 2016; Ricklin et al., 2010). The alternative pathway is the most evolutionary ancient, and similar systems are found also in invertebrates such as ascidians and chordates (Nonaka & Yoshizaki, 2004). Unlike the other pathways, the alternative pathway can initiate spontaneously via hydrolysis of plasma C3. To prevent unintended activation on host tissues, the alternative pathway is controlled by several regulatory mechanisms. One of the regulatory proteins is CFH, which binds to C3b and inhibits the formation and activity of the alternative pathway C3 convertase (C3bBb). This binding preferentially occurs on vertebrate cells due to CFH's affinity for sialic acid residues, which are typically absent on pathogen surfaces, thereby preventing complement activation on self-cells (Meri, 2016; Zipfel & Skerka, 2009).

CFH is part of a gene cluster that includes several factor H-related proteins (FH-R), all sharing conserved function and structure across species (Pouw et al., 2015). These proteins are composed of complement control protein (CCP) domains, also known as short consensus repeats (SCRs) (Norman et al., 1991; Zipfel et al., 2002; Zipfel & Skerka, 2009). In humans and mice CFH consists of 20 CCP domains, while FH-R proteins are shorter, containing 4 to 13 domains (Kristensen & Tack, 1986; Zipfel & Skerka, 1994). The two C-terminal domains, CCP 19 and 20, serve as surface-binding regions and are homologous across CFH and FH-R proteins (Jokiranta et al., 2005; Pouw et al., 2015). These domains are subject to positive selection (Cagliani et al., 2016) and are known binding targets for a range of microbes that hijack CFH to evade complement-mediated destruction, including *Pseudomonas aeruginosa*, *Streptococcus pneumoniae*, *Staphylococcus aureus*, *Plasmodium falciparum*, and *Borrelia burgdorferi*

(Hammerschmidt et al., 2007; Hellwage et al., 2001; Kraiczy et al., 2001; Kunert et al., 2007; Simon et al., 2013; Stevenson et al., 2002; Zhang et al., 2017).

The outer surface protein E (OspE) of *B. afzelii* is one of the bacterial proteins capable of binding host CFH (Hellwage et al., 2001). OspE belongs to the Erp (OspEF-related) protein family, which is encoded on the cp32 plasmids of *Borrelia* (Casjens et al., 2017). A single spirochete can carry multiple cp32 plasmids, allowing it to express on its surface several distinct OspE variants (Stevenson et al., 2002). The expression of OspE increases during tick feeding and is maintained throughout infection in mammalian hosts (Miller et al., 2003). OspE protects *Borrelia* from complement attack by binding host CFH, and enables its persistence in the host bloodstream, highlighting its role in immune evasion and supporting its importance as a target of host innate immunity. These observations serve as the foundation for Chapter II, which examines associations between bank vole CFH and *B. afzelii* OspE.

While prior knowledge of rodent host–*Borrelia* relationships has identified key interacting genes, enabling the testing of specific predictions under host–parasite coevolution scenarios, much of this research is based on laboratory models, particularly house mice. As a result, immune adaptations that have evolved in natural reservoir hosts of *Borrelia* may be overlooked when focusing solely on these model organisms. In particular, understanding how the immune response of wild rodents differs from that of humans or laboratory mice is of great interest, as it may explain why infections in natural hosts tend to be much milder. For example, studies on white-footed mice found little to no effect of *B. burgdorferi* infection on host condition or survival (Schwanz et al., 2011; Voordouw et al., 2015). Similarly, an experimental study in bank voles also reported minimal impact of *B. afzelii* infection on host condition (Sipari et al., 2022). To gain insight into natural reservoir host immunity, Chapter III explores the transcriptomic response of wild-caught bank voles to *B. afzelii* infection using RNA sequencing.

Aims of the thesis

In my thesis I examined associations between host and pathogen genotypes, testing whether different bank vole genotypes are more likely to be infected with particular *B. afzelii* genotypes. The first aim was (1) to test for associations between bank vole MHC class II alleles/supertypes and *B. afzelii* OspC variants at the level of genes and specific anti-OspC antibodies. The second aim was (2) to look for gene-by-gene and spatial relationships between bank vole CFH and *B. afzelii* OspE. In addition to testing associations between candidate genes, I performed RNA sequencing (3) to explore the bank vole's transcriptomic response to *B. afzelii* infection and identify additional genes and pathways potentially involved in host defense. These three objectives correspond to the three chapters of the dissertation.

Chapter I

Interplay between vertebrate adaptive immunity and bacterial infectivity genes: Bank vole MHC versus *Borrelia afzelii* OspC

Różańska-Wróbel, J., Migalska, M., Urbanowicz, A., Grzybek, M., Rego, R. O. M., Bajer, A., Dwuznik-Szarek, D., Alsarraf, M., Behnke-Borowczyk, J., Behnke, J. M., & Radwan, J.

Molecular ecology 33,21 (2024): e17534

https://doi.org/10.1111/mec.17534

Chapter II

Do *Borrelia afzelii* outer surface protein E coevolve with complement factor H of its rodent host? Insights from GxG and spatial associations

Różańska-Wróbel, J., Wasilewska, J., & Radwan, J.

(Unpublished manuscript)

Do *Borrelia afzelii* outer surface protein E coevolve with complement factor H of its rodent host? Insights from GxG and spatial associations.

Joanna Różańska-Wróbel[1], Józefina Wasilewska[1], Jacek Radwan[1]

[1] Evolutionary Biology Group, Faculty of Biology, Adam Mickiewicz University, Poznań, Poland

Abstract

Host-parasite coevolution paradigms assume specific associations between host and parasite genetic variants. Yet, such associations have been rarely demonstrated, particularly among wild vertebrates. In this study, we investigated interactions between *Borrelia afzelii*, a bacterium spread by a tick bite and a major causative agent of Lyme disease in Europe, and its natural rodent host, the bank vole. We focused on genetic variation in the host CFH gene family, which encodes complement regulatory proteins, and in the bacterial OspE gene known to bind CFH to evade destruction by the complement. We tested for genetic associations between variants of the bank vole CFH gene family and infection by *B. afzelii* strains carrying different OspE variants, and the spatial co-structure between these genes in populations sampled across Poland. We found that OspE evolves under positive selection, however, we found no evidence for associations between OspE diversity and host CFH variation. These results suggest that CFH and OspE evolution may be shaped by broader selective pressures, potentially including interactions with multiple host species.

Introduction

Interactions between parasites (in a broad sense, including both micro- and macroparasites) and their hosts shape many evolutionary and ecological processes, including population dynamics (Hochachka et al., 2021) and speciation (Karvonen & Seehausen, 2012). Both parasite and host are expected to continually adapt to one another (Schmid-Hempel, 2011), which may lead either to a series of selective sweeps (arms race) or to the maintenance of high genetic diversity via balancing selection (Woolhouse et al., 2002). In the arms race model, selective sweeps result in the fixation of advantageous alleles (Wilfert & Jiggins, 2013), whereas in the balancing selection model, genetic diversity is maintained through negative frequency-dependent selection (Bodmer, 1972; Ejsmond & Radwan, 2015; Radwan et al., 2020). The balancing selection scenario is more likely when host-parasite interactions are highly genotype-specific (Buckingham & Ashby, 2022; Woolhouse et al., 2002, but see MacPherson et al., 2021).

Both coevolutionary scenarios assume reciprocal interactions between host and parasite genes. However, despite the strong theoretical foundation for models of host-parasite coevolution, empirical evidence for direct genetic associations between vertebrate hosts and their parasites remains relatively limited (Råberg, 2023). Specific genetic interactions have been documented in several systems, such as plant-pathogen relationships (Thrall et al., 2012), bacteria-bacteriophage dynamics (Scanlan et al., 2011), and invertebrates (Routtu & Ebert, 2015). In vertebrates, such examples are largely limited to humans (Ansari et al., 2019; Bartha et al., 2013; Lees et al., 2019), with only a few examples reported from natural populations (Råberg et al., 2022; Różańska-Wróbel et al., 2024).

One of major components of vertebrate immunity is the complement system, which is a part of the innate immune system. It consists of numerous plasma proteins that interact with one another to mark pathogens and trigger inflammatory responses that aid in fighting infections (S. Meri, 2016). The complement system can be activated through three distinct pathways: the classical pathway, triggered by binding of C1q protein to the pathogen surface or an antibody-antigen complex, the lectin pathway, initiated by a serum protein, the mannan-binding lectin, that binds bacteria or viruses, and the alternative pathway, activated directly on pathogen surfaces. The alternative pathway is the most ancient one, playing a key role in protection against microbial infections also

outside vertebrates, as similar pathways are also found in ascidians and chordates (Nonaka & Yoshizaki, 2004). The alternative pathway does not require specific pathogen-binding proteins, as it is spontaneously triggered by hydrolysis of C3 proteins present in plasma. There are, however, several mechanisms that ensure activation occurs only on a surface of pathogen, and not on the host cells. One such regulatory protein, preventing complement activation, is complement factor H (CFH). CFH binds to C3b, which is generated by the cleavage of C3. The binding occurs preferentially on vertebrate cell surfaces due to the affinity of CFH for the sialic acid residues found on those cells, but not on pathogen surfaces. The binding of CFH to C3b inhibits the alternative pathway C3 convertase (C3bBb) formation and activity and thus prevents complement mediated damage of cells (Meri, 2016; Zipfel & Skerka, 2009).

In a genome, CFH is located in a gene cluster along with factor H-related proteins (FH-R). The CFH protein family has a conserved structure and function among species (Pouw et al., 2015). All the family proteins are composed of domains called short consensus repeat (SCR) or complement control protein (CCP) domains (Norman et al., 1991; Zipfel et al., 2002; Zipfel & Skerka, 1994). Human and murine CFH is organized into 20 CCP domains, while FH-R are shorter and contain 4-13 domains (Kristensen & Tack, 1986; Zipfel & Skerka, 1994). The two C-terminal domains of CFH, CCP 19 and CCP 20, function as surface-binding, target-recognition regions and each FH-R also contains CCP domains homologous to CCP 19 and 20 (Jokiranta et al., 2005; Pangburn, 2002; Pouw et al., 2015). Notably, CCP 19-20 domains are under positive selection (Cagliani et al., 2016) and act as binding targets for various microbes (T. Meri et al., 2013; Zipfel et al., 2007). Pathogens capable of recruiting CFH proteins, thereby interfering with the alternative complement pathway, include *Pseudomonas aeruginosa* (Kunert et al., 2007), *Streptococcus pneumoniae* (Hammerschmidt et al., 2007), *Staphylococcus aureus* (Zhang et al., 2017), *Plasmodium falciparum* (Simon et al., 2013) and *Borrelia burgdorferi* (Hellwage et al., 2001; Kraiczy et al., 2001; Stevenson et al., 2002).

One of the proteins that binds host CFH is the outer surface protein E (OspE) of the Lyme disease agent *B. burgdorferi* sensu lato (Hellwage et al., 2001). OspE lipoprotein, together with OspF and Elp proteins, belong to a family known as Erp (OspEF-related) proteins which is encoded on circular plasmids cp32 of *Borrelia* (Casjens et al., 1997). Each single bacterium may carry multiple copies of cp32 plasmids and express different

OspE proteins simultaneously (Stevenson et al., 2002). OspE proteins are exposed on the surface of *Borrelia*, and their expression is upregulated during tick feeding and maintained throughout mammalian infection (Miller et al., 2003). They bind CCP20 domain of host CFH and homologus C-terminal CCP domains of FH-R proteins (Kraiczy & Stevenson, 2013). The binding of CFH protects *Borrelia* from complement-mediated destruction and thus enables its survival in the host serum (Kenedy & Akins, 2011).

Given known function of CFH and OspE in host immunity and parasite infectivity, they are an attractive system to study host-parasite interactions and coevolution. Both genes are polymorphic in natural populations. *Borrelia* strains differ in OspE composition (Brisson et al., 2013; El-Hage & Stevenson, 2002; Stevenson, 2002), and a recent study of the genetic population structure of CFH in wild bank voles (*Clethrionomys glareolus*) from Poland found evidence of positive selection acting on the CCP20 domain. It also identified three distinct CFH groups on a phylogenetic tree, with a clear geographic separation between northeastern and southwestern Poland. This might to some extent reflect the history of expansion from two main refugia, Carpathan and Eastern (Niedziałkowska et al., 2023; Wójcik et al., 2010). However, the population structure of CFH was stronger than the genomic average, implying that local adaptation reduces flow of CFH variants between populations (Notarnicola et al., in preparation). These local pressures could possibly include adaptation to local *Borrelia* genotypes, in which case host population structure should mirror the distribution of *Borrelia* OspE variants.

Although binding between OspE and CFH has been previously demonstrated at the protein level using assays with radiolabeled proteins (Hellwage et al., 2001) and surface plasmon resonance (Alitalo et al., 2004), we are not aware of any study that investigated specificity of associations between variants of this pair of genes in any natural system. In the present study we examined whether specific variants of the bank vole CFH gene family are associated with infection by *B. afzelii* strains carrying different OspE gene variants within populations polymorphic for CFH. Furthermore, we tested whether there is a spatial association between OspE and CFH across populations in Poland.

Materials and methods

Samples and determination of their infection status

This research used the same bank vole samples as our previous study (Różańska-Wróbel et al., 2024) the redundancy analysis examining associations between CFH genotypes and *Borrelia* OspE variants. Bank voles were sampled from 2002 to 2018 at 4-year intervals in northeastern Poland at three study sites spaced approximately 10 km apart: Urwitałt, Tałty and Pilchy (Table S1). Trapping and sampling were performed under approval of the First Warsaw Local Ethics Committee for Animal Experimentation. Tail samples were collected, stored at −20°C, and the total DNA was extracted using the MagJET Genomic DNA Purification Kit (Thermo Fisher Scientific, Waltham, MA, USA). The infection status of bank voles with *B. afzelii* was established by Różańska-Wróbel et al. (2024) using primers targeting 412-421-bp fragment of the OspC gene.

The bank voles were assigned to three age classes: 1—juvenile, 2—young adult, 3—mature adult, using a method originally described in detail by Behnke et al. (2001, 2008); Grzybek, Bajer, Bednarska, et al. (2015); Grzybek, Bajer, Behnke-Borowczyk, et al. (2015) and the resulting data were used in our earlier study (Różańska-Wróbel et al., 2024). Briefly, the age of voles sampled between 2002 and 2010 was established using six morphometric measurements (body weight, dry weight of both lenses, average femur length, nose to anus length, anus to tail end length and skull length) and a principal components analysis (PCA). This data was subsequently used to predict the age of voles sampled in 2014 and 2018 by utilizing a scikit-learn (version 1.5) machine-learning algorithm in Python (https://scikit-learn.org/stable/) to construct a decision tree classifier and assign bank voles to the three age classes (Grzybek et al., 2018).

For the construction of the OspE haplotype network and the co-structure analysis, a different set of samples was used. These samples were collected from 17 populations across Poland (Table S2) between August and October in the years 2020-2023, using live trapping for bank voles and the flagging method for tick collection, under the approval of the Local Ethics Committee for Animal Experimentation in Poznań, decision no. 35/2021. Samples included bank vole ear tissues, ticks collected from bank voles, and questing ticks sampled from the same locations. All samples were stored in ethanol at −20°C prior to DNA extraction. DNA from bank vole ear tissues was extracted using the NucleoMag 96 Tissue Kit (Macherey-Nagel, Duren, Germany), while DNA from ticks was extracted using

the Syngen DNA Mini Kit (Syngen, Wrocław, Poland). *B. afzelii* infection was detected using primers targeting the rrs-rrl (16S-23S) ribosomal intergenetic spacer (Liveris et al., 1999). Nested PCRs were run with the Type-it Microsatellite PCR Kit (Qiagen, Hilden, Germany) for 30 cycles at an annealing temperature of 52°C, with 1 µL of template and 0.4 µM of each primer.

CFH and OspE genotyping

Bank vole CFH and CFH-related genes were amplified using primers CCP20-F: 5'-CATGTGTAATATCAGAAGAGAYCA-3' and CCP20-R: 5'-CATGAATATTTCTMAACAAAGGTTCAT-3', described by Notarnicola et al. (in preparation). The primers were barcoded, and the 'N' (0-3 nucleotide) spacers were added to increase library diversity during sequencing. Barcodes used to identify unique samples have been described by (Kozich et al., 2013). The primers amplified 227-245-bp fragment of the CCP20 domain of CFH family genes. PCR was performed using the Type-it Microsatellite PCR Kit according to the manufacturer's protocol. The reaction mixture (total volume of 25 µL) contained 1 µL of template DNA, 1 µL of CCP20-F primer (final concentration: 0.4 µM), 1 µL of CCP20-R primer (final concentration: 0.4 µM), 12.5 µL of 2x Type-it Multiplex PCR Master Mix, and nuclease-free water up to 25 µL. Annealing temperature was set to 58°C and the PCR was run for 28 cycles.

B. afzelii OspE gene fragment was amplified from the infected bank vole and tick samples with barcoded primes Baf_ospE-F: 5'-GCAATGACTRAAGGYGGATCA-3' and Baf_ospE-R: 5'-TCTTCTARTGRTATTGCAT-3', designed based on alignment of available *B. afzelii* sequences, co-aligned with *B. burdorferi* to identify conserved regions. The length of the PCR product was 200-272 bp and included a region known to interact with CFH (Cagliani et al., 2016). PCR was performed using the Type-it Microsatellite PCR Kit. The reaction mixture (total volume of 25 µL) contained 1 µL of template DNA, 1 µL of Baf_ospE_F primer (final concentration: 0.4 µM), 1 µL of Baf_ospE_R primer (final concentration: 0.4 µM), 12.5 µL of 2x Type-it Multiplex PCR Master Mix, and nuclease-free water up to 25 µL. Annealing temperature was set to 56°C and the PCR was run for 32 cycles.

Sequencing libraries were prepared with NEBNext Ultra II DNA Library Prep Kit for Illumina and NEBNext Multiplex Oligos for Illumina (New England Biolabs, Ipswich, MA,

USA). The CFH and OspE amplicons were sequenced on Illumina MiSeq with MiSeq Reagent Micro Kit v2 (paired-end, 2 × 150) in separate runs. Obtained paired-end reads were merged with the AmpliMERGE tool and genotyped with the AmpliSAS software using the default Illumina parameters for de-multiplexing, clustering and filtering of unique variants (Sebastian et al., 2016).

Statistical analyses

To detect selection at individual sites of the OspE gene variants found in the studied populations, we used the Codeml program from the Phylogenetic Analysis Using Maximum Likelihood (PAML) package, v. 4.10.7 (Yang, 1997, 2007), and the Mixed Effects Model of Evolution (MEME) method (Murrell et al., 2012) from the Hypothesis Testing using Phylogenies (HyPhy) package, v. 2.5.63 (Kosakovsky Pond et al., 2005, 2020). We analyzed a 132 bp-long OspE fragment encoding 44 amino acids (positions 106-149), which was found to contain positively selected sites involved in protein interactions at binding sites (Cagliani et al., 2016). Sequences were aligned using MUSCLE in AliView, v. 1.30 (Larsson, 2014) and a Neighbor-Joining Tree of the aligned OspE variants was constructed using MEGA, v. 11.0.13 (Tamura et al., 2021) with the Maximum Composite Likelihood method. In the PAML package, we applied five codon substitution site models (M0, M1a, M2a, M7 and M8) to detect variation in selective pressure across sites. Model M0 (one-ratio) assumes one ω (dN/dS) value for all sites. M1a (nearly neutral) specifies two site classes: one under purifying selection ($\omega < 1$) and one evolving neutrally ($\omega = 1$), while M2a (positive selection) extends M1a by adding a third class where $\omega > 1$, thus allowing for positive selection. M7 (beta) models ω variation across sites using a beta distribution constrained between 0 and 1 (disallowing positive selection), whereas M8 (beta & ω), includes an additional site class with $\omega > 1$, enabling detection of positive selection. To test for the presence of positively selected sites, we performed Likelihood Ratio Tests (LRTs) in R, v. 4.1.2 (R Core Team, 2021), using Rstudio, v. 2024.4.2.764 (Posit team, 2024), based on log-likelihood values from PAML output, comparing models M1a vs. M2a and M7 vs. M8. Statistical significance was assessed using a chi-squared test, with $P < 0.05$ considered significant. For models allowing for positive selection (M2a and M8), we used the Bayes Empirical Bayes (BEB) method to calculate posterior probabilities for individual codons being under positive selection (Yang, 2007). Sites were considered

to be under positive selection if they had a posterior probability $Pr(\omega > 1) \geq 0.95$. In addition to PAML, we employed the MEME test, using default settings in the HyPhy software package. Unlike PAML, which assumes constant selection pressure across all branches at a given site, MEME detects episodic positive selection acting at individual sites along a subset of lineages (Murrell et al., 2012). For MEME test, sites were considered significant if $P < 0.05$.

Potential associations between bank vole CFH family variants and *B. afzelii* OspE variants were examined with a multivariate statistical method, the redundancy analysis (RDA). We performed partial RDA with the rda function from the vegan R package, v. 2.6-4 (Oksanen et al., 2022). CFH alleles were used as explanatory variables; sampling year, site, host age and sex were included as covariates; and the presence or absence of OspE variants was used as response variables. We included in the analyses only CFH variants that occurred in >10 % of bank voles and OspE variants that were found in at least 10 bank voles. The significance of the RDA models was assessed with permutation tests conducted with the the anova.cca function from the vegan package. The results of permutation tests were averaged over 50 runs of anova.cca.

To examine genetic structure of the OspE gene across populations a haplotype network was constructed using the haploNet function from the pegas R package, v. 1.2 (Paradis, 2010). Haplotype frequencies across populations were calculated using the haploFreq function. The network was visualized with node sizes proportional to haplotype frequencies and pie charts representing the relative contribution of each population.

To predict bank vole CFH family genes composition from *B. afzelii* OspE composition, predictive co-correspondence analysis (CoCA) was performed using the cocorresp R package, v. 0.4-6 (Simpson, 2025). Log-transformed OspE variant abundances in each population were used as the predictor matrix, and log-transformed CFH variant abundances in each population as the response matrix. Predictive accuracy was evaluated with leave-one-out cross-validation using the crossval function. A positive cross-validatory fit indicates that the model predicts the response data better than chance. The significance of each CoCA axis was evaluated with permutation tests using the permutest function from the vegan package, averaged over 50 runs.

Results

In total, we identified 55 OspE variants: 36 variants in our large sample from northeastern subpopulations (Uwritałt, Tałty and Pilchy), including six rare variants not found in the second dataset (each found in no more than six individuals), and 19 additional sequences found exclusively in the remaining sites. The phylogenetic tree of the detected OspE variants, along with *B. burgdorferi* Erp sequences downloaded from the GenBank database, is shown in Figure 1. All *B. afzelii* sequences grouped with *B. burdorferi* Erps A, C and P.

We found evidence of positive selection acting on identified OspE variants. Comparisons of PAML site models M1a vs. M2a (LRT = 10.82, P = 0.004) and M7 vs. M8 (LRT = 12.88, P = 0.002) were statistically significant, indicating that some codons are likely evolving under positive selection. BEB analysis under model M2a identified four positively selected sites ($Pr(\omega > 1) \geq 0.95$): amino acid positions 116 (P = 0.992), 125 (P = 0.978), 138 (P = 0.991), and 145 (P = 0.983). Under model M8, seven positively selected sites were detected: positions 116 (P = 0.999), 120 (P = 0.991), 125 (P = 0.997), 135 (P = 0.979), 138 (P = 0.999), 145 (P = 0.998), and 149 (P = 0.963). The MEME test identified six sites evolving under episodic positive selection ($P < 0.05$): positions 116 (P = 0.022), 119 (P = 0.029), 124 (P = 0.006), 125 (P = 0.021), 139 (P = 0.002), and 145 (P = 0.002). Position numbers refer to amino acid residues in the full-length OspE protein sequence.

Figure 1. Neighbor-Joining tree of *B. afzelii* OspE variants detected in all analyzed bank vole samples (names starting with "ospE"), along with *B. burgdorferi* Erp sequences downloaded from GenBank (names starting with "erp"; accession numbers shown in parentheses). Bootstrap support values (based on 500 replicates) are shown at the nodes.

For RDA, we obtained complete data (host age, sex, CFH/FH-R and OspE variants) for 1166 bank voles, of which 177 (15.18 %) were infected with *B. afzelii* (Table S1). Among infected individuals we identified 60 single-strain infections based on the OspC data from our previous study (Różańska-Wróbel et al. 2024). Each *B. afzelii* strain harbored one to six OspE variants. A histogram of OspE variant counts in single-strain infections is shown in Figure S1. A haplotype combination showing the number of OspE variant co-occurrences within samples carrying a single OspC variant is shown in Table S3. Of 60 single-OspC strain infections, in 38 cases there was at least one other OspC strain carrying exactly the same set of OspE genes. However, most OspC strains had unique OspE combinations, even though in many cases they had some OspE types in common with other representatives of the same OspC type. Furthermore, different OspC types sometimes shared the same OspEs (Table S3). Among the 57 CFH/FH-R variants detected, 10 occurred in more than 10 % of individuals, and among the 36 OspE variants, 23 were found in at least 10 bank voles. These common variants were used in the partial RDA analyses. We examined associations between bank vole and *B. afzelii* genotypes in both the full dataset and the subset of infected voles. However, we found no significant associations in either case (full dataset: adjusted R^2 = 0.001, P = 0.192; infected subset: adjusted R^2 = -0.005, P = 0.713; Table 1).

Table 1. Results of partial redundancy analysis (RDA) performed on bank vole samples collected at three sites: Urwitałt, Tałty, and Pilchy. Bank vole CFH alleles were used as explanatory variables; sampling year, site, host age and sex were included as covariates; and the presence or absence of *B. afzelii* OspE variants was used as response variables.

Dataset	Prop. conditional variance	Prop. contrained variance	Adjused R^2	Df	F	P value
all individuals (n = 1166)	0.056	0.009	0.001	10	1.135	0.192
infected individuals (n = 177)	0.103	0.050	-0.005	10	0.918	0.713

The second dataset, containing samples collected from 17 populations across Poland, included 509 bank vole samples, of which 71 (13.95 %) were infected with *B. afzelii*, and 91 infected ticks (Table S2). In this dataset we identified 70 CFH/FH-R and 49 OspE variants. Haplotype network based on OspE sequences showed that individual OspE variants were broadly shared among populations, with no clear population-specific clustering, which indicates a lack of genetic structure in the OspE gene of *B. afzelii* across the sampled populations (Figure 2). Indeed, the results of predictive CoCA indicated that *Borrelia* OspE composition does not predict vole CFH/FH-R composition. Cross-validatory fit was positive only for the first CoCA axis (1.864), while all remaining axes showed negative fit values, suggesting poor predictive performance of the model (Table 2). Additionally, permutation tests revealed no significant axes except for the axis 16 (P = 0.019), which explained only a minimal proportion of the total variance (3.01 % in the response and 0.49 % in the predictor; Table 2).

Table 2. Results of predictive co-correspondence analysis (CoCA) performed on samples collected from 17 populations across Poland. *B.afzelii* OspE gene variant abundances in each population were used as the predictor matrix and bank vole CFH family gene variant abundances in each population as the response matrix.

Axis	Variance explained in response (CFH)	Variance explained in predictor (OspE)	Cross-validatory % fit	P value
Comp 1	7.089	31.967	1.864	0.547
Comp 2	9.511	18.512	-0.399	0.303
Comp 3	11.119	10.627	-3.426	0.231
Comp 4	6.865	6.131	-4.426	0.996
Comp 5	7.869	4.937	-4.185	0.923
Comp 6	6.190	4.726	-9.120	0.993
Comp 7	6.106	3.910	-9.540	0.991
Comp 8	5.525	3.717	-6.639	0.999
Comp 9	5.012	3.259	-8.684	1.000
Comp 10	4.894	3.170	-9.545	1.000
Comp 11	4.279	2.581	-9.619	0.891
Comp 12	8.335	1.610	-10.853	0.909
Comp 13	3.940	1.969	-13.118	0.994
Comp 14	5.027	1.300	-13.910	0.726
Comp 15	3.973	1.098	-14.536	0.318
Comp 16	3.012	0.487	-19.701	0.019

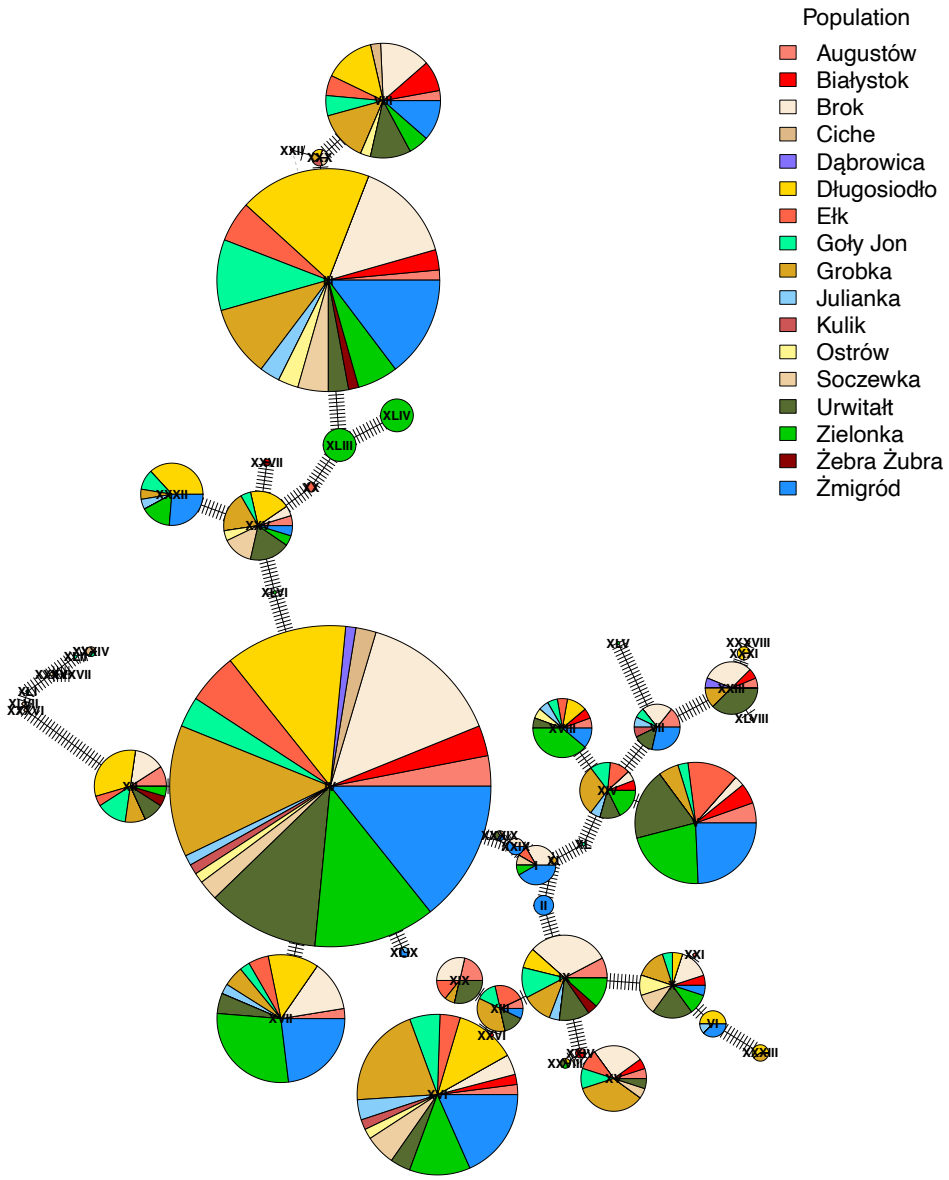


Figure 2. Haplotype network of *B. afzelii* OspE gene variants identified in 17 populations across Poland. Node sizes are proportional to haplotype frequencies, and pie chart colors represent the proportional contribution of each population.

Discussion

In the present study, we tested whether bank vole CFH gene family variants are associated with infection by *B. afzelii* strains carrying different OspE gene variants, and whether host and pathogen genotypes exhibit spatial co-structure across populations in Poland. We

found the evidence that OspE evolves under positive selection, but we did not find significant associations between OspE and CFH genotypes in samples collected over 16 years from three bank vole subpopulations in northeastern Poland: Urwitałt, Tałty and Pilchy. Furthermore, we found no evidence that OspE genetic structure reflects that previously reported to exist for CFH (Notarnicola et al., in preparation).

Our study provides first comprehensive investigation of OspE diversity and molecular evolution in *B. afzelii*. We detected 55 unique OspE variants in a total of 339 *Borrelia* positive samples. The variants clustered on a phylogenetic tree with *B. burgdorferi* Erp genes known to belong to the OspE subfamily (ErpA, ErpC, and ErpP; Brissette et al., 2008; Hillman et al., 2025). The tree also included two distinct clades corresponding to the OspF (ErpG, ErpK, ErpL, ErpY) and Elp (ErpB, ErpD, ErpM, ErpQ) subfamilies (Hillman et al., 2025), but none of our OspE sequences clustered with those loci. Because OspE subfamily is the only Erp group implied in hijacking CFH (Alitalo et al., 2002; Hovis et al., 2006), the extensive diversity we detected could be a result of coevolution with CFH of *Borrelia* hosts, including bank voles.

The high number of distinct OspE variants we identified may, at least partially, be driven by frequent recombination events within the Erp gene family (Brisson et al., 2013). The recombination is facilitated by the presence of multiple copies of OspE variants in bacterial cells. I our dataset, each individual strain, identified based on the OspC data from our previous study (Różańska-Wróbel et al. 2024), harbored between one and six distinct OspE variants, consistent with findings in *B. burgdorferi*, which is known to carry up to nine different Erps (Stevenson & Brissette, 2023). Furthermore, horizontal transfer of plasmids, common in *Borrelia* (Brisson et al., 2013; Stevenson & Brissette, 2023) increases diversity of OspE composition within *Borrelia* strains which can then be recombined to produce novel variants. Indeed, we found that on the same OspC background, which is thought to nearly uniquely characterize *Borrelia* strains (Brisson et al., 2012), different combinations of OspE variants occur. While this data must be interpreted with caution, as it is possible that in some cases we failed to detect coinfection, it seems unlikely that such occasional events could explain multiple cases similar to that of OspC-3a, which, in addition to OspE-20, carried unique combinations of other OspEs which were occasionally shared with different OspC variants. Horizontal transfer of OspEs appears to be the most likely explanation for these complex OspE

sharing patterns (Table S3). Although all Erp genes maintain a nearly identical sequence at their 5' end, the remainder of each gene has accumulated extensive variation. This genomic flexibility likely contributes to the functional versatility of Erp proteins, which are known to bind a variety of host molecules, including CFH, FH-R and plasminogen (Brissette et al., 2009; Seling et al., 2010). These multifunctional roles, coupled with diversification through recombination, suggest that selection on OspE may be shaped by a broad range of host-specific and tissue-specific pressures.

Indeed, PAML and MEME analyses revealed evidence of positive selection acting on the OspE gene, which is consistent with its proposed role in immune evasion (Alitalo et al., 2004). We identified four and seven positively selected sites using the BEB approach under PAML models M2a and M8, respectively. MEME analysis additionally highlighted six sites evolving under episodic positive selection. Notably, three amino acid positions (116, 125 and 145) were identified as positively selected by all three models. Our results partially overlap with findings by Cagliani et al. (2016), who reprted positive selection at positions 120 and 121 of the OspE protein, based on joint analysis of several *Borrelia* species (*B. burgdorferi, B. afzelii, B. garinii, B. andersonii, B. japonica, B. lusitaniae and B. valaisiana*). In our analysis, only PAML model M8 detected positive selection at position 120, which is known to form a hydrogen bond with CFH (Bhattacharjee et al., 2013). Neverthless, the positively selected sites we identified are located at the known CFH-binding interface of OspE (Bhattacharjee et al., 2013), supporting the hypothesis that this region is subject to immune-driven selection and may play a key role in complement evasion.

However, we found no evidence for genetic associations between OspE variants and bank vole CFH/FH-R alleles. Redundancy analysis, testing whether the presence or absence of OspE variants could be predicted by host genotype, revealed no significant associations, either in the full dataset or in the subset of infected voles. This suggests that OspE diversity is unlikely to be shaped by bank vole CFH/FH-R variation alone. One possible explanation is that OspE is evolving under additional selective pressures, such as interactions with other hosts. While bank voles are an important reservoir for *B. afzelii*, other species such as yellow-necked mice (*Apodemus flavicollis*) and common shrews (*Sorex araneus*) may also exert selective pressure on OspE variation (Hanincová et al., 2003; Zhong et al., 2019). Supporting this idea, a study by Stevenson et al. (2002)

demonstrated that Erp proteins of *B. burgdorferi* bind to CFH, but with different specificities depending on the animal species, suggesting that *Borrelia* adapts its complement evasion mechanisms to different hosts. Additionally, a study by Hovis et al. (2006) also reported that *B. burgdorferi* OspE proteins exhibit differential binding to CFH and other serum proteins from humans and various animal species.

It should also be noted that our ability to detect genetic associations may have been limited by the relatively low infection rate in our dataset, with only 177 infected bank voles (15.18%), as well as the possible inclusion of individuals that had not been exposed to *Borrelia*. Earlier work indicated that *B. afzelii* infection can be detected in 19.56 % of voles carrying anti-*Borrelia* antibodies, which were present in 71.10% of the individuals examined (Różańska-Wróbel et al. 2024), suggesting that while exposure rate is high, voles are able to clear the infection. Our work suggests that the clearance rates of strains carrying different sets of OspE is not related to bank vole CFH. Nonetheless, OspE overexpression has been shown to increase the serum resistance of a serum-sensitive *B. burgdorferi* strain *in vitro* (Kenedy & Akins, 2011), and some OspE paralogs have been demonstrated to confer *B. burgdorferi* survival in lizard serum (Nowak et al., 2023), highlighting the relevance of alternative complement pathway evasion mechanisms in host-pathogen interactions. These findings suggest that host-driven selection on OspE may occur in some systems, even if such patterns were not detected in our dataset. However, it is important to note that a recent study demonstrated that *B. burgdorferi* strain lacking all cp32 plasmids, which encode Erp proteins, retains full infectivity in mice (Hillman et al., 2025), indicating that OspE-mediated immune evasion may not be essential for infection in the murine host (although it may increase infection or survival success in natural situations). This also suggests that the loss of Erp proteins may be functionally compensated by other complement regulator acquiring surface proteins (CRASPs), with overlapping roles in complement evasion, such as CspA (Lin et al., 2020).

Furthermore, we analyzed samples from 17 populations across Poland to examine potential genetic structure of the OspE gene. Samples from these populations were included based on a previous study by Notarnicola et al. (in preparation), which identified pronounced genetic structure in the bank vole CFH gene, with a clear separation between northeastern and southwestern Poland, a pattern we hypothesized might reflect local adaptation to *B. afzelii* strains carrying different sets of OspE genes. Such local adaptation

could have evolved either within glacial refugia, or during post-glacial expansion, and could act against the flow of CFH genes between the populations. However, a haplotype network based on OspE sequences showed no clear population structuring, with vast majority of OspE variants segregating in most population examined. Similarly, co-correspondence analysis did not reveal spatial co-structure between bank vole and *B. afzelii* genotypes, as OspE composition did not predict CFH/FH-R composition in the host populations. Together with the lack of individual-level associations, these spatial results further suggest that the bank vole CFH family gene variation does not appear to be shaped by OspE composition in local *Borrelia* strains. However, it is remarkable that OspE shows very little inter-population structure. This is consistent with the idea discussed above that OspE variants represent adaptations to different hosts, and little inter-population differentiation reflects similar composition of host species in different areas. Furthermore, because OspE is encoded on a multi-copy plasmid, it is possible that most *Borrelia* strains may have a toolkit allowing infection of several species. What causes the limited CFH gene flow between populations remains an open question. It could be adaptation to pathogens other than *Borrelia*. Alternatively, the expression of CFH on a genetic background that has evolved independently for a long time might result in hindered binding to self-cells, possibly compromising the function of the alternative complement pathway.

In conclusion, our study provides a comprehensive analysis of interactions between *B. afzelii* OspE and CFH of its rodent host, the bank vole. While we detected signatures of positive selection acting on OspE, we found no evidence for genetic or spatial associations with host CFH/FH-R variation. These results suggest that selective pressures on bank vole CFH responsible for its limited flow among populations include other factors than *B. afzelii* infection. These factors will need to be investigated in future research.

Author contributions

Jacek Radwan designed the study. Jacek Radwan, Joanna Różańska-Wróbel and Józefina Wasilewska were involved in sample collection. Joanna Różańska-Wróbel and Józefina Wasilewska performed molecular work. Joanna Różańska-Wróbel analysed the data, advised by Jacek Radwan. Joanna Różańska wrote the manuscript with input and edits from Jacek Radwan.

Acknowledgements

This work was funded by National Science Centre, Poland (NCN) grant 2019/34/A/NZ8/00231.

Competing Interests

The authors declare no competing interests.

References

Alitalo, A., Meri, T., Lankinen, H., Seppälä, I., Lahdenne, P., Hefty, P. S., Akins, D., & Meri, S. (2002). Complement Inhibitor Factor H Binding to Lyme Disease Spirochetes Is Mediated by Inducible Expression of Multiple Plasmid-Encoded Outer Surface Protein E Paralogs. *The Journal of Immunology*, 169(7), 3847–3853. https://doi.org/10.4049/jimmunol.169.7.3847

Alitalo, A., Meri, T., Chen, T., Lankinen, H., Cheng, Z.-Z., Jokiranta, T. S., Seppälä, I. J. T., Lahdenne, P., Hefty, P. S., Akins, D. R., & Meri, S. (2004). Lysine-Dependent Multipoint Binding of the *Borrelia burgdorferi* Virulence Factor Outer Surface Protein E to the C Terminus of Factor H1. *The Journal of Immunology*, 172(10), 6195–6201. https://doi.org/10.4049/jimmunol.172.10.6195

Behnke, J. M., Bajer, A., Harris, P. D., Newington, L., Pidgeon, E., Rowlands, G., Sheriff, C., Kuliś-Malkowska, K., Siński, E., Gilbert, F. S., & Barnard, C. J. (2008). Temporal and between-site variation in helminth communities of bank voles (*Myodes glareolus*) from N.E. Poland. 2. The infracommunity level. *Parasitology*, *135*(8), 999–1018. https://doi.org/10.1017/S0031182008004484

Behnke, J. M., Barnard, C. J., Bajer, A., Bray, D., Dinmore, J., Frake, K., Osmond, J., Race, T., & Sinski, E. (2001). Variation in the helminth community structure in bank voles (*Clethrionomys glareolus*) from three comparable localities in the Mazury Lake District region of Poland. *Parasitology*, *123*(4), 401–414. https://doi.org/10.1017/S0031182001008605

Bhattacharjee, A., Oeemig, J. S., Kolodziejczyk, R., Meri, T., Kajander, T., Lehtinen, M. J., Iwaï, H., Jokiranta, T. S., & Goldman, A. (2013). Structural Basis for Complement Evasion by Lyme Disease Pathogen *Borrelia burgdorferi*. *Journal of Biological Chemistry*, 288(26), 18685–18695. https://doi.org/10.1074/jbc.M113.459040

Bodmer, W. F. (1972). Evolutionary significance of the HL-A system. *Nature*, *237*(5351). https://doi.org/10.1038/237139a0

Brissette, C. A., Cooley, A. E., Burns, L. H., Riley, S. P., Verma, A., Woodman, M. E., Bykowski, T., & Stevenson, B. (2008). Lyme borreliosis spirochete Erp proteins, their known host ligands, and potential roles in mammalian infection. *International Journal of Medical Microbiology*, 298, 257–267. https://doi.org/10.1016/j.ijmm.2007.09.004

Brissette, C. A., Haupt, K., Barthel, D., Cooley, A. E., Bowman, A., Skerka, C., Wallich, R., Zipfel, P. F., Kraiczy, P., & Stevenson, B. (2009). *Borrelia burgdorferi* Infection-Associated Surface Proteins ErpP, ErpA, and ErpC Bind Human Plasminogen. *Infection and Immunity*, 77(1), 300–306. https://doi.org/10.1128/iai.01133-08

Brisson, D., Drecktrah, D., Eggers, C. H., & Samuels, D. S. (2012). Genetics of *Borrelia burgdorferi*. *Annual Review of Genetics*, 46(1), 515–536. https://doi.org/10.1146/annurev-genet-011112-112140

Brisson, D., Zhou, W., Jutras, B. L., Casjens, S., & Stevenson, B. (2013). Distribution of cp32 Prophages among Lyme Disease-Causing Spirochetes and Natural Diversity of Their Lipoprotein-Encoding *erp* Loci. *Applied and Environmental Microbiology*, *79*(13), 4115–4128. https://doi.org/10.1128/AEM.00817-13

Buckingham, L. J., & Ashby, B. (2022). Coevolutionary theory of hosts and parasites. *Journal of Evolutionary Biology*, *35*(2), 205–224. https://doi.org/10.1111/jeb.13981

Cagliani, R., Forni, D., Filippi, G., Mozzi, A., De Gioia, L., Pontremoli, C., Pozzoli, U., Bresolin, N., Clerici, M., & Sironi, M. (2016). The mammalian complement system as an epitome of host-pathogen genetic conflicts. *Molecular Ecology*, *25*(6), 1324–1339. https://doi.org/10.1111/mec.13558

Casjens, S., van Vugt, R., Tilly, K., Rosa, P. A., & Stevenson, B. (1997). Homology throughout the multiple 32-kilobase circular plasmids present in Lyme disease spirochetes. *Journal of Bacteriology*, *179*(1), 217–227. https://doi.org/10.1128/jb.179.1.217-227.1997

Doherty, P. C., & Zinkernagel, R. M. (1975). Enhanced immunological surveillance in mice heterozygous at the H-2 gene complex. *Nature*, *256*(5512), 50–52. https://doi.org/10.1038/256050a0

Ejsmond, M. J., & Radwan, J. (2015). Red Queen Processes Drive Positive Selection on Major Histocompatibility Complex (MHC) Genes. *PLOS Computational Biology*, *11*(11), e1004627. https://doi.org/10.1371/journal.pcbi.1004627

El-Hage, N., & Stevenson, B. (2002). Simultaneous Coexpression of Borrelia burgdorferi Erp Proteins Occurs through a Specific, erp Locus-Directed Regulatory Mechanism. *Journal of Bacteriology*, *184*(16), 4536–4543. https://doi.org/10.1128/jb.184.16.4536-4543.2002

Grzybek, M., Alsarraf, M., Tołkacz, K., Behnke-Borowczyk, J., Biernat, B., Stańczak, J., Strachecka, A., Guz, L., Szczepaniak, K., Paleolog, J., Behnke, J. M., & Bajer, A. (2018). Seroprevalence of TBEV in bank voles from Poland—A long-term approach. *Emerging Microbes & Infections*, *7*, 145. https://doi.org/10.1038/s41426-018-0149-3

Grzybek, M., Bajer, A., Bednarska, M., Al-Sarraf, M., Behnke-Borowczyk, J., Harris, P. D., Price, S. J., Brown, G. S., Osborne, S.-J., Siński, E., & Behnke, J. M. (2015). Long-term spatiotemporal stability and dynamic changes in helminth infracommunities of bank voles (*Myodes glareolus*) in NE Poland. *Parasitology*, *142*(14), 1722–1743. https://doi.org/10.1017/S0031182015001225

Grzybek, M., Bajer, A., Behnke-Borowczyk, J., Al-Sarraf, M., & Behnke, J. M. (2015). Female host sex-biased parasitism with the rodent stomach nematode Mastophorus muris in wild bank voles (Myodes glareolus). *Parasitology Research*, *114*(2), 523–533. https://doi.org/10.1007/s00436-014-4214-0

Hammerschmidt, S., Agarwal, V., Kunert, A., Haelbich, S., Skerka, C., & Zipfel, P. F. (2007). The host immune regulator factor H interacts via two contact sites with the PspC protein of Streptococcus pneumoniae and mediates adhesion to host epithelial cells. *Journal of Immunology (Baltimore, Md.: 1950)*, *178*(9), 5848–5858. https://doi.org/10.4049/jimmunol.178.9.5848

Hanincová, K., Schäfer, S. M., Etti, S., Sewell, H.-S., Taragelová, V., Ziak, D., Labuda, M., & Kurtenbach, K. (2003). Association of *Borrelia afzelii* with rodents in Europe. *Parasitology*, 126(1), 11–20. https://doi.org/10.1017/S0031182002002548

Hedrick, P. W. (2002). Pathogen Resistance and Genetic Variation at Mhc Loci. *Evolution*, *56*(10), 1902–1908. https://doi.org/10.1111/j.0014-3820.2002.tb00116.x

Hedrick, P. W. (2011). Population genetics of malaria resistance in humans. *Heredity*, *107*(4), 283–304. https://doi.org/10.1038/hdy.2011.16

Hellwage, J., Meri, T., Heikkilä, T., Alitalo, A., Panelius, J., Lahdenne, P., Ilkka Seppälä, J. T., & Meri, S. (2001). The Complement Regulator Factor H Binds to the Surface Protein OspE of Borrelia burgdorferi. *Journal of Biological Chemistry*, *276*(11), 8427–8435. https://doi.org/10.1074/jbc.M007994200

Hochachka, W. M., Dobson, A. P., Hawley, D. M., & Dhondt, A. A. (2021). Host population dynamics in the face of an evolving pathogen. *Journal of Animal Ecology*, *90*(6), 1480–1491. https://doi.org/10.1111/1365-2656.13469

Hovis, K. M., Tran, E., Sundy, C. M., Buckles, E., McDowell, J. V., & Marconi, R. T. (2006). Selective Binding of Borrelia burgdorferi OspE Paralogs to Factor H and Serum Proteins from Diverse Animals: Possible Expansion of the Role of OspE in Lyme Disease Pathogenesis. Infection and Immunity, 74(3), 1967–1972. https://doi.org/10.1128/IAI.74.3.1967-1972.2006

Jokiranta, T. S., Cheng, Z.-Z., Seeberger, H., Jòzsi, M., Heinen, S., Noris, M., Remuzzi, G., Ormsby, R., Gordon, D. L., Meri, S., Hellwage, J., & Zipfel, P. F. (2005). Binding of Complement Factor H to Endothelial Cells Is Mediated by the Carboxy-Terminal Glycosaminoglycan Binding Site. *The American Journal of Pathology*, *167*(4), 1173–1181. https://doi.org/10.1016/S0002-9440(10)61205-9

Karvonen, A., & Seehausen, O. (2012). The Role of Parasitism in Adaptive Radiations—When Might Parasites Promote and When Might They Constrain Ecological Speciation? *International Journal of Ecology*, *2012*, 1–20. https://doi.org/10.1155/2012/280169

Kenedy, M. R., & Akins, D. R. (2011). The OspE-Related Proteins Inhibit Complement Deposition and Enhance Serum Resistance of *Borrelia burgdorferi*, the Lyme Disease Spirochete. *Infection and Immunity*, *79*(4), 1451–1457. https://doi.org/10.1128/IAI.01274-10

Kosakovsky Pond, S. L., Frost, S. D. W., & Muse, S. V. (2005). HyPhy: Hypothesis testing using phylogenies. *Bioinformatics (Oxford, England)*, *21*(5), 676–679. https://doi.org/10.1093/bioinformatics/bti079

Kosakovsky Pond, S. L., Poon, A. F. Y., Velazquez, R., Weaver, S., Hepler, N. L., Murrell, B., Shank, S. D., Magalis, B. R., Bouvier, D., Nekrutenko, A., Wisotsky, S., Spielman, S. J., Frost, S. D. W., & Muse, S. V. (2020). HyPhy 2.5-A Customizable Platform for Evolutionary Hypothesis Testing Using Phylogenies. *Molecular Biology and Evolution*, *37*(1), 295–299. https://doi.org/10.1093/molbev/msz197

Kozich, J. J., Westcott, S. L., Baxter, N. T., Highlander, S. K., & Schloss, P. D. (2013). Development of a Dual-Index Sequencing Strategy and Curation Pipeline for Analyzing Amplicon Sequence Data on the MiSeq Illumina Sequencing Platform. *Applied and Environmental Microbiology*, *79*(17), 5112–5120. https://doi.org/10.1128/AEM.01043-13

Kraiczy, P., Skerka, C., Kirschfink, M., Brade, V., & Zipfel, P. F. (2001). Immune evasion ofBorrelia burgdorferi by acquisition of human complement regulators FHL-1/reconectin and Factor H. *European Journal of Immunology*, *31*(6), 1674–1684. https://doi.org/10.1002/1521-4141(200106)31:6<1674::AID-IMMU1674>3.0.CO;2-2

Kraiczy, P., & Stevenson, B. (2013). Complement regulator-acquiring surface proteins of Borrelia burgdorferi: Structure, function and regulation of gene expression. *Ticks and Tick-Borne Diseases*, *4*(1–2), 26–34. https://doi.org/10.1016/j.ttbdis.2012.10.039

Kristensen, T., & Tack, B. F. (1986). Murine protein H is comprised of 20 repeating units, 61 amino acids in length. *Proceedings of the National Academy of Sciences*, *83*(11), 3963–3967. https://doi.org/10.1073/pnas.83.11.3963

Kunert, A., Losse, J., Gruszin, C., Hühn, M., Kaendler, K., Mikkat, S., Volke, D., Hoffmann, R., Jokiranta, T. S., Seeberger, H., Moellmann, U., Hellwage, J., & Zipfel, P. F. (2007). Immune evasion of the human pathogen Pseudomonas aeruginosa: Elongation factor Tuf is a factor H and plasminogen binding protein. *Journal of Immunology (Baltimore, Md.: 1950)*, *179*(5), 2979–2988. https://doi.org/10.4049/jimmunol.179.5.2979

Larsson, A. (2014). AliView: A fast and lightweight alignment viewer and editor for large datasets. *Bioinformatics*, 30(22), 3276–3278. https://doi.org/10.1093/bioinformatics/btu531

Lin, Y.-P., Frye, A. M., Nowak, T. A., & Kraiczy, P. (2020). New Insights Into CRASP-Mediated Complement Evasion in the Lyme Disease Enzootic Cycle. *Frontiers in Cellular and Infection Microbiology*, 10, 1. https://doi.org/10.3389/fcimb.2020.00001

MacPherson, A., Keeling, M. J., & Otto, S. P. (2021). Coevolution fails to maintain genetic variation in a host–parasite model with constant finite population size. *Theoretical Population Biology*, *137*, 10–21. https://doi.org/10.1016/j.tpb.2020.12.001

Meri, S. (2016). Self-nonself discrimination by the complement system. *FEBS Letters*, *590*(15), 2418–2434. https://doi.org/10.1002/1873-3468.12284

Meri, T., Amdahl, H., Lehtinen, M. J., Hyvärinen, S., McDowell, J. V., Bhattacharjee, A., Meri, S., Marconi, R., Goldman, A., & Jokiranta, T. S. (2013). Microbes Bind Complement Inhibitor Factor H via a Common Site. *PLoS Pathogens*, *9*(4), e1003308. https://doi.org/10.1371/journal.ppat.1003308

Miller, J. C., Von Lackum, K., Babb, K., McAlister, J. D., & Stevenson, B. (2003). Temporal Analysis of *Borrelia burgdorferi* Erp Protein Expression throughout the Mammal-Tick InfectiousCycle. *Infection and Immunity*, *71*(12), 6943–6952. https://doi.org/10.1128/IAI.71.12.6943-6952.2003

Murrell, B., Wertheim, J. O., Moola, S., Weighill, T., Scheffler, K., & Kosakovsky Pond, S. L. (2012). Detecting individual sites subject to episodic diversifying selection. *PLoS Genetics*, *8*(7), e1002764. https://doi.org/10.1371/journal.pgen.1002764

Niedziałkowska, M., Tarnowska, E., Babik, W., Konczal, M., Gharbi, K., Cezard, T., & Jędrzejewska, B. (2023). Different waves of postglacial recolonisation and genomic

structure of bank vole populations in NE Poland. *Heredity*. https://doi.org/10.1038/s41437-023-00600-1

Nonaka, M., & Yoshizaki, F. (2004). Primitive complement system of invertebrates. *Immunological Reviews*, *198*(1), 203–215. https://doi.org/10.1111/j.0105-2896.2004.00118.x

Norman, D. G., Barlow, P. N., Baron, M., Day, A. J., Sim, R. B., & Campbell, I. D. (1991). Three-dimensional structure of a complement control protein module in solution. *Journal of Molecular Biology*, *219*(4), 717–725. https://doi.org/10.1016/0022-2836(91)90666-T

Oksanen, J., Simpson, G., Blanchet, F. G., Kindt, R., Legendre, P., Minchin, P., hara, R., Solymos, P., STEVENS, H., Szöcs, E., Wagner, H., Barbour, M., Bedward, M., Bolker, B., Borcard, D., Carvalho, G., Chirico, M., De Cáceres, M., Durand, S., & Weedon, J. (2022). *Vegan community ecology package version 2.6-2 April 2022*.

Pangburn, M. K. (2002). Cutting Edge: Localization of the Host Recognition Functions of Complement Factor H at the Carboxyl-Terminal: Implications for Hemolytic Uremic Syndrome. *The Journal of Immunology*, *169*(9), 4702–4706. https://doi.org/10.4049/jimmunol.169.9.4702

Paradis, E. (2010). pegas: An R package for population genetics with an integrated–modular approach. *Bioinformatics*, *26*(3), 419–420. https://doi.org/10.1093/bioinformatics/btp696

Posit team. (2024). *RStudio: Integrated Development Environment for R*. Posit Software, PBC, Boston, MA.

Pouw, R. B., Vredevoogd, D. W., Kuijpers, T. W., & Wouters, D. (2015). Of mice and men: The factor H protein family and complement regulation. *Molecular Immunology*, *67*(1), 12–20. https://doi.org/10.1016/j.molimm.2015.03.011

R Core Team. (2021). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Råberg, L. (2023). Human and pathogen genotype-by-genotype interactions in the light of coevolution theory. *PLOS Genetics*, 19(4), e1010685. https://doi.org/10.1371/journal.pgen.1010685

Råberg, L., Clough, D., Hagström, Å., Scherman, K., Andersson, M., Drews, A., Strandh, M., Tschirren, B., & Westerdahl, H. (2022). MHC class II genotype-by-pathogen genotype interaction for infection prevalence in a natural rodent- *Borrelia* system. *Evolution*, 76(9), 2067–2075. https://doi.org/10.1111/evo.14590

Radwan, J., Babik, W., Kaufman, J., Lenz, T. L., & Winternitz, J. (2020). Advances in the Evolutionary Understanding of MHC Polymorphism. *Trends in Genetics*, *36*(4), 298–311. https://doi.org/10.1016/j.tig.2020.01.008

Routtu, J., & Ebert, D. (2015). Genetic architecture of resistance in Daphnia hosts against two species of host-specific parasites. *Heredity*, 114(2), 241–248. https://doi.org/10.1038/hdy.2014.97

Różańska-Wróbel, J., Migalska, M., Urbanowicz, A., Grzybek, M., Rego, R. O. M., Bajer, A., Dwuznik-Szarek, D., Alsarraf, M., Behnke-Borowczyk, J., Behnke, J. M., & Radwan,

J. (2024). Interplay between vertebrate adaptive immunity and bacterial infectivity genes: Bank vole MHC versus *Borrelia afzelii* OspC. *Molecular Ecology*, 33, e17534. https://doi.org/10.1111/mec.17534

Scanlan, P. D., Hall, A. R., Lopez-Pascua, L. D. C., & Buckling, A. (2011). Genetic basis of infectivity evolution in a bacteriophage. *Molecular Ecology*, 20(5), 981–989. https://doi.org/10.1111/j.1365-294X.2010.04903.x

Schmid-Hempel, P. (2011). *Evolutionary parasitology: The integrated study of infections, immunology, ecology, and genetics*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199229482.001.0001

Sebastian, A., Herdegen, M., Migalska, M., & Radwan, J. (2016). AMPLISAS: A web server for multilocus genotyping using next-generation amplicon sequencing data. *Molecular Ecology Resources*, *16*(2), 498–510. https://doi.org/10.1111/1755-0998.12453

Seling, A., Siegel, C., Fingerle, V., Jutras, B. L., Brissette, C. A., Skerka, C., Wallich, R., Zipfel, P. F., Stevenson, B., & Kraiczy, P. (2010). Functional Characterization of *Borrelia spielmanii* Outer Surface Proteins That Interact with Distinct Members of the Human Factor H Protein Family and with Plasminogen. *Infection and Immunity*, 78(1), 39–48. https://doi.org/10.1128/IAI.00691-09

Simon, N., Lasonder, E., Scheuermayer, M., Kuehn, A., Tews, S., Fischer, R., Zipfel, P. F., Skerka, C., & Pradel, G. (2013). Malaria Parasites Co-opt Human Factor H to Prevent Complement-Mediated Lysis in the Mosquito Midgut. *Cell Host & Microbe*, *13*(1), 29–41. https://doi.org/10.1016/j.chom.2012.11.013

Simpson, G. (2025). cocorresp: Co-correspondence analysis ordination methods. *R Package Version 0.4-6*.

Stevenson, B. (2002). Borrelia burgdorferi erp (ospE-Related) Gene Sequences Remain Stable during Mammalian Infection. *Infection and Immunity*, *70*(9), 5307–5311. https://doi.org/10.1128/IAI.70.9.5307-5311.2002

Stevenson, B., & Brissette, C. A. (2023). Erp and Rev Adhesins of the Lyme Disease Spirochete's Ubiquitous cp32 Prophages Assist the Bacterium during Vertebrate Infection. *Infection and Immunity*, 91(3), e00250-22. https://doi.org/10.1128/iai.00250-22

Stevenson, B., El-Hage, N., Hines, M. A., Miller, J. C., & Babb, K. (2002). Differential Binding of Host Complement Inhibitor Factor H by *Borrelia burgdorferi* Erp Surface Proteins: A Possible Mechanism Underlying the Expansive Host Range of Lyme Disease Spirochetes. *Infection and Immunity*, *70*(2), 491–497. https://doi.org/10.1128/IAI.70.2.491-497.2002

Tamura, K., Stecher, G., & Kumar, S. (2021). MEGA11: Molecular Evolutionary Genetics Analysis Version 11. *Molecular Biology and Evolution*, *38*(7), 3022–3027. https://doi.org/10.1093/molbev/msab120

Thrall, P. H., Laine, A.-L., Ravensdale, M., Nemri, A., Dodds, P. N., Barrett, L. G., & Burdon, J. J. (2012). Rapid genetic change underpins antagonistic coevolution in a natural host-pathogen metapopulation. *Ecology Letters*, 15(5), 425–435. https://doi.org/10.1111/j.1461-0248.2012.01749.x

Wilfert, L., & Jiggins, F. M. (2013). The Dynamics of Reciprocal Selective Sweeps of Host Resistance and a Parasite Counter-Adaptation in Drosophila. *Evolution*, 67(3), 761–773. https://doi.org/10.1111/j.1558-5646.2012.01832.x

Woolhouse, M. E. J., Webster, J. P., Domingo, E., Charlesworth, B., & Levin, B. R. (2002). Biological and biomedical implications of the co-evolution of pathogens and their hosts. *Nature Genetics*, *32*(4), 569–577. https://doi.org/10.1038/ng1202-569

Wójcik, J. M., Kawałko, A., Marková, S., Searle, J. B., & Kotlík, P. (2010). Phylogeographic signatures of northward post-glacial colonization from high-latitude refugia: A case study of bank voles using museum specimens. Journal of Zoology, 281(4), 249–262. https://doi.org/10.1111/j.1469-7998.2010.00699.x

Yang, Z. (1997). PAML: A program package for phylogenetic analysis by maximum likelihood. *Computer Applications in the Biosciences: CABIOS*, *13*(5), 555–556. https://doi.org/10.1093/bioinformatics/13.5.555

Yang, Z. (2007). PAML 4: Phylogenetic analysis by maximum likelihood. *Molecular Biology and Evolution*, *24*(8), 1586–1591. https://doi.org/10.1093/molbev/msm088

Zhang, Y., Wu, M., Hang, T., Wang, C., Yang, Y., Pan, W., Zang, J., Zhang, M., & Zhang, X. (2017). Staphylococcus aureus SdrE captures complement factor H's C-terminus via a novel 'close, dock, lock and latch' mechanism for complement evasion. *The Biochemical Journal*, *474*(10), 1619–1631. https://doi.org/10.1042/BCJ20170085

Zhong, X., Nouri, M., & Råberg, L. (2019). Colonization and pathology of *Borrelia afzelii* in its natural hosts. *Ticks and Tick-Borne Diseases*, 10(4), 822–827. https://doi.org/10.1016/j.ttbdis.2019.03.017

Zipfel, P. F., & Skerka, C. (1994). Complement factor H and related proteins: An expanding family of complement-regulatory proteins? *Immunology Today*, *15*(3), 121–126. https://doi.org/10.1016/0167-5699(94)90155-4

Zipfel, P. F., & Skerka, C. (2009). Complement regulators and inhibitory proteins. *Nature Reviews Immunology*, *9*(10), 729–740. https://doi.org/10.1038/nri2620

Zipfel, P. F., Skerka, C., Hellwage, J., Jokiranta, S. T., Meri, S., Brade, V., Kraiczy, P., Noris, M., & Remuzzi, G. (2002). Factor H family proteins: On complement, microbes and human diseases. *Biochemical Society Transactions*, *30*(6), 971–978. https://doi.org/10.1042/bst0300971

Zipfel, P. F., Würzner, R., & Skerka, C. (2007). Complement evasion of pathogens: Common strategies are shared by diverse organisms. *Molecular Immunology*, *44*(16), 3850–3857. https://doi.org/10.1016/j.molimm.2007.06.149

Supplementary Figure

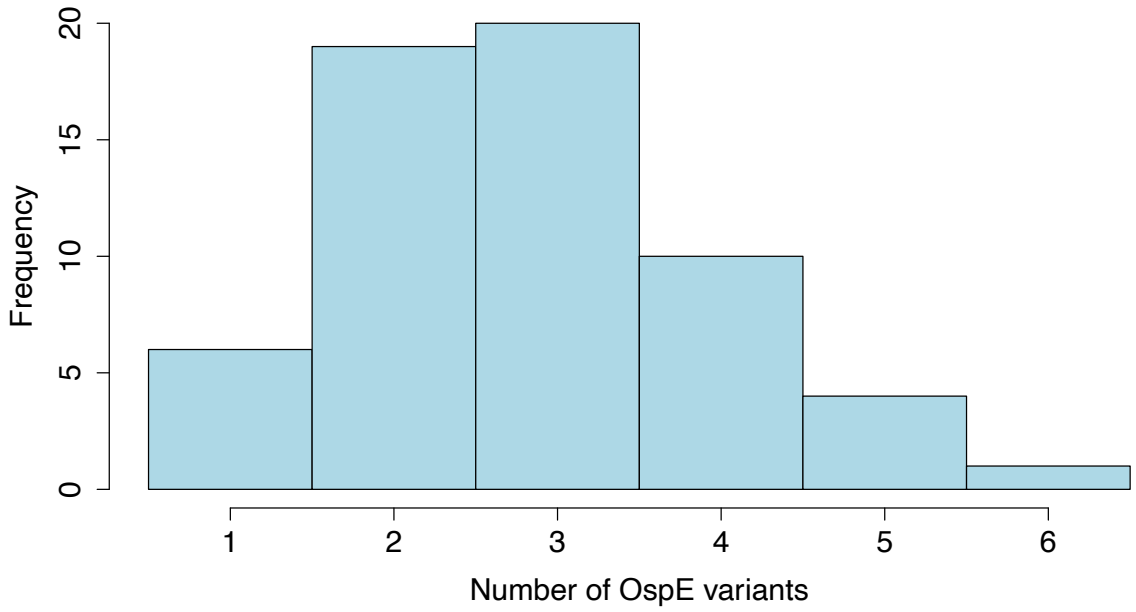


Figure S1. Histogram of OspE variant counts in single-strain *B. afzelii* infections identified in the dataset from northeastern Poland.

Supplementary Tables

Table S1. Populations used for the redundancy analysis, including geographic coordinates, sampling years and sample sizes.

Population	Coordinates	Year	# bank voles	# infected	% infected
Urwitałt	N53.48153, E21.39784	2002	72	8	11.11
		2006	100	11	11.00
		2010	80	17	21.25
		2014	83	5	6.02
		2018	60	5	8.33
Tałty	N53.53644, E21.33.049	2002	69	5	7.25
		2006	62	11	17.74
		2010	88	15	17.05
		2014	75	5	6.67
		2018	71	7	9.86
Pilchy	N53.42228, E21.48499	2002	74	20	27.03
		2006	87	31	35.63
		2010	97	24	24.74
		2014	94	4	4.26
		2018	54	9	16.67

Table S2. Populations used for co-correspondence analysis and haplotype network construction, including geographic coordinates and sample sizes.

Population	Coordinates	# bank voles	# infected bank voles	# infected ticks
Augustów	N53.796200, E23.150766	34	5	1
Białystok	N53.279194, E23.373034	23	4	0
Brok	N52.700283, E21.918842	81	12	5
Ciche	N53.424610, E19.393917	14	0	4
Dąbrowica	N50.475753, E22.359139	26	0	1
Długosiodło	N52.763774, E21.632352	32	10	8
Ełk	N53.801549, E22.399108	18	7	0
Goły Jon	N53.693694, E18.151994	29	5	4
Grobka	N53.516669, E20.624690	117	15	7
Julianka	N50.773350, E19.465696	11	3	1
Kulik	N53.567483, E21.707750	4	0	1
Ostrów	N52.824982, E21.953556	3	1	1
Soczewka	N52.534904, E19.598152	27	0	3
Urwitałt	N53.481530, E21.397840	31	3	15
Zielonka	N52.582626, E17.152514	31	1	23
Żebra Żubra	N52.705555, E23.825724	5	0	1
Żmigród	N51.508057, E17.056494	23	5	16

Table S3. Haplotype combination table showing the number of OspE variant co-occurrences within samples carrying a single OspC variant, indicative of single-strain *B. afzelii* infections.

OspC variant	OspE-1	OspE-2	OspE-3	OspE-4	OspE-5	OspE-6	OspE-7	OspE-8	OspE-9	OspE-10	OspE-11	OspE-12	OspE-13	OspE-14	OspE-16	OspE-17	OspE-18	OspE-19	OspE-20	OspE-21	OspE-23	OspE-26	OspE-43	OspE-55	OspE-101	OspE-105	#
OspC-10b		+	+						+											+	+						1
OspC-10b		+	+																	+							1
OspC-10b		+	+						+											+							6
OspC-10b		+	+						+																		2
OspC-10b		+	+						+														+				1
OspC-10b	+	+		+					+											+	+						1
OspC-19a					+					+				+								+					1
OspC-19a					+					+				+													3
OspC-19a					+					+	+										+						2
OspC-1b			+					+				+						+	+								1
OspC-2a										+				+								+					1
OspC-2a			+								+			+													1
OspC-3a	+															+			+								2
OspC-3a																+			+								1
OspC-3a			+																+								2
OspC-3a													+						+								1
OspC-3a																			+							+	1
OspC-3a	+												+						+		+						1
OspC-3a			+																+		+						1
OspC-3a																			+								2
OspC-4c			+					+	+							+				+							1
OspC-4c		+	+																								1
OspC-4c			+																								2
OspC-4c										+					+									+			1
OspC-4c										+					+												1
OspC-4c			+					+																			2
OspC-4c	+						+	+	+								+								+		1
OspC-4c			+												+												1
OspC-4c															+									+			1
OspC-4c			+									+															1
OspC-7b	+					+	+																				6
OspC-7b	+					+																					4
OspC-7b						+	+																				2
OspC-9c	+																+										1
OspC-9c	+			+													+										3

Chapter III

Transcriptomic response to *Borrelia afzelii* infection in the skin of wild bank voles

Różańska-Wróbel, J., Konczal, M., Notarnicola, R. F., & Radwan, J.

(Unpublished manuscript)

Transcriptomic response to *Borrelia afzelii* infection in the skin of wild bank voles

Joanna Różańska-Wróbel[1]#, Mateusz Konczal[1], Rocco F. Notarnicola[1], Jacek Radwan[1]

[1] Evolutionary Biology Group, Institute of Environmental Biology, Faculty of Biology, Adam Mickiewicz University, Poznań, Poland

#Corresponding author: Joanna Różańska-Wróbel, joanna.rozanska@amu.edu.pl

Abstract

Bank voles are one of the main reservoirs of tick-transmitted spirochete *Borrelia afzelii*, a causative agent of Lyme disease in humans in Europe. How the immune system deals with infection at the site of entry, i.e. skin, has not been explored in this species. Here, we used RNA-seq to explore the transcriptomic response in the ear skin of wild bank voles infected with *B. afzelii*. We identified 73 differentially expressed genes, of which 58 showed upregulation and 15 showed downregulation in infected voles compared to uninfected ones. Weighted Gene Co-expression Network Analysis (WGCNA) identified six gene modules, which were positively or negatively correlated with infection status. Enrichment analysis revealed numerous biological processes and pathways related to immune response, extracellular matrix organization, metabolism, energy production, gene expression, and cell cycle regulation. Among immunity-related genes, pathways related to B-cell activity and antibody production were particularly upregulated. However, we found that the pro-inflammatory response is suppressed compared to that reported in humans, and we identified changes in the expression of genes related to the extracellular matrix, whose products are bound and colonized by *Borrelia*. These findings indicate a complex response of bank voles to *B. afzelii* infection and provide valuable information on the molecular mechanisms underlying adaptation to *Borrelia* infection in wild reservoirs, shedding light on why vole infections are milder than in humans.

Importance

Lyme disease is a common infectious disease in Europe and North America caused by *Borrelia burgdorferi* sensu lato spirochetes, which are transmitted through tick bites. While the infection can lead to severe symptoms in humans, including fatigue, fever, joint pain, and neurological disorders, natural reservoir hosts such as rodents typically remain asymptomatic, providing an important model for uncovering the molecular basis of infection tolerance. By comparing gene expression differences between *Borrelia*-infected and uninfected individuals of a wild rodent species, the bank vole, we identified molecular pathways involved in the early response at the site of infection, the skin. Our findings revealed a reduced pro-inflammatory response, enhanced adaptive immune activation, particularly involving B-cell–mediated processes, and changes in extracellular

matrix organization. These results provide insight into the immune strategy of reservoir hosts and may help explain why Lyme disease causes more severe symptoms in humans.

Keywords: Bank vole, *Borrelia afzelii*, Lyme disease, RNA-seq, WGCNA

Highlights:

- RNA-seq was used to investigate the transcriptomic response of wild bank voles to infection with *Borrelia afzelii* at the site of entry, the ear skin.
- 58 upregulated and 15 downregulated DEGs were identified in infected bank voles compared to uninfected ones.
- Six WGCNA gene modules were positively or negatively correlated with infection status.
- The infected bank voles showed differences in the expression of genes related to immune response, metabolism, gene expression, cell cycle regulation, and extracellular matrix organization.

Introduction

Infectious diseases are a major global health concern, affecting both humans and animals, with zoonotic pathogens being a key driver of emerging diseases (Morens et al., 2004; Rahman et al., 2020; Kane et al., 2023). Many pathogens can infect multiple host species, although they can persist only in selected reservoir hosts (Haydon et al., 2002). These natural reservoir hosts, crucial for the transmission and persistence of pathogens, often establish mild or asymptomatic infections. To understand the mechanisms that facilitate pathogen spread and survival, it is essential to study the coevolutionary dynamics of host-pathogen interactions in such systems and the molecular basis of host immune responses.

Lyme disease is one of the vector-borne zoonosis that has spread in the northern hemisphere over the last few decades (Mead, 2015). It is caused by spirochetes from the *Borrelia burgdorferi* sensu lato (s.l.) species complex, which are transferred to vertebrate hosts by *Ixodes* ticks (Steere et al. 2016). In humans, Lyme disease can develop with varying severity and symptoms, such as skin manifestations, carditis, arthritis, and neurological disorders (Stanek et al., 2012). However, humans are only incidental hosts of *B. burgdorferi* s.l., whereas reservoir hosts include rodents and birds. In Europe, the primary causative agents of Lyme disease are the spirochetes *B. afzelii* and *B. garinii*, which exhibit host specialization. *B. afzelii* is primarily associated with rodents and *B. garinii* with birds (Steere et al., 2016). One of the main reservoir hosts of *B. afzelii* is the bank vole (*Clethrionomys glareolus*) (Humair et al., 1999; van Duijvendijk et al., 2015), a small rodent found in forests throughout Europe (Kotlík et al., 2022). Bank voles are typically not affected by *B.afzelii* infection in terms of condition or survival (Sipari et al., 2022), although it has been shown that infection can reduce their reproductive success (Cayol et al., 2018). Considering the role of bank voles as a key reservoir for *B. afzelii*, examining their molecular response to infection is essential to understand the broader dynamics of the ecology of Lyme disease, but also to understand how the spirochete is maintained in natural ecosystems and how it adapts to the immune responses of hosts.

Here, we performed RNA sequencing (RNA-seq) to analyze the responses of the whole transcriptome to *B. afzelii* infection and identify candidate genes involved in host-pathogen interactions. While previous research investigated transcriptomic responses to

Borrelia infection in bank voles (Zhong et al., 2020) and other rodents (Long et al., 2019, Gaber et al., 2023), only one of them focused on skin tissue (Long et al., 2019), which is the primary site of *Borrelia* infection and replication (Bernard et al., 2020). However, this study was conducted on *Peromyscus leucopus* and involved an experimental infection with *B. burgdorferi* sensu stricto (s.s.). Given its role as the primary site of infection, effective immune responses in the skin are key to controlling *Borrelia* infection and preventing its dissemination to other tissues. No investigation of skin expression patterns in response to *B. afzelii* has been carried out in bank voles. Here, we fill this gap by analyzing the whole transcriptome of 44 wild-caught individuals, comparing those infected with *B. afzelii* to uninfected ones.

Methods

Sample collection

Bank voles were sampled in August-September 2023 at three sites in Poland: Brok (52°42'01.0"N, 21°55'07.8"E), Długosiodło (52°45'49.6"N, 21°37'56.5"E), and Grobka (53°31'00.0"N, 20°37'28.9"E) under approval of the Local Ethics Committee for Animal Experimentation in Poznań, decision no. 35/2021. All sites were forests with similar habitats characterized by dense undergrowth and alders (genus *Alnus*) as dominant trees. We captured bank voles in wooden traps, weighed them, and sampled one ear biopsy of each individual before releasing them. The ear samples were then split; one half was stored in ethanol for DNA extraction, while the other was preserved in RNAlater Stabilization Solution (Thermo Fisher Scientific, Waltham, MA, USA) for RNA extraction. After 4-8 hours, RNAlater samples were frozen and stored at -20 °C.

DNA extraction and *Borrelia* screening

DNA was extracted from bank vole ear samples with the NucleoMag 96 Tissue Kit (Macherey-Nagel, Duren, Germany) and used to determine *Borrelia* infection status. Primers used for *Borrelia* screening targeted the rrs-rrl (16S-23S) ribosomal intergenetic spacer (Liveris et al., 1999). Nested PCR reactions were performed using the Type-it Microsatellite PCR Kit (Qiagen, Hilden, Germany) according to the manufacturer's protocol. The first reaction mixture contained 1 µL of template DNA, 1 µL of PA primer (final concentration: 0.4 µM), 1 µL of P95 primer (final concentration: 0.4 µM), 12.5 µL of 2x Type-it Multiplex PCR Master Mix, and nuclease-free water up to 25 µL. The second

reaction mixture contained 1 µL of PCR product from the first reaction, 1 µL of PB primer (final concentration: 0.4 µM), 1 µL of P97 primer (final concentration: 0.4 µM), 12.5 µL of 2x Type-it Multiplex PCR Master Mix, and nuclease-free water up to 25 µL. Both reactions were carried out for 30 cycles using the same annealing temperature of 52 °C. PCR products were stained with GelRed (Biotium, Fremont, CA, USA) and resolved on a 1.5% agarose gel. Bank voles were considered infected if a visible band of the expected size was observed.

RNA extraction, sequencing, and quality control

RNA was obtained from 21 infected and 23 uninfected bank vole samples and used for mRNA-Seq. The samples, along with their corresponding sampling sites, body mass, and *Borrelia* infection statuses, are listed in Table S1. Total RNA was extracted with the ReliaPrep RNA Tissue Miniprep System (Promega, Madison, WI, USA). RNA concentration, integrity, and purity was assessed using a 4200 TapeStation System (Agilent Technologies, Santa Clara, CA, USA) and a NanoPhotometer N120 (Implen, Munich, Germany). mRNA sequencing was performed by Novogene (Munich, Germany) using the polyA enrichment method for library preparation, followed by sequencing on the Illumina NovaSeq X Plus platform (PE150). The quality of the reads was assessed using FastQC v.0.12.1 (Andrews, 2010) and the pre-processing of the FASTQ files, including adapter trimming and quality filtering, was performed with fastp v.0.23.4 (S. Chen et al., 2018).

Expression quantification

Gene expression was estimated using Salmon v.1.10.0 (Patro et al., 2017) and the accessioned RNA products annotated on the bank vole reference genome assembly (NCBI RefSeq accession number: GCF_902806735.1) as an index. All the following analyses were performed in R v.4.4.1 (R Core Team, 2024). To create a table with gene counts, Salmon quantification results and a transcript-to-gene mapping file (created from a GTF file using the rtracklayer package v.1.54.0 (Lawrence et al., 2009)) were imported using the tximport package v.1.30.0 (Soneson et al., 2016). Subsequently, low-count genes were filtered out using the filterByExpr function from the edgeR package v.4.0.16 (Y. Chen et al., 2024; Robinson et al., 2010). Normalization factors were calculated using the calcNormFactors function from edgeR, which employs the Trimmed Mean of M-

values (TMM) method (Robinson & Oshlack, 2010). Normalized $\log_2$-transformed counts per million (CPM) were calculated using the cpm function from the edgeR package.

Differential gene expression

Differential gene expression (DGE) analysis was performed in the edgeR package on the filtered and normalized count data. A quasi-likelihood negative binomial generalized linear model (GLM) was fitted to the count data using the glmQLFit function. The design matrix included infection status as a factor, as well as sampling site and body mass as covariates, to account for any between-population and body mass/age-related variation. A gene-wise quasi-likelihood F-test was then applied with the glmQLFTest function. The results are presented as $\log_2$ fold changes (logFC) in gene expression levels of *Borrelia*-infected bank voles compared to uninfected individuals. The false discovery rate (FDR) was calculated by applying the Benjamini-Hochberg method on the p-values to adjust for multiple testing. Genes with an FDR < 0.05 and |logFC| > 0.58 (corresponding to approximately 1.5-fold change in gene expression) were considered significantly differentially expressed between infected and uninfected bank voles. A heatmap of gene expression for differentially expressed genes (DEGs) across all individuals was generated using the pheatmap v.1.0.12 package (Kolde, 2019).

WGCNA

Weighted Gene Co-expression Network Analysis (WGCNA) was performed on the normalized, log-transformed count data to identify potential differences in coexpressed gene clusters between infected and uninfected bank voles. Prior to the analysis, the data were adjusted for the sampling site and body mass covariates using the removeBatchEffect function from the limma v.3.58.1 package (Ritchie et al., 2015). WGCNA involved creating a network of genes with similar expression patterns, clustering them into modules, associating these modules with *Borrelia* infection status, and identifying driver genes in each significant module. While DGE analysis tests each gene individually and applies corrections for multiple testing, WGCNA focuses on collective expression patterns within coexpressed modules. This approach is particularly valuable for detecting more subtle, but consistent within modules changes in gene expression that might not reach significance in gene-by-gene DGE analysis (Langfelder & Horvath, 2008).

First, hierarchical clustering was performed using the hclust function from the fastcluster v.1.2.6 package (Müllner, 2013) to detect outlier samples and exclude them from the analysis. Subsequently, WGCNA was performed using the blockwiseModules function from the WGCNA R package v.1.73 (Langfelder & Horvath, 2008, 2012). The adjacency matrix was constructed using Pearson's correlation, signed network type, and a soft-threshold power determined using the scale-free topology criterion. The results of the scale-free analysis were visualized using ggplot2 v.3.5.1 (Wickham, 2016). Modules were identified using the signed Topological Overlap Matrix (TOM) and the dynamic tree cut algorithm, with a minimum module size of 50 genes. Similar modules were merged based on eigengene similarity using a cut height threshold of 0.25. The results of module clustering were visualized with the plotDendroAndColors function from the WGCNA package.

To associate the modules with *Borrelia* infection status, we correlated the module eigengenes with the infection data using Pearson's correlation and calculated the corresponding Student's asymptotic p-values using the cor and corPvalueStudent functions from the WGCNA package. The results of module-trait correlations were visualized using the pheatmap package. For subsequent analyses, we focused on modules that showed significant associations with infection status ($P < 0.05$), considering these modules as potentially biologically relevant.

To identify driver genes (hub genes) within significant modules, we performed intramodular analysis by examining module membership (MM) and gene significance (GS). MM reflects the correlation between the expression profile of an individual gene and the respective module eigengene. GS represents the correlation between individual gene expression and infection status and can have positive or negative values. We selected hub genes by applying thresholds of $MM > 0.8$ and $|GS| > 0.2$, focusing on genes with strong connections within the module and relevance to infection status.

Functional enrichment analysis

Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analyses were performed to explore the biological functions of DEGs and hub genes in each significant module. The set of all expressed genes identified with edgeR in our bank vole samples was used as the background for the analyses. GO enrichment was

conducted for the Biological Process (BP) ontology using the Gene Annotation File for the bank vole reference genome and the topGO package v.2.46.0 (Alexa & Rahnenfuhrer, 2021), which applied the default weight01 algorithm and Fisher's exact test. KEGG Orthology (KO) assignments were performed using the GhostKOALA automatic annotation server. Subsequently, the enrichKEGG function from the clusterProfiler package v.4.10.0 (Wu et al., 2021; Yu et al., 2012) was used to perform enrichment analysis using the KO database, applying a q-value cutoff of 0.05.

Results

RNA sequencing and gene expression

The sequencing of 44 samples resulted in approximately 1.52 billion read pairs. The number of read pairs per sample ranged from 29.79 million to 86.25 million, with a mean of 34.47 million read pairs per sample. After adapter trimming and quality filtering, the total number of read pairs was reduced to approximately 1.51 billion across all samples. The number of read pairs per sample after filtering ranged from 29.55 million to 85.65 million, with a mean of 34.21 million read pairs per sample. Gene expression analysis revealed 27,048 genes across all samples. After filtering out genes with low counts, 14,318 genes were retained for DGE analysis and WGCNA.

Differential Gene Expression analysis

DGE analysis identified 73 genes with significant differential expression between infected and uninfected bank voles (Table S2). 58 genes were upregulated and 15 genes were downregulated in infected bank voles compared to uninfected ones. Normalized logarithmically transformed gene expression counts for all DEGs across analyzed samples are shown in Figure 1. Notably, three genes – Cxcl13, Jchain, and LOC125407534 – showed particularly distinct expression patterns between the two bank vole groups, with all three being upregulated in infected individuals (Figure 1). The Cxcl13 gene encodes C-X-C motif chemokine ligand 13, a B lymphocyte chemoattractant that promotes B cell migration to infection sites. The Jchain gene encodes the joining chain of multimeric IgA and IgM, a protein that regulates the polymerization of IgM and IgA antibodies. The LOC125407534 gene encodes the immunoglobulin lambda-1 light chain-like protein, a subunit of antibodies. These genes highlight the transcriptional differences in immune-related genes involved in B cell recruitment and antibody function.

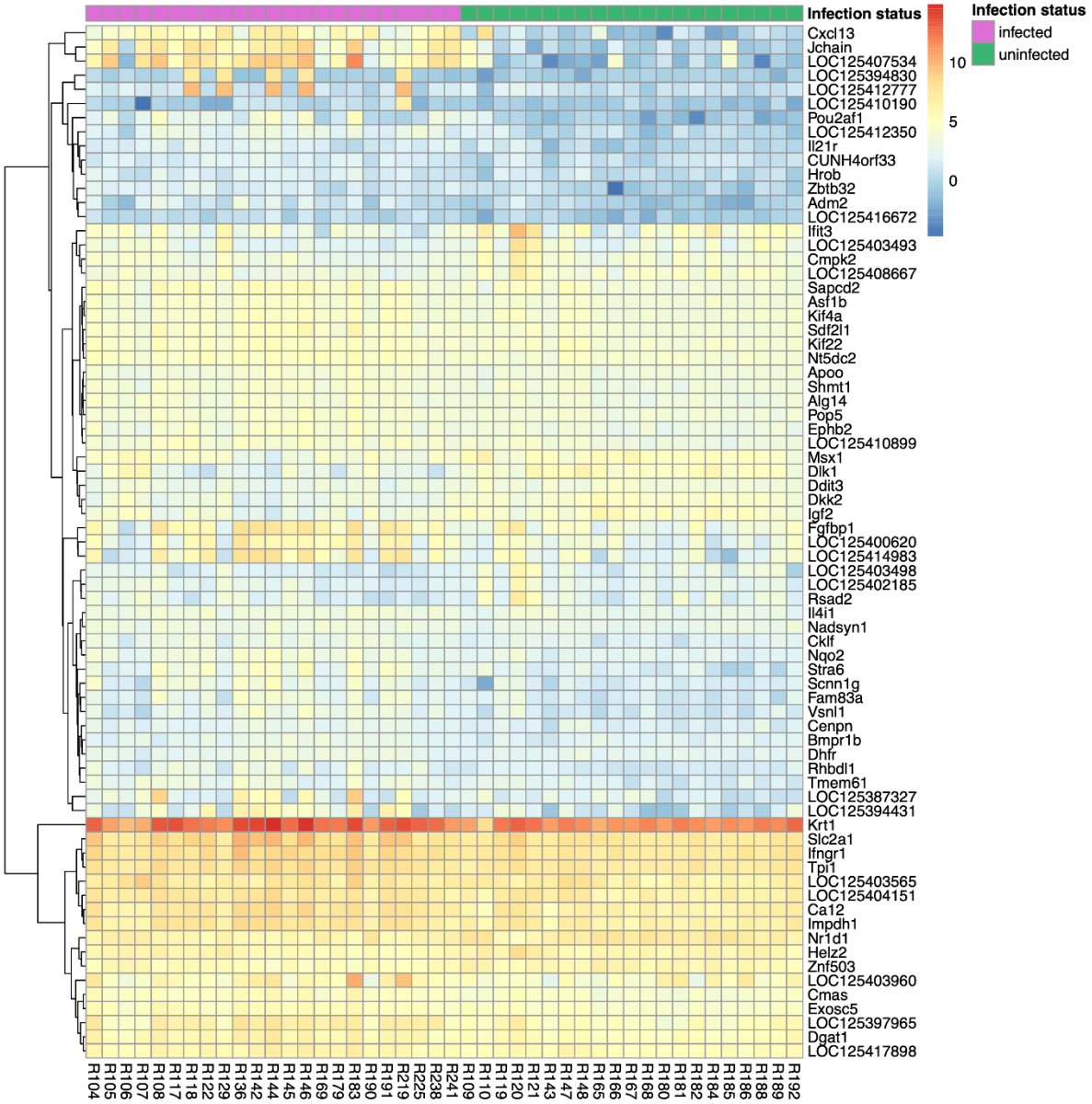


Figure 1. Heatmap of normalized log-transformed gene expression counts for differentially expressed genes across analyzed bank vole samples.

Enrichment analysis of DEGs identified multiple significant GO biological processes (Table S3), but no significant KEGG pathways. Among the enriched GO terms, several were related to the immune response. The defense response to virus (GO:0051607, P = 5.4e-07) included downregulated genes, such as Ifit3, which encodes an interferon-induced protein with tetratricopeptide repeats 3 involved in the antiviral innate immune response and exhibiting antiproliferative activity, LOC125403493, which encodes an interferon-induced protein with tetratricopeptide repeats 2-like, and Rsad2,

which encodes an interferon-inducible antiviral protein. The terms antimicrobial humoral immune response mediated by antimicrobial peptide (GO:0061844, P = 0.005) and chemokine-mediated signaling pathway (GO:0070098, P = 0.004) included upregulated Cxcl13 chemokine and LOC125394431, which encodes platelet basic protein-like, a growth factor involved in neutrophil chemotaxis and antimicrobial immune response. The cellular response to biotic stimulus (GO:0071216, P = 0.014) again included upregulated Cxcl13 and LOC125394431, as well as downregulated Ddit3, which encodes the DNA damage inducible transcript 3, a multifunctional transcription factor involved in response to cell stress and apoptosis. Lastly, the peptide antigen assembly with MHC class II protein complex (GO:0002503, P = 0.043) included upregulated LOC125407534, which encodes immunoglobulin lambda-1 light chain-like.

The enrichment analysis also identified several GO terms associated with metabolic processes, such as the glycine metabolic process (GO:0006544, P = 7.1e-04), the water-soluble vitamin metabolic process (GO:0006767, P = 0.002), and the cellular amino acid catabolic process (GO:0009063, P = 0.018). Furthermore, GO enrichment highlighted terms related to cellular structure and organization, such as the keratinization (GO:0031424, P = 0.003) and intermediate filament organization (GO:0045109, P = 0.008), as well as terms associated with cell cycle regulation and gene expression.

WGCNA

Based on the hierarchical clustering dendrogram of samples (Figure S1), a cut height of 110 was selected, resulting in three samples (R110, R143, and R165) being identified as outliers and removed from the analyses. For WGCNA, a soft threshold power of 14 was selected based on the scale-free topology model fit (Figure S2). WGCNA identified 19 gene modules, represented by different colors, excluding the gray module, which represents unassigned genes. The cluster dendrogram of the detected modules is shown in Figure 2. Of the modules discovered, six were significantly correlated with infection status ($P < 0.05$, Figure 3), including the blue, yellow, greenyellow, salmon, brown, and turquoise modules. The blue, yellow and greenyellow modules contained 1714, 1493, and 272 genes, respectively, and were positively correlated with infection, meaning that the genes in these modules had higher expression levels in response to infection. The salmon, brown and turquoise modules contained 263, 1521, and 2537 genes,

respectively, and were negatively correlated with infection. After selecting hub genes based on their module membership and gene significance, the gene counts were as follows: 275 in the blue module, 294 in the yellow, 24 in the greenyellow, 69 in the salmon, 190 in the brown, and 495 in the turquoise.

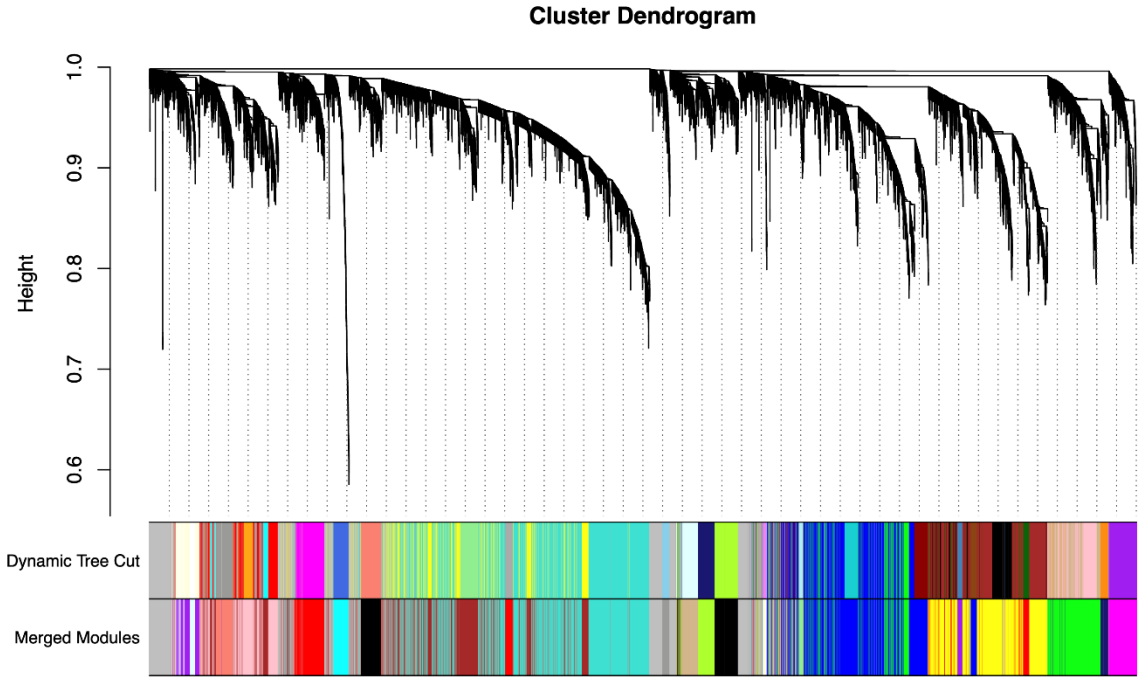


Figure 2. Hierarchical clustering dendrogram of genes, constructed to identify co-expression modules. Different colors represent different modules, with grey indicating unassigned genes. The Dynamic Tree Cut represents the initial clustering, while the final module assignment, based on eigengene similarity, is shown as Merged Modules.

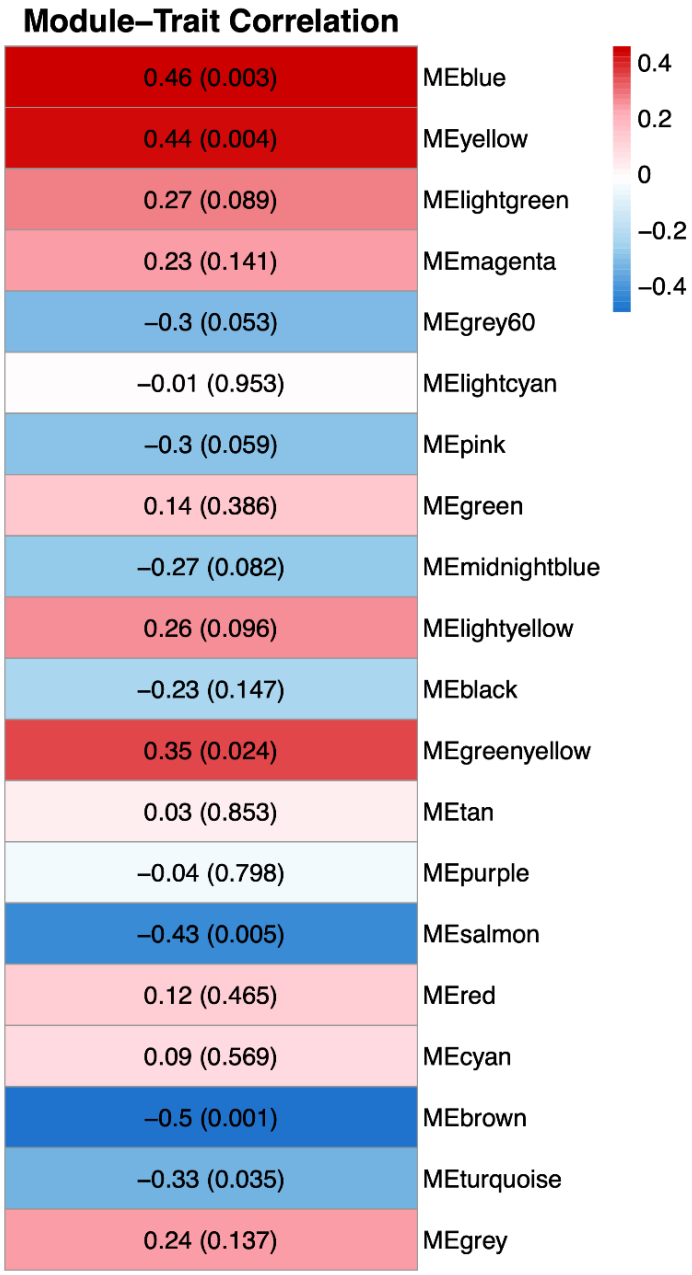


Figure 3. Module-trait correlation heatmap visualizing the correlation between WGCNA module eigengenes and *B. afzelii* infection status. Each row represents a module, and each cell contains the correlation coefficient with the corresponding p-value in parentheses. Red indicates a positive correlation between the module and infection, while blue indicates a negative correlation.

Functional enrichment analysis revealed GO terms and KEGG pathways enriched in the hub genes present in the significant modules (Tables S4 – S13). Three modules that were positively correlated with infection status, the blue, yellow, and greenyellow module, showed significant enrichment in biological processes related to cellular metabolism,

energy production, regulation of gene expression, protein synthesis and folding, neuronal functions, transport and signaling, as well as cellular structure, organization and division. GO terms related to metabolism covered various functions, encompassing carbohydrate, lipid, amino acid, nucleotide, heme and vitamin metabolism, with the glycolytic process (GO:0006096, P = 1.9e-12) being the most significant term. In addition to metabolism, particularly interesting terms included the cellular response to hypoxia, positive regulation of TOR signaling, and keratinization (Tables S4, S6 and S8). Additionally, KEGG enrichment identified several pathways associated with neurodegenerative diseases, such as amyotrophic lateral sclerosis, prion disease, Parkinson's, Huntington's, and Alzheimer's disease, as well as pathways related to metabolic processes and other significant pathways, such as the HIF-1 signaling pathway, pathogenic *Escherichia coli* infection, oxidative phosphorylation, thermogenesis, diabetic cardiomyopathy, and chemical carcinogenesis – reactive oxygen species (Tables S5 and S7).

Three modules that were negatively correlated with infection status, the salmon, brown and turquoise modules, were enriched in GO terms related to gene expression, cellular homeostasis, protein processing, vesicle transport and endocytosis, as well as extracellular matrix organization, cytoskeletal dynamics, collagen metabolic process, and angiogenesis. Furthermore, enriched terms were associated with cell migration, adhesion, tissue development, nervous system development, myelination, and signaling. (Tables S9, S11, and S12). Moreover, several significant KEGG pathways were identified, including pathways related to cell adhesion, platelets, cytoskeleton, and extracellular matrix, as well as signaling pathways, such as the TGF-beta signaling pathway, and pathways associated with various diseases, including different types of cancer, cardiovascular disease, and infectious diseases, all of which can be linked to changes in cell signaling and adhesion (Tables S10 and S13).

Discussion

In the present study, we explored changes in gene expression in wild bank vole populations in response to B. *afzelii* infection, with a focus on the skin of the ear as the infection site. While the data may be noisy due to the complexities of studying wild animals, including environmental factors, different stages of *Borrelia* infection, and potential coinfections with other pathogens, it provides valuable insights into host-

pathogen interactions within a natural ecological context. This is particularly relevant as we examined the tick bite site, the initial site of infection, where the host response will determine whether the spirochete disseminates to other tissues. By focusing on the site of entry, we expected a clearer signal for *Borrelia*-specific responses, as *Borrelia* is likely to be more abundant here compared to other potential pathogens.

We conducted DGE analysis and WGCNA, followed by enrichment analysis, and uncovered a broad set of genes and pathways involved in the bank vole response to *B. afzelii* infection. Our findings highlighted diverse responses of the bank vole transcriptome, as indicated by 73 differentially expressed genes and six gene modules correlated with infection. The results of our analysis revealed several immune-related genes, as well as numerous genes involved in metabolic processes, cellular structure, gene expression, and cell cycle regulation.

Notably, most immune-related genes were associated with B cell activity and antibody production. For example, we identified upregulated antibody-related genes: Jchain and immunoglobulin lambda-1 light chain-like, as well as Il21r, which encodes the interleukin 21 receptor, involved in the development of B cells and antibody production (Pérez-Mazliah et al., 2015) and Cxcl13, which encodes a B cell chemoattractant, the presence of which in cerebrospinal fluid is known to be a marker of neuroborreliosis (Pilz et al., 2020; Gudowska-Sawczuk & Mroczko, 2020). Moreover, among upregulated DEGs, we found the Il4i1 gene, which encodes an immunosuppressive enzyme that inhibits T cell proliferation (Scarlata et al., 2015). These results highlight the role of adaptive immune response to infection with *Borrelia*, consistent with the findings of high seroprevalence and association of strain-specific infections with major histocompatibility complex (MHC) genotype (Råberg et al., 2022; Różańska-Wróbel et al., 2024).

Studies in humans with Lyme disease report a strong pro-inflammatory response to *Borrelia* infection, including upregulation of genes encoding interferons, pro-inflammatory cytokines (TNF, IL1B, IL6), toll-like receptor 2 (TLR2), as well as C-C and C-X-C motif chemokines (Bouquet et al., 2016; Marques et al., 2018; Steere et al., 2016; Thompson et al., 2021). In contrast, we did not observe such a pronounced pro-inflammatory immune response in bank voles. While the interferon-gamma receptor 1 (Ifngr1) was among the upregulated DEGs, genes induced by type I interferons, such as Ifit3 and Rsad2, were downregulated. This is particularly interesting, as type I interferons

and interferon-stimulated genes, such as Ifit3, have previously been shown to be upregulated in erythema migrans skin lesions from untreated patients with Lyme disease (Marques et al., 2018), in human monocytes incubated with live *Borrelia* spirochetes (Salazar et al., 2009), and in *Borrelia*-infected mice (Petzke et al., 2016).

Notably, in our analysis, TNF superfamily member 12 (Tnfsf12), a pro-inflammatory and pro-apoptotic cytokine (Sidler et al., 2017; Wiley & Winkles, 2003), was found in the turquoise module, which was negatively correlated with infection. In addition, two receptors for pro-inflammatory cytokines (Il11ra and Il17rd) were also present in modules negatively correlated with infection, further suggesting a suppression of inflammation in bank voles. In the blue module, which was positively correlated with infection, we also found the interleukin-1 receptor antagonist (Il1rn), which inhibits the activities of pro-inflammatory cytokines IL1A and IL1B. We also identified interleukin 33 (Il33), an alarmin released after tissue injury. These results suggest that responses in reservoir hosts, such as bank voles, differ from responses in non-competent, accidental hosts, such as humans, whereby reservoir hosts regulate their immune responses to prevent excessive inflammation and tissue damage. This modulation may explain why Zhong et al. (2019) did not observe pathological changes in infected bank vole joints similar to those seen in human Lyme disease, despite *Borrelia* dissemination to joints occurring in a larger proportion of infected bank voles.

Another potential mechanism by which bank voles modulate *Borrelia* infection, limit its dissemination, and prevent the establishment of chronic infection is through changes in the expression of genes related to the extracellular matrix. *Borrelia* is known to bind to laminin (Verma et al., 2009) and colonize and degrade collagen fibers, facilitating its spread and invasion into connective tissues of various organs, such as joints, heart, and nervous system (Zambrano et al., 2004). In line with this, we found that the expression of bank vole genes encoding collagens (Col4a6, Col6a1, Col8a2, Col14a1, Col16a1) and laminins (Lama2, Lama4, Lamb1, Lamb2, Lamc1) was negatively correlated with *B. afzelii* infection. Interestingly, a study on human dermal fibroblasts found that genes encoding collagen (Col8a1) and laminin (Lama1) were upregulated by *Borrelia* (Schramm et al., 2012). An additional example is matrix metalloproteinases (MMPs), enzymes involved in the degradation of the extracellular matrix, influencing processes such as tissue remodeling. We found matrix metallopeptidases (Mmp2,

Mmp14, Mmp17) in the turquoise module, which was negatively correlated with infection, while studies in humans reported elevated levels of different MMPs in response to *Borrelia* (Behera et al., 2005; Lin et al., 2001; Schramm et al., 2012). These differences highlight the distinct responses of reservoir hosts and accidental end-point hosts, which may have consequences for different courses of infection: while reservoirs do not exhibit clinical manifestations of *Borrelia* infection (Rogovskyy et al., 2024; Voordouw et al., 2015), accidental hosts such as humans often develop chronic multi-organ disease with symptoms including fatigue, fever, headaches, joint pain and swelling, heart palpitations, and neurological disorders (Stanek et al., 2012).

Interestingly, we also identified several KEGG pathways related to neurodegenerative diseases (e.g., Parkinson's, Huntington's, and Alzheimer's disease) enriched in the yellow module positively correlated with *Borrelia* infection. This may reflect a response to increased oxidative stress, associated with the energetic costs of the immune response and the need to produce antimicrobial reactive oxygen species (Mastroeni et al., 2000). Mitochondrial reprogramming in infected voles was indicated by significant GO terms such as oxidative phosphorylation, ATP synthesis coupled proton transport, mitochondrial electron transport, and proteasome assembly identified in this module. Furthermore, the yellow module included the Park7 gene, which encodes a protein that protects neurons against oxidative stress, and the Txn gene, which is involved in the regulation of cellular redox homeostasis. These pathways suggest a shift toward increased mitochondrial activity, enhanced energy production, and protein turnover. Indeed, in our analysis, several metabolic pathways were upregulated in the DGE analysis and WGCNA, particularly in the blue module, which was positively correlated with infection, where the 'glycolytic process' was the most significant term. In human immune cells, *Borrelia* infection was shown to alter glucose metabolism by enhancing glycolysis through the mTOR/HIF-1a pathway, which is crucial for the induction of cytokine responses (Oosting et al., 2016). This suggests a potential link between cellular metabolism and immune responses, although in our study we did not observe a concurrent upregulation of cytokines. In addition to the increased energy demands associated with the immune response, the changes in cellular metabolism that we detected in infected individuals may result from manipulation by *Borrelia*, which has restricted metabolic capacity and relies on the host for essential nutrients (Kerstholt et

al., 2020; Steere et al., 2016). Taken together, these findings may indicate that *Borrelia* infection induces metabolic and oxidative stress, resulting in heightened metabolic demands, including a shift toward increased glycolysis, increased mitochondrial respiration, protein turnover, and neuroprotective responses, potentially indicating the involvement of neurodegenerative processes in bank voles infected with *B. afzelii*.

Our results are broadly consistent with the findings of previous studies on the transcriptional responses of reservoir hosts to *Borrelia*, confirming key patterns observed in earlier research. A previous study that compared the spleen transcriptome responses of *B. afzelii* infected vs. uninfected wild bank voles identified eight differentially expressed genes and 28 Hallmark gene sets uncovered by gene set enrichment analysis, including upregulated processes such as oxidative phosphorylation and heme metabolism, as well as downregulated processes, such as inflammatory response, TGF beta signaling, and Notch signaling (Zhong et al., 2020), which were also identified by our enrichment analysis. Additionally, among the downregulated Hallmark gene sets identified by Zhong et al. (2020) there was the epithelial mesenchymal transition set, which is involved in wound healing and fibrosis. Similarly, our analysis revealed downregulation of GO terms and KEGG pathways related to extracellular matrix organization, involving genes such as laminins, collagens, and metalloproteinases, suggesting shared changes in skin tissue as a potential defensive mechanism to restrict *Borrelia* dissemination. However, by focusing on the skin as the infection site, we identified more differentially expressed genes compared to the analysis of the spleen transcriptome, which is consistent with the findings of a previous study of white-footed mice (*Peromyscus leucopus*) experimentally infected with *B. burgdorferi* s.s., which found that during persistent infection, more differentially expressed genes were present in the skin than in the blood (Long et al., 2019). The study by Long et al. (2019) and another experimental study by Gaber et al. (2023), which compared spleen transcriptome responses to *Borrelia* infection in *P. leucopus* and mice (*Mus musculus*), showed similar transcriptional patterns in response to infection as observed in our results. The study by Long et al. revealed that upregulated DEGs in the skin included immunoglobulin light chain genes and several keratin genes, while downregulated genes were related to the cytoskeleton or extracellular matrix, such as laminin, similar to our findings. They also found no genes associated with inflammation, which aligns with our results. The study by Gaber et al.

identified pathways such as hydroxycarboxylic acid-binding receptors, auto-degradation of Cdh1 by Cdh1:APC/C, hepatitis C, protein processing in endoplasmic reticulum, adipocytokine signaling pathway, and peroxisome proliferator-activated receptor (PPAR) signaling pathway. Our results are similar as we identified several metabolic pathways, cell cycle regulation processes, and immune responses to infection. These pathways suggest a shared biological response to *Borrelia* infection, reflecting both metabolic and immunoregulatory mechanisms. However, our study stands out from previous work by highlighting the role of adaptive B-cell mediated immune response, and by suggesting transcriptional changes related to neural health as a consequence of infection.

It is important to note that some changes in bank vole gene expression may not be caused by *Borrelia* itself, but rather by tick saliva, which is necessary to establish *Borrelia* infection (Bernard et al., 2020; Hovius, 2009). Tick saliva plays a crucial role in transmitting tick-borne pathogens and facilitating infection by suppressing the immune response of the host, thereby enabling pathogen survival. (Kazimirova & Stibraniova, 2013). Noteworthy, *Borrelia*, which is present in tick salivary glands, can modify the protein content of tick saliva to enhance its survival at the feeding site (Kim et al., 2021). Nevertheless, our findings are consistent with previous experimental studies, discussed earlier, in which *P. leucopus* were infected via inoculation, bypassing the influence of tick saliva. While tick saliva is known to suppress the host's immune response, these studies suggest that *Borrelia* itself plays a more significant role in altering host gene expression than tick saliva-mediated immune modulation.

The broad change that we observed in the expression of genes in the skin of infected individuals may contribute to the relatively mild or even asymptomatic course of *Borrelia* infection. Studies on bank voles reported no severe effects on their condition and survival (Sipari et al., 2022). An effective immune response, including the engagement of the adaptive immune response at the site of infection, which was evident in our data, may help to clear systemic infections. Indeed, earlier work on voles and *Peromyscus* indicates the presence of anti-*Borrelia* antibodies in a proportion of rodents that is much larger than the proportion of those infected at the sampling time (Bunikis et al., 2004; Różańska-Wróbel et al., 2024), indicating that the infection can be effectively cleared by most individuals. Yet, effective clearance may involve costs, as experimental infection with *Borrelia* altered the reproductive success of bank voles, particularly in males in a low-

density environment, where increased exploratory behavior is required (Cayol et al., 2018). Our analysis indicated that these costs may be mediated by energy metabolism, which was significantly affected by infection. It is thus possible that combating infection is energetically costly, either directly or indirectly through the production of highly reactive metabolites or the restructuring of metabolic pathways, which, in turn, limits resources that could otherwise be devoted to energetically demanding tasks, such as competition for mates. Additionally, our results indicate that altered metabolism may have consequences for nervous system functioning, possibly also contributing to decreased competitiveness of infected individuals.

To conclude, our results revealed the involvement of numerous genes related to immune responses, metabolism, cellular structure, gene expression, and cell cycle regulation, which aligns with previous studies on rodents (Gaber et al., 2023; Long et al., 2019; Zhong et al., 2020). Our results highlight the role of adaptive immune response to infection, but at the same time support the potential importance of modulation of inflammatory responses and changes in extracellular matrix composition in bank voles, which may help them reduce cell damage and prevent the establishment of chronic multi-organ disease. Additionally, we identified several GO terms and KEGG pathways associated with the nervous system, related to increased energy production, protein turnover, and responses to oxidative stress, suggesting potential involvement of neurodegenerative processes in infected bank voles, which has not been previously reported in similar studies. These findings provide new insights into host-pathogen interactions in reservoir hosts and offer potential directions for further investigation into the molecular mechanisms underlying *Borrelia* infection.

Author contributions

Joanna Różańska-Wróbel: Writing – original draft, Conceptualization, Methodology, Investigation, Formal analysis, Visualization. **Mateusz Konczal**: Writing – review and editing, Methodology, Investigation. **Rocco F. Notarnicola**: Writing – review and editing, Investigation. **Jacek Radwan**: Writing – review and editing, Conceptualization, Funding acquisition, Methodology, Investigation.

Acknowledgments

This work was funded by the National Science Centre, Poland, grant 2019/34/A/NZ8/00231.

Declaration of competing interest

The authors declare no competing interests.

Data availability statement

The data have been deposited in the Open Science Framework (OSF) repository and are publicly available at https://osf.io/4deyj (DOI: 10.17605/OSF.IO/4DEYJ).

References

Alexa, A., & Rahnenfuhrer, J. (2021). topGO: Enrichment Analysis for Gene Ontology. R package version 2.46.0

Andrews, S. (2010). FastQC: A Quality Control Tool for High Throughput Sequence Data. Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc

Bean, A. G. D., Baker, M. L., Stewart, C. R., Cowled, C., Deffrasnes, C., Wang, L.-F., & Lowenthal, J. W. (2013). Studying immunity to zoonotic diseases in the natural host—Keeping it real. Nature Reviews Immunology, 13(12), 851–861. https://doi.org/10.1038/nri3551

Behera, A. K., Hildebrand, E., Scagliotti, J., Steere, A. C., & Hu, L. T. (2005). Induction of host matrix metalloproteinases by Borrelia burgdorferi differs in human and murine lyme arthritis. Infection and immunity, 73(1), 126–134. https://doi.org/10.1128/IAI.73.1.126-134.2005

Bernard, Q., Grillon, A., Lenormand, C., Ehret-Sabatier, L., & Boulanger, N. (2020). Skin Interface, a Key Player for Borrelia Multiplication and Persistence in Lyme Borreliosis. Trends in Parasitology, 36(3), 304–314. https://doi.org/10.1016/j.pt.2019.12.017

Bouquet, J., Soloski, M. J., Swei, A., Cheadle, C., Federman, S., Billaud, J. N., Rebman, A. W., Kabre, B., Halpert, R., Boorgula, M., Aucott, J. N., & Chiu, C. Y. (2016). Longitudinal Transcriptome Analysis Reveals a Sustained Differential Gene Expression Signature in Patients Treated for Acute Lyme Disease. mBio, 7(1), e00100–e116. https://doi.org/10.1128/mBio.00100-16

Bunikis, J., Tsao, J., Luke, C. J., Luna, M. G., Fish, D., & Barbour, A. G. (2004). Borrelia burgdorferi infection in a natural population of Peromyscus Leucopus mice: a longitudinal study in an area where Lyme Borreliosis is highly endemic. The Journal of infectious diseases, 189(8), 1515–1523. https://doi.org/10.1086/382594

Cayol, C., Giermek, A., Gomez-Chamorro, A., Hytönen, J., Kallio, E. R., Mappes, T., Salo, J., Voordouw, M. J., & Koskela, E. (2018). Borrelia afzelii alters reproductive success in a rodent host. Proceedings of the Royal Society B: Biological Sciences, 285(1884), 20181056. https://doi.org/10.1098/rspb.2018.1056

Chen, S., Zhou, Y., Chen, Y., & Gu, J. (2018). fastp: An ultra-fast all-in-one FASTQ preprocessor. Bioinformatics, 34(17), i884–i890. https://doi.org/10.1093/bioinformatics/bty560

Chen, Y., Chen, L., Lun, A. T. L., Baldoni, P. L., & Smyth, G. K. (2024). edgeR v4: Powerful differential analysis of sequencing data with expanded functionality and improved support for small counts and larger datasets (p. 2024.01.21.576131). bioRxiv. https://doi.org/10.1101/2024.01.21.576131

Gaber, A. M., Mandric, I., Nitirahardjo, C., Piontkivska, H., Hillhouse, A. E., Threadgill, D. W., Zelikovsky, A., & Rogovskyy, A. S. (2023). Comparative transcriptome analysis of Peromyscus leucopus and C3H mice infected with the Lyme disease pathogen. Frontiers in Cellular and Infection Microbiology, 13. https://doi.org/10.3389/fcimb.2023.1115350

Gudowska-Sawczuk, M., & Mroczko, B. (2020). Chemokine Ligand 13 (CXCL13) in Neuroborreliosis and Neurosyphilis as Selected Spirochetal Neurological Diseases: A Review of Its Diagnostic Significance. International Journal of Molecular Sciences, 21(8), 2927. https://doi.org/10.3390/ijms21082927

Haydon, D. T., Cleaveland, S., Taylor, L. H., & Laurenson, M. K. (2002). Identifying Reservoirs of Infection: A Conceptual and Practical Challenge. Emerging Infectious Diseases, 8(12), 1468–1473. https://doi.org/10.3201/eid0812.010317

Hovius, J. W. R. (2009). Spitting Image: Tick Saliva Assists the Causative Agent of Lyme Disease in Evading Host Skin's Innate Immune Response. Journal of Investigative Dermatology, 129(10), 2337–2339. https://doi.org/10.1038/jid.2009.202

Humair, P. F., Rais, O., & Gern, L. (1999). Transmission of Borrelia afzelii from Apodemus mice and Clethrionomys voles to Ixodes ricinus ticks: Differential transmission pattern and overwintering maintenance. Parasitology, 118(1), 33–42. https://doi.org/10.1017/S0031182098003564

Kane, Y., Wong, G., & Gao, G. F. (2023). Animal Models, Zoonotic Reservoirs, and Cross-Species Transmission of Emerging Human-Infecting Coronaviruses. Annual review of animal biosciences, 11, 1–31. https://doi.org/10.1146/annurev-animal-020420-025011

Kazimirova, M., & Stibraniova, I. (2013). Tick salivary compounds: Their role in modulation of host defences and pathogen transmission. Frontiers in Cellular and Infection Microbiology, 3. https://doi.org/10.3389/fcimb.2013.00043

Kerstholt, M., Netea, M. G., & Joosten, L. A. B. (2020). Borrelia burgdorferi hijacks cellular metabolism of immune cells: Consequences for host defense. Ticks and Tick-Borne Diseases, 11(3), 101386. https://doi.org/10.1016/j.ttbdis.2020.101386

Kim, T. K., Tirloni, L., Bencosme-Cuevas, E., Kim, T. H., Diedrich, J. K., Yates, J. R., & Mulenga, A. (2021). Borrelia burgdorferi infection modifies protein content in saliva of Ixodes scapularis nymphs. BMC Genomics, 22(1), 152. https://doi.org/10.1186/s12864-021-07429-0

Kolde, R. (2019). Pheatmap: Pretty Heatmaps. R package version 1.0.12.

Langfelder, P., & Horvath, S. (2008). WGCNA: An R package for weighted correlation network analysis. BMC Bioinformatics, 9(1), 559. https://doi.org/10.1186/1471-2105-9-559

Langfelder, P., & Horvath, S. (2012). Fast R Functions for Robust Correlations and Hierarchical Clustering. Journal of Statistical Software, 46, 1–17. https://doi.org/10.18637/jss.v046.i11

Lawrence, M., Gentleman, R., & Carey, V. (2009). rtracklayer: An R package for interfacing with genome browsers. Bioinformatics, 25(14), 1841–1842. https://doi.org/10.1093/bioinformatics/btp328

Lin, B., Kidder, J. M., Noring, R., Steere, A. C., Klempner, M. S., & Hu, L. T. (2001). Differences in synovial fluid levels of matrix metalloproteinases suggest separate mechanisms of pathogenesis in Lyme arthritis before and after antibiotic

treatment. The Journal of infectious diseases, 184(2), 174–180. https://doi.org/10.1086/322000

Liveris, D., Varde, S., Iyer, R., Koenig, S., Bittker, S., Cooper, D., McKenna, D., Nowakowski, J., Nadelman, R. B., Wormser, G. P., & Schwartz, I. (1999). Genetic Diversity of Borrelia burgdorferi in Lyme Disease Patients as Determined by Culture versus Direct PCR with Clinical Specimens. Journal of Clinical Microbiology, 37(3), 565–569. https://doi.org/10.1128/JCM.37.3.565-569.1999

Long, A. D., Baldwin-Brown, J., Tao, Y., Cook, V. J., Balderrama-Gutierrez, G., Corbett-Detig, R., Mortazavi, A., & Barbour, A. G. (2019). The genome of Peromyscus leucopus, natural host for Lyme disease and other emerging infections. Science Advances, 5(7), eaaw6441. https://doi.org/10.1126/sciadv.aaw6441

Marques, A., Schwartz, I., Wormser, G. P., Wang, Y., Hornung, R. L., Demirkale, C. Y., Munson, P. J., Turk, S. P., Williams, C., Lee, C. R., Yang, J., & Petzke, M. M. (2017). Transcriptome Assessment of Erythema Migrans Skin Lesions in Patients With Early Lyme Disease Reveals Predominant Interferon Signaling. The Journal of infectious diseases, 217(1), 158–167. https://doi.org/10.1093/infdis/jix563

Marques, A. R., Strle, F., & Wormser, G. P. (2021). Comparison of Lyme Disease in the United States and Europe. Emerging Infectious Diseases, 27(8), 2017–2024. https://doi.org/10.3201/eid2708.204763

Mastroeni, P., Vazquez-Torres, A., Fang, F. C., Xu, Y., Khan, S., Hormaeche, C. E., & Dougan, G. (2000). Antimicrobial Actions of the Nadph Phagocyte Oxidase and Inducible Nitric Oxide Synthase in Experimental Salmonellosis. II. Effects on Microbial Proliferation and Host Survival in Vivo. The Journal of Experimental Medicine, 192(2), 237–248.

Mead, P. S. (2015). Epidemiology of Lyme Disease. Infectious Disease Clinics of North America, 29(2), 187–210. https://doi.org/10.1016/j.idc.2015.02.010

Morens, D. M., Folkers, G. K., & Fauci, A. S. (2004). The challenge of emerging and re-emerging infectious diseases. Nature, 430(6996), 242–249. https://doi.org/10.1038/nature02759

Müllner, D. (2013). fastcluster: Fast Hierarchical, Agglomerative Clustering Routines for R and Python. Journal of Statistical Software, 53, 1–18. https://doi.org/10.18637/jss.v053.i09

Oosting, M., Kerstholt, M., Horst, R. ter, Li, Y., Deelen, P., Smeekens, S., Jaeger, M., Lachmandas, E., Vrijmoeth, H., Lupse, M., Flonta, M., Cramer, R. A., Kullberg, B. J., Kumar, V., Xavier, R., Wijmenga, C., Netea, M. G., & Joosten, L. A. B. (2016). Functional and Genomic Architecture of Borrelia burgdorferi-Induced Cytokine Responses in Humans. Cell Host & Microbe, 20(6), 822–833. https://doi.org/10.1016/j.chom.2016.10.006

Patro, R., Duggal, G., Love, M. I., Irizarry, R. A., & Kingsford, C. (2017). Salmon provides fast and bias-aware quantification of transcript expression. Nature Methods, 14(4), 417–419. https://doi.org/10.1038/nmeth.4197

Pérez-Mazliah, D., Ng, D. H. L., Rosário, A. P. F. do, McLaughlin, S., Mastelic-Gavillet, B., Sodenkamp, J., Kushinga, G., & Langhorne, J. (2015). Disruption of IL-21 Signaling

Affects T Cell-B Cell Interactions and Abrogates Protective Humoral Immunity to Malaria. PLOS Pathogens, 11(3), e1004715. https://doi.org/10.1371/journal.ppat.1004715

Petzke, M. M., Iyer, R., Love, A. C., Spieler, Z., Brooks, A., & Schwartz, I. (2016). Borrelia burgdorferi induces a type I interferon response during early stages of disseminated infection in mice. BMC Microbiology, 16(1), 29. https://doi.org/10.1186/s12866-016-0644-4

Pilz, G., Sakic, I., Wipfler, P., Kraus, J., Haschke-Becher, E., Hitzl, W., Trinka, E., & Harrer, A. (2020). Chemokine CXCL13 in serum, CSF and blood–CSF barrier function: Evidence of compartment restriction. Fluids and Barriers of the CNS, 17(1), 7. https://doi.org/10.1186/s12987-020-0170-5

R Core Team. (2024). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Rahman, M. T., Sobur, M. A., Islam, M. S., Ievy, S., Hossain, M. J., El Zowalaty, M. E., Rahman, A. T., & Ashour, H. M. (2020). Zoonotic Diseases: Etiology, Impact, and Control. Microorganisms, 8(9), 1405. https://doi.org/10.3390/microorganisms8091405

Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., & Smyth, G. K. (2015). Limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Research, 43(7), e47. https://doi.org/10.1093/nar/gkv007

Robinson, M. D., McCarthy, D. J., & Smyth, G. K. (2010). edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics, 26(1), 139–140. https://doi.org/10.1093/bioinformatics/btp616

Robinson, M. D., & Oshlack, A. (2010). A scaling normalization method for differential expression analysis of RNA-seq data. Genome Biology, 11(3), R25. https://doi.org/10.1186/gb-2010-11-3-r25

Rogovskyy, A. S., Pliasas, V. C., Buhrer, R., Lewy, K., Wiener, D. J., Jung, Y., Bova, J., Rogovska, Y., Kim, S. J., & Jeon, E. “Grace”. (2024). Do white-footed mice, the main reservoir of the Lyme disease pathogen in the United States, clinically respond to the borrelial tenancy? Infection and Immunity, 92(12), e00382-24. https://doi.org/10.1128/iai.00382-24

Różańska-Wróbel, J., Migalska, M., Urbanowicz, A., Grzybek, M., Rego, R. O. M., Bajer, A., Dwuznik-Szarek, D., Alsarraf, M., Behnke-Borowczyk, J., Behnke, J. M., & Radwan, J. (2024). Interplay between vertebrate adaptive immunity and bacterial infectivity genes: Bank vole MHC versus Borrelia afzelii OspC. Molecular ecology, 33(21), e17534. https://doi.org/10.1111/mec.17534

Salazar, J. C., Duhnam-Ems, S., Vake, C. L., Cruz, A. R., Moore, M. W., Caimano, M. J., Velez-Climent, L., Shupe, J., Krueger, W., & Radolf, J. D. (2009). Activation of Human Monocytes by Live Borrelia burgdorferi Generates TLR2-Dependent and -Independent Responses Which Include Induction of IFN-β. PLOS Pathogens, 5(5), e1000444. https://doi.org/10.1371/journal.ppat.1000444

Scarlata, C.-M., Celse, C., Pignon, P., Ayyoub, M., & Valmori, D. (2015). Differential expression of the immunosuppressive enzyme IL4I1 in human induced Aiolos+,

but not natural Helios+, FOXP3+ Treg cells. European Journal of Immunology, 45(2), 474–479. https://doi.org/10.1002/eji.201444897

Schramm, F., Kern, A., Barthel, C., Nadaud, S., Meyer, N., Jaulhac, B., & Boulanger, N. (2012). Microarray analyses of inflammation response of human dermal fibroblasts to different strains of Borrelia burgdorferi sensu stricto. PloS one, 7(6), e40046. https://doi.org/10.1371/journal.pone.0040046

Sipari, S., Hytönen, J., Pietikäinen, A., Mappes, T., & Kallio, E. R. (2022). The effects of Borrelia infection on its wintering rodent host. Oecologia, 200(3–4), 471–478. https://doi.org/10.1007/s00442-022-05272-y

Soneson, C., Love, M. I., & Robinson, M. D. (2016). Differential analyses for RNA-seq: Transcript-level estimates improve gene-level inferences. F1000Research, 4, 1521. https://doi.org/10.12688/f1000research.7563.2

Stanek, G., Wormser, G. P., Gray, J., & Strle, F. (2012). Lyme borreliosis. The Lancet, 379(9814), 461–473. https://doi.org/10.1016/S0140-6736(11)60103-7

Steere, A. C., Strle, F., Wormser, G. P., Hu, L. T., Branda, J. A., Hovius, J. W. R., Li, X., & Mead, P. S. (2016). Lyme borreliosis. Nature Reviews Disease Primers, 2(1), 16090. https://doi.org/10.1038/nrdp.2016.90

Thompson, D., Watt, J. A., & Brissette, C. A. (2021). Host transcriptome response to Borrelia burgdorferi sensu lato. Ticks and Tick-Borne Diseases, 12(2), 101638. https://doi.org/10.1016/j.ttbdis.2020.101638

van Duijvendijk, G., Sprong, H., & Takken, W. (2015). Multi-trophic interactions driving the transmission cycle of Borrelia afzelii between Ixodes ricinus and rodents: A review. Parasites & Vectors, 8(1), 643. https://doi.org/10.1186/s13071-015-1257-8

Verma, A., Brissette, C. A., Bowman, A., & Stevenson, B. (2009). Borrelia burgdorferi BmpA is a laminin-binding protein. Infection and immunity, 77(11), 4940–4946. https://doi.org/10.1128/IAI.01420-08

Wickham, H. (2016). ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York.

Wu, T., Hu, E., Xu, S., Chen, M., Guo, P., Dai, Z., Feng, T., Zhou, L., Tang, W., Zhan, L., Fu, X., Liu, S., Bo, X., & Yu, G. (2021). clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. The Innovation, 2(3), 100141. https://doi.org/10.1016/j.xinn.2021.100141

Yu, G., Wang, L.-G., Han, Y., & He, Q.-Y. (2012). clusterProfiler: An R Package for Comparing Biological Themes Among Gene Clusters. OMICS: A Journal of Integrative Biology, 16(5), 284–287. https://doi.org/10.1089/omi.2011.0118

Zambrano, M. C., Beklemisheva, A. A., Bryksin, A. V., Newman, S. A., & Cabello, F. C. (2004). Borrelia burgdorferi binds to, invades, and colonizes native type I collagen lattices. Infection and immunity, 72(6), 3138–3146. https://doi.org/10.1128/IAI.72.6.3138-3146.2004

Zhong, X., Lundberg, M., & Råberg, L. (2020). Comparison of spleen transcriptomes of two wild rodent species reveals differences in the immune response against Borrelia

afzelii. Ecology and Evolution, 10(13), 6421–6434. https://doi.org/10.1002/ece3.6377

Zhong, X., Nouri, M., & Råberg, L. (2019). Colonization and pathology of Borrelia afzelii in its natural hosts. Ticks and tick-borne diseases, 10(4), 822–827. https://doi.org/10.1016/j.ttbdis.2019.03.017

Supplementary Figures

Transcriptomic response to *Borrelia afzelii* infection in the skin of wild bank voles

Joanna Różańska-Wróbel[1], Mateusz Konczal[1], Rocco F. Notarnicola[1], Jacek Radwan[1]

[1] Evolutionary Biology Group, Institute of Environmental Biology, Faculty of Biology, Adam Mickiewicz University, Poznań, Poland

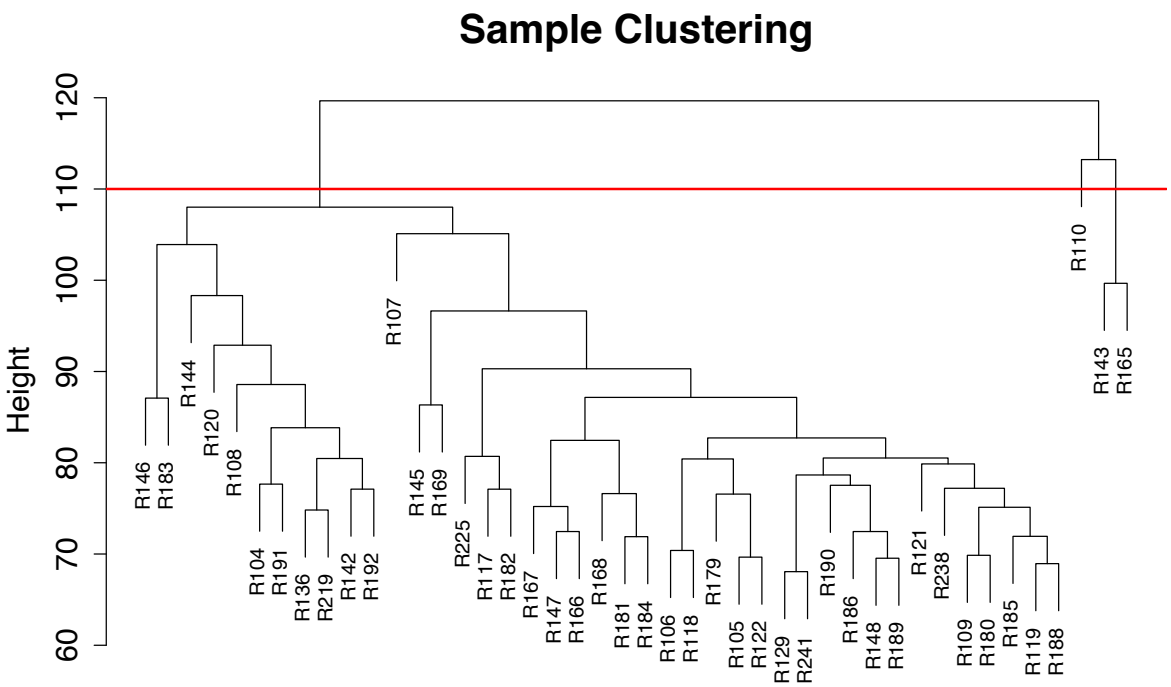


Figure S1. Hierarchical clustering dendrogram of samples. Samples above the cut height of 110 (R110, R143 and R165) were identified as outliers and removed from the WGCNA analysis.

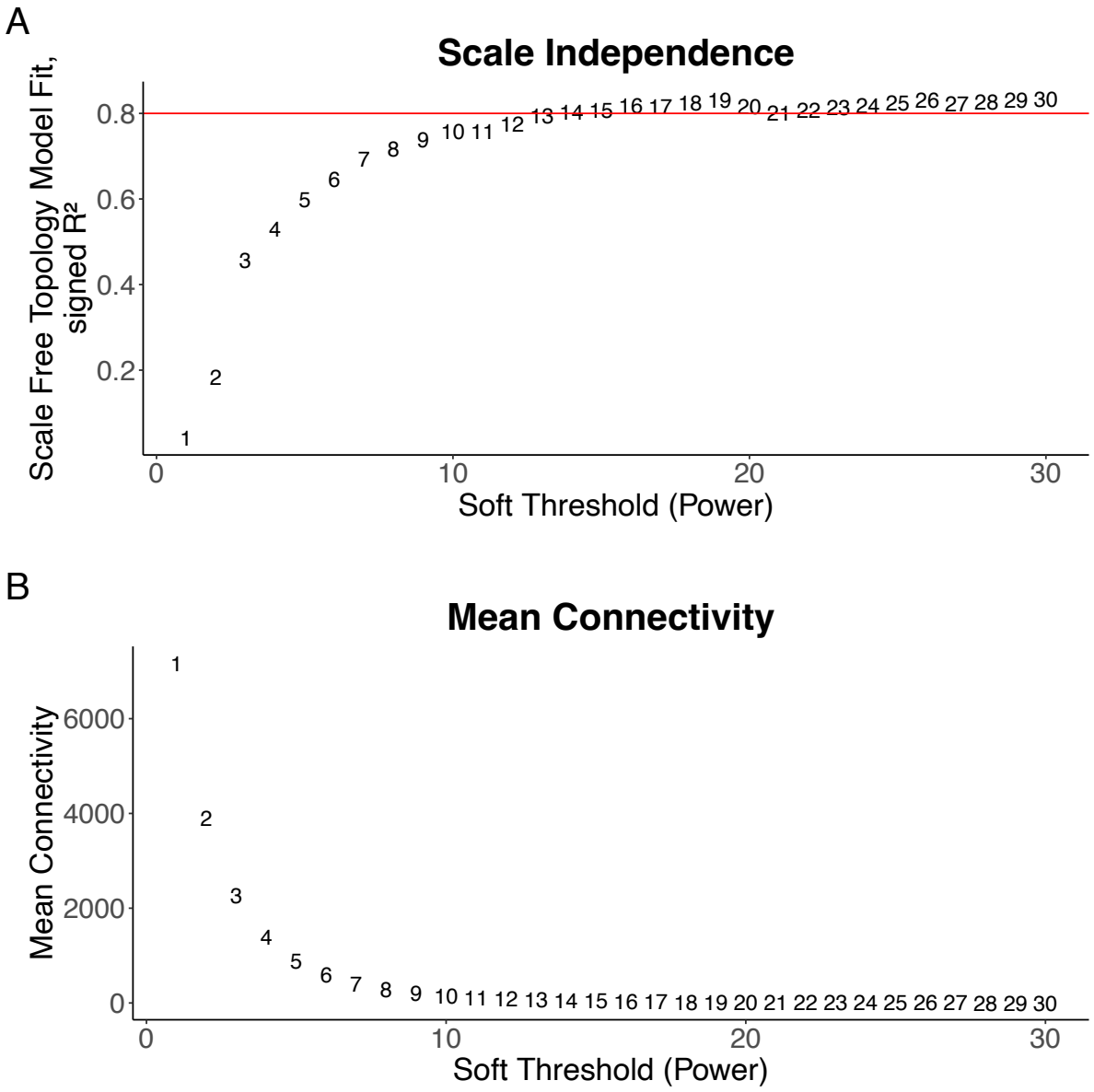


Figure S2. Scale-free topology analysis under different soft thresholding powers (1 – 30) performed to determine the optimal power for WGCNA network construction: A) scale-free topology model fit; B) mean connectivity. A power of 14 was chosen, resulting in an R^2 of 0.80 and a mean connectivity of 55.80, which ensures a reasonable approximation to a scale-free topology while maintaining sufficient network connectivity.

Supplementary Tables

Transcriptomic response to *Borrelia afzelii* infection in the skin of wild bank voles

Joanna Różańska-Wróbel[1], Mateusz Konczal[1], Rocco F. Notarnicola[1], Jacek Radwan[1]

[1] Evolutionary Biology Group, Institute of Environmental Biology, Faculty of Biology, Adam Mickiewicz University, Poznań, Poland

Table S1. Bank vole samples used for RNA sequencing.

Bank vole ID	Sampling site	Body mass [g]	*B. afzelii* infection status
R104	Brok	21.0	infected
R105	Brok	25.0	infected
R106	Brok	12.0	infected
R107	Brok	22.0	infected
R108	Brok	21.0	infected
R109	Brok	21.0	uninfected
R110	Brok	23.0	uninfected
R117	Długosiodło	24.0	infected
R118	Długosiodło	26.0	infected
R119	Długosiodło	17.0	uninfected
R120	Długosiodło	22.0	uninfected
R121	Długosiodło	16.0	uninfected
R122	Długosiodło	25.0	infected
R129	Długosiodło	20.0	infected
R136	Długosiodło	25.0	infected
R142	Długosiodło	24.0	infected
R143	Długosiodło	15.0	uninfected

R144	Długosiodło	23.0	infected
R145	Długosiodło	24.0	infected
R146	Długosiodło	23.0	infected
R147	Długosiodło	12.0	uninfected
R148	Długosiodło	11.0	uninfected
R165	Grobka	17.0	uninfected
R166	Grobka	26.0	uninfected
R167	Grobka	19.0	uninfected
R168	Grobka	18.0	uninfected
R169	Grobka	16.0	infected
R179	Grobka	20.5	infected
R180	Grobka	16.0	uninfected
R181	Grobka	16.0	uninfected
R182	Grobka	16.0	uninfected
R183	Grobka	19.0	infected
R184	Grobka	17.0	uninfected
R185	Grobka	18.0	uninfected
R186	Grobka	15.5	uninfected
R188	Grobka	18.0	uninfected
R189	Grobka	17.0	uninfected
R190	Grobka	18.0	infected
R191	Grobka	18.5	infected
R192	Grobka	15.5	uninfected
R219	Grobka	17.5	infected
R225	Grobka	20.5	infected
R238	Grobka	25.5	infected
R241	Grobka	21.0	infected

Table S2. Differentially expressed genes between bank voles infected with *Borrelia afzelii* and uninfected ones. A positive logFC value indicates that a gene is upregulated in infected voles compared to uninfected ones, while negative logFC value indicates that a gene is downregulated in infected voles compared to uninfected.

Gene	logFC	logCPM	F	PValue	FDR
LOC125407534	6.16	7.58	44.52	3.88E-08	5.55E-04
Jchain	4.14	5.94	41.56	8.25E-08	5.91E-04
LOC125394830	4.90	4.27	31.45	1.36E-06	0.006
LOC125412777	5.44	6.59	29.99	2.11E-06	0.008
Pou2af1	2.74	2.38	26.65	5.96E-06	0.017
LOC125412350	1.54	2.05	23.93	1.44E-05	0.021
Nt5dc2	0.73	4.52	23.93	1.45E-05	0.021
Cxcl13	3.24	5.20	23.76	1.53E-05	0.021
Nadsyn1	0.70	3.15	23.74	1.54E-05	0.021
Cmpk2	-1.23	4.12	23.31	1.78E-05	0.021
LOC125414983	2.94	5.66	22.80	2.12E-05	0.023
LOC125410190	3.04	1.83	22.48	2.36E-05	0.023
Apoo	0.62	3.76	22.46	2.38E-05	0.023
LOC125387327	3.59	5.00	21.22	3.64E-05	0.033
Cmas	0.72	4.83	20.50	4.69E-05	0.037
Stra6	1.58	3.29	19.89	5.81E-05	0.040
LOC125403498	-1.55	2.90	19.72	6.19E-05	0.040
LOC125394431	2.83	4.07	19.61	6.43E-05	0.040
Nqo2	0.96	3.15	19.37	7.00E-05	0.040
LOC125417898	0.61	5.82	19.08	7.78E-05	0.040
Ifit3	-2.33	5.48	18.85	8.48E-05	0.040
Pop5	0.61	4.04	18.67	9.03E-05	0.040

Sdf2l1	0.85	4.52	18.53	9.52E-05	0.040
Exosc5	0.58	5.01	18.36	1.01E-04	0.040
Bmpr1b	0.70	2.43	18.36	1.01E-04	0.040
Helz2	-0.70	6.16	18.30	1.03E-04	0.040
Znf503	-0.60	5.77	18.27	1.05E-04	0.040
Dhfr	0.66	2.67	18.17	1.09E-04	0.040
Alg14	0.64	3.83	17.93	1.19E-04	0.040
Igf2	-1.10	4.37	17.88	1.21E-04	0.040
Rsad2	-1.66	3.40	17.41	1.44E-04	0.043
LOC125404151	0.81	7.10	17.33	1.48E-04	0.043
LOC125400620	1.93	4.89	17.32	1.49E-04	0.043
Cenpn	0.80	2.29	17.20	1.56E-04	0.043
Vsnl1	1.19	2.55	17.18	1.57E-04	0.043
Kif22	0.59	4.52	16.89	1.75E-04	0.045
Ca12	0.97	7.26	16.79	1.81E-04	0.046
Asf1b	0.61	4.16	16.68	1.90E-04	0.048
Dkk2	-0.93	4.01	16.55	1.99E-04	0.049
Ddit3	-0.62	3.50	16.44	2.07E-04	0.049
Slc2a1	1.08	8.26	16.41	2.10E-04	0.049
Fgfbp1	2.16	5.90	16.22	2.26E-04	0.049
Impdh1	0.89	7.18	16.21	2.26E-04	0.049
Dgat1	0.76	5.76	16.08	2.38E-04	0.049
Il21r	0.99	1.44	15.95	2.51E-04	0.049
Rhbdl1	1.12	2.34	15.90	2.56E-04	0.049
Cklf	0.82	2.70	15.66	2.80E-04	0.049
Dlk1	-1.35	4.50	15.58	2.89E-04	0.049
Shmt1	0.60	3.93	15.55	2.92E-04	0.049

Fam83a	1.26	2.72	15.43	3.06E-04	0.049
LOC125403493	-1.71	4.27	15.31	3.21E-04	0.049
Il4i1	0.63	3.29	15.29	3.23E-04	0.049
LOC125403960	2.13	6.56	15.27	3.25E-04	0.049
LOC125416672	1.02	0.27	15.27	3.26E-04	0.049
LOC125403565	0.91	7.19	15.21	3.34E-04	0.049
Scnn1g	1.48	2.82	15.09	3.50E-04	0.049
Sapcd2	0.83	4.57	15.08	3.51E-04	0.049
Ephb2	0.74	3.86	15.04	3.56E-04	0.049
Nr1d1	-0.69	6.72	15.04	3.57E-04	0.049
LOC125410899	0.61	4.04	14.97	3.66E-04	0.049
Zbtb32	1.11	1.23	14.91	3.74E-04	0.049
Hrob	0.97	1.36	14.89	3.77E-04	0.049
LOC125408667	-1.11	4.22	14.87	3.82E-04	0.049
Krt1	1.44	12.62	14.79	3.93E-04	0.049
Tpi1	0.72	7.61	14.75	4.00E-04	0.049
Adm2	1.38	1.08	14.74	4.01E-04	0.049
Msx1	-0.87	4.78	14.65	4.16E-04	0.049
LOC125402185	-0.99	3.01	14.63	4.19E-04	0.049
Tmem61	0.99	2.72	14.62	4.21E-04	0.049
Ifngr1	0.81	7.89	14.57	4.28E-04	0.049
LOC125397965	1.13	6.12	14.57	4.29E-04	0.049
Kif4a	0.62	4.31	14.55	4.32E-04	0.049
CUNH4orf33	0.79	1.47	14.51	4.38E-04	0.049

Table S3. Summary of GO biological process enrichment analysis of differentially expressed genes.

GO ID	Term	Annotated	Significant	Expected	weightedFisher
GO:0051607	defense response to virus	44	5	0.16	5.40E-07
GO:0006544	glycine metabolic process	11	2	0.04	7.10E-04
GO:0046653	tetrahydrofolate metabolic process	13	2	0.05	0.001
GO:0009070	serine family amino acid biosynthetic process	13	2	0.05	0.001
GO:0006767	water-soluble vitamin metabolic process	17	2	0.06	0.002
GO:0031424	keratinization	23	2	0.08	0.003
GO:0070098	chemokine-mediated signaling pathway	27	2	0.1	0.004
GO:0061844	antimicrobial humoral immune response mediated by antimicrobial peptide	29	2	0.11	0.005
GO:0030593	neutrophil chemotaxis	34	2	0.12	0.007
GO:0043648	dicarboxylic acid metabolic process	34	2	0.12	0.007
GO:0045109	intermediate filament organization	37	2	0.14	0.008
GO:0071222	cellular response to lipopolysaccharide	41	2	0.15	0.010
GO:0043603	cellular amide metabolic process	964	5	3.54	0.013
GO:0071216	cellular response to biotic stimulus	48	3	0.18	0.014
GO:0007052	mitotic spindle organization	52	2	0.19	0.016

GO:0009063	cellular amino acid catabolic process	56	2	0.21	0.018
GO:0006334	nucleosome assembly	57	2	0.21	0.019
GO:0072522	purine-containing compound biosynthetic process	109	2	0.4	0.029
GO:0045944	positive regulation of transcription by RNA polymerase II	192	3	0.71	0.033
GO:0033993	response to lipid	80	3	0.29	0.035
GO:0006490	oligosaccharide-lipid intermediate biosynthetic process	10	1	0.04	0.036
GO:0006984	ER-nucleus signaling pathway	10	1	0.04	0.036
GO:0009263	deoxyribonucleotide biosynthetic process	10	1	0.04	0.036
GO:0070059	intrinsic apoptotic signaling pathway in response to endoplasmic reticulum stress	10	1	0.04	0.036
GO:0042559	pteridine-containing compound biosynthetic process	10	1	0.04	0.036
GO:0071025	RNA surveillance	10	1	0.04	0.036
GO:0009132	nucleoside diphosphate metabolic process	41	2	0.15	0.039
GO:0015749	monosaccharide transmembrane transport	11	1	0.04	0.040
GO:0034427	nuclear-transcribed mRNA catabolic process, exonucleolytic, 3'-5'	11	1	0.04	0.040
GO:0001682	tRNA 5'-leader removal	11	1	0.04	0.040
GO:0006183	GTP biosynthetic process	11	1	0.04	0.040

GO:0006144	purine nucleobase metabolic process	11	1	0.04	0.040
GO:0009071	serine family amino acid catabolic process	11	1	0.04	0.040
GO:0031055	chromatin remodeling at centromere	11	1	0.04	0.040
GO:0009394	2'-deoxyribonucleotide metabolic process	11	1	0.04	0.040
GO:0006636	unsaturated fatty acid biosynthetic process	11	1	0.04	0.040
GO:0019432	triglyceride biosynthetic process	12	1	0.04	0.043
GO:0031126	sno(s)RNA 3'-end processing	12	1	0.04	0.043
GO:0002503	peptide antigen assembly with MHC class II protein complex	12	1	0.04	0.043
GO:0009435	NAD biosynthetic process	12	1	0.04	0.043
GO:0006730	one-carbon metabolic process	12	1	0.04	0.043
GO:0006695	cholesterol biosynthetic process	12	1	0.04	0.043
GO:0009148	pyrimidine nucleoside triphosphate biosynthetic process	13	1	0.05	0.047
GO:0042832	defense response to protozoan	13	1	0.05	0.047
GO:0007007	inner mitochondrial membrane organization	13	1	0.05	0.047
GO:0019886	antigen processing and presentation of exogenous peptide antigen via MHC class II	13	1	0.05	0.047

Table S4. Summary GO biological process enrichment analysis of the blue gene module identified by WGCNA.

GO ID	Term	Annotated	Significant	Expected	weightedFisher
GO:0006096	glycolytic process	28	10	0.38	1.90E-12
GO:0006418	tRNA aminoacylation for protein translation	42	10	0.57	1.80E-10
GO:0046519	sphingoid metabolic process	11	3	0.15	3.80E-04
GO:0046475	glycerophospholipid catabolic process	15	3	0.20	0.001
GO:0006094	gluconeogenesis	18	3	0.25	0.002
GO:0031424	keratinization	23	3	0.31	0.004
GO:0034314	Arp2/3 complex-mediated actin nucleation	25	3	0.34	0.005
GO:0071456	cellular response to hypoxia	10	2	0.14	0.008
GO:0006596	polyamine biosynthetic process	10	2	0.14	0.008
GO:0044282	small molecule catabolic process	148	5	2.02	0.009
GO:0006783	heme biosynthetic process	12	2	0.16	0.011
GO:0031126	sno(s)RNA 3'-end processing	12	2	0.16	0.011
GO:0046031	ADP metabolic process	29	11	0.40	0.013
GO:0009070	serine family amino acid biosynthetic process	13	2	0.18	0.013
GO:0034312	diol biosynthetic process	13	2	0.18	0.013
GO:0072331	signal transduction by p53 class mediator	13	2	0.18	0.013
GO:0072698	protein localization to microtubule cytoskeleton	13	2	0.18	0.013

GO:2001022	positive regulation of response to DNA damage stimulus	14	2	0.19	0.014
GO:0030834	regulation of actin filament depolymerization	30	2	0.41	0.014
GO:0040011	locomotion	455	6	6.21	0.014
GO:0045109	intermediate filament organization	37	3	0.50	0.014
GO:0006508	proteolysis	1034	23	14.11	0.014
GO:0046112	nucleobase biosynthetic process	15	2	0.20	0.017
GO:0006767	water-soluble vitamin metabolic process	17	2	0.23	0.022
GO:0002183	cytoplasmic translational initiation	18	2	0.25	0.025
GO:0044275	cellular carbohydrate catabolic process	18	2	0.25	0.025
GO:0051156	glucose 6-phosphate metabolic process	19	2	0.26	0.027
GO:0000278	mitotic cell cycle	326	13	4.45	0.028
GO:0007052	mitotic spindle organization	52	3	0.71	0.034
GO:0006606	protein import into nucleus	64	4	0.87	0.035
GO:0015931	nucleobase-containing compound transport	80	5	1.09	0.038
GO:0034311	diol metabolic process	16	3	0.22	0.040
GO:0033013	tetrapyrrole metabolic process	20	3	0.27	0.040
GO:0098969	neurotransmitter receptor transport to postsynaptic membrane	13	2	0.18	0.040

GO:1901264	carbohydrate derivative transport	24	2	0.33	0.042
GO:0030833	regulation of actin filament polymerization	102	6	1.39	0.044

Table S5. Summary of KEGG pathway enrichment analysis of the blue gene module identified by WGCNA.

ID	Description	GeneRatio	BgRatio	pvalue	p.adjust	qvalue
map01110	Biosynthesis of secondary metabolites	28/149	294/5881	6.86E-10	1.26E-07	1.14E-07
map00970	Aminoacyl-tRNA biosynthesis	10/149	29/5881	1.07E-09	1.26E-07	1.14E-07
map00010	Glycolysis / Gluconeogenesis	10/149	35/5881	8.60E-09	6.76E-07	6.12E-07
map01230	Biosynthesis of amino acids	9/149	45/5881	1.37E-06	8.11E-05	7.34E-05
map01200	Carbon metabolism	11/149	79/5881	3.92E-06	1.85E-04	1.68E-04
map04066	HIF-1 signaling pathway	10/149	75/5881	1.63E-05	6.41E-04	5.80E-04
map00680	Methane metabolism	5/149	15/5881	2.39E-05	8.05E-04	7.29E-04
map01120	Microbial metabolism in diverse environments	12/149	122/5881	5.30E-05	0.002	0.001
map05230	Central carbon metabolism in cancer	7/149	46/5881	1.36E-04	0.003	0.003
map00051	Fructose and mannose metabolism	5/149	21/5881	1.43E-04	0.003	0.003
map00270	Cysteine and methionine metabolism	6/149	38/5881	3.38E-04	0.007	0.007
map01232	Nucleotide metabolism	6/149	51/5881	0.002	0.033	0.030
map00450	Selenocompound metabolism	3/149	11/5881	0.002	0.041	0.037
map04217	Necroptosis	8/149	94/5881	0.002	0.042	0.038
map04922	Glucagon signaling pathway	6/149	56/5881	0.003	0.043	0.039
map05130	Pathogenic Escherichia coli infection	9/149	120/5881	0.003	0.048	0.043

Table S6. Summary GO biological process enrichment analysis of the yellow gene module idenfitied by WGCNA.

GO ID	Term	Annotated	Significant	Expected	weightedFisher
GO:0015986	ATP synthesis coupled proton transport	23	12	0.37	2.40E-16
GO:0006457	protein folding	168	16	2.68	2.30E-09
GO:0006120	mitochondrial electron transport, NADH to ubiquinone	16	7	0.26	2.40E-09
GO:0006122	mitochondrial electron transport, ubiquinol to cytochrome c	19	6	0.30	3.50E-07
GO:0000387	spliceosomal snRNP assembly	30	7	0.48	3.60E-07
GO:0032543	mitochondrial translation	33	7	0.53	7.20E-07
GO:0006123	mitochondrial electron transport, cytochrome c to oxygen	32	6	0.51	9.90E-06
GO:0051603	proteolysis involved in cellular protein catabolic process	435	26	6.93	1.00E-05
GO:0000413	protein peptidyl-prolyl isomerization	35	6	0.56	1.70E-05
GO:1990542	mitochondrial transmembrane transport	82	12	1.31	3.20E-05
GO:0030150	protein import into mitochondrial matrix	24	5	0.38	3.30E-05
GO:0006413	translational initiation	67	8	1.07	9.70E-05
GO:0006119	oxidative phosphorylation	72	22	1.15	3.50E-04
GO:0006098	pentose-phosphate shunt	10	3	0.16	4.40E-04
GO:0006886	intracellular protein transport	411	20	6.55	8.70E-04

GO:0006414	translational elongation	39	4	0.62	0.003
GO:0072583	clathrin-dependent endocytosis	20	3	0.32	0.004
GO:0006099	tricarboxylic acid cycle	24	3	0.38	0.006
GO:0006511	ubiquitin-dependent protein catabolic process	357	15	5.69	0.009
GO:0033962	P-body assembly	10	2	0.16	0.010
GO:0006183	GTP biosynthetic process	11	2	0.18	0.013
GO:0006241	CTP biosynthetic process	11	2	0.18	0.013
GO:0006901	vesicle coating	11	2	0.18	0.013
GO:0006465	signal peptide processing	11	2	0.18	0.013
GO:0043248	proteasome assembly	11	2	0.18	0.013
GO:0006415	translational termination	11	2	0.18	0.013
GO:0006412	translation	766	34	12.21	0.016
GO:0007005	mitochondrion organization	201	14	3.20	0.017
GO:0045116	protein neddylation	13	2	0.21	0.018
GO:0048199	vesicle targeting, to, from or within Golgi	13	2	0.21	0.018
GO:0006406	mRNA export from nucleus	36	3	0.57	0.019
GO:0006487	protein N-linked glycosylation	50	4	0.80	0.021
GO:0033617	mitochondrial cytochrome c oxidase assembly	15	2	0.24	0.023
GO:2000045	regulation of G1/S transition of mitotic cell cycle	16	2	0.26	0.026
GO:0071230	cellular response to amino acid stimulus	16	2	0.26	0.026
GO:0032233	positive regulation of actin filament bundle assembly	17	2	0.27	0.029

GO:0000398	mRNA splicing, via spliceosome	212	13	3.38	0.030
GO:0043161	proteasome-mediated ubiquitin-dependent protein catabolic process	230	8	3.67	0.031
GO:0032511	late endosome to vacuole transport via multivesicular body sorting pathway	18	2	0.29	0.033
GO:0006446	regulation of translational initiation	19	2	0.30	0.036
GO:0032008	positive regulation of TOR signaling	21	2	0.33	0.044
GO:0050821	protein stabilization	21	2	0.33	0.044
GO:0032970	regulation of actin filament-based process	137	6	2.18	0.047
GO:0050658	RNA transport	57	4	0.91	0.047
GO:0032981	mitochondrial respiratory chain complex I assembly	22	2	0.35	0.047
GO:0006891	intra-Golgi vesicle-mediated transport	22	2	0.35	0.047

Table S7. Summary of KEGG pathway enrichment analysis of the yellow gene module identified by WGCNA.

ID	Description	GeneRatio	BgRatio	pvalue	p.adjust	qvalue
map05012	Parkinson disease	60/180	192/5881	6.85E-47	9.23E-45	8.30E-45
map05016	Huntington disease	62/180	211/5881	1.04E-46	9.23E-45	8.30E-45
map05020	Prion disease	57/180	193/5881	7.76E-43	4.61E-41	4.14E-41
map00190	Oxidative phosphorylation	40/180	97/5881	2.13E-36	9.49E-35	8.53E-35
map05014	Amyotrophic lateral sclerosis	57/180	262/5881	9.01E-35	3.21E-33	2.88E-33
map05010	Alzheimer disease	58/180	277/5881	1.94E-34	5.75E-33	5.17E-33
map05022	Pathways of neurodegeneration - multiple diseases	59/180	345/5881	6.30E-30	1.60E-28	1.44E-28
map05415	Diabetic cardiomyopathy	41/180	146/5881	1.84E-29	4.09E-28	3.67E-28
map05208	Chemical carcinogenesis - reactive oxygen species	40/180	155/5881	3.66E-27	7.24E-26	6.51E-26
map04714	Thermogenesis	41/180	171/5881	1.78E-26	3.17E-25	2.85E-25
map03050	Proteasome	20/180	42/5881	5.17E-20	8.36E-19	7.52E-19
map04932	Non-alcoholic fatty liver disease	29/180	120/5881	6.71E-19	9.95E-18	8.94E-18
map05017	Spinocerebellar ataxia	20/180	114/5881	1.44E-10	1.97E-09	1.77E-09
map04260	Cardiac muscle contraction	15/180	63/5881	3.66E-10	4.66E-09	4.19E-09
map03010	Ribosome	17/180	120/5881	1.03E-07	1.23E-06	1.10E-06
map04723	Retrograde endocannabinoid signaling	15/180	111/5881	1.13E-06	1.25E-05	1.13E-05
map00710	Carbon fixation by Calvin cycle	5/180	15/5881	5.95E-05	6.23E-04	5.60E-04
map03040	Spliceosome	12/180	110/5881	1.20E-04	0.001	0.001
map04141	Protein processing in endoplasmic reticulum	12/180	130/5881	5.78E-04	0.005	0.005
map05110	Vibrio cholerae infection	5/180	34/5881	0.003	0.031	0.027

Table S8. Summary of GO biological process enrichment analysis of the greenyellow gene module identified by WGCNA.

GO ID	Term	Annotated	Significant	Expected	weightedFisher
GO:0048489	synaptic vesicle transport	10	1	0.01	0.010
GO:0008053	mitochondrial fusion	10	1	0.01	0.010
GO:0015936	coenzyme A metabolic process	10	1	0.01	0.010
GO:1905952	regulation of lipid localization	11	1	0.01	0.011
GO:0006183	GTP biosynthetic process	11	1	0.01	0.011
GO:0046519	sphingoid metabolic process	11	1	0.01	0.011
GO:0045022	early endosome to late endosome transport	11	1	0.01	0.011
GO:0006241	CTP biosynthetic process	11	1	0.01	0.011
GO:0006275	regulation of DNA replication	12	1	0.01	0.012
GO:0007007	inner mitochondrial membrane organization	13	1	0.01	0.013
GO:0034312	diol biosynthetic process	13	1	0.01	0.013
GO:0051402	neuron apoptotic process	14	1	0.01	0.014
GO:0010833	telomere maintenance via telomere lengthening	15	1	0.01	0.015
GO:0007032	endosome organization	16	1	0.02	0.016
GO:0019915	lipid storage	16	1	0.02	0.016
GO:0034030	ribonucleoside bisphosphate biosynthetic process	22	1	0.02	0.021
GO:0034033	purine nucleoside bisphosphate biosynthetic process	22	1	0.02	0.021
GO:0043112	receptor metabolic process	27	1	0.03	0.026
GO:0032543	mitochondrial translation	33	1	0.03	0.032

GO:0048812	neuron projection morphogenesis	183	2	0.18	0.035
GO:0000082	G1/S transition of mitotic cell cycle	40	1	0.04	0.039
GO:0042147	retrograde transport, endosome to Golgi	40	1	0.04	0.039
GO:0006418	tRNA aminoacylation for protein translation	42	1	0.04	0.041
GO:0030148	sphingolipid biosynthetic process	47	1	0.05	0.045
GO:0044271	cellular nitrogen compound biosynthetic process	3094	5	3.04	0.048

Table S9. Summary of GO biological process enrichment analysis of the salmon gene module identified by WGCNA.

GO ID	Term	Annotated	Significant	Expected	weightedFisher
GO:0000381	regulation of alternative mRNA splicing, via spliceosome	30	4	0.10	2.40E-06
GO:0048268	clathrin coat assembly	13	2	0.04	7.70E-04
GO:0072583	clathrin-dependent endocytosis	20	2	0.06	0.002
GO:0048488	synaptic vesicle endocytosis	29	2	0.09	0.004
GO:0006357	regulation of transcription by RNA polymerase II	1429	11	4.59	0.012
GO:0006355	regulation of transcription, DNA-templated	1936	16	6.22	0.013
GO:0007254	JNK cascade	29	2	0.09	0.019
GO:0006353	DNA-templated transcription, termination	10	1	0.03	0.032
GO:0048489	synaptic vesicle transport	10	1	0.03	0.032
GO:0018023	peptidyl-lysine trimethylation	10	1	0.03	0.032
GO:0060996	dendritic spine development	11	1	0.04	0.035
GO:0016556	mRNA modification	12	1	0.04	0.038
GO:0048814	regulation of dendrite morphogenesis	12	1	0.04	0.038
GO:0035518	histone H2A monoubiquitination	12	1	0.04	0.038
GO:0008089	anterograde axonal transport	14	1	0.04	0.044
GO:0048024	regulation of mRNA splicing, via spliceosome	46	5	0.15	0.046
GO:0061462	protein localization to lysosome	15	1	0.05	0.047

GO:0006482	protein demethylation	15	1	0.05	0.047

Table S10. Summary of KEGG pathway enrichment analysis of the salmon gene module identified by WGCNA.

ID	Description	GeneRatio	BgRatio	pvalue	p.adjust	qvalue
map04144	Endocytosis	6/28	166/5881	1.04E-04	0.016	0.012
map04137	Mitophagy - animal	4/28	82/5881	5.57E-04	0.033	0.026
map04722	Neurotrophin signaling pathway	4/28	88/5881	7.28E-04	0.033	0.026
map04350	TGF-beta signaling pathway	4/28	92/5881	8.62E-04	0.033	0.026
map05135	Yersinia infection	4/28	105/5881	0.001	0.043	0.034

Table S11. Summary of GO biological process enrichment analysis of the brown gene module identified by WGCNA.

GO ID	Term	Annotated	Significant	Expected	weightedFisher
GO:0007264	small GTPase mediated signal transduction	223	11	2.24	3.10E-04
GO:0006357	regulation of transcription by RNA polymerase II	1429	34	14.34	5.50E-04
GO:0000122	negative regulation of transcription by RNA polymerase II	253	9	2.54	0.001
GO:0071526	semaphorin-plexin signaling pathway	28	3	0.28	0.003
GO:0043547	positive regulation of GTPase activity	102	5	1.02	0.005
GO:0006470	protein dephosphorylation	157	6	1.58	0.005
GO:0060395	SMAD protein signal transduction	40	3	0.40	0.007
GO:0007266	Rho protein signal transduction	52	5	0.52	0.008
GO:0045944	positive regulation of transcription by RNA polymerase II	192	6	1.93	0.013
GO:0009755	hormone-mediated signaling pathway	32	3	0.32	0.015
GO:0035025	positive regulation of Rho protein signal transduction	20	2	0.20	0.017
GO:0048843	negative regulation of axon extension involved in axon guidance	20	2	0.20	0.017
GO:0050919	negative chemotaxis	21	2	0.21	0.019
GO:0001755	neural crest cell migration	25	2	0.25	0.026

GO:0006355	regulation of transcription, DNA-templated	1936	44	19.43	0.026
GO:0061512	protein localization to cilium	26	2	0.26	0.028
GO:0051085	chaperone cofactor-dependent protein refolding	27	2	0.27	0.030
GO:0030036	actin cytoskeleton organization	322	7	3.23	0.031
GO:0030509	BMP signaling pathway	69	3	0.69	0.032
GO:0030335	positive regulation of cell migration	70	3	0.70	0.033
GO:0007411	axon guidance	93	5	0.93	0.036
GO:0007179	transforming growth factor beta receptor signaling pathway	34	2	0.34	0.046

Table S12. Summary GO biological process enrichment analysis of the turquoise gene module identified by WGCNA.

GO ID	Term	Annotated	Significant	Expected	weightedFisher
GO:0030198	extracellular matrix organization	75	14	1.71	1.30E-09
GO:0002040	sprouting angiogenesis	11	6	0.25	5.70E-08
GO:0016477	cell migration	301	32	6.87	3.90E-07
GO:0050919	negative chemotaxis	21	6	0.48	5.50E-06
GO:0007169	transmembrane receptor protein tyrosine kinase signaling pathway	223	23	5.09	6.30E-06
GO:0007399	nervous system development	553	34	12.62	1.40E-05
GO:0034446	substrate adhesion-dependent cell spreading	15	5	0.34	1.50E-05
GO:0001755	neural crest cell migration	25	6	0.57	1.70E-05
GO:0030036	actin cytoskeleton organization	322	21	7.35	3.70E-05
GO:0021782	glial cell development	10	4	0.23	5.00E-05
GO:0001936	regulation of endothelial cell proliferation	11	4	0.25	7.80E-05
GO:0042552	myelination	21	5	0.48	9.10E-05
GO:0007160	cell-matrix adhesion	49	7	1.12	1.10E-04
GO:0033674	positive regulation of kinase activity	88	9	2.01	1.80E-04
GO:0030178	negative regulation of Wnt signaling pathway	49	5	1.12	2.20E-04
GO:0007275	multicellular organism development	1155	73	26.36	7.10E-04

GO:0048843	negative regulation of axon extension involved in axon guidance	20	4	0.46	9.70E-04
GO:0007417	central nervous system development	116	7	2.65	0.001
GO:0045766	positive regulation of angiogenesis	10	3	0.23	0.001
GO:0032963	collagen metabolic process	24	5	0.55	0.002
GO:0050679	positive regulation of epithelial cell proliferation	11	3	0.25	0.002
GO:0006468	protein phosphorylation	687	28	15.68	0.002
GO:0048856	anatomical structure development	1411	86	32.21	0.002
GO:1905475	regulation of protein localization to membrane	13	3	0.30	0.003
GO:0009790	embryo development	90	8	2.05	0.003
GO:0060070	canonical Wnt signaling pathway	106	9	2.42	0.003
GO:0071526	semaphorin-plexin signaling pathway	28	4	0.64	0.004
GO:0030335	positive regulation of cell migration	70	8	1.60	0.004
GO:0010810	regulation of cell-substrate adhesion	15	3	0.34	0.004
GO:0007155	cell adhesion	467	34	10.66	0.004
GO:0097529	myeloid leukocyte migration	51	4	1.16	0.005
GO:0072659	protein localization to plasma membrane	72	6	1.64	0.006
GO:0002690	positive regulation of leukocyte chemotaxis	17	3	0.39	0.006
GO:0007517	muscle organ development	18	3	0.41	0.007

GO:0006629	lipid metabolic process	662	14	15.11	0.009
GO:0090263	positive regulation of canonical Wnt signaling pathway	20	3	0.46	0.010
GO:0007156	homophilic cell adhesion via plasma membrane adhesion molecules	106	7	2.42	0.011
GO:0007165	signal transduction	3803	111	86.81	0.011
GO:0051016	barbed-end actin filament capping	21	3	0.48	0.012
GO:0030574	collagen catabolic process	21	3	0.48	0.012
GO:0001666	response to hypoxia	21	3	0.48	0.012
GO:0061564	axon development	153	11	3.49	0.013
GO:0052652	cyclic purine nucleotide metabolic process	22	4	0.50	0.013
GO:0043086	negative regulation of catalytic activity	130	6	2.97	0.017
GO:0016339	calcium-dependent cell-cell adhesion via plasma membrane cell adhesion molecules	24	3	0.55	0.017
GO:0030334	regulation of cell migration	135	13	3.08	0.019
GO:0030032	lamellipodium assembly	10	2	0.23	0.021
GO:0038084	vascular endothelial growth factor signaling pathway	10	2	0.23	0.021
GO:0042982	amyloid precursor protein metabolic process	10	2	0.23	0.021
GO:0050918	positive chemotaxis	10	2	0.23	0.021
GO:0048009	insulin-like growth factor receptor signaling pathway	10	2	0.23	0.021
GO:0007219	Notch signaling pathway	47	4	1.07	0.022

GO:0050768	negative regulation of neurogenesis	22	5	0.50	0.023
GO:0030835	negative regulation of actin filament depolymerization	29	5	0.66	0.023
GO:0120034	positive regulation of plasma membrane bounded cell projection assembly	11	2	0.25	0.025
GO:0030204	chondroitin sulfate metabolic process	11	2	0.25	0.025
GO:0051341	regulation of oxidoreductase activity	11	2	0.25	0.025
GO:0035269	protein O-linked mannosylation	12	2	0.27	0.029
GO:0048704	embryonic skeletal system morphogenesis	12	2	0.27	0.029
GO:0034332	adherens junction organization	30	3	0.68	0.030
GO:0000902	cell morphogenesis	261	21	5.96	0.033
GO:0001525	angiogenesis	62	11	1.42	0.034
GO:0030282	bone mineralization	13	2	0.30	0.034
GO:0050982	detection of mechanical stimulus	13	2	0.30	0.034
GO:0031638	zymogen activation	13	2	0.30	0.034
GO:0006509	membrane protein ectodomain proteolysis	13	2	0.30	0.034
GO:0030837	negative regulation of actin filament polymerization	42	6	0.96	0.039
GO:0050650	chondroitin sulfate proteoglycan biosynthetic process	14	2	0.32	0.039
GO:0043542	endothelial cell migration	14	2	0.32	0.039

GO:0006182	cGMP biosynthetic process	14	2	0.32	0.039
GO:0099536	synaptic signaling	252	6	5.75	0.040
GO:0007179	transforming growth factor beta receptor signaling pathway	34	3	0.78	0.042
GO:0006906	vesicle fusion	58	4	1.32	0.043
GO:0030308	negative regulation of cell growth	23	5	0.52	0.045
GO:0006024	glycosaminoglycan biosynthetic process	15	2	0.34	0.045
GO:0055074	calcium ion homeostasis	118	5	2.69	0.045
GO:0008543	fibroblast growth factor receptor signaling pathway	35	3	0.80	0.045
GO:0071496	cellular response to external stimulus	41	2	0.94	0.045

Table S13. Summary of KEGG pathway enrichment analysis of the turquoise gene module identified by WGCNA.

ID	Description	GeneRatio	BgRatio	pvalue	p.adjust	qvalue
map04510	Focal adhesion	27/210	145/5881	6.46E-13	1.73E-10	1.31E-10
map04820	Cytoskeleton in muscle cells	24/210	153/5881	5.75E-10	7.71E-08	5.81E-08
map04151	PI3K-Akt signaling pathway	28/210	253/5881	6.26E-08	5.59E-06	4.22E-06
map05200	Pathways in cancer	36/210	396/5881	1.21E-07	8.09E-06	6.10E-06
map04512	ECM-receptor interaction	13/210	64/5881	2.78E-07	1.49E-05	1.12E-05
map04015	Rap1 signaling pathway	19/210	155/5881	2.07E-06	9.26E-05	6.98E-05
map05146	Amoebiasis	12/210	67/5881	3.33E-06	1.27E-04	9.61E-05
map05205	Proteoglycans in cancer	18/210	158/5881	1.11E-05	3.71E-04	2.80E-04
map04514	Cell adhesion molecules	14/210	104/5881	1.64E-05	4.89E-04	3.69E-04
map04072	Phospholipase D signaling pathway	13/210	104/5881	7.32E-05	0.002	0.001
map04926	Relaxin signaling pathway	12/210	91/5881	8.22E-05	0.002	0.001
map05165	Human papillomavirus infection	20/210	218/5881	8.71E-05	0.002	0.001
map05222	Small cell lung cancer	10/210	70/5881	1.65E-04	0.003	0.003
map04933	AGE-RAGE signaling pathway in diabetic complications	10/210	73/5881	2.36E-04	0.005	0.003
map05215	Prostate cancer	10/210	78/5881	4.09E-04	0.007	0.006
map04520	Adherens junction	10/210	81/5881	5.56E-04	0.009	0.007
map05224	Breast cancer	11/210	100/5881	8.04E-04	0.013	0.010
map04540	Gap junction	8/210	57/5881	8.46E-04	0.013	0.010
map05416	Viral myocarditis	7/210	45/5881	9.57E-04	0.013	0.010
map04020	Calcium signaling pathway	16/210	188/5881	0.001	0.013	0.010
map05226	Gastric cancer	11/210	103/5881	0.001	0.013	0.010
map04014	Ras signaling pathway	15/210	175/5881	0.001	0.016	0.012

map05218	Melanoma	7/210	48/5881	0.001	0.016	0.012
map05412	Arrhythmogenic right ventricular cardiomyopathy	9/210	76/5881	0.001	0.016	0.012
map04924	Renin secretion	7/210	49/5881	0.002	0.016	0.012
map05414	Dilated cardiomyopathy	10/210	93/5881	0.002	0.016	0.012
map01521	EGFR tyrosine kinase inhibitor resistance	8/210	63/5881	0.002	0.016	0.012
map04970	Salivary secretion	7/210	53/5881	0.003	0.024	0.018
map04810	Regulation of actin cytoskeleton	14/210	168/5881	0.003	0.024	0.018
map04391	Hippo signaling pathway - fly	6/210	40/5881	0.003	0.024	0.018
map05144	Malaria	6/210	41/5881	0.003	0.026	0.020
map04928	Parathyroid hormone synthesis, secretion and action	9/210	85/5881	0.003	0.026	0.020
map04022	cGMP-PKG signaling pathway	11/210	119/5881	0.003	0.027	0.020
map05410	Hypertrophic cardiomyopathy	9/210	87/5881	0.004	0.029	0.022
map04611	Platelet activation	9/210	88/5881	0.004	0.029	0.022
map04725	Cholinergic synapse	9/210	88/5881	0.004	0.029	0.022
map04392	Hippo signaling pathway - multiple species	4/210	19/5881	0.004	0.029	0.022
map05100	Bacterial invasion of epithelial cells	7/210	58/5881	0.004	0.030	0.023
map04973	Carbohydrate digestion and absorption	4/210	20/5881	0.005	0.033	0.025
map04935	Growth hormone synthesis, secretion and action	9/210	91/5881	0.005	0.033	0.025
map05142	Chagas disease	8/210	75/5881	0.005	0.033	0.025
map04976	Bile secretion	7/210	60/5881	0.005	0.033	0.025

map04270	Vascular smooth muscle contraction	9/210	93/5881	0.006	0.035	0.027
map04915	Estrogen signaling pathway	8/210	77/5881	0.006	0.036	0.027
map04974	Protein digestion and absorption	7/210	62/5881	0.006	0.037	0.028
map05145	Toxoplasmosis	8/210	78/5881	0.006	0.037	0.028
map04390	Hippo signaling pathway	10/210	113/5881	0.007	0.039	0.029
map04925	Aldosterone synthesis and secretion	7/210	64/5881	0.007	0.041	0.031
map04916	Melanogenesis	7/210	66/5881	0.009	0.048	0.036

Conclusions and future prospects

This thesis investigated gene-level associations between bank voles and the tick-borne bacterium *B. afzelii*, with a focus on host immune genes and pathogen surface proteins. Additionally, it provides broader insight into the immune response of bank voles to *B. afzelii* infection, highlighting potential pathways that may confer resistance and/or tolerance to the pathogen. By combining gene-by-gene analyses and transcriptomic profiling, it provides a comprehensive perspective on host-pathogen coevolution in a natural rodent-pathogen system.

First, evidence was found that the ospC gene of *B. afzelii* evolves under balancing selection. While bank vole MHC class II alleles and supertypes did not significantly predict infection with *B. afzelii* strains carrying specific OspC variant at the gene level, MHC class II alleles were significantly associated with variation in IgG antibody levels against two common OspC variants among seropositive individuals. These results suggest that MHC alleles differ in the ability to induce variant-specific antibody responses, which may contribute to the maintenance of OspC polymorphism.

Second, it was demonstrated that the ospE gene of *B. afzelii* is subject to positive selection. However, no gene-by-gene associations were found between OspE variants and bank vole CFH/FH-R alleles. Furthermore, OspE sequences showed little population structuring, and spatial analyses revealed no co-structure between bank vole and *B. afzelii* genotypes. These findings suggest that the evolution of CFH and OspE is influenced by broader selective pressures, highlighting the need for further research.

Third, the transcriptomic response of bank voles to *B. afzelii* infection was characterized. The results identified upregulation of immune pathways, especially those linked to B-cell activity and antibody production. Notably, the pro-inflammatory response was suppressed compared to humans, suggesting a complex host adaptation that may underlie the milder infections observed in voles. Additionally, enrichment analysis highlighted significant changes in pathways related to extracellular matrix organization, metabolism, and energy production. This broadens our understanding of the reservoir host's defense response and identifies additional molecular components that may play a role in tolerance to infection.

Taken together, the results reveal complex interactions between the bank vole and *B. afzelii* and provide a basis for future studies investigating the mechanisms shaping host-pathogen coevolution and infection dynamics. These findings, while focusing on distinct aspects of the host-pathogen interaction, reveal important connections that highlight how immune responses and pathogen strategies may co-shape their evolution. For instance, the contrast between OspC and OspE illustrates how differing levels of immunogenicity may shape the genomic architecture and functional roles of surface proteins. OspC's strong antigenicity, reflected in the specific MHC-linked antibody responses and likely contributing to plasmid copy number constraints, suggests that too many OspC variants could severely limit infection success in populations with diverse immunological memory. Consistent with this, OspC expression is downregulated after initial infection in vertebrate hosts, likely as a mechanism to evade immune detection. In contrast, the lack of variant-specific immune associations for OspE supports a more generalist function potentially suited to infecting hosts with diverse genotypes. Its continued expression throughout vertebrate infection further supports a role in persistent immune evasion rather than acute-stage invasion. This functional differentiation in surface protein roles may allow *B. afzelii* to balance immune evasion and host range. Moreover, the transcriptomic data reinforce the immunological relevance of these findings, especially the upregulation of B cell- and antibody-related pathways, which aligns with the functional role of MHC in mediating specific responses to OspC. Together, this evidence supports a model in which selective pressures act on both host immune recognition and pathogen strategy, shaping the evolutionary interplay observed in this natural system.

Looking ahead, several directions emerge for future research. One possible approach is to investigate the drivers of CFH and OspE variation across multiple host species to better understand selective pressures shaping their evolution. Additionally, the RNA sequencing results reveal numerous novel pathways, opening new directions for exploring the molecular mechanisms underlying host defense against *B. afzelii*, both through studies in natural populations and controlled experimental infections. Future work could investigate whether variation in gene expression depends on specific combinations of bank vole and *B. afzelii* genotypes, similar to findings in other systems, where both host and parasite genotypes shaped transcriptional responses (Barribeau et

al., 2014). Finally, an intriguing idea emerging from this work is to examine whether the host can modulate the expression of CFH and FH-related genes as a defense strategy. Since FH-R proteins lack the regulatory domains found in CFH, they may bind antigens like OspE without protecting the pathogen from complement attack. This may allow the host to deceive the pathogen, enhancing its clearance while limiting the pathogen's capacity for immune evasion.

References

Ashby, B. (2020). When does parasitism maintain sex in the absence of Red Queen Dynamics? *Journal of Evolutionary Biology*, *33*(12), 1795–1805. https://doi.org/10.1111/jeb.13718

Barbour, A. G., & Travinsky, B. (2010). Evolution and Distribution of the *ospC* Gene, a Transferable Serotype Determinant of Borrelia burgdorferi. *mBio*, *1*(4), e00153-10. https://doi.org/10.1128/mBio.00153-10

Barribeau, S. M., Sadd, B. M., du Plessis, L., & Schmid-Hempel, P. (2014). Gene expression differences underlying genotype-by-genotype specificity in a host–parasite system. *Proceedings of the National Academy of Sciences*, 111(9), 3496–3501. https://doi.org/10.1073/pnas.1318628111

Baum, E., Hue, F., & Barbour, A. G. (2012). Experimental Infections of the Reservoir Species Peromyscus leucopus with Diverse Strains of Borrelia burgdorferi, a Lyme Disease Agent. *mBio*, *3*(6), e00434-12. https://doi.org/10.1128/mBio.00434-12

Bhattacharjee, A., Oeemig, J. S., Kolodziejczyk, R., Meri, T., Kajander, T., Lehtinen, M. J., Iwaï, H., Jokiranta, T. S., & Goldman, A. (2013). Structural Basis for Complement Evasion by Lyme Disease Pathogen Borrelia burgdorferi. *Journal of Biological Chemistry*, *288*(26), 18685–18695. https://doi.org/10.1074/jbc.M113.459040

Bodmer, W. F. (1972). Evolutionary significance of the HL-A system. *Nature*, *237*(5351). https://doi.org/10.1038/237139a0

Brisson, D., & Dykhuizen, D. E. (2004). ospC Diversity in Borrelia burgdorferi. *Genetics*, *168*(2), 713–722. https://doi.org/10.1534/genetics.104.028738

Buckingham, L. J., & Ashby, B. (2022). Coevolutionary theory of hosts and parasites. *Journal of Evolutionary Biology*, *35*(2), 205–224. https://doi.org/10.1111/jeb.13981

Cagliani, R., Forni, D., Filippi, G., Mozzi, A., De Gioia, L., Pontremoli, C., Pozzoli, U., Bresolin, N., Clerici, M., & Sironi, M. (2016). The mammalian complement system as an epitome of host-pathogen genetic conflicts. *Molecular Ecology*, *25*(6), 1324–1339. https://doi.org/10.1111/mec.13558

Caine, J. A., Lin, Y.-P., Kessler, J. R., Sato, H., Leong, J. M., & Coburn, J. (2017). *Borrelia burgdorferi* outer surface protein C (OspC) binds complement component C4b and confers bloodstream survival. *Cellular Microbiology*, *19*(12), e12786. https://doi.org/10.1111/cmi.12786

Casjens, S. R., Gilcrease, E. B., Vujadinovic, M., Mongodin, E. F., Luft, B. J., Schutzer, S. E., Fraser, C. M., & Qiu, W.-G. (2017). Plasmid diversity and phylogenetic consistency in the Lyme disease agent Borrelia burgdorferi. *BMC Genomics*, *18*(1), 165. https://doi.org/10.1186/s12864-017-3553-5

Cayol, C., Giermek, A., Gomez-Chamorro, A., Hytönen, J., Kallio, E. R., Mappes, T., Salo, J., Voordouw, M. J., & Koskela, E. (2018). *Borrelia afzelii* alters reproductive success in a rodent host. *Proceedings of the Royal Society B: Biological Sciences*, *285*(1884), 20181056. https://doi.org/10.1098/rspb.2018.1056

Doherty, P. C., & Zinkernagel, R. M. (1975). Enhanced immunological surveillance in mice heterozygous at the H-2 gene complex. *Nature*, *256*(5512), 50–52. https://doi.org/10.1038/256050a0

Ejsmond, M. J., & Radwan, J. (2015). Red Queen Processes Drive Positive Selection on Major Histocompatibility Complex (MHC) Genes. *PLOS Computational Biology*, *11*(11), e1004627. https://doi.org/10.1371/journal.pcbi.1004627

Grimm, D., Tilly, K., Byram, R., Stewart, P. E., Krum, J. G., Bueschel, D. M., Schwan, T. G., Policastro, P. F., Elias, A. F., & Rosa, P. A. (2004). Outer-surface protein C of the Lyme disease spirochete: A protein induced in ticks for infection of mammals. *Proceedings of the National Academy of Sciences*, *101*(9), 3142–3147. https://doi.org/10.1073/pnas.0306845101

Hammerschmidt, S., Agarwal, V., Kunert, A., Haelbich, S., Skerka, C., & Zipfel, P. F. (2007). The host immune regulator factor H interacts via two contact sites with the PspC protein of Streptococcus pneumoniae and mediates adhesion to host epithelial cells. *Journal of Immunology (Baltimore, Md.: 1950)*, *178*(9), 5848–5858. https://doi.org/10.4049/jimmunol.178.9.5848

Hanincová, K., Schäfer, S. M., Etti, S., Sewell, H.-S., Taragelová, V., Ziak, D., Labuda, M., & Kurtenbach, K. (2003). Association of *Borrelia afzelii* with rodents in Europe. *Parasitology*, *126*(1), 11–20. https://doi.org/10.1017/S0031182002002548

Hedrick, P. W. (2011). Population genetics of malaria resistance in humans. *Heredity*, *107*(4), 283–304. https://doi.org/10.1038/hdy.2011.16

Hellwage, J., Meri, T., Heikkilä, T., Alitalo, A., Panelius, J., Lahdenne, P., Ilkka Seppälä, J. T., & Meri, S. (2001). The Complement Regulator Factor H Binds to the Surface Protein OspE of Borrelia burgdorferi. *Journal of Biological Chemistry*, *276*(11), 8427–8435. https://doi.org/10.1074/jbc.M007994200

Hofmeister, E. K., Ellis, B. A., Glass, G. E., & Childs, J. E. (1999). Longitudinal study of infection with Borrelia burgdorferi in a population of Peromyscus leucopus at a Lyme disease-enzootic site in Maryland. *The American Journal of Tropical Medicine and Hygiene*, *60*(4), 598–609. https://doi.org/10.4269/ajtmh.1999.60.598

Jacquet, M., Durand, J., Rais, O., & Voordouw, M. J. (2015). Cross-reactive acquired immunity influences transmission success of the Lyme disease pathogen, Borrelia afzelii. *Infection, Genetics and Evolution*, *36*, 131–140. https://doi.org/10.1016/j.meegid.2015.09.012

Jokiranta, T. S., Cheng, Z.-Z., Seeberger, H., Jòzsi, M., Heinen, S., Noris, M., Remuzzi, G., Ormsby, R., Gordon, D. L., Meri, S., Hellwage, J., & Zipfel, P. F. (2005). Binding of Complement Factor H to Endothelial Cells Is Mediated by the Carboxy-Terminal Glycosaminoglycan Binding Site. *The American Journal of Pathology*, *167*(4), 1173–1181. https://doi.org/10.1016/S0002-9440(10)61205-9

Juarez-Estrada, M. A., Graham, D., Hernandez-Velasco, X., & Tellez-Isaias, G. (2023). Editorial: Parasitism: the good, the bad and the ugly. *Frontiers in Veterinary Science*, *10*, 1304206. https://doi.org/10.3389/fvets.2023.1304206

Karvonen, A., & Seehausen, O. (2012). The Role of Parasitism in Adaptive Radiations—When Might Parasites Promote and When Might They Constrain Ecological Speciation? *International Journal of Ecology*, *2012*, 1–20. https://doi.org/10.1155/2012/280169

Kaufman, J. (2018). Unfinished Business: Evolution of the MHC and the Adaptive Immune System of Jawed Vertebrates. *Annual Review of Immunology*, *36*(1), 383–409. https://doi.org/10.1146/annurev-immunol-051116-052450

Kotlík, P., Deffontaine, V., Mascheretti, S., Zima, J., Michaux, J. R., & Searle, J. B. (2006). A northern glacial refugium for bank voles (*Clethrionomys glareolus*). *Proceedings of the National Academy of Sciences*, *103*(40), 14860–14864. https://doi.org/10.1073/pnas.0603237103

Kotlík, P., Marková, S., Horníková, M., Escalante, M. A., & Searle, J. B. (2022). The Bank Vole (Clethrionomys glareolus) as a Model System for Adaptive Phylogeography in the European Theater. *Frontiers in Ecology and Evolution*, *10*, 866605. https://doi.org/10.3389/fevo.2022.866605

Kraiczy, P., Skerka, C., Kirschfink, M., Brade, V., & Zipfel, P. F. (2001). Immune evasion ofBorrelia burgdorferi by acquisition of human complement regulators FHL-1/reconectin and Factor H. *European Journal of Immunology*, *31*(6), 1674–1684. https://doi.org/10.1002/1521-4141(200106)31:6<1674::AID-IMMU1674>3.0.CO;2-2

Kristensen, T., & Tack, B. F. (1986). Murine protein H is comprised of 20 repeating units, 61 amino acids in length. *Proceedings of the National Academy of Sciences*, *83*(11), 3963–3967. https://doi.org/10.1073/pnas.83.11.3963

Kunert, A., Losse, J., Gruszin, C., Hühn, M., Kaendler, K., Mikkat, S., Volke, D., Hoffmann, R., Jokiranta, T. S., Seeberger, H., Moellmann, U., Hellwage, J., & Zipfel, P. F. (2007). Immune evasion of the human pathogen Pseudomonas aeruginosa: Elongation factor Tuf is a factor H and plasminogen binding protein. *Journal of Immunology (Baltimore, Md.: 1950)*, *179*(5), 2979–2988. https://doi.org/10.4049/jimmunol.179.5.2979

Kurtenbach, K., De Michelis, S., Etti, S., Schäfer, S. M., Sewell, H.-S., Brade, V., & Kraiczy, P. (2002). Host association of Borrelia burgdorferi sensu lato – the key role of host complement. *Trends in Microbiology*, *10*(2), 74–79. https://doi.org/10.1016/S0966-842X(01)02298-3

Lively, C. M. (2010). A Review of Red Queen Models for the Persistence of Obligate Sexual Reproduction. *Journal of Heredity*, *101*(suppl_1), S13–S20. https://doi.org/10.1093/jhered/esq010

Mayer, A., Mora, T., Rivoire, O., & Walczak, A. M. (2016). Diversity of immune strategies explained by adaptation to pathogen statistics. *Proceedings of the National Academy of Sciences*, *113*(31), 8630–8635. https://doi.org/10.1073/pnas.1600663113

Meri, S. (2016). Self-nonself discrimination by the complement system. *FEBS Letters*, *590*(15), 2418–2434. https://doi.org/10.1002/1873-3468.12284

Migalska, M., Przesmycka, K., Alsarraf, M., Bajer, A., Behnke-Borowczyk, J., Grzybek, M., Behnke, J. M., & Radwan, J. (2022). Long term patterns of association between MHC and helminth burdens in the bank vole support Red Queen dynamics. *Molecular Ecology*, mec.16486. https://doi.org/10.1111/mec.16486

Miller, J. C., Von Lackum, K., Babb, K., McAlister, J. D., & Stevenson, B. (2003). Temporal Analysis of *Borrelia burgdorferi* Erp Protein Expression throughout the Mammal-Tick InfectiousCycle. *Infection and Immunity*, *71*(12), 6943–6952. https://doi.org/10.1128/IAI.71.12.6943-6952.2003

Nonaka, M., & Yoshizaki, F. (2004). Primitive complement system of invertebrates. *Immunological Reviews*, *198*(1), 203–215. https://doi.org/10.1111/j.0105-2896.2004.00118.x

Norman, D. G., Barlow, P. N., Baron, M., Day, A. J., Sim, R. B., & Campbell, I. D. (1991). Three-dimensional structure of a complement control protein module in solution. *Journal of Molecular Biology*, *219*(4), 717–725. https://doi.org/10.1016/0022-2836(91)90666-T

Pal, U., Yang, X., Chen, M., Bockenstedt, L. K., Anderson, J. F., Flavell, R. A., Norgard, M. V., & Fikrig, E. (2004). OspC facilitates Borrelia burgdorferi invasion of Ixodes scapularis salivary glands. *The Journal of Clinical Investigation*, *113*(2), 220–230. https://doi.org/10.1172/JCI19894

Penn, D. J., Damjanovich, K., & Potts, W. K. (2002). MHC heterozygosity confers a selective advantage against multiple-strain infections. *Proceedings of the National Academy of Sciences*, 99(17), 11260–11264. https://doi.org/10.1073/pnas.162006499

Phillips, K. P., Cable, J., Mohammed, R. S., Herdegen-Radwan, M., Raubic, J., Przesmycka, K. J., van Oosterhout, C., & Radwan, J. (2018). Immunogenetic novelty confers a selective advantage in host–pathogen coevolution. *Proceedings of the National Academy of Sciences*, *115*(7), 1552–1557. https://doi.org/10.1073/pnas.1708597115

Pouw, R. B., Vredevoogd, D. W., Kuijpers, T. W., & Wouters, D. (2015). Of mice and men: The factor H protein family and complement regulation. *Molecular Immunology*, *67*(1), 12–20. https://doi.org/10.1016/j.molimm.2015.03.011

Prado, F., Sheih, A., West, J. D., & Kerr, B. (2009). Coevolutionary cycling of host sociality and pathogen virulence in contact networks. *Journal of Theoretical Biology*, *261*(4), 561–569. https://doi.org/10.1016/j.jtbi.2009.08.022

Råberg, L., Hagström, Å., Andersson, M., Bartkova, S., Scherman, K., Strandh, M., & Tschirren, B. (2017). Evolution of antigenic diversity in the tick-transmitted bacterium *Borrelia afzelii*: A role for host specialization? *Journal of Evolutionary Biology*, *30*(5), 1034–1041. https://doi.org/10.1111/jeb.13075

Radwan, J., Babik, W., Kaufman, J., Lenz, T. L., & Winternitz, J. (2020). Advances in the Evolutionary Understanding of MHC Polymorphism. *Trends in Genetics*, *36*(4), 298–311. https://doi.org/10.1016/j.tig.2020.01.008

Ricklin, D., Hajishengallis, G., Yang, K., & Lambris, J. D. (2010). Complement—A key system for immune surveillance and homeostasis. *Nature Immunology*, *11*(9), 785–797. https://doi.org/10.1038/ni.1923

Schwan, T. G., & Piesman, J. (2000). Temporal Changes in Outer Surface Proteins A and C of the Lyme Disease-Associated Spirochete, Borrelia burgdorferi, during the Chain of Infection in Ticks and Mice. *Journal of Clinical Microbiology*, *38*(1), 382–388. https://doi.org/10.1128/jcm.38.1.382-388.2000

Schwan, T. G., Piesman, J., Golde, W. T., Dolan, M. C., & Rosa, P. A. (1995). Induction of an outer surface protein on Borrelia burgdorferi during tick feeding. *Proceedings of the National Academy of Sciences*, *92*(7), 2909–2913. https://doi.org/10.1073/pnas.92.7.2909

Schwanz, L. E., Voordouw, M. J., Brisson, D., & Ostfeld, R. S. (2011). *Borrelia burgdorferi* Has Minimal Impact on the Lyme Disease Reservoir Host *Peromyscus leucopus*. *Vector-Borne and Zoonotic Diseases*, *11*(2), 117–124. https://doi.org/10.1089/vbz.2009.0215

Schwartz, I., Margos, G., Casjens, S. R., Qiu, W.-G., & Eggers, C. H. (2022). Multipartite Genome of Lyme Disease *Borrelia*: Structure, Variation and Prophages. *Current Issues in Molecular Biology*, 409–454. https://doi.org/10.21775/cimb.042.409

Simon, N., Lasonder, E., Scheuermayer, M., Kuehn, A., Tews, S., Fischer, R., Zipfel, P. F., Skerka, C., & Pradel, G. (2013). Malaria Parasites Co-opt Human Factor H to Prevent Complement-Mediated Lysis in the Mosquito Midgut. *Cell Host & Microbe*, *13*(1), 29–41. https://doi.org/10.1016/j.chom.2012.11.013

Sipari, S., Hytönen, J., Pietikäinen, A., Mappes, T., & Kallio, E. R. (2022). The effects of Borrelia infection on its wintering rodent host. *Oecologia*, *200*(3–4), 471–478. https://doi.org/10.1007/s00442-022-05272-y

Stanek, G., Wormser, G. P., Gray, J., & Strle, F. (2012). Lyme borreliosis. *The Lancet*, *379*(9814), 461–473. https://doi.org/10.1016/S0140-6736(11)60103-7

Steere, A. C., Strle, F., Wormser, G. P., Hu, L. T., Branda, J. A., Hovius, J. W. R., Li, X., & Mead, P. S. (2016). Lyme borreliosis. *Nature Reviews Disease Primers*, *2*(1), 16090. https://doi.org/10.1038/nrdp.2016.90

Stevenson, B., El-Hage, N., Hines, M. A., Miller, J. C., & Babb, K. (2002). Differential Binding of Host Complement Inhibitor Factor H by *Borrelia burgdorferi* Erp Surface Proteins: A Possible Mechanism Underlying the Expansive Host Range of Lyme Disease Spirochetes. *Infection and Immunity*, *70*(2), 491–497. https://doi.org/10.1128/IAI.70.2.491-497.2002

Tarnowska, E., Niedziałkowska, M., Gerc, J., Korbut, Z., Górny, M., & Jędrzejewska, B. (2016). Spatial distribution of the Carpathian and Eastern mtDNA lineages of the bank vole in their contact zone relates to environmental conditions. *Biological Journal of the Linnean Society*, *119*(3), 732–744. https://doi.org/10.1111/bij.12764

Tarnowska, E., Niedziałkowska, M., & Jędrzejewska, B. (2019). Genetic structure of bank vole populations in the contact zone of two lineages in north-eastern Poland. *Mammalian Biology*, *96*, 93–101. https://doi.org/10.1016/j.mambio.2018.10.011

Tilly, K., Krum, J. G., Bestor, A., Jewett, M. W., Grimm, D., Bueschel, D., Byram, R., Dorward, D., VanRaden, M. J., Stewart, P., & Rosa, P. (2006). *Borrelia burgdorferi* OspC Protein Required Exclusively in a Crucial Early Stage of Mammalian Infection. *Infection and Immunity*, *74*(6), 3554–3564. https://doi.org/10.1128/IAI.01950-05

van Duijvendijk, G., Sprong, H., & Takken, W. (2015). Multi-trophic interactions driving the transmission cycle of Borrelia afzelii between Ixodes ricinus and rodents: A review. *Parasites & Vectors*, *8*(1), 643. https://doi.org/10.1186/s13071-015-1257-8

Voordouw, M. J., Lachish, S., & Dolan, M. C. (2015). The Lyme Disease Pathogen Has No Effect on the Survival of Its Rodent Reservoir Host. *PLOS ONE*, *10*(2), e0118265. https://doi.org/10.1371/journal.pone.0118265

Woolhouse, M. E. J., Webster, J. P., Domingo, E., Charlesworth, B., & Levin, B. R. (2002). Biological and biomedical implications of the co-evolution of pathogens and their hosts. *Nature Genetics*, *32*(4), 569–577. https://doi.org/10.1038/ng1202-569

Zhang, Y., Wu, M., Hang, T., Wang, C., Yang, Y., Pan, W., Zang, J., Zhang, M., & Zhang, X. (2017). Staphylococcus aureus SdrE captures complement factor H's C-terminus via a novel 'close, dock, lock and latch' mechanism for complement evasion. *The Biochemical Journal*, *474*(10), 1619–1631. https://doi.org/10.1042/BCJ20170085

Zhong, X., Nouri, M., & Råberg, L. (2019). Colonization and pathology of Borrelia afzelii in its natural hosts. *Ticks and Tick-Borne Diseases*, *10*(4), 822–827. https://doi.org/10.1016/j.ttbdis.2019.03.017

Zipfel, P. F., & Skerka, C. (1994). Complement factor H and related proteins: An expanding family of complement-regulatory proteins? *Immunology Today*, *15*(3), 121–126. https://doi.org/10.1016/0167-5699(94)90155-4

Zipfel, P. F., & Skerka, C. (2009). Complement regulators and inhibitory proteins. *Nature Reviews Immunology*, *9*(10), 729–740. https://doi.org/10.1038/nri2620

Zipfel, P. F., Skerka, C., Hellwage, J., Jokiranta, S. T., Meri, S., Brade, V., Kraiczy, P., Noris, M., & Remuzzi, G. (2002). Factor H family proteins: On complement, microbes and human diseases. *Biochemical Society Transactions*, *30*(6), 971–978. https://doi.org/10.1042/bst0300971

Authorship statements

Poznań, 24.06.2025

Authorship statement

I. Różańska-Wróbel, J., Migalska, M., Urbanowicz, A., Grzybek, M., Rego, R. O. M., Bajer, A., Dwuznik-Szarek, D., Alsarraf, M., Behnke-Borowczyk, J., Behnke, J. M., & Radwan, J. (2024). Interplay between vertebrate adaptive immunity and bacterial infectivity genes: Bank vole MHC versus *Borrelia afzelii* OspC. *Molecular ecology*, 33(21), e17534. https://doi.org/10.1111/mec.17534

I declare that I have contributed to molecular analyses, including DNA extraction, *Borrelia* screening, OspC genotyping, qPCR and ELISA, as well as data analysis and visualization, interpretation of results, and manuscript writing. My contribution to the study was 60 %.

II. Różańska-Wróbel, J., Wasilewska, J., Radwan, J. (2025). Do *Borrelia afzelii* outer surface protein E coevolve with complement factor H of its rodent host? Insights from GxG and spatial associations

I declare that I have contributed to methodology development, sample collection, molecular analyses, including bank vole DNA extraction, *Borrelia* screening, CFH and OspE genotyping, as well as data analysis and visualization, interpretation of results, and manuscript writing. My contribution to the study was 80 %.

III. Różańska-Wróbel, J., Konczal, M., Notarnicola R. F., Radwan, J. (2025). Transcriptomic response to *Borrelia afzelii* infection in the skin of wild bank voles

I declare that I have contributed to study design, methodology development, sample collection, *Borrelia* screening, data analysis and visualization, interpretation of results, and manuscript writing. My contribution to the study was 80 %.

Joanna Różańska-Wróbel

Joanna Różańska-Wróbel

Supervisor: prof. dr hab. Jacek Radwan

Poznań, 24.06.2025

I hereby confirm that I am a co-author of the paper: Różańska-Wróbel, J., Migalska, M., Urbanowicz, A., Grzybek, M., Rego, R. O. M., Bajer, A., Dwuznik-Szarek, D., Alsarraf, M., Behnke-Borowczyk, J., Behnke, J. M., & Radwan, J. (2024). Interplay between vertebrate adaptive immunity and bacterial infectivity genes: Bank vole MHC versus *Borrelia afzelii* OspC. *Molecular ecology*, 33(21), e17534. https://doi.org/10.1111/mec.17534

I declare that I have contributed to study design, funding acquisition, methodology development, supervision of molecular and data analysis, and manuscript writing.

prof. dr hab. Jacek Radwan

I hereby confirm that I am a co-author of the manuscript: Różańska-Wróbel, J., Wasilewska, J., Radwan, J. (2025). Do *Borrelia afzelii* outer surface protein E coevolve with complement factor H of its rodent host? Insights from GxG and spatial associations.

I declare that I have contributed to study design, funding acquisition, methodology development, sample collection, data analysis supervision, and manuscript revision.

prof. dr hab. Jacek Radwan

I hereby confirm that I am a co-author of the manuscript: Różańska-Wróbel, J., Konczal, M., Notarnicola R. F., Radwan, J. (2025). Transcriptomic response to *Borrelia afzelii* infection in the skin of wild bank voles.

I declare that I have contributed to study design, funding acquisition, methodology development, sample collection, data analysis supervision, and manuscript revision.

prof. dr hab. Jacek Radwan

Kraków, 05.06.2025

I hereby confirm that I am a co-author of the paper: Różańska-Wróbel, J., Migalska, M., Urbanowicz, A., Grzybek, M., Rego, R. O. M., Bajer, A., Dwuznik-Szarek, D., Alsarraf, M., Behnke-Borowczyk, J., Behnke, J. M., & Radwan, J. (2024). Interplay between vertebrate adaptive immunity and bacterial infectivity genes: Bank vole MHC versus *Borrelia afzelii* OspC. *Molecular ecology*, 33(21), e17534. https://doi.org/10.1111/mec.17534

I declare that I have contributed to molecular work (MHC genotyping), data analysis supervision, and manuscript revision.

dr Magdalena Migalska

Poznań, 08.06.2025

I hereby confirm that I am a co-author of the paper: Różańska-Wróbel, J., Migalska, M., Urbanowicz, A., Grzybek, M., Rego, R. O. M., Bajer, A., Dwuznik-Szarek, D., Alsarraf, M., Behnke-Borowczyk, J., Behnke, J. M., & Radwan, J. (2024). Interplay between vertebrate adaptive immunity and bacterial infectivity genes: Bank vole MHC versus *Borrelia afzelii* OspC. *Molecular ecology*, 33(21), e17534. https://doi.org/10.1111/mec.17534

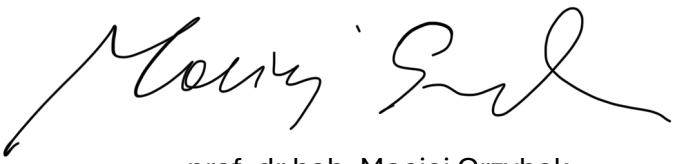
I declare that I have contributed to OspC protein production, supervision of molecular work, and manuscript revision.

dr hab. Anna Urbanowicz, prof. ICHB PAN

Gdynia, 16th June 2025

I hereby confirm that I am a co-author of the paper: Różańska-Wróbel, J., Migalska, M., Urbanowicz, A., Grzybek, M., Rego, R. O. M., Bajer, A., Dwuznik-Szarek, D., Alsarraf, M., Behnke-Borowczyk, J., Behnke, J. M., & Radwan, J. (2024). Interplay between vertebrate adaptive immunity and bacterial infectivity genes: Bank vole MHC versus *Borrelia afzelii* OspC. *Molecular ecology*, 33(21), e17534. https://doi.org/10.1111/mec.17534

I declare that I have contributed to sample collection, bank vole age assignment, and manuscript revision.



prof. dr hab. Maciej Grzybek

České Budějovice, 5.6.2025

I hereby confirm that I am a co-author of the paper: Różańska-Wróbel, J., Migalska, M., Urbanowicz, A., Grzybek, M., Rego, R. O. M., Bajer, A., Dwuznik-Szarek, D., Alsarraf, M., Behnke-Borowczyk, J., Behnke, J. M., & Radwan, J. (2024). Interplay between vertebrate adaptive immunity and bacterial infectivity genes: Bank vole MHC versus *Borrelia afzelii* OspC. *Molecular ecology*, 33(21), e17534. https://doi.org/10.1111/mec.17534

I declare that I have contributed to providing positive controls for ELISA and manuscript revision.

Ryan O. M. Rego Ph.D.

Warszawa, dn. 05.06.2025

I hereby confirm that I am a co-author of the paper: Różańska-Wróbel, J., Migalska, M., Urbanowicz, A., Grzybek, M., Rego, R. O. M., Bajer, A., Dwuznik-Szarek, D., Alsarraf, M., Behnke-Borowczyk, J., Behnke, J. M., & Radwan, J. (2024). Interplay between vertebrate adaptive immunity and bacterial infectivity genes: Bank vole MHC versus *Borrelia afzelii* OspC. *Molecular ecology*, 33(21), e17534. https://doi.org/10.1111/mec.17534

I declare that I have contributed to sample collection and manuscript revision.

prof. dr hab. Anna Bajer

Place, Warsaw, Date 05.06.2025

I hereby confirm that I am a co-author of the paper: Różańska-Wróbel, J., Migalska, M., Urbanowicz, A., Grzybek, M., Rego, R. O. M., Bajer, A., Dwuznik-Szarek, D., Alsarraf, M., Behnke-Borowczyk, J., Behnke, J. M., & Radwan, J. (2024). Interplay between vertebrate adaptive immunity and bacterial infectivity genes: Bank vole MHC versus *Borrelia afzelii* OspC. *Molecular ecology*, 33(21), e17534. https://doi.org/10.1111/mec.17534

I declare that I have contributed to sample collection.

dr Dorota Dwużnik-Szarek

Warsaw, 05.06.2025

I hereby confirm that I am a co-author of the paper: Różańska-Wróbel, J., Migalska, M., Urbanowicz, A., Grzybek, M., Rego, R. O. M., Bajer, A., Dwuznik-Szarek, D., Alsarraf, M., Behnke-Borowczyk, J., Behnke, J. M., & Radwan, J. (2024). Interplay between vertebrate adaptive immunity and bacterial infectivity genes: Bank vole MHC versus *Borrelia afzelii* OspC. *Molecular ecology*, 33(21), e17534. https://doi.org/10.1111/mec.17534

I declare that I have contributed to sample collection.

dr. Mohammed Alsarraf

Poznań, 05.06. 2025 r.

Place, Date

I hereby confirm that I am a co-author of the paper: Różańska-Wróbel, J., Migalska, M., Urbanowicz, A., Grzybek, M., Rego, R. O. M., Bajer, A., Dwuznik-Szarek, D., Alsarraf, M., Behnke-Borowczyk, J., Behnke, J. M., & Radwan, J. (2024). Interplay between vertebrate adaptive immunity and bacterial infectivity genes: Bank vole MHC versus *Borrelia afzelii* OspC. *Molecular ecology*, 33(21), e17534. https://doi.org/10.1111/mec.17534

I declare that I have contributed to sample collection.

Jolanta Behnke-Borowczyk

dr hab. Jolanta Behnke-Borowczyk, prof. UPP

Nottingham, U.K.

5th June 2025

I hereby confirm that I am a co-author of the paper: Różańska-Wróbel, J., Migalska, M., Urbanowicz, A., Grzybek, M., Rego, R. O. M., Bajer, A., Dwuznik-Szarek, D., Alsarraf, M., Behnke-Borowczyk, J., Behnke, J. M., & Radwan, J. (2024). Interplay between vertebrate adaptive immunity and bacterial infectivity genes: Bank vole MHC versus *Borrelia afzelii* OspC. *Molecular ecology*, 33(21), e17534. https://doi.org/10.1111/mec.17534

I declare that I have contributed to sample collection and manuscript revision.

prof. Jerzy M. Behnke

Poznań, 24.06.2025

I hereby confirm that I am a co-author of the manuscript: Różańska-Wróbel, J., Konczal, M., Notarnicola R. F., Radwan, J. (2025). Transcriptomic response to *Borrelia afzelii* infection in the skin of wild bank voles.

I declare that I have contributed to methodology development, sample collection, data analysis supervision, and manuscript revision.

dr hab. Mateusz Konczal

Poznań, 19.06.2025

I hereby confirm that I am a co-author of the manuscript: Różańska-Wróbel, J., Konczal, M., Notarnicola R. F., Radwan, J. (2025). Transcriptomic response to *Borrelia afzelii* infection in the skin of wild bank voles.

I declare that I have contributed to sample collection and manuscript revision.

dr. Rocco F. Notarnicola

Poznań, 24.06.2025

I hereby confirm that I am a co-author of the manuscript: Różańska-Wróbel, J., Wasilewska, J., Radwan, J. (2025). Do *Borrelia afzelii* outer surface protein E coevolve with complement factor H of its rodent host? Insights from GxG and spatial associations.

I declare that I have contributed to sample collection and molecular analyses, including tick DNA extraction and *Borrelia* screening.

Józefina Wasilewska

mgr Józefina Wasilewska