Josip Juraj Strossmayer University of Osijek,

University of Dubrovnik,

Ruđer Bošković Institute, Zagreb

University Postgraduate Interdisciplinary Doctoral Study Molecular Biosciences

Nina Čorak

# DIFFERENTIAL EXPRESSION AND EVOLUTIONARY SIGNATURES OF *Borrelia burgdorferi* MORPHOTYPES

DOCTORAL THESIS

Osijek, 2022.

**TEMELJNA DOKUMENTACIJSKA KARTICA**

**Sveučilište Josipa Jurja Strossmayera u Osijeku** **Doktorska disertacije**
**Sveučilište u Dubrovniku**
**Institut Ruđer Bošković**
**Poslijediplomski interdisciplinarni sveučilišni studij Molekularne bioznanosti**

**Znanstveno područje:** Prirodne znanosti
**Znanstveno polje:** Biologija

**Diferencijalna ekspresija gena i evolucijska obilježja morfotipova bakterije *Borrelia burgdorferi***

Nina Čorak

**Disertacija je izrađena u:** Laboratorij za evolucijsku genetiku, Zavod za molekularnu biologiju, Institut Ruđer Bošković, Bijenička cesta 54, 10000 Zagreb
**Mentor:** dr. rer. nat. Tomislav Domazet-Lošo, izvanredni profesor i viši znanstveni suradnik

**Kratki sažetak doktorske disertacije:**

*Borrelia burgdorferi* u laboratorijskim uvjetima može poprimiti oblik spiroheta, okruglih tjelešaca, mjehurastih formi ili biofilma. Rezultati proizašli iz ovog istraživanja upućuju na to da je tranzicija iz spiroheta u okrugla tjelešca regulirana malim brojem evolucijski konzerviranih gena, dok je prijelaz iz spiroheta u mjehuraste forme ili biofilm popraćen velikim ekspresijskim promjenama koje uključuju pojačanu ekspresiju evolucijski mlađih gena.

**Broj stranica: 162**
**Broj slika: 16**
**Broj tablica: 4**
**Broj literaturnih navoda: 186**
**Jezik izvornika:** engleski
**Ključne riječi:** *Borrelia burgdorferi*, Lajmska bolest, diferencijalna ekspresija, bakterijski morfotipovi, morfološka plastičnost, genomska filostratigrafija

**Datum obrane: 11.7.2022.**

**Stručno povjerenstvo za obranu:**
1. dr. sc. Đurđica Ugarković, znanstveni savjetnik u trajnom zvanju
2. dr. sc. Damjan Franjević, izvanredni profesor
3. dr. sc. Ivana Ivančić Baće, izvanredni profesor
4. prof. dr. sc. Dubravka Hranilović, redoviti profesor (zamjena)

**Disertacija je pohranjena u:** Nacionalnoj i sveučilišnoj knjižnici Zagreb, Ul. Hrvatske bratske zajednice 4, Zagreb; Gradskoj i sveučilišnoj knjižnici Osijek, Europska avenija 24, Osijek; Sveučilištu Josipa Jurja Strossmayera u Osijeku, Trg sv. Trojstva 3, Osijek

**BASIC DOCUMENTATION CARD**

**Josip Juraj Strossmayer University of Osijek** **Doctoral thesis**
**University of Dubrovnik**
**Ruđer Bošković Institute**
**University Postgraduate Interdisciplinary Doctoral Study of Molecular biosciences**

**Scientific Area:** Natural sciences
**Scientific Field:** Biology

# Differential expression and evolutionary signatures of *Borrelia burgdorferi* morphotypes

Nina Čorak

**Thesis performed at:** Laboratory of Evolutionary Genetics, Department of Molecular Biology, Ruđer Bošković Institute, Bijenička cesta 54, 10000 Zagreb

**Supervisor:** Tomislav Domazet-Lošo, dr. rer. nat., Associate Professor and Senior Research Associate

**Short abstract:**
In laboratory conditions, *Borrelia burgdorferi*, the causative agent of Lyme disease, exists as a spirochete, round body, bleb, or biofilm culture. This doctoral thesis suggests that spirochete to round body transition relies on the delicate regulation of a relatively small number of evolutionary conserved genes. In contrast, spirochete to bleb and biofilm transition includes significant reshaping of transcription profiles towards evolutionary young genes.

**Number of pages: 162**
**Number of figures: 16**
**Number of tables: 4**
**Number of references: 186**
**Original in:** English
**Key words:** *Borrelia burgdorferi*, Lyme disease, Differential expression, Bacterial morphotypes, Morphological plasticity, Genomic phylostratigraphy

**Date of the thesis defense: 11.7.2022.**

**Reviewers:**
1. dr. sc. Đurđica Ugarković, scientific adviser
2. dr. sc. Damjan Franjević, associate professor
3. dr. sc. Ivana Ivančić Baće, associate professor
4. prof. dr. sc. Dubravka Hranilović, full professor (substitute)

**Thesis deposited in:** National and University Library in Zagreb, Ul. Hrvatske bratske zajednice 4, Zagreb; City and University Library of Osijek, Europska avenija 24, Osijek; Josip Juraj Strossmayer University of Osijek, Trg sv. Trojstva 3, Osijek

Research for this doctoral thesis was performed in the Laboratory for Evolutionary Genetics, Department of Molecular Biology, Ruđer Bošković Institute, Zagreb, under the supervision of associate professor Tomislav Domazet-Lošo, dr. rer. nat. This work was supported by the Croatian National Science Foundation under the project "Phylostratigraphy of gene gain and loss" (HRZZ-IP-2016-06-5925) and "Young Researchers' Career Development Project – Training New Doctoral Students" (DOK-2018-01-6673).

# Acknowledgments

I would like to express my gratitude to my supervisor, dr. rer. nat. Tomislav Domazet-Lošo for allowing me to work in the Laboratory for evolutionary genetics and for the guidance, patience, and motivation he provided me with throughout my training as a doctoral student.

I thank all who contributed to the research of differential expression and evolutionary signatures of *Borrelia burgdorferi* morphotypes, especially to my collaborators at the BCA Clinic - Christina Daschkin, Viktoria Krey, Sirli Anniko, Innokenty Woichansky, Horst-Günter Maxeiner, and Carsten Nicolaus. The making and completion of my doctoral thesis would not have been possible without each of their essential contributions.

Especially, I thank my dear colleagues, Sara, Momir, Antonio, and Anja, for sharing their knowledge and experiences while making my first research job enjoyable and full of laughter.

Most of all, I would like to thank my family for their unconditional love and support throughout writing this thesis and my entire education and life in general.

**TABLE OF CONTENTS**

# 1. Introduction

*Borrelia burgdorferi* is the causative agent of Lyme disease, the most prevalent vector-borne disease in the Northern Hemisphere (Rosenberg et al., 2018). The zoonotic cycle of this spirochete includes a mammalian host and the hard tick of the genus *Ixodes* (Barbour & Hayes, 1986; Radolf, Caimano, Stevenson & Hu, 2012), whose rapid spread across natural ecosystems caused the increase in Lyme disease prevalence (Stafford, Cartter, Magnarelli, Ertel & Mshar, 1998; Dumic & Severnini, 2018). About 65,000 cases are reported annually in Europe, while in the United States 300,000 cases are registered each year (Mead, 2015). Unfortunately, the diagnosis of this zoonosis remains challenging (Berndtson, 2013). Some studies suggest that less than half of those infected with Lyme disease were diagnosed within five years after developing symptoms (Hündersen, Forst & Kasten, 2021). If left untreated, Lyme disease can manifest with severe symptoms such as encephalitis, chronic neuroborreliosis, facial paralysis, chronic arthritis, carditis, acrodermatitis chronica atrophicans, and lymphocytomas (Lindgren, Jaenson & Menne, 2006; Bamm, Ko, Mainprize, Sanderson & Wills, 2019). Consequently, *B. burgdorferi* is increasingly recognized as an escalating public health problem that requires a more profound understanding of its complex features (Bamm, Ko, Mainprize, Sanderson & Wills, 2019).

## 1.1. Infectious life cycle of *Borrelia burgdorferi*

*B. burgdorferi* is an obligate parasite, relying heavily on its arthropod and mammalian host (Drecktrah et al., 2015). The life cycle of the tick vector includes the egg, larval, nymphal, and adult stage (Radolf, Caimano, Stevenson & Hu, 2012; Kurokawa et al., 2020). Since *B. burgdorferi* is not transovarially transmitted (transmitted from a parent to its offspring), the pathogen is acquired by larval ticks during feeding on infected mammals (Drecktrah et al., 2015). After entering the larval midgut, *B. burgdorferi* starts to divide, drastically increasing its population size (Piesman, Schneider & Zeidner, 2001). As the midgut becomes nutrient-depleted, *B. burgdorferi* adapts to the new environment through changes in morphology and gene expression that lead to dormancy (Drecktrah et al., 2015). In this form, bacteria persist until the nymphal tick takes another blood meal that triggers *B. burgdorferi* transmission to a new mammalian host (Schwan & Piesman, 2002; Kung, Anguita & Pal, 2013). Although all three tick stages can feed on humans, nymphs are

responsible for the majority of spirochete transmission to humans (Radolf, Caimano, Stevenson & Hu, 2012). After the bite, spirochetes are deposited into the wound along with the tick saliva proteins (Drecktrah et al., 2015). There they develop complex interactions with the mammalian host, enabling successful persistence, dissemination, and, consequently, acute or sometimes chronic infection (Bamm, Ko, Mainprize, Sanderson & Wills, 2019). The enzootic cycle is continued when a naive tick acquires *B. burgdorferi* from its vertebrate host (Drecktrah et al., 2015).

## 1.2. Phylogeny of *Borrelia burgdorferi*

The *Borreliacea* and *Treponema* phylum, along with other free-living spirochetes and spirochetes which are symbionts of termite guts, are part of the *Spirochaetales* family (Fig. 1) (Paster & Dewhirst, 2000). *Borreliaceae* are tick-borne pathogens that often parasite on various vertebrates (Kurtenbach et al., 2002). They include species not causing human disease, species belonging to the relapsing fever group, and notably, the Lyme disease branch (Barbour & Gupta, 2021, Biesiada, Czepiel, Leśniak, Garlicki & Mach, 2012). According to Barbour and Gupta (2021), *Borreliaceae* may have initially been symbionts of the ticks. This notion is supported by the fact that *Centruroides limpidus* and *Stegodyphus dumicola*, the closest known relatives to the members of the *Borreliaceae* phylum, are found to be symbionts in the Arachnida gut (Barbour & Gupta, 2021). Additionally, for some *Borreliaceae* species, transovarial transmission occurs at high enough efficiency that a population of spirochete might be maintained through several generations of ticks without the need for passing through a vertebrate host (Barbour & Hayes, 1986).

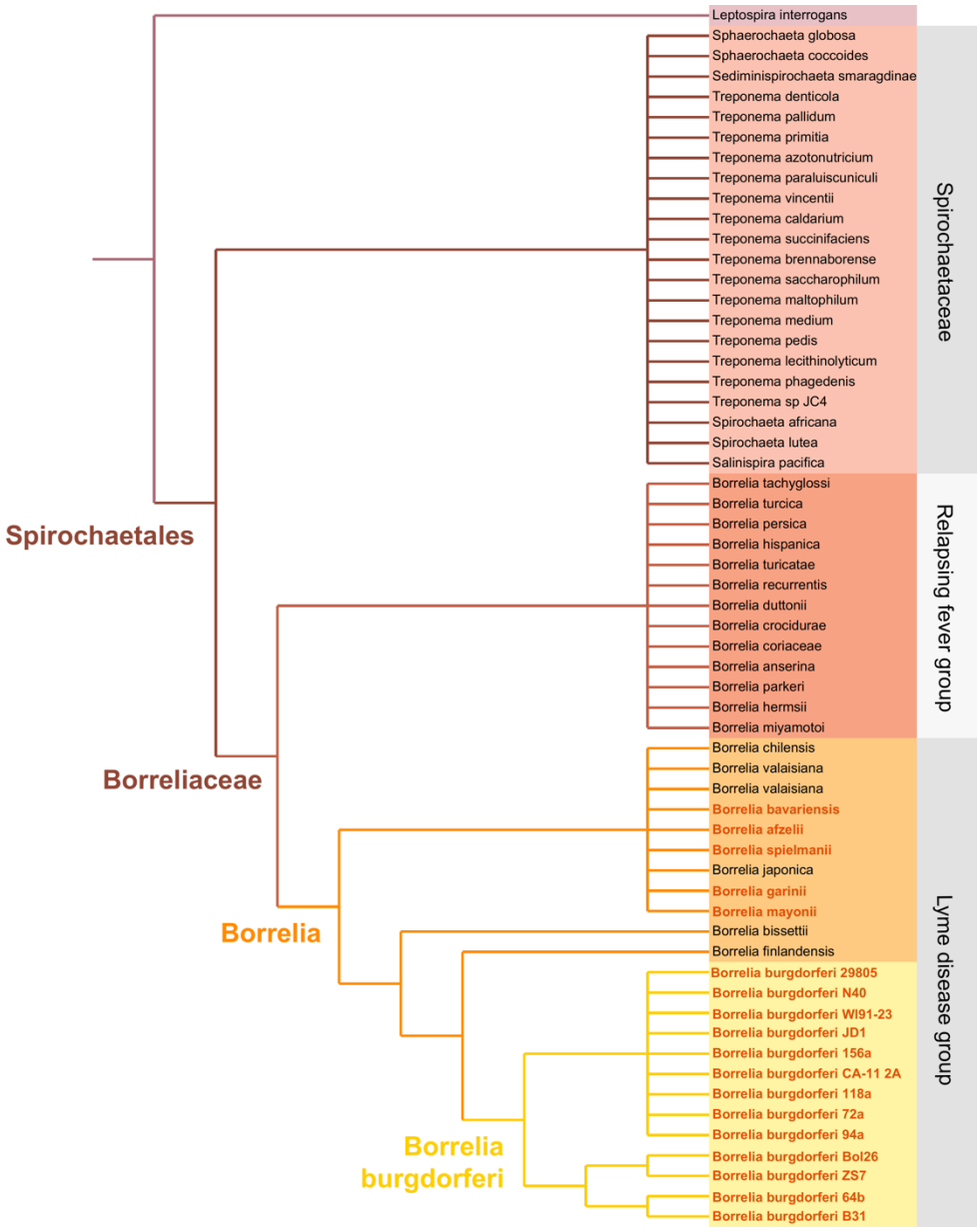



**Figure 1. Phylogeny of the Spirochaetales phylum.** Phylogenetic relationships are based on relevant phylogenetic literature (Margos et al., 2018; Parks et al., 2018; Mukherjee et al., 2017; Hug et al., 2016; Paster & Dewhirst, 2000; Raymann, Brochier-Armanet & Gribaldo, 2015; Di et al., 2014; Gupta, Mahmood & Adeolu, 2013; Rinke et al., 2013; Wu et al., 2009; Richter et al., 2006). Species known to cause Lyme disease are written in orange and bold (Barbour & Gupta, 2021; Biesiada, Czepiel, Leśniak, Garlicki & Mach, 2012).

In contrast to species belonging to the Relapsing fever group, where transovarian transmission may occur, species in the Lyme disease group obligatory rely on the mammalian host to maintain the population (Bamm, Ko, Mainprize, Sanderson & Wills, 2019). These species are present at low densities in the vertebrate blood, maintaining high numbers in the tissues and causing persistent infections (Barbour & Gupta, 2009). Other pathogens of the *Borreliaceae* phylum retain high concentrations in the blood during the infection, but after a few series of relapsing infections, evidence for residual infection is hard to find (Assous & Wilamowski, 2009). Since different species in the *Borrelia* genus have adapted to be resistant only to the components of the complement system in their specific vertebrate host and sensitive to components of the complement system found in other vertebrate species, it is proposed that specialization in the Borrelia genus is mediated by components of the vertebrate complement system (Radolf, Caimano, Stevenson & Hu, 2012).

*B. burgdorferi* is an interesting exception, as it is highly promiscuous in its selection of both the tick and the vertebrate species (Kurtenbach et al., 2002). Namely, *B. burgdorferi* is transmitted by *Ixodes scapularis* on the east, *Ixodes pacificus* on the west coast USA, *Ixodes ricinus* in Europe, and *Ixodes persulcatus* in Asia and Europe (Steere, Coburn & Glickstein, 2004). Also, this pathogen is carried and transmitted by a larger variety of birds and rodents, while deers, cattle, canines, and humans are dead-end hosts. Since *B. burgodrferi* parasites on a broad range of vertebrate and tick species, and the species belonging to the Borreliaceae phylum cumulatively also parasites on a broad range of ticks and vertebrates, *B. burgdorferi* is a good choice for studying adaptations evolved in the *Borreliacea* phylum.

## 1.3. Lyme disease

The first official account of Lyme disease was authored by Steere et al. in 1977, describing a condition that was initially thought to be an outbreak of juvenile rheumatoid arthritis in and around the town of Lyme in the United States of America. The cause of the syndrome was a previously unknown spirochete named *B. burgdorferi* after its discoverer in 1982 (Burgdorfer et al., 1982; Steere et al., 1983). Today, although a considerable amount of knowledge is gained, the progression of Lyme disease remains a topic of controversy (Bamm, Ko, Mainprize, Sanderson &

Wills, 2019). Based on a paper published by Berndson, Lyme disease can be described in three stages: early localized, early disseminated, and late Lyme disease (Berndson 2013).

The hallmark of early, localized infection is erythema migrans (Berndston 2013), an expanding skin lesion with central clearing at the tick bite site (Coumou, van der Poll, Speelman & Hovius, 2011). This mobile circular skin blemish develops two to three weeks after the tick bite (Steere et al., 2016) in 70-80% of infected patients (Aguero-Rosenfeld, Wang, Schwartz & Wormser, 2005). Besides erythema migrans, the acute phase is characterized by non-specific clinical signals such as fever, muscle aches, headache, nausea, and fatigue (Berndston 2013; Kurokawa et al., 2020), which hinders early diagnosis and treatment of patients without the typical skin rash (Bamm, Ko, Mainprize, Sanderson & Wills, 2019).

After two to three weeks, the second stage of infection begins (Berndston 2013). Bacteria can evade innate immune recognition, enter the bloodstream (Hyde, 2017; Bamm, Ko, Mainprize, Sanderson & Wills, 2019), and disseminate to the heart, urinary bladder, joint tissues, and central nervous system (Kurokawa et al., 2020). This phase is characterized by worsening symptoms such as migratory joint pains, fatigue, myocarditis, atrioventricular heart block, synovitis, acute neuroborreliosis, and borrelial lymphocitoma (Berndston 2013; Bamm, Ko, Mainprize, Sanderson & Wills, 2019).

If left untreated, the infection progresses into the third phase characterized by multisystemic manifestations, including rheumatologic disease, cardiac disease, and neuroborreliosis (Kurokawa et al., 2020). Neuroborreliosis occurs in about 20% of chronic Lyme disease cases (Kurokawa et al., 2020), causing irreversible neuronal damage (Peters & Benach, 1997). It is generally manifested as a painful meningoradiculitis known as Bannwarth syndrome, facial nerve palsy, encephalitis, segmental myelitis, cranial neuritis, radiculoneuritis, vasculitis, and intracranial hypertension (Koedel, Fingerle & Pfister, 2015; Ogrinc et al., 2016; Uldry, Regli & Bogousslavsky, 1987; Stanek et al., 2011). Another common manifestation of late-stage is Lyme arthritis, which affects 10% of infected patients (Kurokawa et al., 2020). The main feature of this syndrome is joint swelling caused by the inflammatory response in synovial tissue, consisting of synovial hypertrophy, vascular proliferation, and infiltration of mononuclear cells (Puius & Kalish, 2008).

Although antibiotic treatment in early diagnosed patients is mostly successful in preventing the infection from entering the third stage (Bamm, Ko, Mainprize, Sanderson & Wills, 2019), many cases report ongoing symptoms despite the inability to confirm the presence of *B. burgdorferi* by standardized diagnostic protocols (Bamm, Ko, Mainprize, Sanderson & Wills, 2019). This stage, often referred to as Post-Treatment Lyme Disease Syndrome (PTLDS) or Chronic Lyme Disease (CLD), is a topic that continues to be controversial for some members of the medical profession (Bamm, Ko, Mainprize, Sanderson & Wills, 2019). While the cause of PTLDS and CLD is still unknown, some prominent theories include the possibility of co-infection (Swanson, Neitzel, Reed & Belongia, 2006), autoimmune response (Steere, Gross, Meyer & Huber, 2001), immune response to the continued presence of antigenic debris (Bockenstedt, Gonzalez, Haberman & Belperron, 2012), and most notably the presence of *B. burgdorferi* persisters not killed by antibiotics (Hodzic, Feng, Holden, Freet & Barthold, 2008; Hodzic, Imai, Feng & Barthold, 2014; Embers et al., 2012).

## 1.4. Cellular and molecular biology of *Borrelia burgdorferi*

The cell envelope of *B. burgdorferi* is made of the protoplasmic cylinder surrounded by an outer membrane, below which is the peptidoglycan layer. The periplasmic space between these two membranes accommodates numerous flagella (Vancová et al., 2017) wrapped around the cell, confining the spiral shape of the spirochete (Motaleb et al., 2000). It is suggested that this spiral shape provides the bacteria with its specific motility that allows efficient dissemination and tissue penetration (Bernardson, 2013; Yang, Blair & Salama, 2016; Harman et al., 2012). At the same time, the location of the flagella prevents the exposure of flagellar antigens to immune system effectors and, consequently, protects the bacteria from host immune system recognition (Charon et al., 2012).

Swimming through environments such as hemolymph, blood, cerebrospinal fluid (Charon et al., 2012), and especially the highly-dense extracellular matrix network in the dermis of mammals (Berndtson, 2013) requires high metabolic activity (Meriläinen, Herranen, Schwarzbach & Gilbert, 2015). Since *B. burgdorferi* possesses a limited capability of *de novo* biosynthesis (Groshong & Blevins, 2014), energy production relies heavily on the host and transportation systems to scavenge nutrients from the environment (Kurokawa et al., 2020). The carbon source

used by *B. burgdorferi* changes during different stages of its life cycle. Namely, in the mammalian host, the primary carbon source utilized by *B. burgdorferi* is glucose (Hoon-Hanks et al., 2012), while in a tick, the bacteria predominantly uses glycerol (Pappas et al., 2012) and, to a lesser extent, chibitose (Tilly, Grimm, Bueschel, Krum & Rosa, 2004). Because *B. burgdorferi* takes part in the two-host enzootic cycle, nutrient acquisition mechanisms must be regulated depending on the nutrients available in these diverse environments (Groshong & Blevins, 2014).

Along with the nutrient acquisition, *B. burgdorferi* must delicately control immune evasion and tissue colonization, both in the tick and human host (Bamm, Ko, Mainprize, Sanderson & Wills, 2019). However, not much is known regarding how the bacteria avoids the arthropod's innate immune response during initial acquisition and throughout the tick molting period (Groshong & Blevins, 2014). On the other hand, the virulence factors necessary for infection in mammals are better understood (Bamm, Ko, Mainprize, Sanderson & Wills, 2019). *B. burgdorferi* lacks classic lipopolysacharrides in the outer membrane (Kurokawa et al., 2020). Instead of them, a large proportion of *B. burgodrferi* genomic resources are devoted to producing outer surface proteins (Berndson 2013) essential for spirochete survival and navigation through physiologically and immunologically hostile host environments (Kurokawa et al., 2020).

Environmental conditions such as temperature, pH, osmolarity, oxidative stress, cell density, carbon starvation (Kazmierczak, Wiedmann & Boor, 2005), and the host in which the *B. burgdorferi* parasites influence the expression of a variety of genes (Kurokawa et al., 2020). One of the most crucial mechanisms controlling gene expression in *B. burgdorferi* is the RpoN-RpoS alternative sigma factor pathway (Ouyang, Blevins & Norgard, 2008). It secures the successful transmission from tick to vertebrate host and regulates more than 100 genes involved in survival and stress response (Ouyang, Narasimhan & Neelakanta, 2012). Some of those genes are well-described virulence factors, such as outer surface proteins OspA, OspB, and OspC, as well as decorin-binding and fibronectin-binding proteins (Ouyang, Blevins & Norgard, 2008).

Another bacterial stress control system is stringent response (Cabello, Godfrey, Bugrysheva & Newman, 2017), which activates during periods of starvation (Haseltine & Block, 1973). It is characterized by increased levels of guanosine tetraphosphate (ppGpp) and guanosine pentaphosphate (pppGpp), which act as triggers for downstream reactions (Drecktrah et al., 2015). *In B. burgdorferi*, it regulates global gene expression in ticks between the larval and nymph blood

meal and during *in vitro* starvation (Drecktrah et al., 2015). Activation of the stringent response is associated with low expression of the *rpoS* gene. Namely, RpoS is an alternative sigma factor whose expression is increased during a nymphal blood meal. It is required for transcription of genes involved in tick to mammal transmission and mammalian infection, and repression of σ70-dependent genes expressed by *B. burgdorferi* in the tick phase of the enzootic cycle. The expression of *rpoS* is correlated with the expression of another sigma factor, RpoN, which is taking part in the transcription of the *rpoS* gene (Caimano et al., 2019).

Another important bacterial communication system is quorum sensing, which regulates gene expression in response to population density (De Keersmaecker, Sonck & Vanderleyden, 2006). As bacterial population density rises, molecules called autoinducers accumulate in the extracellular solution (De Keersmaecker, Sonck & Vanderleyden, 2006), causing changes in bacterial gene expression and biofilm formation (De Keersmaecker, Sonck & Vanderleyden, 2006). In *B. burgdorferi*, LuxS/autoinducer-2 is utilized during quorum sensing response (Stevenson et al., 2003). Described differential gene expression control systems are crucial for maintaining *B. burgdorferi* enzootic life cycle and survival in adverse environmental conditions (Stevenson et al., 2003; Cabello, Godfrey, Bugrysheva & Newman, 2017).

## 1.5. *Borrelia burgdorferi* chromosome and plasmids

The genome of *B. burgdorferi* harbors a linear chromosome of about 900 kb in length and a plethora of circular and linear plasmids (Fraser et al., 1997; Casjens et al., 2000). Most genes on the chromosome are bacterial orthologues with known housekeeping functions (Kurokawa et al., 2020), such as replication, transcription, translation, energy metabolism, and transmembrane transport (Berndtson, 2013). However, because of its parasitic lifestyle, the *B. burgdorferi* chromosome has no genes for cellular biosynthesis of amino acids, fatty acids, enzyme cofactors, or nucleotides (Berndtson, 2013), making it one of the smallest genomes found among bacteria (Schwartz, Margos, Casjens, Qiu & Eggers, 2020).

Additionally, species belonging to the *Borrelia* taxa carry more plasmids than other bacteria (Schwartz, Margos, Casjens, Qiu & Eggers, 2020). Specifically, 10 circular and 12 linear plasmids, ranging from 5 to 84 kbp in size (Casjens et al., 2000), are present in the cell at low copy numbers (Casjens et al., 2017). Some circular plasmids are potential prophages (Schwartz, Margos, Casjens, Qiu & Eggers, 2020), while linear plasmids have covalently-closed hairpin ends (Casjens et al., 2017). The plasmids have unusual characteristics such as low density of protein-coding genes, many paralogous sequences, and a large number of pseudogenes (Schwartz, Margos, Casjens, Qiu & Eggers, 2020). Some carry essential genes, but most plasmid genes are coding for differentially expressed surface proteins important in the interactions between bacteria and their arthropod and vertebrate hosts (Schwartz, Margos, Casjens, Qiu & Eggers, 2020). In addition, many of these genes are essential for accomplishing immune evasion (Berndtson, 2013) and represent potential vaccine and detection targets (Casjens et al., 2017).

The plasmids are generally not required for growth in culture, but the plasmid cp26 is an exception since it carries genes essential for *B. burgdorferi* survival (Schwartz, Margos, Casjens, Qiu & Eggers, 2020). In contrast, plasmid lp28-1 is not essential for survival in laboratory conditions but is crucial for pathogenesis in both the tick and vertebrate host (Schwartz, Margos, Casjens, Qiu & Eggers, 2020). It codes for the VlsE antigen, which is highly variable and shields the bacteria from the host's immune response (Verhey, Castellanos & Chaconas, 2019). This protein's high degree of heterogeneity is gained by antigenic switching through recombination in the *vlsE* locus and *vls* silent cassettes, representing a challenge for developing effective vaccines (Verhey, Castellanos & Chaconas, 2019). Another plasmid important in mouse infection is plasmid lp54, whose one-third of genes are regulated by the transcription factor RpoS (Caimano et al., 2007). Those genes include major surface proteins OspA and OspB (Tilly, Checroun & Rosa, 2012) and decorin binding proteins DbpA and DbpB (Salo 2015). Finally, although not essential for maintenance in hosts, cp32 plasmids are interesting since they contain genes such as RevA, ErpM, and ErpY, which are coding for fibronectin, plasminogen, lamin, and complement H binding products, making them potential adhesion regulating plasmids (Schwartz, Margos, Casjens, Qiu & Eggers, 2020).

## 1.6. Bacterial dormancy spectrum

In some cases, several morphological forms can be simultaneously present in a bacterial culture at a given time. This population-level phenomenon is often referred to as pleomorphism (Caccamo et al., 2019) and is well described in *Escherichia coli*, *Pseudomonas aeruginosa*, *Mycobacterium tuberculosis*, *Salmonella enterica*, and *Staphylococcus aureus* (Helaine & Kugelberg, 2014; Harms, Maisonneuve & Gerdes, 2016; Michiels, Van den Bergh, Verstraeten & Michiels, 2016). Such morphological heterogeneity, which often includes cells on different parts of the dormancy spectrum, enables selective benefits to the bacterial populations under stressful conditions (Ayrapetyan, Williams & Oliver, 2018).

Bacterial dormancy is a state characterized by low metabolic activity, extended periods without replication, and different gene expression profiles than those found in fully active replicating cells (Rudenko, Golovchenko, Kybicova & Vancova, 2019). In many species, the formation of dormant cells is a dynamic, stepwise process that can happen stochastically or due to environmental cues (Ayrapetyan, Williams & Oliver, 2018). For example, the number of active reproductive cells decreases during exposure to environmental stress, while the number of persisters and viable but unculturable cells rises (Ayrapetyan, Williams & Oliver, 2018).

Persister bacteria are cells on the dormancy spectrum that tolerate antibiotics, not by gene mutation, but because the efficiency of most antibiotics depends on the presence of actively growing cells (Bernatdtson 2013). Most authors agree that persisters comprise around 1% of growing bacterial populations (Bernatdtson 2013). Similarly, viable but nonculturable cells (VBNC) are a functionally viable subpopulation in bacterial cultures transiently unable to grow on media on which the fully active form of bacteria grows (Xu et al., 1982). These cells exhibit significantly lower, although present, metabolic activity than their actively growing counterparts, and they continue to maintain membrane integrity and produce proteins (Mali et al., 2017; Oliver, 2010).

When the stressor is removed, these persisters and VBNC can revert to the most common morphological form under optimal culturing conditions (Balaban, Merrin, Chait, Kowalik & Leibler, 2004; Ayrapetyan, Williams & Oliver, 2018). In addition, these stress-tolerant cell types are often present in low numbers and characterized by low metabolic activity and low replication

rates (Fisher, Gollan & Helaine, 2017). For this reason, commonly used antibiotics targeting metabolic production have a low impact on the fitness of bacterial populations showing pleomorphism (Fisher, Gollan & Helaine, 2017).

## 1.7. Persisters of *Borrelia burgdorferi*

The persistence of tissue spirochetes has been suggested since they were first reported by Dutton in the 19th century (Brorson et al., 2009). Since then, subpopulations of *B. burgdorferi* cells, which remain viable despite antibiotic therapy and revert into motile spirochetal forms under favorable conditions, have been frequently reported *in vitro* (Timmaraju et al., 2015; Vancová et al., 2017; Rudenko, Golovchenko, Kybicova & Vancova, 2019; Feng, Zhang, Shi & Zhang, 2016; Nocton et al., 1994; Li et al., 2011; Zhang, 2014). Even *in vivo*, it was demonstrated that antibiotic treatments currently viewed as adequate for achieving complete eradication could not clear persisting *B. burgdorferi* from mice (Barthold et al., 2010) and nonhuman primate tissues (Embers et al., 2012). This phenotypic drug tolerance of persistent subpopulations is associated with stringent response, the primary bacterial stress response mechanism (Rudenko, Golovchenko, Kybicova & Vancova, 2019). Evidence suggests that *B. burgdorferi* is adapted to persist in immune-competent hosts and remain infective despite aggressive antibiotic challenges (Berndtson, 2013). Because of that, PTLDS is often explained by the presence of persisters in the bacterial population, and it is hypothesized that these cells correspond to alternative morphotypes of *B. burgdorferi* (Bamm, Ko, Mainprize, Sanderson & Wills, 2019; Berndtson, 2013).

## 1.8. *Borrelia burgdorferi* morphotypes

Morphological plasticity is the ability of individual bacterial cells to dynamically change their shape in response to environmental conditions (Caccamo et al., 2019). This feature can be found among various bacteria taxa, including pathogens, where the colonization of distinct tissues, transmission between hosts, and transit through environmental reservoirs are often accompanied by morphological transformations of bacterial cells (Yang, Blair & Salama, 2016).

Similar to other spirochetes (Vesey & Kuramitsu, 2004; Ristow et al., 2008; Umemoto et al., 1984), *B. burgdorferi* evokes morphological alterations to respond to hostile environmental signals (Rudenko, Golovchenko, Kybicova & Vancova, 2019). Several of these morphologies, namely, spirochetes, round bodies, blebs, and biofilms, are simultaneously present in *B. burgdorferi* cultures grown in the BSK-II medium — the most common medium used in *B. burgdorferi* cultivation (Meriläinen, Herranen, Schwarzbach & Gilbert, 2015). The spirochete morphotype, in particular, is corkscrew-shaped and represents the dominant morphotype in BSK-II medium-raised cultures (Barbour, 1984; Meriläinen, Herranen, Schwarzbach & Gilbert, 2015; Vancová et al., 2017). Compared to alternative pleomorphic forms, spirochetes are relatively easy to cultivate in laboratory conditions, and thus they are the most commonly studied *B. burgdorferi* morphotype (Rudenko, Golovchenko, Kybicova & Vancova, 2019). Consequently, most of our knowledge regarding *B. burdgforeri* refers to the spirochete morphotype.

### 1.8.1. Round bodies

Spherical *B. burgdorferi* cells with intact, flexible cell envelopes enclosing numerous flagella are named in various ways, e.g., round bodies, spheroplasts, cystic forms, spherules, coccoid forms, protoplasts, and propagules (Vancová et al., 2017; Domingue & Woody, 1997; Stricker & Johnson, 2011; Meriläinen, Herranen, Schwarzbach & Gilbert, 2015; Bamm, Ko, Mainprize, Sanderson & Wills, 2019). Nonetheless, all of these labels describe the same spherical structures, and as suggested by Meriläinen et al. (2015), I will refer to them as round bodies. Round bodies are viable, slowly reproductive morphologies (Margulis, Maniotis & MacAllister, 2009), with a mean size of approximately 2.8 μm (Meriläinen, Herranen, Schwarzbach & Gilbert, 2015). Although they are less motile than typical helical-shaped spirochetes, round bodies twitch and possess lateral mobility (Brorson et al., 2009). During the transformation of the spirochetes into the round body pleomorphic variant, the flexibility of the outer membrane facilitates the expansion of the outer membrane, and the loose links between the inner and outer membrane enable the folding of the protoplasmic cylinder within the confinements of the outer envelope (Alban, Johnson & Nelson, 2010; Meriläinen, Herranen, Schwarzbach & Gilbert, 2015). Unlike other morphotypes, the outer membrane of round bodies stains positively for N-acetylglucosamine polysaccharide, both *in vitro* (Meriläinen, Herranen, Schwarzbach & Gilbert, 2015) and in human Langerhans cells

(Hulínská et al., 1994). The elasticity of the outer membrane and reorganization of the membrane components during round body formation could provide a potential explanation for N-acetylglucosamine membrane exposure and subsequent staining (Meriläinen, Herranen, Schwarzbach & Gilbert, 2015).

Round bodies make up a small subpopulation in BSK-II medium-raised cultures (Meriläinen, Herranen, Schwarzbach & Gilbert, 2015; Alban, Johnson & Nelson, 2010; Vancová et al., 2017). Only 0.4% of bacteria can be found in this form after four days of cultivation (Meriläinen, Herranen, Schwarzbach & Gilbert, 2015). On a molecular level, the transition of spirochetes to round body morphotype is dependent on the production of alarmones guanosine tetraphosphate and guanosine pentaphosphate, general effectors of the stringent response (Drecktrah et al., 2015). Furthermore, deletion of the $Rel_{Bbu}$ gene involved in the *B. burgdorferi* stringent response decreases cell survival and increases the number of round bodies present under starvation conditions (Drecktrah et al., 2015). Based on that, it is not surprising that round body rich cultures can be obtained under conditions that limit bacterial growth or induce bacterial cell desiccation (Margulis, Maniotis & MacAllister, 2009). These conditions include culture aging, changes in acidity-alkalinity, viscosity, temperature, salt concentration, gas composition, concentrations of antibiotics, sugars, amino acids or exposure to oxygen gas, total anoxia, and sulfide (Alban, Johnson & Nelson, 2010; Brorson & Brorson, 1998; Murgia, Piazzetta & Cinco, 2002; Feng, Shi, Zhang & Zhang, 2015; Meriläinen, Herranen, Schwarzbach & Gilbert, 2015; Vancová et al., 2017; Brorson et al., 2009; Sapi et al., 2011; Bamm, Ko, Mainprize, Sanderson & Wills, 2019; Brorson et al., 2009).

The exposure of spirochetes to distilled water is the most commonly used method for the induction of round body morphotype under laboratory conditions (Bamm, Ko, Mainprize, Sanderson & Wills, 2019; Miklossy et al., 2008; Meriläinen, Herranen, Schwarzbach & Gilbert, 2015). This is not surprising since, after only 10 min of incubation, almost 85% of cells obtain a round body morphology (Meriläinen, Herranen, Schwarzbach & Gilbert, 2015). Cultivating *B. burgdorferi* in the BSK-II medium in which rabbit serum is replaced with human serum also induces round body formation in conditions mimicking those found in the human host (Meriläinen, Herranen, Schwarzbach & Gilbert, 2015). Since the rabbit serum typically supplementing the *B. burgdorferi culture* medium has the same osmolarity as the human serum (Meriläinen, Herranen,

Schwarzbach & Gilbert, 2015) and round bodies induced by human serum share morphological features with those induced with distilled water, it is clear that osmotic stress is not the only factor triggering the transformation (Meriläinen, Herranen, Schwarzbach & Gilbert, 2015). Moreover, research has shown that the change in morphology can be induced by the complement system or antibody exposure, which is clinically interesting and worthy of further studies (Meriläinen, Herranen, Schwarzbach & Gilbert, 2015). Another method frequently used for round body induction is serum starvation (Bamm, Ko, Mainprize, Sanderson & Wills, 2019). When *B. burgdorferi* is exposed to BSK-II medium without rabbit serum, the lack of nutrients triggers spirochetes to convert into round bodies (Meriläinen, Herranen, Schwarzbach & Gilbert, 2015; Alban, Johnson & Nelson, 2010; Murgia et al. 2002). Round bodies are metabolically less active than motile, reproducible spirochetes (Meriläinen, Herranen, Schwarzbach & Gilbert, 2015), and since *Borrelia* is an obligate parasite that lacks genes required for *de novo* amino acid synthesis (Groshong, Dey, Bezsonova, Caimano & Radolf, 2017; Fraser et al., 1997), the round body morphotype provides a path to bacterial persistence when its nutritional needs are not met (Margulis, Maniotis & MacAllister, 2009).

Interestingly, antibiotics commonly used for Lyme disease treatment (Bamm, Ko, Mainprize, Sanderson & Wills, 2019), such as ceftriaxone (Kersten, Poitschek, Rauch & Aberer, 1995), amoxicillin (Feng, Zhang, Shi & Zhang, 2016), and doxycycline (Sapi et al., 2011), are known to trigger round body formation. For instance, although its usage reduces the number of spiral-shaped cells by 90%, doxycycline almost doubles the number of round bodies in the culture (Sapi et al., 2011). Similarly, after three days of incubation with amoxicillin, 96% of *B. burgdorferi* cells are present in the culture as round bodies (Feng, Zhang, Shi & Zhang, 2016). Round bodies induced by these two antibiotics enable spirochete survival by downregulation of outer membrane lipoprotein gene expression, most probably as a mean of drug target reduction (Feng, Shi, Zhang & Zhang, 2015). Most importantly, doxycycline, amoxicillin, tigecycline, metronidazole, and tinidazole show reduced effect on round body termination (Sapi et al., 2011). Although it is not certain if round bodies formed in response to antibiotic exposure are the same as those present in other *in vitro* culturing conditions (Feng, Zhang, Shi & Zhang, 2016), reduced effectiveness of antibiotics coupled with a significant round body induction potential goes hand in hand with the idea that round bodies indeed are persisters (Bamm, Ko, Mainprize, Sanderson & Wills, 2019).

Cultivation of spirochetes in chicken neurons, rat neuron and astrocyte cultures, human monocyte and astrocyte culture (Miklossy et al., 2008), human cerebrospinal fluid (Brorson & Brorson, 1998), and tonsillar tissue laboratory cultures (Duray et al., 2005) are additional methods for round body induction. Additionally, non-motile spherical *B. burgdorferi* cells have been visualized within the midgut of unfed *Ixodes scapularis* nymphs, indicating that the formation of round bodies is a potential survival strategy during conditions of limited nutritional availability (Dunham-Ems, Caimano, Eggers & Radolf, 2012). Most importantly, spherical structures with round body morphology were also found *in vivo*, in histopathological samples of the dogs with myocarditis (Janus et al., 2014), in the cerebral cortex of patients with chronic Lyme neuroborreliosis (Miklossy et al., 2008), and in the skin tissues of patients with erythema migrans (Aberer, Kersten, Klade, Poitschek & Jurecka, 1996; Hulínská et al., 1994).

The presence of persisters in *B. burgdorferi* cultures indicates their importance in the development of PTLDS and CLD (Bamm, Ko, Mainprize, Sanderson & Wills, 2019). Since round bodies possess many typical persistent features, they are generally considered persisters (Rudenko, Golovchenko, Kybicova & Vancova, 2019; Brorson et al., 2009; Alban, Johnson & Nelson, 2010; Feng, Shi, Zhang & Zhang, 2015; Strnad, Grubhoffer & Rego, 2020). Because of that, heterogeneous populations of spirochetes and round bodies cannot be killed by antibiotics currently used for Lyme disease treatment (Feng, Shi, Zhang & Zhang, 2015). Moreover, not even daptomycin, a persistent targeting drug, affects round body eradication (Feng, Shi, Zhang & Zhang, 2015). Instead, only a combination of a drug that kills persisters and a drug that eradicates growing forms is effective enough to kill cultures containing both spirochetes and round bodies (Feng, Shi, Zhang & Zhang, 2015). Crucial round body traits enabling this type of persistence are low metabolic activity (Meriläinen, Herranen, Schwarzbach & Gilbert, 2015) and the ability of the population to grow in size (Ayrapetyan, Williams & Oliver, 2018). A decrease in metabolic activity in round body morphotypes is proved by measuring almost non-existent amounts of ATP (Meriläinen, Herranen, Schwarzbach & Gilbert, 2015) and the unchanged composition of the BSK-II medium used for round body induction (Bernardson 2013). Low metabolic activity is accompanied by reduced nutrient requirements and decreased protein production, making round bodies less susceptible to medium composition and antibiotic exposure (Ayrapetyan, Williams & Oliver, 2018). Besides, round bodies on their own are incapable of growth by reproduction (Margulis, Maniotis & MacAllister, 2009). Because of that, a reversion back into reproductive

spirochetes after re-introduction into the standard growth medium is an essential precondition for round body persistence (Bamm, Ko, Mainprize, Sanderson & Wills, 2019). This reversible morphological plasticity was confirmed for round bodies gained by exposure to distilled water (Meriläinen, Herranen, Schwarzbach & Gilbert, 2015), nutrient deprivation (Alban, Johnson & Nelson, 2010), cultivation in the spinal fluid (Brorson & Brorson, 1998), and neuronal cultures (Miklossy et al., 2008). Despite round bodies' lower metabolic production once they revert to spirochetes, cells become fully metabolically active and capable of reproduction (Meriläinen, Herranen, Schwarzbach & Gilbert, 2015).

Regarding their potential clinical relevance, many authors have suggested that the transformation of *B. burgdorferi* from spirochetes to round bodies may enhance immune system evasion (Al-Robaiy et al., 2010; Brorson & Brorson, 1998; Lawrence, Lipton, Lowy & Coyle, 1995). This notion was tested *in vitro* by Meriläinen et al. (2016). Their research demonstrated that macrophages internalized more spirochetes per cell and possessed higher lysosomal processing capacity than when they internalized round bodies (Meriläinen, Brander, Herranen, Schwarzbach & Gilbert, 2016). Also, the macrophage association with spirochetes and the macrophage association with round bodies caused a significant difference in the expression of seven macrophage immune-modulating mediators (Meriläinen, Brander, Herranen, Schwarzbach & Gilbert, 2016). More precisely, spirochetes induced higher secretion of IL-1b, IL-1ra, IL-6, MIF, MIP-1b, and RANTES, while round bodies induced a significantly elevated level of MCP-1 (Meriläinen, Brander, Herranen, Schwarzbach & Gilbert, 2016). MCP-1 regulates the migration and infiltration of monocytes, T-cells, and NK-cells, and it potentially has a role in the polarization of naïve T cells (Gu et al., 2000). Furthermore, its expression is required for the development of experimental Lyme arthritis in mice (Brown, Blaho & Loiacono, 2003), and it is thought to be associated with other autoimmune diseases as well (Deshmane, Kremlev, Amini & Sawaya, 2009), proposing a potential role of round bodies in modulating macrophage immune response (Meriläinen, Brander, Herranen, Schwarzbach & Gilbert, 2016).

### 1.8.2. Blebs

The least studied pleomorphic form, the so-called "bleb" morphotype, is characterized by the formation of large bulges (Vancová et al., 2017) of the outer membrane of otherwise spiral-shaped *B. burgdorferi* cells. These bulges bud into small outer membrane vesicles (OMVs) (Vancová et al., 2017; Berndtson, 2013) which shed off the bacterial surface. In *B. burgdorferi* and various bacterial species, these vesicles carry secretory products such as metabolites, nucleic acids, proteins, and endotoxic lipopolysaccharides (Jan, 2017), of which many are virulence factors (Toledo, Coleman, Kuhlow, Crowley & Benach, 2012). Furthermore, a significant representation of cytosolic and inner membrane molecules inside OMVs (Li, Clarke & Beveridge, 1998) and biased mRNA transcript distribution between the bacterial cell and its OMV (Malge et al., 2018) indicates an active sorting process in the bleb morphotype.

The bleb morphotype makes up to 4% of *B. burgdorferi* cells raised in the BSK-II culture at 37 °C (Meriläinen, Herranen, Schwarzbach & Gilbert, 2015). However, a more significant percentage of bleb cells can be induced *in vitro* by other environmental triggers like antibiotics, components of the complement system, and culture aging (Kersten, Poitschek, Rauch & Aberer, 1995; Barbour & Hayes, 1986; Meriläinen, Herranen, Schwarzbach & Gilbert, 2015). Few bleb cells can be found in chicken and rat neuron cultures and rat and human astrocytes cultures (Miklossy et al., 2008)**.** Additionally, bleb morphotype was also observed *in vivo* in cell cultures isolated from erythema migrans lesions on the skin of Lyme disease patients (Kersten, Poitschek, Rauch & Aberer, 1995; Aberer, Kersten, Klade, Poitschek & Jurecka, 1996) and cerebral cortex of patients with Lyme neuroborreliosis (Miklossy et al., 2008). Also, it is confirmed that *B. burgdorferi* blebs form during the blood-feeding events in ticks (Malge et al., 2018).

The precise function of blebs in *B. burgdorferi* is still unknown (Meriläinen, Herranen, Schwarzbach & Gilbert, 2015), but several functions are proposed based on previous research. It is known that OMVs shed from the bleb surface are abundant with plasminogen receptors, which potentiate proteolytic degradation of the vertebrate extracellular matrix (Toledo, Coleman, Kuhlow, Crowley & Benach, 2012). In this way, the bleb morphotype initiates dissemination throughout the host tissue and expands the nutrient availability (Toledo, Coleman, Kuhlow, Crowley & Benach, 2012). Bleb OMVs are also a standard way for bacteria to communicate with

each other and their environment (Malge et al., 2018). They are also involved in biofilm formation, interspecies and intraspecies delivery of molecules, resistance against antibiotics, and modulation of the host immune response (Jan, 2017). In particular, OMVs shed from the *B. burgdorferi* bleb morphotype are enriched with transcripts associated with nucleic acid/DNA metabolism, integration, and recombination (Malge et al., 2018). These OMVs can bind to human endothelial cells in culture (Shoberg & Thomas, 1993) and induce B-cell response in mouse models (Whitmire & Garon, 1993) and are thus considered important for Lyme disease initiation, progression, and persistence (Malge et al., 2018).

### 1.8.3. Biofilms

Bacterial biofilms are multicellular assemblies composed of cells embedded in a self-produced extracellular polysaccharide matrix characterized by fine-tuned physiology, ordered structural organization, and interactive social behavior (Meriläinen, Herranen, Schwarzbach & Gilbert, 2015). The complex mixture of polysaccharides, nucleic acids, lipid proteins, and other molecules present in the extracellular matrix (Song, Duperthuy & Wai, 2016) keeps the integrity of the biofilm, holds the community together (Vlamakis, Aguilar, Losick & Kolter, 2008; Vlamakis, Chai, Beauregard, Losick & Kolter, 2013), and represents a perfect hiding place for individual cells (Donlan, 2002). Environmental triggers such as a change in media composition, nutrition availability, temperature, pH, osmolarity, iron exposure, oxygen exposure, and other stressful conditions might induce cell-surface and cell-cell interactions of free-living bacteria, which initiates the formation of true a biofilm (O'Toole, Kaplan & Kolter, 2000; Rudenko, Golovchenko, Kybicova & Vancova, 2019). These conditions can be obtained *in vitro* where biofilms can be cultivated at the air-agar interface, floating in at the air-liquid interface, submerged, or surface-adhered at the liquid-solid interface (Vlamakis, Chai, Beauregard, Losick & Kolter, 2013). While biofilms are the predominant form of bacteria in almost all natural and man-made habitats (Flemming & Wuertz, 2019), their association with antibiotic resistance, embryo-like features, bacterial persistence, and chronic infection makes them one of the most relevant topics in contemporary medical microbiology (Flemming & Wuertz, 2019; Futo et al., 2021).

*B. burgdorferi* biofilms are communities of spirochetes, round bodies, and blebs encased in the extracellular matrix (Meriläinen, Herranen, Schwarzbach & Gilbert, 2015) made of alginate, calcium, eDNA, and many other molecules (Sapi et al., 2012). Alginate is a form of viscous gum which provides the bacteria with a source of nutrition and hydration, while calcium contributes to the density of the outer biofilm shell (Bernardson 2013). As in many other bacterial species, alginate and calcium associate together, forming insoluble calcium alginate (Sapi, Theophilus, Pham, Burugu & Luecke, 2016). Extracellular DNA (eDNA) adsorbs to and extends from the cell surface, promoting adhesion to abiotic surfaces through acid-base interactions (Okshevsky & Meyer, 2013). The existence of *B. burgdorferi* biofilms *in vitro* was confirmed precisely by the detection of alginate, calcium, and eDNA, which are considered typical biofilm markers (Sapi et al., 2012). Additionally, atomic force microscopy showed that structural rearrangements occur at different stages of biofilm development and that channel-like structures are present in *B. burgdorferi* biofilms (Sapi et al., 2012). In comparison, these features are a signature of a true developmental process in *Bacillus subtilis*, a well-established biofilm model (Futo et al., 2021).

Genes governing bacterial biofilm formation are generally involved in adhesion, quorum sensing, cell wall synthesis, metabolism, stress response division, and motility (Jefferson, 2004). In *B. burgdorferi*, genes taking part in gene regulation and quorum sensing are experimentally established as necessary for the development of true biofilms *in vitro* (Sapi, Theophilus, Pham, Burugu & Luecke, 2016). *B. burgdorferi* cells lacking RpoN and RpoS transcriptional factors and LuxS protein included in quorum sensing fail to form robust biofilms (Sapi, Theophilus, Pham, Burugu & Luecke, 2016). Although all tree mutants form biofilm-like structures in the stationary phase of growth, these aggregates are loose, dispersed, and much smaller than the wild type (Sapi, Theophilus, Pham, Burugu & Luecke, 2016). All three mutants showed a 20-60% reduction in extracellular matrix mass and a higher sensitivity to the antibiotic doxycycline, but this effect is most prominent in mutants lacking the *luxS* gene (Sapi, Theophilus, Pham, Burugu & Luecke, 2016).

Although *B. burgdorferi* biofilms make less than 2% of overall *B. burgdorferi* cells in the exponential phase culture (Meriläinen, Herranen, Schwarzbach & Gilbert, 2015) in a physiologically relevant culturing environment, a high proportion of biofilms could be raised by increasing the cell density in the culture (Srivastava & de Silva, 2009). *B. burgdorferi* biofilms are also observed *in vivo* in the cerebral cortex of patients with Lyme neuroborreliosis (Miklossy et al., 2008; Sapi et al., 2019), skin biopsies isolated from patients with lymphocytomas after a tick bite (Aberer, Kersten, Klade, Poitschek & Jurecka, 1996; Sapi et al., 2016), and heart, liver and kidney tissue of Lyme disease infected patients (Sapi et al., 2019). Their high prevalence in tissues affected in the later phases of Lyme disease hints at the relevance of biofilms in tissue colonization (Sapi et al., 2019). Some authors even propose that the epithelial cell-associated network of non-motile *B. burgdorferi* cells progressing through the nymphal midgut during blood-feeding are indeed *B. burgdorferi* biofilms which form as a response to changes in temperature and pH introduced by the inflow of mammalian blood (Dunham-Els et al. 2012, Rudenko, Golovchenko, Kybicova & Vancova, 2019).

While free-floating, planktonic forms or bacteria are usually associated with acute infections (Bernardson 2013), *B. burgdorferi* biofilms are often considered the causative agents of CLD (Di Domenico et al., 2018). Their *in vivo* presence (Miklossy et al., 2008; Sapi et al., 2019; Aberer, Kersten, Klade, Poitschek & Jurecka, 1996; Sapi et al., 2016), antibiotic resistance (Feng, Zhang, Shi & Zhang, 2016; Feng et al., 2018), and general biofilm characteristics (Meriläinen, Herranen, Schwarzbach & Gilbert, 2015; Sapi et al., 2012) are indicators of their role in bacterial dissemination, persistence and immune evasion (Bernardson 2013). Biofilms are an advantageous residing location for *B. burgdorferi* cells, allowing them to avoid phagocytosis (Meriläinen, Herranen, Schwarzbach & Gilbert, 2015) and remain less visible to host immune defense (Bernardson 2013). The presence of collagen-like proteins in the extracellular matrix may enhance the binding of bacteria to mammalian tissues (Meriläinen, Herranen, Schwarzbach & Gilbert, 2015) or contribute to the successful transmission from the tick to the vertebrate host (Grothe, 2019).

*B. burgdorferi* biofilms have been shown to accommodate persister cells (Feng, Zhang, Shi & Zhang, 2016; Rudenko, Golovchenko, Kybicova & Vancova, 2019; Strnad, Grubhoffer & Rego, 2020). Incomplete penetration of certain antibiotics deep inside the matrix and inactivation of antibiotics by altering the microenvironment within the biofilm (Song, Duperthuy & Wai, 2016) may result in increased biofilm resistance to various types of antibiotics (Sapi et al., 2011). Significant killing was shown by doxycycline, amoxicillin, tigecycline, metronidazole, tinidazole (Sapi et al., 2011), and pulse-dosed ceftriaxone (Feng, Zhang, Shi & Zhang, 2016) in the case of spirochetes and round bodies, but neither one of the studied drugs was able to reduce biofilm formation by more than 55%. Complete eradication of *B. burgdorferi* biofilms *in vitro* was confirmed for a combination of daptomycin, doxycycline, and cefoperazone, but clinical applications of this drug combination remain to be validated (Feng, Zhang, Shi & Zhang, 2016).

### 1.8.4. Medical relevance of *Borrelia burgdorferi* morphotypes

Although the role of alternative pleomorphic variants did not gain universal recognition by the general scientific community (Lantos, Auwaerter & Wormser, 2014; Onwuamaegbu, Belcher & Soare, 2005; Schnell et al., 2014), there is a growing number of research papers supporting the clinical relevance of round bodies, blebs, and biofilms in the progression of Lyme disease (Bamm, Ko, Mainprize, Sanderson & Wills, 2019; Di Domenico et al., 2018; Margulis, Maniotis & MacAllister, 2009; Meriläinen, Herranen, Schwarzbach & Gilbert, 2015; Meriläinen, Brander, Herranen, Schwarzbach & Gilbert, 2016; Bernardtson 2013; Vancová et al., 2017; Rudenko, Golovchenko, Kybicova & Vancova, 2019; Sapi et al., 2019). Round bodies and biofilms are persisters in *in vitro* culturing conditions (Feng, Zhang, Shi & Zhang, 2016; Rudenko, Golovchenko, Kybicova & Vancova, 2019; Strnad, Grubhoffer & Rego, 2020; Feng, Shi, Zhang & Zhang, 2015; Sapi et al., 2011; Bamm, Ko, Mainprize, Sanderson & Wills, 2019; Rudenko, Golovchenko, Kybicova & Vancova, 2019; Alban, Johnson & Nelson, 2010), while all three alternative morphotypes have been found in various human tissues such as cerebral cortex (Miklossy et al., 2008; Sapi et al., 2019), skin (Aberer, Kersten, Klade, Poitschek & Jurecka, 1996; Sapi et al., 2016; Hulínská et al., 1994; Kersten, Poitschek, Rauch & Aberer, 1995), heart (Sapi et al., 2019), liver (Sapi et al., 2019) and kidney (Sapi et al., 2019).

Antibiotics commonly used in Lyme disease treatment are showing reduced activity against alternative morphotypes (Sapi et al., 2011). Moreover, most of them induce round body, bleb, and biofilm formation *in vitro* (Kersten, Poitschek, Rauch & Aberer, 1995; Feng, Zhang, Shi & Zhang, 2016; Sapi et al., 2011; Feng et al., 2018). The formation of different pleomorphic variants could potentially explain the persistence of *B. burgdorferi* infection or the presence of unusual symptoms of PTLDS and CLD (Kersten, Poitschek, Rauch & Aberer, 1995; Rudenko, Golovchenko, Kybicova & Vancova, 2019; Sapi et al., 2016).

Spirochetes, round bodies, blebs, and biofilms are present simultaneously in morphologically heterogenous *B. burgdorferi* cultures (Meriläinen, Herranen, Schwarzbach & Gilbert, 2015). Furthermore, the proportion of each pleomorphic variant depends on culturing conditions (Meriläinen, Herranen, Schwarzbach & Gilbert, 2015), indicating there is a connection between cell morphology and cellular habitat. Since the change in environmental conditions is a feature of the *B. burgdorferi* life cycle (Drecktrah et al., 2015), morphotypes may play a role in adapting bacteria to various stages of transmission and infection.

Although the role of each morphotype is not precisely known, several papers suggest they may modulate the immune response (Meriläinen, Brander, Herranen, Schwarzbach & Gilbert, 2016; Whitmire & Garon, 1993) and promote dissemination inside the human host (Meriläinen, Herranen, Schwarzbach & Gilbert, 2015; Toledo, Coleman, Kuhlow, Crowley & Benach, 2012; Shoberg & Thomas, 1993). Unfortunately, a fundamental lack of understanding of the molecular mechanism governing morphotype formation resulted in the inability to decisively differentiate viable alternative morphotypes from debris and non-specific staining (Bamm, Ko, Mainprize, Sanderson & Wills, 2019). Besides, traditional micrograph analysis protocols may be optimized for detecting spirochetes at the expense of alternative morphotypes (Bamm, Ko, Mainprize, Sanderson & Wills, 2019). Considering that, further biochemical and functional characterization of *B. burgdorferi* morphotypes could contribute to the development of improved diagnostic techniques.

## 1.9. Open questions

Despite accumulated evidence that *B. burgdorferi* pleomorphic forms are a biological reality, their role in the enzootic cycle and pathogenesis of Lyme disease remains unclear (Lantos, Auwaerter & Wormser, 2014; Bamm, Ko, Mainprize, Sanderson & Wills, 2019). Previous studies directed toward the biological characterization of spirochetes, round body, bleb, and biofilm *B. burgdorferi* morphotypes examined morphotype-specific induction methods, the share of each morphotype in heterogeneous *B. burgdorferi* cultures, viability and antibiotics sensitivity, morphological features of the membrane, and protein content by two-dimensional gel electrophoresis (Meriläinen, Herranen, Schwarzbach & Gilbert, 2015; Al-Robaiy et al., 2010; Alban, Johnson & Nelson, 2010; Drecktrah et al., 2015; Sapi et al., 2012; Vancová et al., 2017; Sapi, Theophilus, Pham, Burugu & Luecke, 2016; Toledo, Coleman, Kuhlow, Crowley & Benach, 2012; Malge et al., 2018). However, along with a vast majority of other cellular and molecular features, global gene expression analyses of *B. burgdorferi* morphotypes are essentially non-existent, except for protein profiling of spirochetes and round bodies by two-dimensional gel electrophoresis (Meriläinen, Brander, Herranen, Schwarzbach & Gilbert, 2016; Alban, Johnson & Nelson, 2010).

# 2. Aims and hypothesis

The main goal of this study was to address the void in the understanding of transcriptional changes associated with *B. burgdorferi* pleomorphic forms. The working hypothesis adopted as the basis for this research is that *B. burgdorferi* pleomorphic forms show differences at the transcriptome level.

To accomplish this general goal, I define several specific aims:

1) To explore differences in gene expression between *B. burgdorferi* morphotypes at the global level;

2) To explore functional differences among differentially expressed genes in *B. burgdorferi* morphotypes;

3) To explore the genomic localization of differentially expressed genes in *B. burgdorferi* morphotypes;

4) To explore evolutionary imprints in the differentially expressed genes of *B. burgdorferi* morphotypes using the phylostratigraphic approach;

5) To explore the expression profiles of *B. burgdorferi* genes known to be involved in zoonotic life cycle regulation and persistence in the mammalian host.

# 3. Materials and Methods

Culturing of *B. burgdorferi* morphotype rich culture, imaging, and RNA isolatiolation were performed by my collaborators at the BCA Clinic located in Augsburg, Germany. I organized the RNA sequencing and performed the mapping of RNA sequences, quantification of mapped reads, transcriptome data analysis, phylostratigraphic analysis, enrichment analysis, and all other bioinformatic analyses in this work.

## 3.1. Culturing conditions

BSK-H medium containing 6% rabbit serum (bio&sell, Germany, Feucht) was used for culturing *Borrelia burgdorferi* B31 (DSMZ, Germany, Brunswick) precultures. A preculture consists of sprichete morphotype *B. burgdorferi* cells. Inoculations from this preculture were added to four different culturing media to gain four specific *B. burgdorferi* morphotype cultures – spirochete, round body, bleb, and biofilm rich culture. All samples, excluding blebs, were taken in three biological replicates per morphotype, while the bleb morphotype was represented with two biological replicates. Each day cells were counted using a C-Chip Disposable Haemocytometer (Neubauer Improved system, DHC-N01, Merck Millipore/Biochrom, Germany, Berlin) and Leica DM6 B fluorescence microscope with phase-contrast (PH) setting and x 40 objective.

Spirochete-rich cultures were gained by adding 40 mL of preculture with $10^7$ cell/mL in 50 mL conical tubes with a tightly closed lid. Cultures rich with round body morphotype were raised by resuspending and incubating 5 x $10^8$ preculture cells in molecular-biology grade water for 10-30 minutes. A high percentage (80%) of bleb-containing cultures were gained by adding 6 mL of the preculture with $10^7$ cell/mL in 15 mL conical tubes with a vented lid. Finally, biofilms were induced by cultivating 5 mL of preculture with $10^8$ cell/mL. For biofilm visualization, the preculture was raised within tissue-culture dishes (Eppendorf, Germany, Hamburg), while biofilms grown in 15 mL tubes with vented lids were used for RNA extraction. All cultures, including the precultures, were raised at 37°C. Once a yield of 5 x $10^8$ bacterial cells was reached, cultures were visualized and used for RNA extraction.

## 3.2. Imaging of pleomorphic forms

For visualization (Fig. 2) of different pleomorphic forms, a 10 µl sample from each culture tube or well was prepared on a microscope slide and imaged using a Leica DM6 B fluorescence microscope with PH setting and x 40 objective (400 x magnification).

## 3.3. RNA extraction

Aliquots of 5 x $10^8$ *B. burgdorferi* cells were harvested (5,000 x g, 5 min) for RNA extraction. The cell concentration was determined by counting shortly before harvesting. The cell pellets were resuspended in 300 µL peqGOLD TriFastTM reagent (VWR Peqlab, Germany, Darmstadt) and either directly processed or frozen at -20 °C. Direct-zolTM RNA Miniprep Plus Kit (Zymo Research, Germany, Freiburg) was used to extract and process RNA samples. After loading the samples, an on-column DNA digest was performed with the RNase-free DNase set (Qiagen, Germany, Hilden). The RNA was diluted in RNAse-free water and stored at -80°C. The RNA quantity was measured spectroscopically, and the integrity was assessed by agarose gel electrophoresis.

## 3.4. RNA sequencing

Ribo-Zero rRNA Removal Kit (Illumina) was used for Ribosomal RNA removal from the total RNA samples. Illumina TruSeq RNA Sample Preparation v2 Kit (Illumina) was used for RNA-seq libraries preparation. Additionally, Illumina NextSeq 500 platform at the EMBL Genomics Core Facility (Heidelberg, Germany) for bidirectional RNA sequencing generated 450 million reads per run. Using BBMap V37.66 (Bushnell, 2014), 927,047,716 paired-end sequences (75 bp) were mapped onto the *B. burgdorferi* reference genome (NCBI Assembly accession: ASM868v2; GCF_000008685.2) with an average of 94.32% mapped reads per sample (supplementary data Table S1). On average, 84 million reads per replicate were mapped with low variation between the samples (supplementary data Table S1). The mapping was performed using the standard settings with the option of trimming the read names after the first whitespace was

enabled. For generating, sorting, and indexing BAM files, SAMtools package V2.0.3 (Li et al., 2009) was used, and downstream RNAseq data processing was analyzed in R V3.6.0 (R Development Core Team, 2019) using custom-made scripts. Rsamtools package V2.0.3 (Morgan, Pagès, Obenchain & Hayden, 2019) was used for mapped reads quantification per each *B. burgdorferi* open reading frame, and GenomicAlignments R package V1.20.1 (Lawrence et al., 2013) was used for retrieving raw counts for 1544 open reading frames. Expression similarity across morphotypes and replicates was assessed using principal component analysis (PCA) (Fig. 3a) implemented in the R package DESeq2 V1.24.0 (Love, Huber & Anders, 2014) and visualized using the R package ggplot2 V3.3.2 (Wickham, 2016) (Fig. 3a).

## 3.5. Transcriptome data analyses

The genome of *B. burgdorferi* consists of 1,544 genes, of which 1,347 are coding for proteins. In total, 1,344 protein coding genes passed the phylostratigraphic procedure (Table 4). Raw counts of 1,544 genes were normalized by calculating the fraction of transcripts based on feature-length and sequencing depth (Li, Ruotti, Stewart, Thomson & Dewey, 2010), and replicates were resolved by calculating the median of all nonzero transcription values. To generate more comparable gene expression profiles, genes that had zero expression values in two or more morphotypes were discarded, and gene expression profiles were brought to the same scale by normalization to the median and $\log_2$ transformation of obtained values (standardized expression values). Afterward, genes were clustered based on standardized expression values per morphotype using the DP_GP_cluster (McDowell et al., 2018), with the maximum Gibbs sampling iterations set to 500 (supplementary data Table S3). Additionally, average standardized gene expression per morphotype was calculated for each gene cluster. Standardized gene expression values and their average in morphotypes were visualized for each cluster by using the R ggplot2 package V3.3.2 (Fig. 11). The statistical significance of expressional changes in four different morphotypes per gene was assessed by the Likelihood Ratio Test (LRT) implemented in the DESeq2 V1.24.0 package. The expression profile of each differentially expressed gene was determined by the LRT test, and the expression profiles of specific differentially expressed genes and each gene

upregulated or downregulated by RpoS (supplementary data Table S36 and S37) were visualized using the R ggplot2 package V3.3.2 (Fig. 13, Fig. 14).

Pairwise differential gene expression between *B. burgdorferi* round body, bleb, and biofilm morphotype compared to spirochete morphotype was estimated using DESeq2 V1.24.0 package. Differences in expression between round body, bleb, and biofilm morphotypes compared to spirochetes were visualized by plotting the negative value of $\log_{10}$ FDR p-value in relationship to $\log_2$ fold change value (Fig. 3b, Fig. 3c, Fig. 3d) using the ggplot2 V3.3.2 package (Wickham, 2016). Two criteria were used to define which genes were counted among differentially expressed. Under permissive criteria, the FDR p-value had to be below 0.05 for a gene to be assigned as differentially expressed. Under the stringent criteria, in addition to the FRD p-value, an added cutoff value was considered. Namely, for a gene to be considered differentially expressed, the $\log_2$ fold change value had to be greater than 1 for upregulated genes and below -1 for downregulated genes. Both differential expression criteria were used in enrichment analysis.

## 3.6. Phylostratigraphic analysis

The standard phylostratigraphic procedure was performed as described previously (Domazet-Loso, Brajković & Tautz, 2007). Using the relevant phylogenetic literature (Margos et al., 2018; Parks et al., 2018; Mukherjee et al., 2017; Hug et al., 2016; Paster & Dewhirst, 2000; Raymann, Brochier-Armanet & Gribaldo, 2015; Di et al., 2014; Gupta, Mahmood & Adeolu, 2013; Rinke et al., 2013; Wu et al., 2009; Richter et al., 2006), we constructed a consensus phylogeny covering divergence from the last common ancestor of all cellular organisms to the *B. burgdorferi* (Fig. 6, supplementary data Fig. S24). Nodes were chosen based on previously mentioned phylogenetic literature, the importance of evolutionary transition, and annotation completeness estimated using Benchmarking Universal Single-Copy Orthologs (BUSCO) scores (Simão, Waterhouse, Ioannidis, Kriventseva & Zdobnov, 2015).

Protein sequences for 926 terminal taxa were retrieved from ENSEMBL (Yates et al., 2020) (719) and NCBI (24) databases and used for protein sequence database preparation. In the case of eukaryotic organisms, only the longest splicing variant per gene was used. To construct the phylostratigraphic map (Domazet-Loso, Brajković & Tautz, 2007) of *B. burgdorferi*, I compared 1347 *B. burgdorferi* proteins with the protein sequence database using the script developed by Futo et al. (2021) which implement the blastp algorithm V2.8.1 (Altschul, Gish, Miller, Myers & Lipman, 1990) and varies e-value thresholds (1, $10^{-1}$, $10^{-2}$, $10^{-3}$, $10^{-5}$, $10^{-10}$, $10^{-15}$, $10^{-20}$, $10^{-30}$; supplementary data Table S3). Proteins that did not return their own sequence as a match were discarded. The remaining protein sequences were mapped on the eight phylostrata of the consensus phylogeny using a pipeline developed by Futo et al. (2021), and the oldest phylostratum on the phylogeny where a protein still had a match was assigned to that protein (Domazet-Loso, Brajković & Tautz, 2007).

## 3.7. Enrichment analysis

1,347 protein sequences were annotated to Clusters of Orthologous Genes (COG) (Tatusov, Galperin, Natale & Koonin, 2000) by searching the eggNOG V5.0. (Huerta-Cepas et al., 2019) database using the eggNOG-mapper V2 server (Huerta-Cepas et al., 2017). Annotations were transferred from any ortholog in the Bacteria taxa (taxID:2), the minimum hit e-value was 0.001, the minimum hit bit-score was 60, and the minimum 20% of the query was covered. Functional enrichment of upregulated and downregulated round body, bleb, and biofilm protein annotations among all *B. burgdorferi* proteome annotations, and the functional enrichment of cluster-specific protein annotation among all *B. burgdorferi* proteome annotations were estimated using the two-way hypergeometric test (supplementary data Table S4-S13, Fig. 4). Enrichment of phylostratum-specific upregulated and downregulated round body, bleb, and biofilm genes that passed the phylostratigraphic procedure among all *B. burgdorferi* genes, as well as the enrichment of location-specific upregulated and downregulated round body, bleb and biofilm genes located on either the bacterial chromosome or one of 21 plasmids among all *B. burgdorferi* genes (1544 genes) were also estimated using the two-way hypergeometric test (supplementary data Table S14-S23, supplementary data Table S25-S35, Fig. 5, Fig.7., Fig. 10, Fig. 12). Enrichment of phylostratum-

specific (e-value cut off 10, 1, $10^{-1}$, $10^{-2}$, $10^{-3}$, $10^{-5}$, $10^{-10}$, $10^{-15}$, $10^{-20}$, $10^{-30}$) genes among all COG annotated genes compared to phylostratum-specific genes in the *B. burgdorferi* genome was also calculated using the two-way hypergeometric test (Fig. 8, Fig. 9). Two-way hypergeometric tests were performed by scripts adapted from those used by Futo et al. (2021).

Additional enrichment analyses were performed on genes determined by Caimano et al. (2019) to be upregulated by RNA polymerase sigma factor RpoS. Enrichments of genes upregulated or downregulated by RpoS among phylostratum-specific genes were tested by two-way hypergeometric tests (supplementary data Table S38 and S39, Fig. 16). Similarly, enrichments of genes upregulated or downregulated by RpoS among cluster-specific genes were estimated using the same test (supplementary data Table S40 and S41, Fig. 15). In all enrichment analyses, p values were adjusted for multiple comparisons using the Benjamini and Hochberg procedure (1995). Functional enrichment of COG annotations, enrichment of phylostratum-specific proteins, and enrichment of location-specific genes were visualized using custom-made scripts based on the R package ggplot2 V3.3.2 (Wickham, 2016) (Fig. 3, Fig. 4, Fig. 7, Fig. 8, Fig. 9, Fig. 10, Fig. 11, Fig. 12, Fig. 13, Fig. 14, Fig. 15, Fig. 16).

# 4. Results

## 4.1. Expression profiles of *Borrelia burgdorferi* morphotypes

To obtain transcriptome expression levels of *B. burgdorferi* B31 pleiomorphic forms, spirochete (SP), round body (RB), bleb (BL), and biofilm (BF) morphotypes were sampled (Fig. 2). When cumulatively considered, the evidence of transcription among these morphotypes was found for 1,370 (89%) predicted *B. burgdorferi* genes and for 1,306 (92%) predicted protein-coding genes. These numbers were comparable to previous transcriptomic studies in *B. burgdorferi* (Arnold et al., 2016; Wu et al., 2015; Malge et al., 2018). A principal component analysis (PCA) revealed a fairly resolved pattern where biofilm and bleb morphotypes show distinct transcriptomes compared to spirochete and round body morphotypes, which cluster together (Fig. 3a).

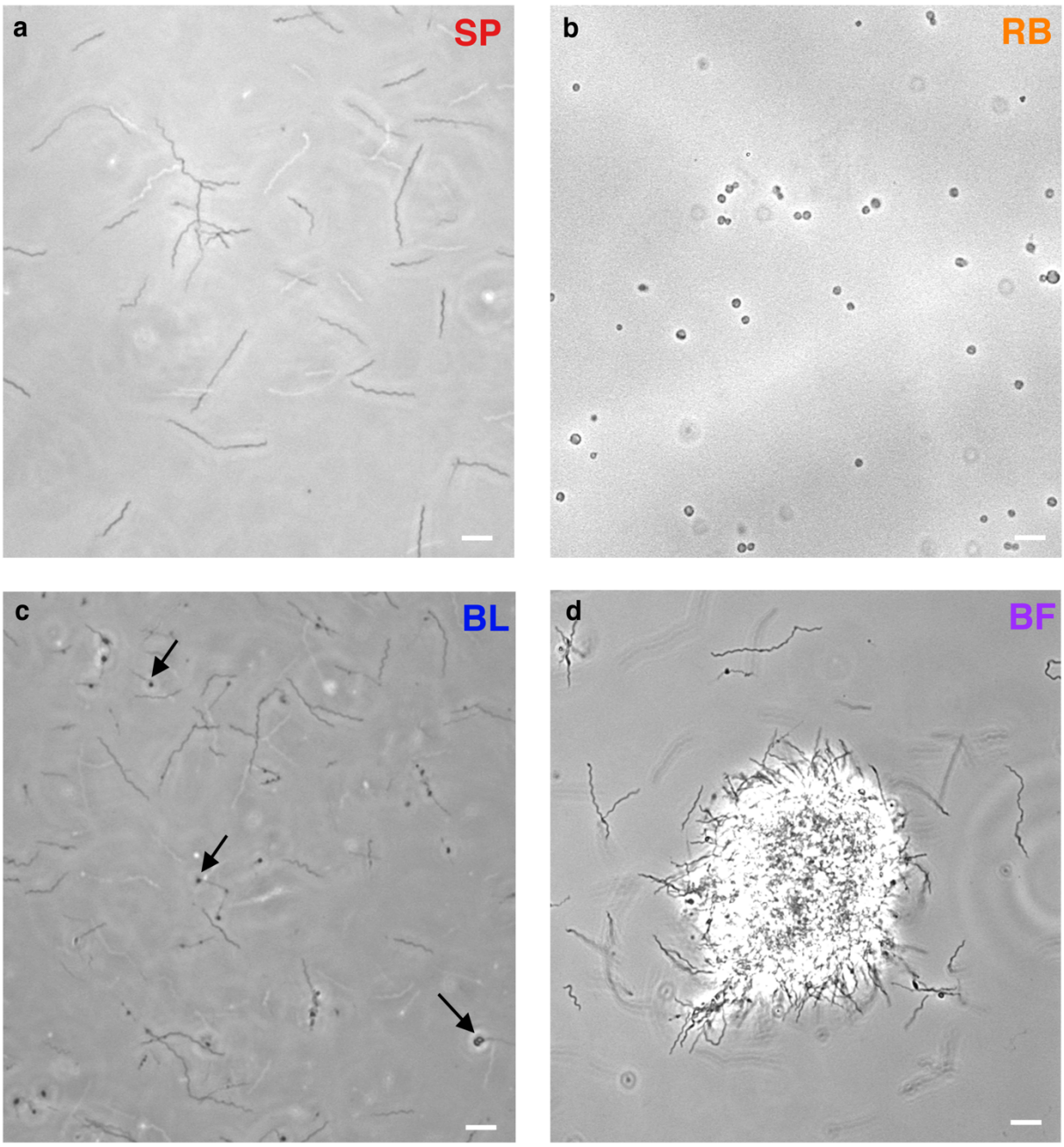



**Figure. 2. Representative images of *B. burgdorferi* B31 morphotypes.** Phase-contrast images of *B. burgdorferi* live cell cultures: **(a)** spirochetes (SP), **(b)** $H_2O$-induced round bodies (RB), **(c)** blebs on spirochetes (BL) marked by black arrows, and **(d)** biofilm (BF). White bars - 10 µm (400 x magnification). These images were taken by my collaborators from the BCA Clinic located in Augsburg, Germany.

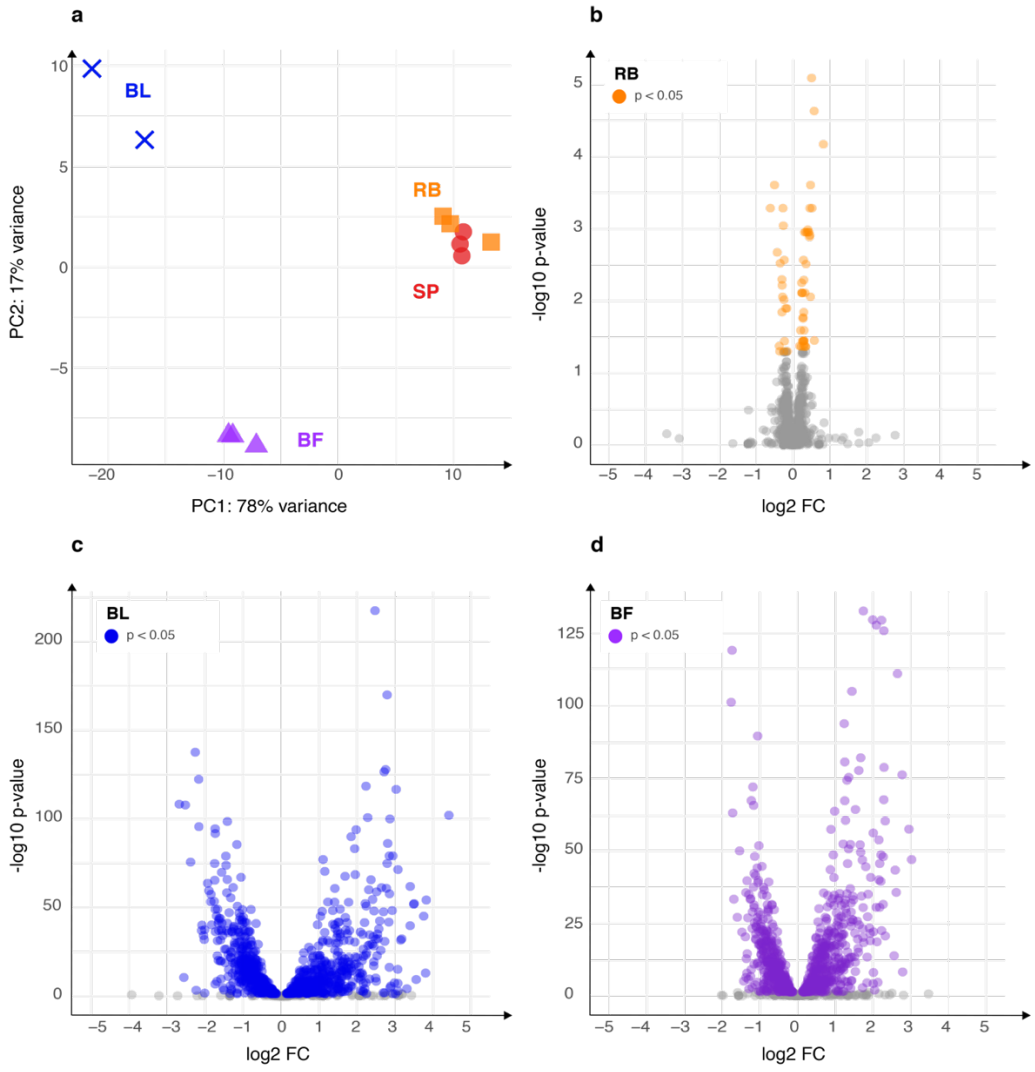



**Figure. 3. *B. burgdorferi* morphotypes are showing differential gene expression.** Spirochetes (SP) and round bodies (RB) share similar expression profiles. At the same time, the bleb (BL) morphotype and biofilms (BF) show a noticeable difference in expression, both between each other and when compared to spirals and round bodies. (**a)** Principal component analysis (PCA) of *B. burgdorferi* B31 transcriptome data. The replicates of one morphotype have the same color and symbol. **(b-d)** Volcano plots show differentially expressed genes in pairwise comparisons. (**b)** The round body (RB) morphotype is compared to spirochetes. (**c**) The bleb morphotype (BL) in comparison to spirochetes. (**d)** Biofilms (BF) in comparison to spirochetes. Significantly differentially expressed genes (p-value < 0.05) are shown in orange (RB), blue (BL), or purple (RB). Genes that are not significantly differentially expressed (p-value >= 0.05) are shown in gray. Abbreviation FC stands for fold change.

To identify differentially expressed genes, round body, bleb, and biofilm morphotypes were compared against spirochetes as a reference. These types of pairwise comparisons were chosen since all alternative morphotypes studied here were derived from spirochete cultures after implementing changes in growth conditions (see Methods). The comparison of the fold-change and p-values in volcano plots reveals that round bodies have a relatively small number of differentially transcribed genes (4.3%, Fig. 3b, Table 1). Moreover, the magnitude of fold-change for these differentially transcribed genes is below twofold (Table 1). These values reflect previous work that detected only 77 differentially expressed proteins by 2D gel electrophoresis during spirochete to round body transition (Meriläinen, Brander, Herranen, Schwarzbach & Gilbert, 2016) and suggest that the transcriptional profile of *B. burgdorferi* round bodies greatly resembles the profile observed in spirochetes.

In contrast to round bodies, a high number of differentially transcribed genes were detected in bleb (68%) and biofilm (60%) morphotypes (Fig. 3c and d, Table 1). When a more stringent criteria was applied, which considers only differentially expressed genes with the magnitude of fold-change above twofold, a substantial number of differentially expressed genes was still detectable (27% blebs, 14% biofilms, Fig. 3 c and d, Table 1). The PCA analysis of all genes (Fig. 3a) indicated that the bleb and biofilm morphotype express different transcriptomes. Based on that, a test of how many differentially expressed genes are shared between the two morphotypes was performed. The results showed that roughly 70% of differentially expressed genes in biofilms were also differentially expressed in the same direction in the bleb morphotype (Table 2 and 3). This indicates that although the overall transcriptomic profile of these two morphotypes is different from each other, blebs and biofilms share a significant proportion of differentially expressed genes.

**Table 1.** The number of differentially expressed genes among *B. burgdorferi* morphotypes. Total number ($N_{genes}$) of upregulated genes (up), downregulated genes (down), and overall differentially expressed genes (total) between round bodies (RB) and spirochetes, blebs (BL) and spirochetes and between biofilms (BF) and spirochetes based on a permissive ($p < 0.05$) and stringent ($p < 0.05$, $|\log2\ FC| > 1$) criteria.

| DE cutoff | permissive | | | stringent | | |
|---|---|---|---|---|---|---|
| $N_{genes}$ (%) | up | down | total | up | down | total |
| **RB** | 44 (2.85) | 23 (1.49) | 67 (4.34) | 0 (0) | 0 (0) | 0 (0) |
| **BL** | 529 (34.26) | 522 (33.81) | 1051 (68.07) | 274 (17.75) | 142 (9.20) | 416 (26.94) |
| **BF** | 467 (30.25) | 464 (30.01) | 931 (60.30) | 156 (10.10) | 60 (3.89) | 216 (13.99) |

**Table 2.** The percentage of differentially expressed shared between biofilm and blebs from the perspective of blebs. Genes were categorized as upregulated or downregulated based on two types of criteria. Under the permissive criteria, the adjusted p-value (p) calculated by DeSeq2 pairwise comparison analysis had to be below 0.05. Under the stringent criteria, the adjusted p-value (p) calculated by DeSeq2 pairwise comparison analysis had to be below 0.05, while the absolute value of log2 fold change calculated by DeSeq2 had to be greater than 1. Based on the number of genes that are differentially expressed in the same direction in both blebs and biofilms (shared genes), and the total number of genes which are differentially expressed in blebs (BL genes), the percentage of shared genes between biofilms and blebs was calculated. The upregulated genes (up) and downregulated genes (down) are considered separately.

| criteria | permissive | | | stringent | | |
|---|---|---|---|---|---|---|
| | up | down | total | up | down | total |
| N (shared genes) | 329 | 360 | 689 | 117 | 34 | 151 |
| N (BL genes) | 529 | 522 | 1051 | 274 | 142 | 416 |
| % shared genes | 62 | 69 | 66 | 43 | 24 | 36 |

**Table 3.** The percentage of differentially expressed shared between biofilm and blebs from the perspective of biofilms. Genes were categorized as upregulated or downregulated based on two types of criteria. Under the permissive criteria, the adjusted p-value (p) calculated by DeSeq2 pairwise comparison analysis had to be below 0.05. Under the stringent criteria, the adjusted p-value (p) calculated by DeSeq2 pairwise comparison analysis had to be below 0.05, while the absolute value of log2 fold change calculated by DeSeq2 had to be greater than 1. Based on the number of genes that are differentially expressed in the same direction in both blebs and biofilms (shared genes), and the total number of genes which are differentially expressed in biofilms (BF genes), the percentage of shared genes between biofilms and blebs was calculated. The upregulated genes (up) and downregulated genes (down) are considered separately.

| criteria | permissive | | | stringent | | |
|---|---|---|---|---|---|---|
| | up | down | total | up | down | total |
| N (shared genes) | 329 | 360 | 689 | 117 | 34 | 151 |
| N (BF genes) | 467 | 464 | 931 | 156 | 60 | 216 |
| % BF shared genes | 70 | 78 | 74 | 75 | 57 | 70 |

Based on PCA (Fig. 3a) and pairwise comparison between spirochetes and alternative morphotypes (Fig. 3a-c), it is obvious that the transcriptional profile of *B. burgdorferi* round bodies resembles the profile observed in spirochetes. On the other hand, the expression of genes in bleb and biofilm greatly differs from gene expression in spirochetes. Results obtained by both PCA and pairwise comparison reveal that while different from each other, bleb and biofilm share a significant proportion of differentially expressed genes.

## 4.2. Morphotype-specific functional enrichment

To determine the function of *B. burgdorferi* genes, each gene was paired with its associated COG annotation (see Methods). This revealed that only 631 (47%) genes have some functional annotation. Despite that, functional enrichment analysis revealed that a significant number of genes upregulated in round bodies participate in "translational, ribosomal structure, and biogenesis" function (Fig. 4, supplementary data Table S4). More precisely, 22 (47%) of genes upregulated in round bodies are paired with "translational, ribosomal structure, and biogenesis" COG annotation (supplementary data Table S4). All these genes are coding for structural components of the bacterial ribosome, although when compared with the list of constitutive riboproteins in bacteria provided by Schuwirth et al. (2005), they make only 44% of the 30S and 50S constitutive riboproteins.

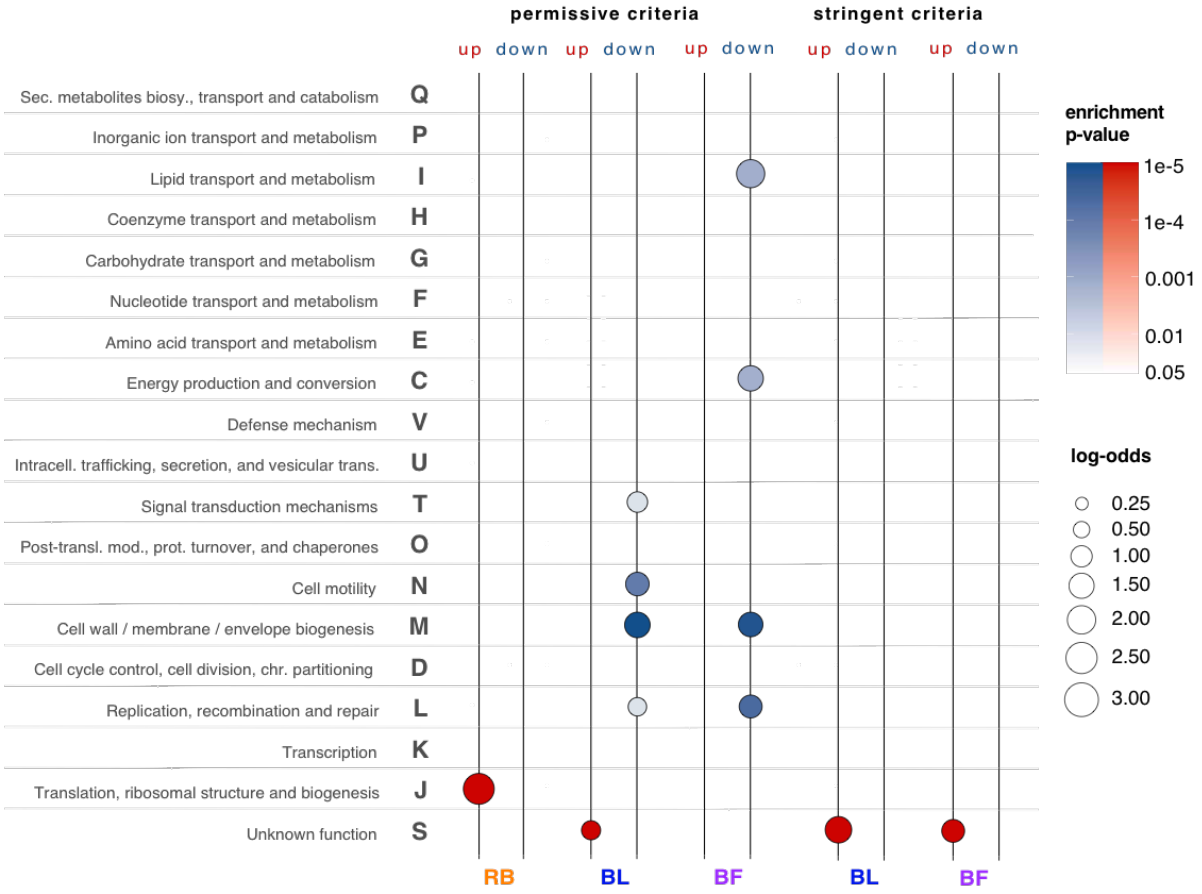



**Figure. 4. The functional enrichment analysis of differentially expressed genes in *B. burgdorferi* morphotypes.** Enrichment of COG annotations among upregulated (up) and downregulated (down) genes in the round body (RB), bleb (BL), and biofilm (BF) morphotypes was tested by two-tailed hypergeometric test corrected for multiple comparisons at 0.05 level. Differentially expressed genes were determined in reference to spirochetes using DeSeq2 pairwise comparisons. Under permissive criteria, a gene was considered differentially expressed if the shift in expression, regardless of its magnitude, was statistically significant ($p < 0.05$). Under stringent criteria, it was additionally required that the magnitude of change was at least twofold. Under the stringent criteria, there were no differentially expressed genes in round bodies, thus the enrichment analysis was not performed. Enrichment of COG functional categories was shown by log-odds. Log-odds levels were shown by circles of different sizes and p-values by color shades.

Functional enrichment analysis of genes overexpressed in blebs (Fig. 4, supplementary data Table S5) revealed a significant proportion of genes of unknown function (82%). When a more stringent criteria that considered only differentially expressed genes with the magnitude of fold-change above twofold was applied (Fig. 4, supplementary data Table S6), the number of functionally uncharacterized genes remained high (76%). Based only on permissive criteria, genes overexpressed in biofilms are not significantly enriched with any functional annotation (Fig. 4, supplementary data Table S7). However, when stringent criteria is applied by considering only genes with fold-change above twofold, enrichment with genes of unknown function turns out to be significant and made 64% of upregulated biofilm genes (Fig. 4, supplementary data Table S8).

Genes taking part in COG terms "replication/recombination/repair" and "cell wall/membrane/envelope biogenesis" functions make up a significant proportion of genes downregulated in both bleb and biofilm morphotype (Fig. 4, supplementary data Table S10-14). Additionally, blebs significantly downregulate the expression of genes taking part in "cell motility" and "signal transduction mechanisms" (Fig. 4, supplementary data Table S10, supplementary data Table S12), while genes involved in "energy production and conversion" and "lipid transport and metabolism" are noticeably downregulated in biofilms (Fig. 4, supplementary data Table S11, supplementary data Table 14). Almost 94% of bleb downregulated genes taking part in "replication, recombination and repair" and "cell wall/membrane/envelope biogenesis" function are also found among biofilm downregulated genes. Similarly, 85% of biofilm downregulated genes taking part in "replication, recombination and repair" and "cell wall/membrane/envelope biogenesis" functions are found among bleb downregulated genes. This means that blebs and biofilms have a similar downregulation profile of genes taking part in "replication, recombination and repair" and "cell wall/membrane/envelope biogenesis."

## 4.3. Genomic distribution of differentially expressed genes

Almost 45% of all *B. burgdorferi* genes and 37% of genes differentially expressed in round body, bleb, and biofilm morphotypes are located on circular or linear plasmids. Locations significantly contributing to the pool of round body, bleb and biofilm differentially expressed genes were found using the hypergeometric test (Fig. 5, supplementary data Table S14-24). Around 98% of genes upregulated in round bodies are located on the main chromosome (Fig. 5, supplementary data Table S14), emphasizing the importance of main chromosome located genes in the regulation of round body formation. On the other hand, genes expressed from the main chromosome make up a significant proportion of genes downregulated in blebs (93%) and biofilms (88%) (Fig. 5, supplementary data Table S20-24).

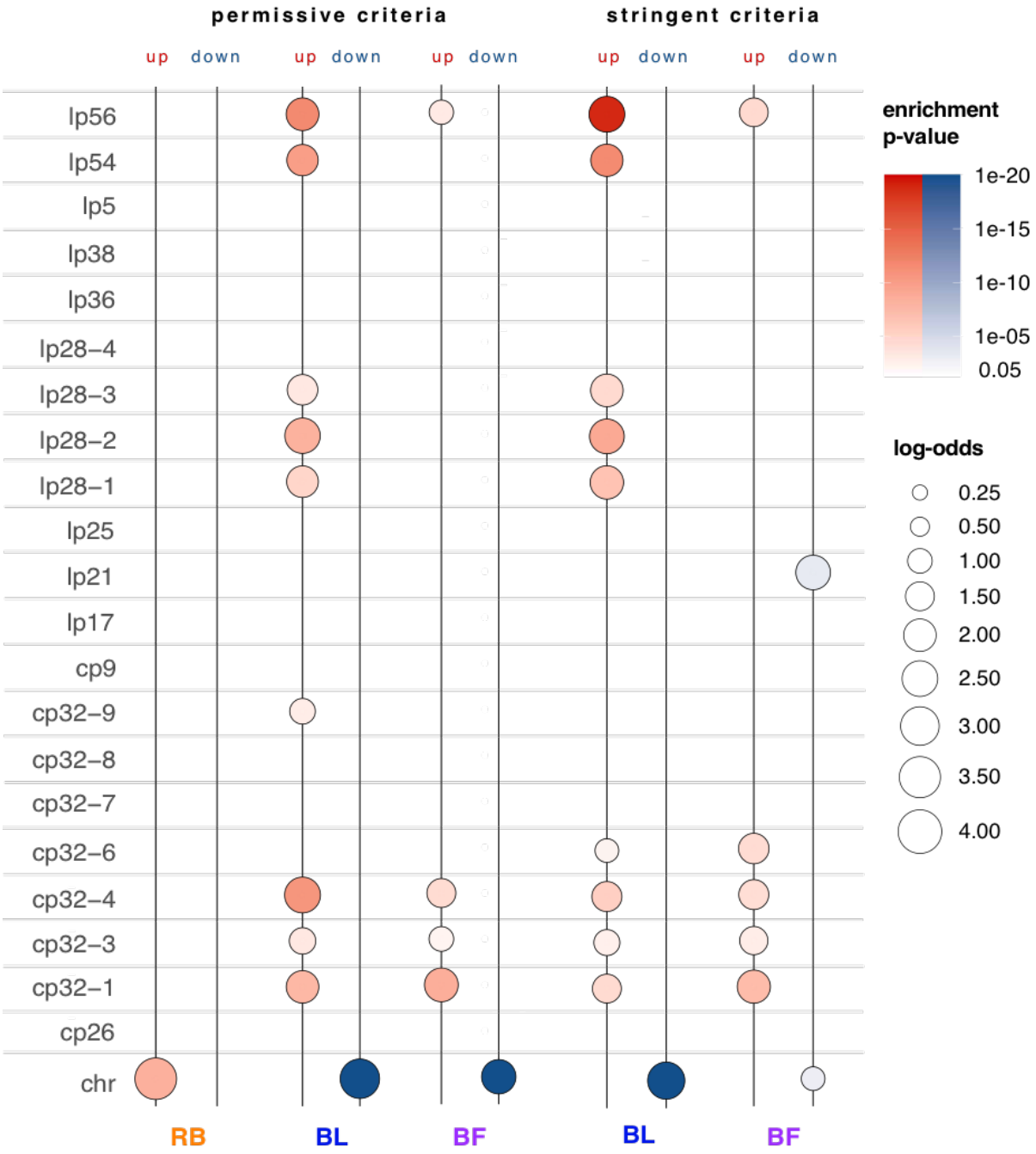

permissive criteria
stringent criteria
up
down
lp56
lp54
lp5
lp38
lp36
lp28–4
lp28–3
lp28–2
lp28–1
lp25
lp21
lp17
cp9
cp32–9
cp32–8
cp32–7
cp32–6
cp32–4
cp32–3
cp32–1
cp26
chr
RB
BL
BF
BL
BF
enrichment p-value
1e-20
1e-15
1e-10
1e-05
0.05
log-odds
0.25
0.50
1.00
1.50
2.00
2.50
3.00
3.50
4.00

**Figure. 5. The genomic location enrichment analysis of differentially expressed genes in *B. burgdorferi* morphotypes.** Frequencies of round body (RB), bleb (BL) and biofilm (BF) upregulated (up) and downregulated (down) genes across *B. burgdorferi* B31 bacterial chromosome (chr) and plasmids (cp26, cp32-1, cp32-3, cp32-4, cp32-6, cp32-7, cp32-8, cp32-9, cp9, lp17, lp21, lp25, lp28-1, lp28-2, lp28-3, lp28-4, lp36, lp38, lp5, lp54, lp56) were compared by two-tailed hypergeometric test corrected for multiple comparisons at 0.05 level. Differentially expressed genes were determined in reference to spirochetes using DeSeq2 pairwise comparisons. Under permissive criteria, a gene was considered differentially expressed if the shift in expression, regardless of its magnitude, was statistically significant ($p < 0.05$). Under stringent criteria, it was additionally required that the magnitude of change was at least twofold. Under the stringent criteria, there were no differentially expressed genes in round bodies, thus the enrichment analysis was not performed. Deviations from the expected frequencies were shown by log-odds. Log-odds levels were shown by circles of different sizes and p-values by color shades.

Out of 21 *B. burgdorferi* plasmids, 11 accommodate a significant number of genes differentially expressed in blebs or biofilms. Genes upregulated in blebs predominantly derive from plasmids lp56, lp54, lp28-1, lp28-2, lp28-3, cp32-1, cp32-3, cp32-4, cp32-6, and cp32-9 (Fig. 5, supplementary data Table S15, supplementary data Table S17), while genes upregulated in biofilms are mainly found on plasmids lp56, cp32-1, cp32-3, cp32-4 and cp32-6 (Fig. 5, supplementary data Table S16, supplementary data Table S18). Plasmids lp54, lp28-1, lp28-2, lp28-3, and cp32-9 contain a significant number of genes upregulated in blebs, but they do not contain a significant number of genes upregulated in biofilms. In contrast to that, plasmids lp56, cp32-1, cp32-3, cp32-4, and cp32-6 contain both a significant number of genes upregulated in blebs and a significant number of genes upregulated in biofilms. Overall, based on the distribution of differentially expressed genes across *B. burgdorferi* plasmids and the chromosome, we can conclude that the transition from spirochetes into round bodies is primarily associated with expressional changes of bacterial chromosome genes, while plasmid genes are significantly differentially expressed during the conversion of spirochetes into blebs or biofilms.

## 4.4. Evolutionary signatures of *Borrelia burgdorferi* morphotypes

To pinpoint the evolutionary origin of genes regulating the morphotype switch in *B. burgdorferi*, I performed the phylogenetic analysis (Domazet-Loso, Brajković & Tautz, 2007). The phylogenetic position of *B. burgdorferi* among other cellular organisms was described by defining the consensus phylogeny and eight different phylostrata (Fig. 6, supplementary data Fig. S24). Using the blastp e-value thresholds of $10^{-3}$, I used the phylostratigraphic approach to calculate the relative ages of 1,344 *B. burgdorferi* protein sequences (Table 4). Almost 42% of *B. burgdorferi* genes occupy the oldest phylostratum, while 33% of genes originated during the formation of the *Borreliaceae* phylum. The remaining 25% of genes evolved during other periods in the evolutionary history of *B. burgdorferi*.

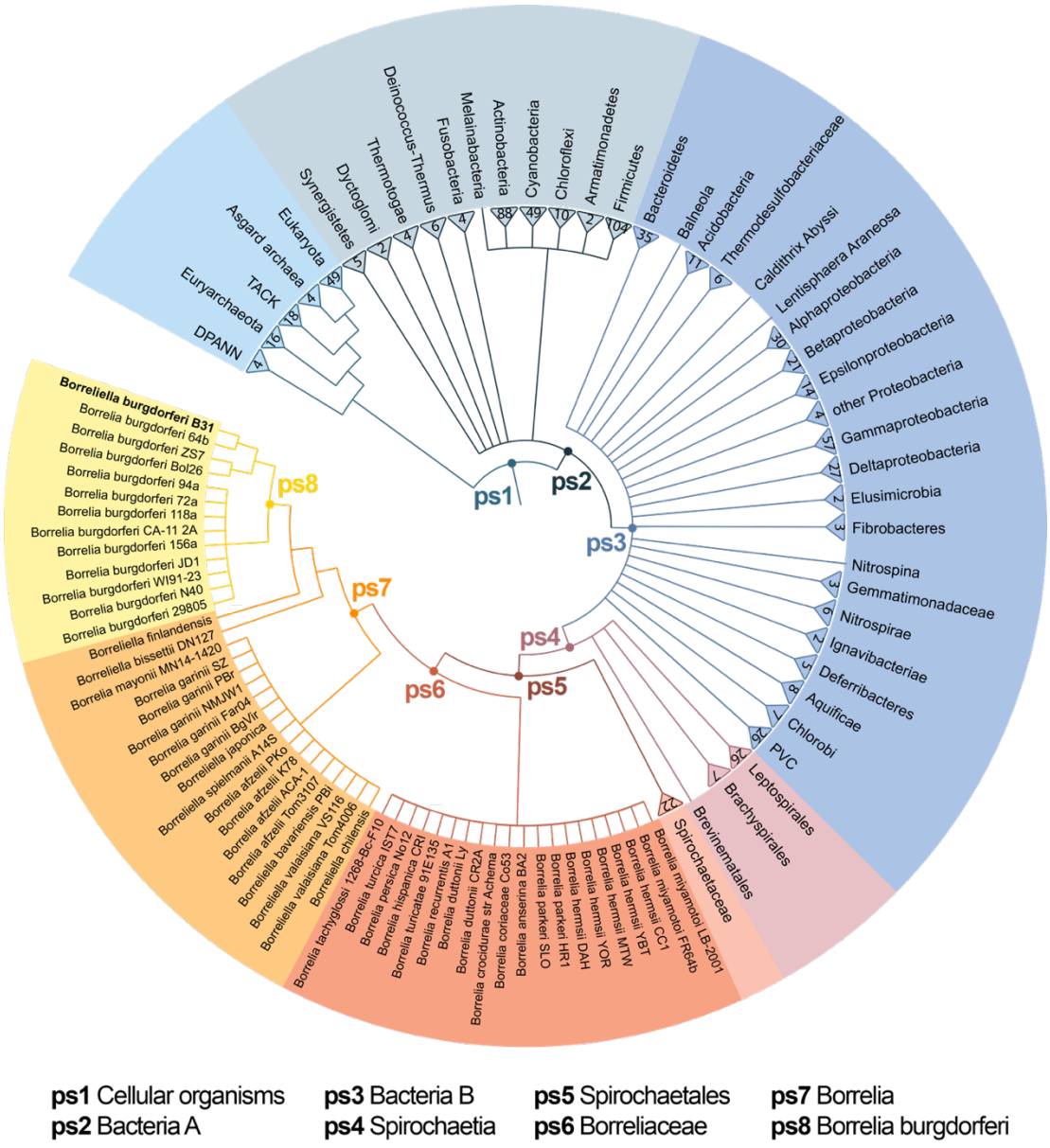



**Figure 6. The consensus phylogeny used in the phylostratigraphic analysis**. The consensus tree covers divergence from the last common ancestor of cellular organisms to *B. burgdorferi* B31 as a focal organism (see supplementary data Fig. S24 for a fully resolved tree). It is constructed based on the importance of evolutionary transitions, availability of reference genomes, and their completeness estimated using BUSCO scores. Eight phylostrata defined in the phylostratigraphic analysis are marked by ps1-ps8. Numbers at the top of terminal nodes represent the number of species in the fully resolved tree and correspond to the genomes used to populate the reference database for sequence similarity searches.

**Table 4.** Distribution of *Borrelia burgdorferi* B31 genes on the phylostratigraphic map and summary statistics, 1e-3 e-value cutoff.

| Phylostratum number | Phylostratum name | Number of genes | Percentage of genes |
|---|---|---|---|
| 1 | Cellular organisms | 557 | 41.44% |
| 2 | Bacteria A | 148 | 11.01% |
| 3 | Bacteria B | 24 | 1.79% |
| 4 | Spirochaetia | 18 | 1.34% |
| 5 | Spirochaetales | 46 | 3.42% |
| 6 | Borreliaceae | 446 | 33.18% |
| 7 | Borrelia | 88 | 6.55% |
| 8 | Borrelia burgdorferi | 17 | 1.26% |
| | **Total:** | **1344** | **100,00%** |

The evolutionary age of genes combined with differential expression data was used for calculating the enrichments of s differentially expressed genes along phylostrata. The distribution of differentially expressed genes across phylostrata (Fig. 7a, supplementary data Table S25) shows that a significant number of genes upregulated in round bodies morphotype (87%) are genes occupying the evolutionary oldest phylostratum (ps1). Contrary to that, enrichment analysis revealed that genes downregulated in round bodies are not present in any of the eight phylostrata at a significantly high number (Fig. 7a, supplementary data Table S30).

Genes upregulated in blebs (48%) preferentially originate in *Borreliaceae* (ps6) (Fig. 7a, supplementary data Table S26), while genes downregulated in blebs are enriched in Cellular organisms (ps1), Bacteria A (ps2), Bacteria B (ps3) and *Spirochaetales* (ps5) (Fig. 7a, supplementary data Table S31). When only genes with a magnitude of fold-change above twofold were considered, genes upregulated in both blebs (68%) and biofilms (63%) contained a significant number of genes originating in the *Borreliaceae* (ps6) (Fig. 7b, supplementary data Table S28 and S29).On the other hand, genes downregulated in blebs (39%) contained a significant number of genes evolved in the *Spirochaetales* phylum (ps5), while genes downregulated in biofilms aren't significantly enriched (Fig. 7b, supplementary data Table S33 and S34).

A significant number of genes upregulated in biofilms (40%) occupy the phylostratum corresponding to *Borreliaceae* (ps6) (Fig. 7a, supplementary data Table S27), while a significant number of genes downregulated in biofilms evolved in the Cellular organisms (ps1), Bacteria A (ps2) and Bacteria B (ps3) phylostrata (Fig. 7a, supplementary data Table S32). When the stringent criteria was applied, genes upregulated in biofilms (63%) included a significant number of genes evolved in the *Borreliaceae* phylum (ps6) (Fig. 7b, supplementary data Table S29), while there was no significant enrichment of phylostratum-specific genes among genes downregulated in biofilms (Fig. 7b, supplementary data Table S34).

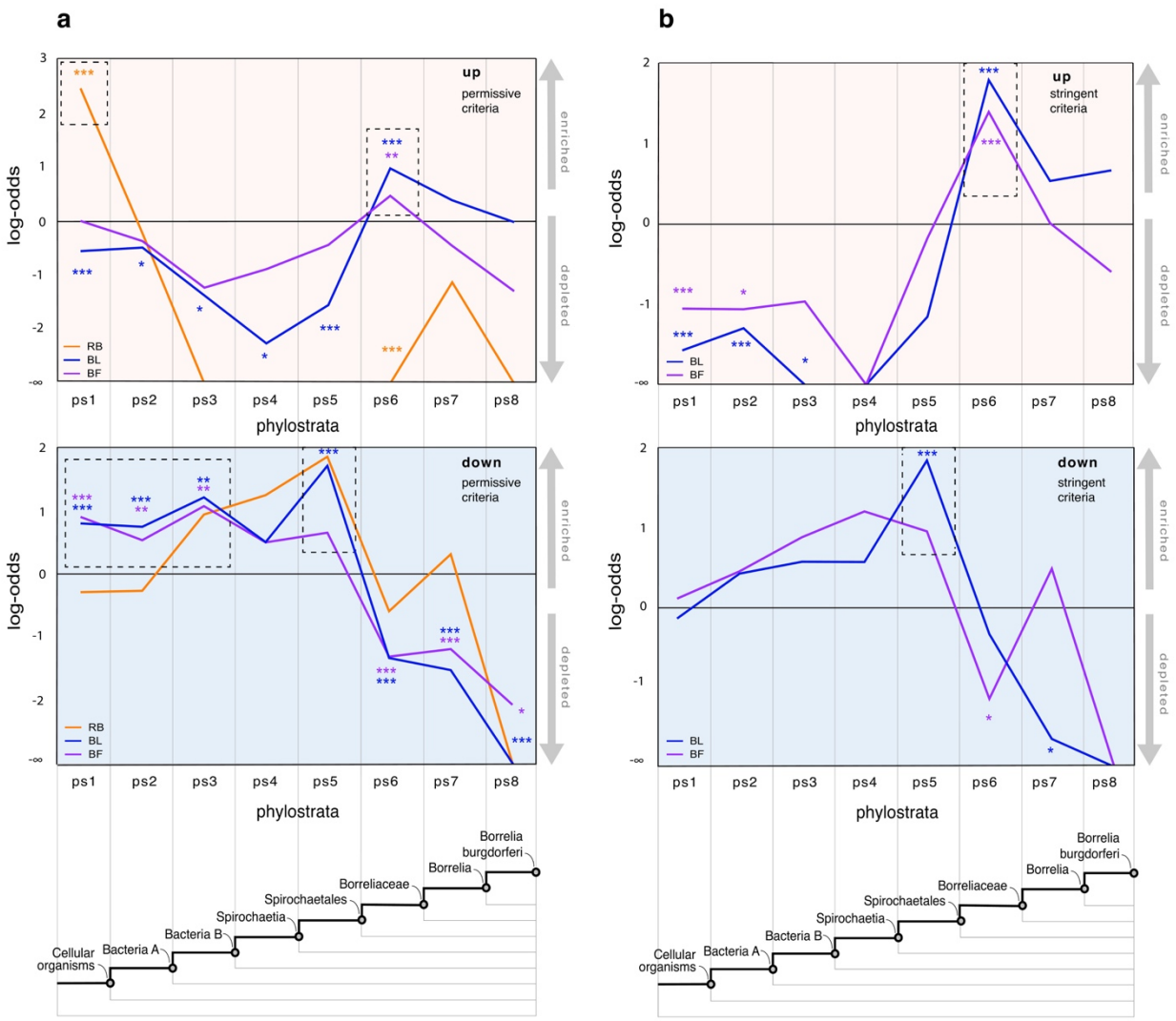

a
b
up
permissive criteria
stringent criteria
down
enriched
depleted
log-odds
RB
BL
BF
ps1
ps2
ps3
ps4
ps5
ps6
ps7
ps8
phylostrata
Cellular organisms
Bacteria A
Bacteria B
Spirochaetia
Spirochaetales
Borreliaceae
Borrelia
Borrelia burgdorferi

**Figure 7. Phylostratigraphic analysis of differentially expressed genes in *B. burgdorferi* B31 morphotypes.** A horizontal grid depicts the eight phylostrata (ps1-ps8) assigned using the e-value $10^{-3}$ BLASTp cutoff value. The frequency of phylostratum-specific genes among upregulated (up) genes (upper panel, light red background) or downregulated (down) genes (lower panel, light blue background) in the round body (RB), bleb (BL), and biofilm (BF) morphotype is compared to the frequency of phylostratum-specific genes in the complete genome and deviations are shown by log-odds (y-axis). The log-odds of zero marks that the frequency of upregulated or downregulated genes in a phylostratum equals the expected frequency estimated from the distribution of all genes across phylostrata. Differentially expressed genes were determined in reference to spirochetes using DeSeq2 pairwise comparisons. Under permissive criteria (a), a gene was considered differentially expressed if the shift in expression, regardless of its magnitude, was statistically significant ($p < 0.05$). Under stringent criteria (b), it was additionally required that the magnitude of change was at least twofold. Under the stringent criteria, there were no differentially expressed genes in round bodies, thus the enrichment analysis was not performed. Alterations from the expected frequencies were tested by a two-tailed hypergeometric test corrected for multiple comparisons at 0.05 level (*$P < 0.05$; **$P < 0.01$; ***$P < 0.001$).

To test the robustness of the phylostratigraphic approach, the analysis was repeated with a range of e-value thresholds (1, $10^{-1}$, $10^{-2}$, $10^{-5}$, $10^{-10}$, $10^{-15}$, $10^{-20}$, $10^{-30}$). Using the permissive criteria when defining differentially expressed genes, it was determined that genes occupying the oldest phylostratum (ps1) are present in a significant number among genes upregulated in round bodies, regardless of the used e-value (Fig. 8a). In contrast, the e-value threshold of $10^{-1}$ is the only one that resulted in a significant enrichment of phylostratum-specific genes (*Spirochaetales*, ps5) among genes downregulated in round bodies (Fig. 8b). Genes originating in *Borreliaceae* (ps6) consistently make a significant proportion of genes upregulated by blebs (Fig. 8c). Genes downregulated in blebs are enriched with genes evolved at the origin of life (ps1) for all used e-values, with genes originating in the Bacteria A phylum (ps2) for all e-values other than e-value 1, with genes originating in the Bacteria B (ps3) for e-value thresholds of $10^{-3}$ or less and with genes evolved in the *Spirochaetales* phylum (ps5) for all used e-value thresholds (Fig. 8d). Genes upregulated by biofilms contain a significant number of genes originating in the *Borreliaceae* phylum (ps6) for e-values of $10^{-2}$ and less (Fig. 8e), while genes downregulated in biofilms contain a significant number of genes placed at the origin of life (ps1) for all used e-values, genes evolved

in the Bacteria A phylum (ps2) for e-values $10^{-3}$ and less, genes originating in the Bacteria B (ps3) for e-value thresholds of $10^{-10}$, $10^{-5}$ $10^{-3}$ and $10^{-2}$ and with genes evolved in the *Spirochaetales* phylum (ps5) for e-value thresholds of $10^{-5}$ and less (Fig. 8f).

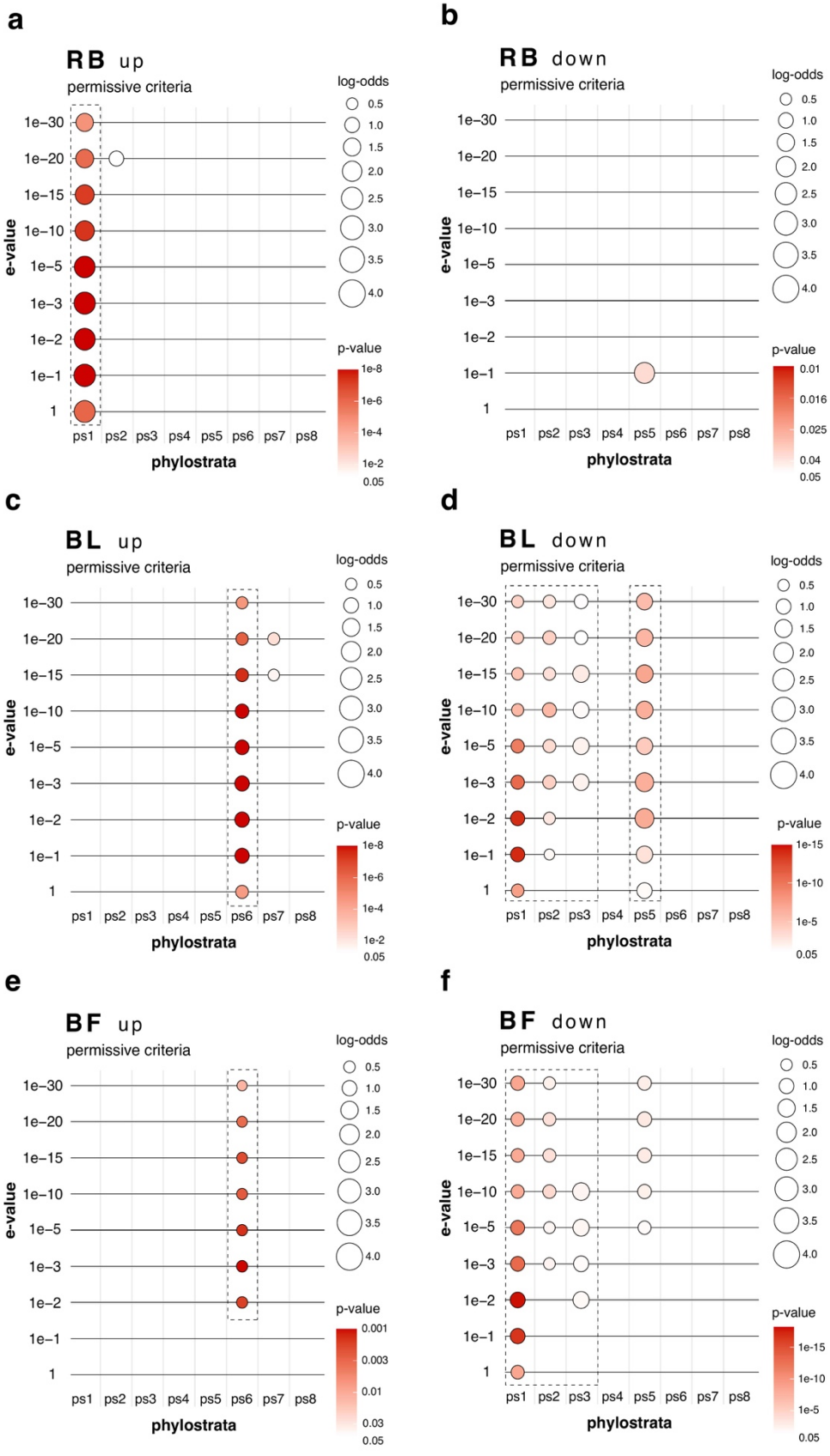

a
RB up
permissive criteria
b
RB down
permissive criteria
c
BL up
permissive criteria
d
BL down
permissive criteria
e
BF up
permissive criteria
f
BF down
permissive criteria
log-odds
p-value
e-value
1e−30
1e−20
1e−15
1e−10
1e−5
1e−3
1e−2
1e−1
1
ps1 ps2 ps3 ps4 ps5 ps6 ps7 ps8
phylostrata

**Figure 8. Robustness of phylostratigraphic analysis of differentially expressed genes in *B. burgdorferi* B31 morphotypes (permissive).** A horizontal grid depicts the eight phylostrata (ps1-ps8). The vertical grid represents used e-value thresholds (1, $10^{-1}$, $10^{-2}$, $10^{-3}$, $10^{-5}$, $10^{-10}$, $10^{-15}$, $10^{-20}$, $10^{-30}$). The frequency of phylostratum-specific genes among (a,c,e) upregulated (up) genes or (b,d,f) downregulated (down) in (a,b) round body (RB), (c,d) bleb (BL), and (e,f) biofilm (BF) morphotype is compared to the frequency of phylostratum-specific genes in the complete genome and deviations are shown by log-odds (circle size). The log-odds of zero marks that the frequency of upregulated or downregulated genes in a phylostratum equals the expected frequency estimated from the distribution of all genes across phylostrata. Differentially expressed genes were determined in reference to spirochetes using DeSeq2 pairwise comparisons and the permissive criteria. Under permissive criteria, a gene was considered differentially expressed if the shift in expression, regardless of its magnitude, was statistically significant ($p < 0.05$). Alterations from the expected frequencies were tested by a two-tailed hypergeometric test corrected for multiple comparisons at a 0.05 level. The shades of red color reflect p-values.

When the stringent criteria was used for testing the robustness of the phylostratigraphic approach, the blebs morphotype was enriched with genes originating in the *Borreliaceae* phylum (ps6) for all used e-value thresholds, with genes evolved in the *Borrelia* genus (ps7) for e-value thresholds $10^{-30}$, $10^{-20}$, $10^{-15}$ and $10^{-5}$, and with genes originating in the *Borrelia burgdorferi* species (ps8) when the used e-value was $10^{-10}$ (Fig. 9a). Genes downregulated in blebs were enriched with genes evolved in the *Spirochaetales* phylum (ps5) when e-values $10^{-10}$ and above were used and with genes originating in the *Borreliaceae* (ps6) when the e-value $10^{-20}$ was applied (Fig. 9b). Genes upregulated in biofilms contained a significant number of genes evolved in the *Borreliaceae* phylum (ps6) for all used e-values (Fig. 9c), while genes downregulated in the biofilm morphotype didn't reveal a significant enrichment in any phylostratum regardless of the applied e-value (Fig. 9d).

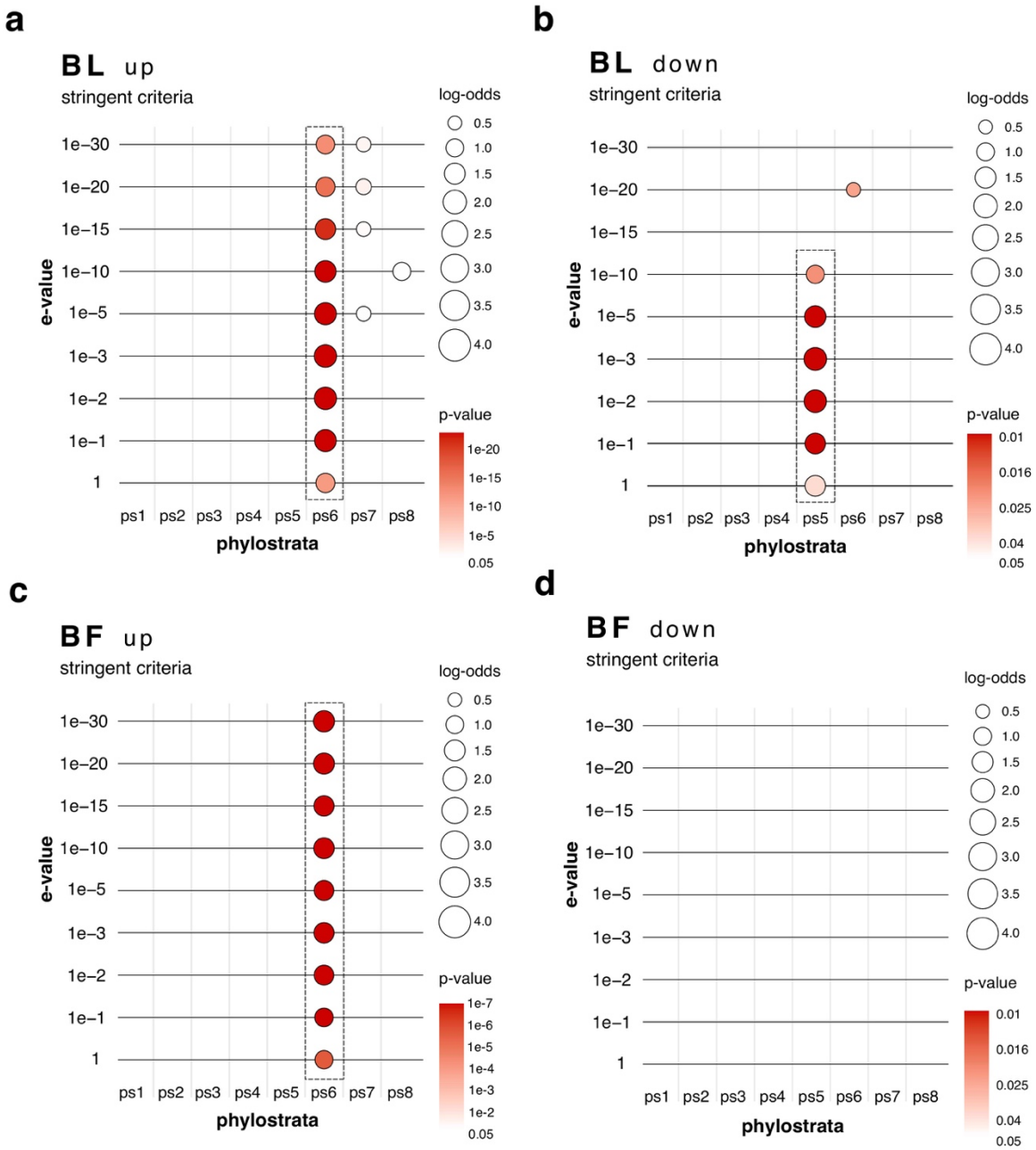

a
BL up
stringent criteria
log-odds
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
p-value
1e-20
1e-15
1e-10
1e-5
0.05
e-value
1e−30
1e−20
1e−15
1e−10
1e−5
1e−3
1e−2
1e−1
1
ps1 ps2 ps3 ps4 ps5 ps6 ps7 ps8
phylostrata
b
BL down
stringent criteria
log-odds
p-value
0.01
0.016
0.025
0.04
0.05
c
BF up
stringent criteria
1e-7
1e-6
1e-5
1e-4
1e-3
1e-2
d
BF down
stringent criteria

**Figure 9. Robustness of phylostratigraphic analysis of differentially expressed genes in *B. burgdorferi* B31 morphotypes (stringent).** A horizontal grid depicts the eight phylostrata (ps1-ps8). The vertical grid represents used e-value thresholds (1, $10^{-1}$, $10^{-2}$, $10^{-3}$, $10^{-5}$, $10^{-10}$, $10^{-15}$, $10^{-20}$, $10^{-30}$). The frequency of phylostratum-specific genes among (a,c) upregulated (up) genes or (b,d) downregulated (down) in (a,b) bleb (BL) and (c,d) biofilm (BF) morphotype is compared to the frequency of phylostratum-specific genes in the complete genome, and deviations are shown by log-odds (circle size). The log-odds of zero marks that the frequency of upregulated or downregulated genes in a phylostratum equals the expected frequency estimated from the distribution of all genes across phylostrata. Differentially expressed genes were determined in reference to spirochetes using DeSeq2 pairwise comparisons and the stringent criteria. Under stringent criteria, a gene was considered differentially expressed if the shift in expression, regardless of its magnitude, was statistically significant ($p < 0.05$) and the magnitude of change was at least twofold. Under the stringent criteria, there were no differentially expressed genes in round bodies, thus the enrichment analysis was not performed. Alterations from the expected frequencies were tested by a two-tailed hypergeometric test corrected for multiple comparisons at a 0.05 level. The shades of red color reflect p-values.

To gain information about the distribution of COG annotated *B. burgdorferi* B31 genes across phylostrata, I calculated the percentage of COG annotated genes in each phylostratum, and tested if the difference in the distribution of COG annotated genes and the distribution of COG annotated genes in the *B. burgdorferi* genome are significant (Fig. 10, supplementary data Table S35). The analysis revealed that the amount of COG annotated genes drops as we approach younger phylostrata. In accordance with that, the majority of genes originating in the first and second phylostratum are significantly well annotated. In contrast, annotation of younger genes is still lacking.

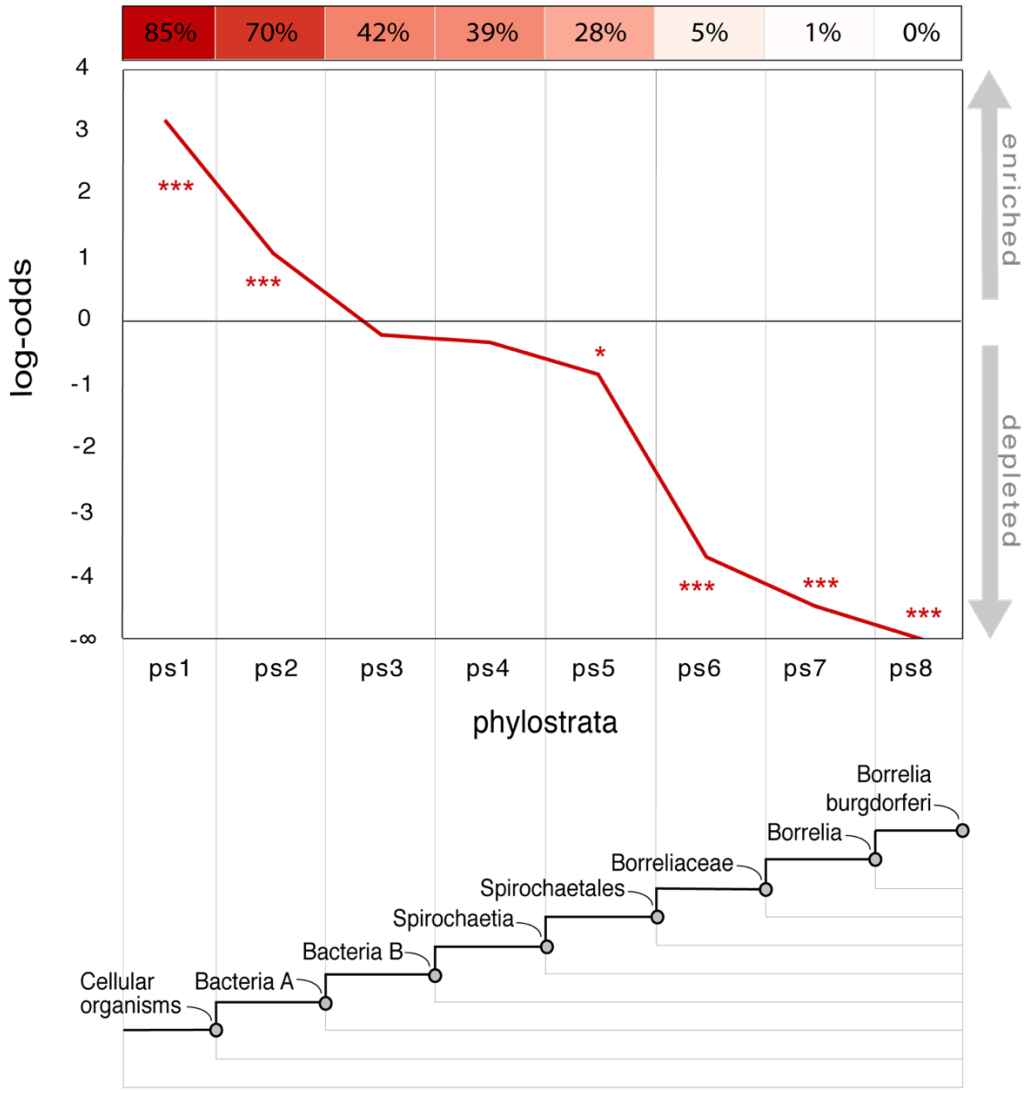



**Figure 10. The distribution of COG annotated *B. burgdorferi* B31 genes across phylostrata shows a lack of the annotation of younger genes.** A vertical grid depicts the eight phylostrata assigned using the e-value $10^{-3}$ BLASTp cutoff value corresponding to the lower panel's phylogeny. The frequency of phylostratum-specific genes in COG annotated *B. burgdorferi* B31 genes is compared to the frequency of phylostratum-specific genes in the genome. Deviations are shown by log-odds (y-axis). The log-odds of zero marks that the frequency of phylostratum-specific genes in COG annotated *B. burgdorferi* B31 genes equals the expected frequency estimated from the distribution of all genes in the genome across phylostrata. Alterations from the expected frequencies were tested by a two-tailed hypergeometric test corrected for multiple comparisons at 0.05 level (*P < 0.05; **P < 0.01; ***P < 0.001). The heatmap above the coordinate grid shows the percentage of COG annotated *B. burgdorferi* B31 genes in each phylostrata.

## 4.5. Clustering of *Borrelia burgdorferi* expression profiles

To gain insight into which genes share a similar transcription profile, I performed clustering of differentially expressed genes based on their transcription profiles across *B. burgdorferi* morphotypes, and 22 gene clusters were obtained (supplementary data Table S3, Fig. 11). These clusters vary greatly in the number of genes occupying the cluster and their expression profiles. For example, the most occupied cluster, cluster 9, contains around 13% of all *B. burgdorferi* genes, while the least populated cluster, cluster 19, contains only 0.4% of all *B. burgdorferi* genes. Also, while some clusters are well annotated, some contain a high proportion of genes of unknown function. The percentage of COG annotated genes ranges from 29.6% in cluster 11 to 90% in cluster 17. Notably, genes populating cluster 11 are more expressed in blebs and biofilms when compared to spirochetes and round bodies, while the contrary is true for cluster 17 (supplementary data Table S3, Fig. 11).

To gain insight into which functions are significantly present among genes occupying each cluster, a functional enrichment analysis of COG annotations was performed (Fig. 12). Significant enrichment of COG annotations was absent in 18 out of 22 clusters, but the remaining 6 clusters have shown a distinct distribution of significantly enriched COG functions. Genes involved in "cell wall/membrane/envelope biogenesis" are significantly present in cluster 1, while genes involved in "replication, recombination and repair" and "lipid transport and metabolism" make up a significant proportion among genes in cluster 5. Both clusters are populated with genes more expressed in spirochetes and round bodies than in biofilms and blebs. Cluster 16 is enriched with genes involved in "translation, ribosomal structure and biogenesis", with an average expression value higher in round bodies, blebs, and biofilms than in spirochetes. Lastly, clusters 3, 9, and 11 are enriched with genes of unknown functions. All these clusters contain genes more expressed in blebs and biofilms than in spirochetes and round bodies and collectively account for 27,6% of all *B. burgdrorferi* genes.

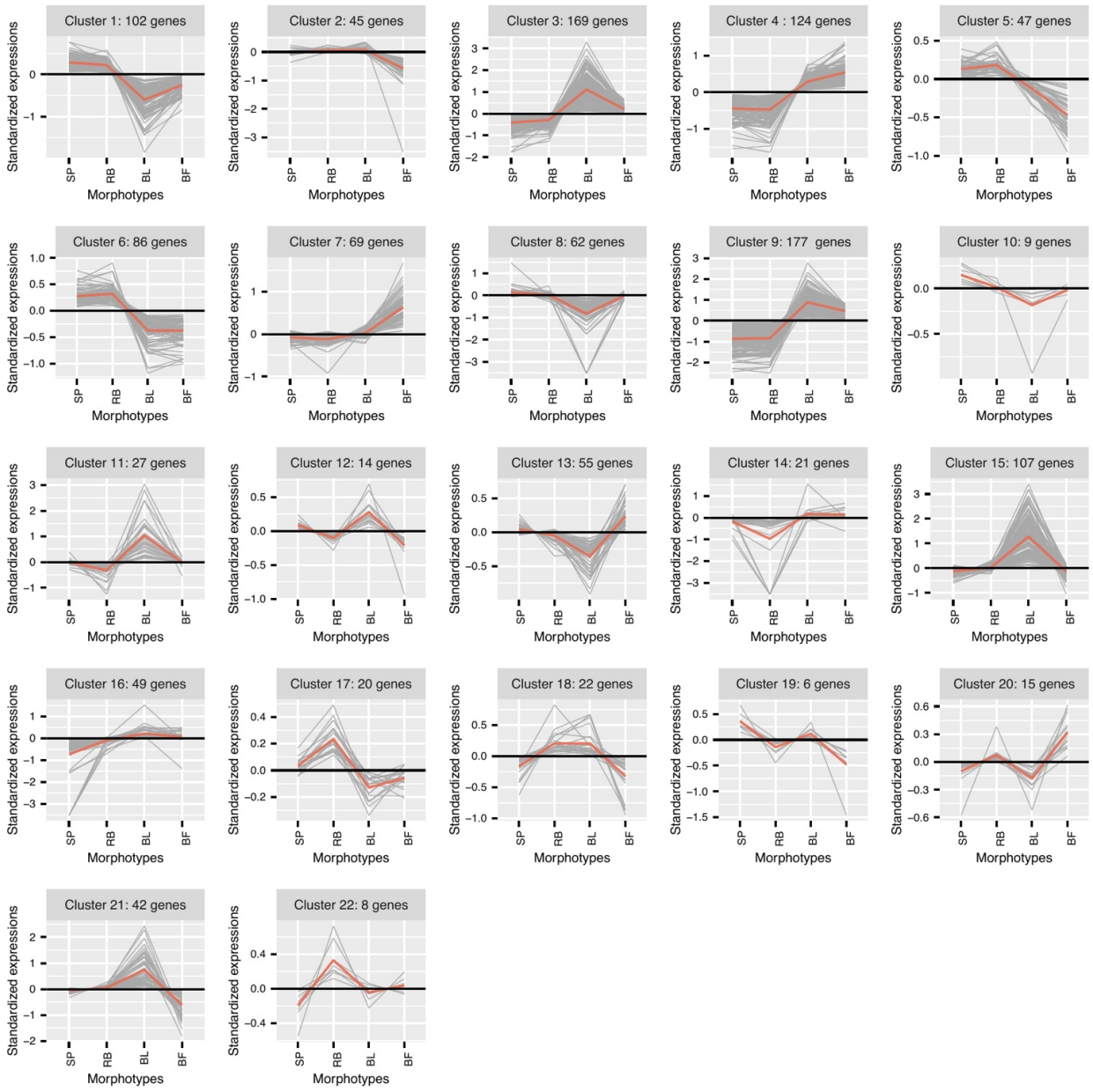



**Figure 11. Standardized expression profiles of gene clusters across *B. burgdrorferi* morphotypes.** Clusters were generated with the DP_GP_cluster algorithm at maximum Gibbs sampling iterations set to 500 based on standardized expression across *B. burgdroferi* morphotypes (see Methods). Standardized expression profiles of all genes in a cluster are shown as grey lines, while the arithmetic mean of their expression is shown in orange.

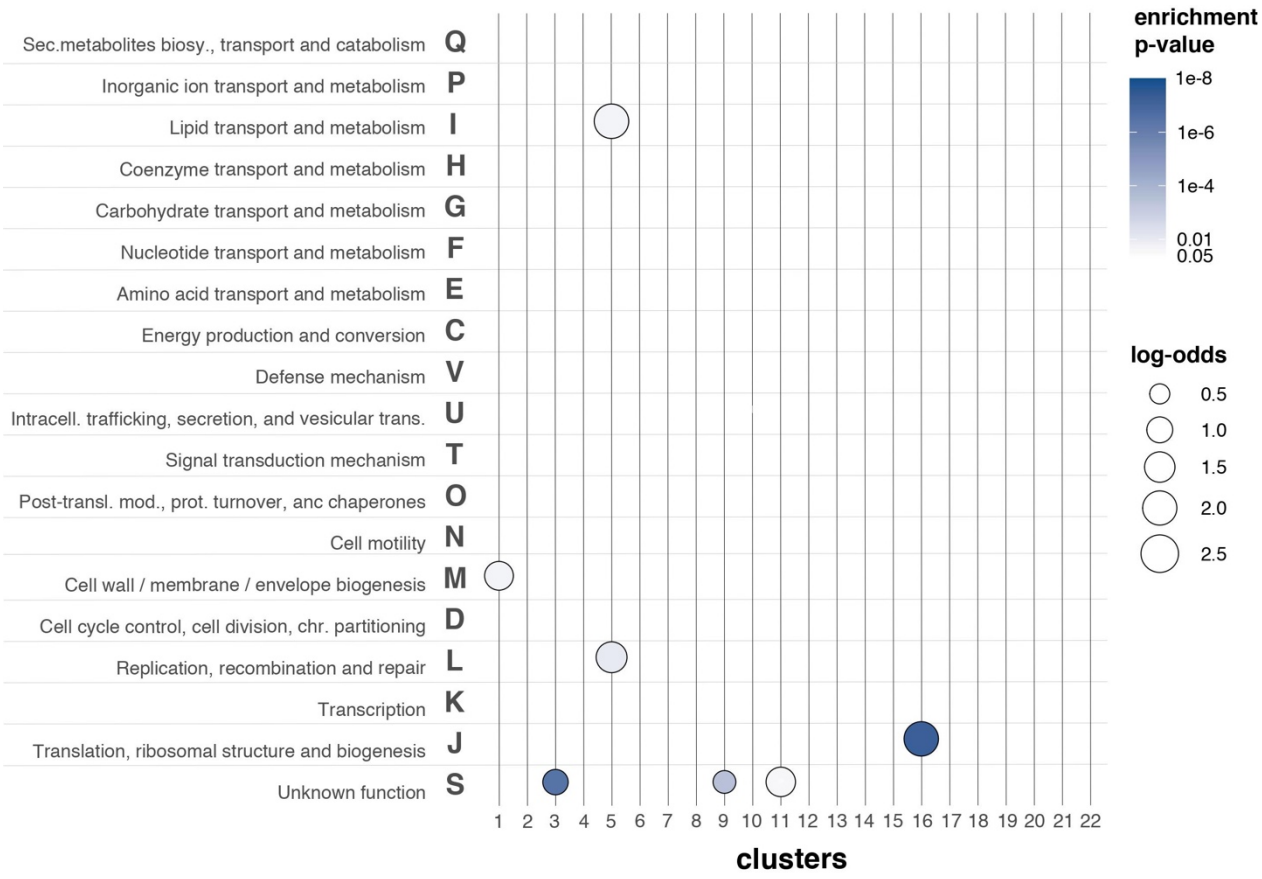



**Figure. 12. The functional enrichment analysis of genes occupying clusters with similar expression across *B. burgdorferi* morphotypes.** Clusters were generated with the DP_GP_cluster algorithm at maximum Gibbs sampling iterations set to 500 based on standardized expression across *B. burgdroferi* morphotypes. Enrichment of COG annotations among clusters was tested by two-tailed hypergeometric test corrected for multiple comparisons at 0.05 level. Enrichment of COG functional categories was shown by log-odds. Log-odds levels were shown by circles of different sizes and p-values by color shades.

## 4.6. Expression profiles of genes involved in the life cycle and persistence

To gain insight into the difference in expression of genes taking part in the zoonotic life cycle and persistence in the mammalian host in different *B. burgdorferi* morphotypes, standardized gene expressions across morphotypes were shown by expression profiles, and the significance of expressional difference across spirochete, round body, bleb, and biofilm morphotypes were calculated by LRT analysis for each gene (Fig. 13). Based on the LRT analysis, it is evident that genes coding for outer surface protein OspC, RNA polymerase sigma factor-54 (RpoN), variable surface antigen (VlsE), decorin-binding proteins (DbpA and DbpB), S-ribosylhomocisteine lyase (LuxS), fibronectin-binding protein (RevA) and plasminogen-binding proteins (ErpM and ErpY) have significantly different expression across different *B. burgdorferi* morphotypes. On the other hand, the expression of proteins coding for outer surface lipoproteins OspA and OspB, and RNA polymerase sigma factor RpoS is not significantly different between spirochetes, round bodies, blebs, and biofilms.

Genes coding for proteins VlsE, DbpA, DbpB, RevA, and ErpY are more expressed in blebs than in any other morphotype. Additionally, their expression is reduced in spirochetes and round bodies when compared to biofilms. On the other hand, the *ospC* gene reaches the maximum expression in spirochetes, while it is the least expressed in the bleb morphotype. Gene *luxS* is highly expressed in spirochetes and round bodies and has reduced expression in blebs and biofilms. The opposite is true in the case of gene ErpM, which is more expressed in blebs and biofilms than in spirochetes and round bodies. Overall, some genes involved in the enzootic life cycle and persistence of *B. burgdorferi* are differentially expressed between spirochetes, round bodies, blebs, and biofilms.

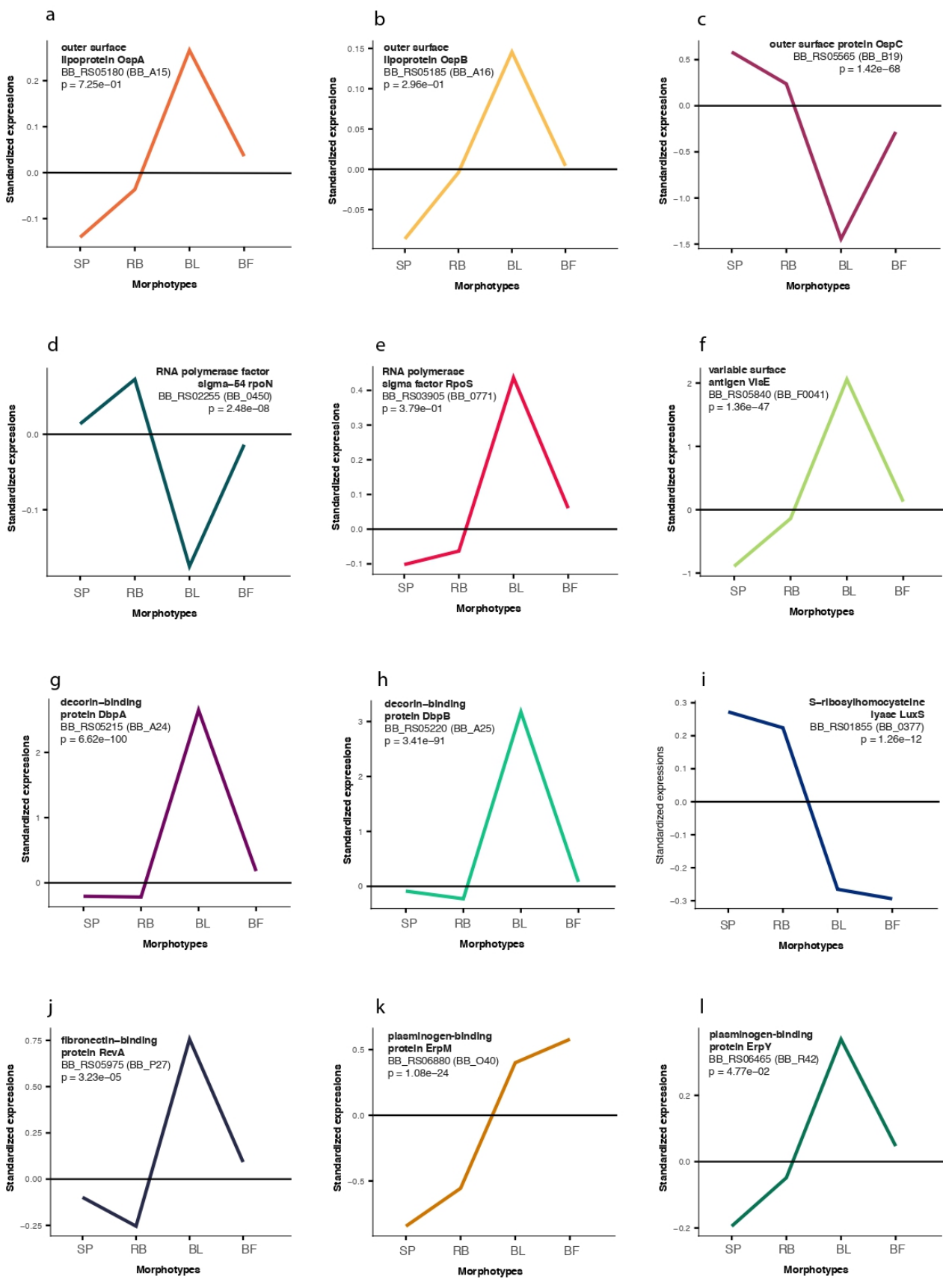

a
outer surface lipoprotein OspA
BB_RS05180 (BB_A15)
p = 7.25e−01
b
outer surface lipoprotein OspB
BB_RS05185 (BB_A16)
p = 2.96e−01
c
outer surface protein OspC
BB_RS05565 (BB_B19)
p = 1.42e−68
d
RNA polymerase factor sigma−54 rpoN
BB_RS02255 (BB_0450)
p = 2.48e−08
e
RNA polymerase sigma factor RpoS
BB_RS03905 (BB_0771)
p = 3.79e−01
f
variable surface antigen VlsE
BB_RS05840 (BB_F0041)
p = 1.36e−47
g
decorin−binding protein DbpA
BB_RS05215 (BB_A24)
p = 6.62e−100
h
decorin−binding protein DbpB
BB_RS05220 (BB_A25)
p = 3.41e−91
i
S−ribosylhomocysteine lyase LuxS
BB_RS01855 (BB_0377)
p = 1.26e−12
j
fibronectin−binding protein RevA
BB_RS05975 (BB_P27)
p = 3.23e−05
k
plasminogen−binding protein ErpM
BB_RS06880 (BB_O40)
p = 1.08e−24
l
plasminogen−binding protein ErpY
BB_RS06465 (BB_R42)
p = 4.77e−02
Standardized expressions
SP
RB
BL
BF
Morphotypes

**Figure 13.** Standardized expression profile of genes taking part in zoonotic life cycle regulation and persistence in the mammalian host across *B. burgdrorferi* morphotypes. Standardized expression profiles were marked with colored lines and associated p-value calculated by LRT analysis (see Methods). Graphs are showing standardised expression profiles of genes: (a) outer surface lipoprotein OspA, (b) outer surface lipoprotein OspB, (c) outer surface lipoprotein OspC, (d) RNA polymerase factor sigma-54 RpoN, (e) RNA polymerase factor sigma RpoS, (f) variable surface antigen VlsE, (g) decorin-binding protein DbpA, (h) decorin-binding protein DbpB, (i) S-ribosylhomocisteine lyase LuxS, (j) fibronectin-binding protein RevA, (k) plasminogen-binding protein ErpM, (l) plasminogen-binding protein ErpY.

## 4.7. RpoS-regulated genes

In *B. burgdorferi*, the alternative sigma factor RpoS acts as a transcriptional regulator influencing the expression of genes involved in the stress response mechanism (Caimano et al., 2019). Its expression is increased during the nymphal blood meal, during the transmission from a tick to a mammal, and during the mammalian infection. As such, RpoS is regulating the shift in gene expression which occurs during *B. burgdorferi* life cycle, and most importantly, it influences the course of the infection in mammals. Based on research performed by Caimano et al. (2019), the alternative sigma factor RpoS upregulates the expression of 52 (supplementary data Table S36) and downregulates the expression of 38 *B. burgdorferi* genes (supplementary data Table S37). The LRT analysis I performed has shown that 90% of genes upregulated by RpoS are differentially expressed among morphotypes (supplementary data Table S36). On the other hand, only 47% of genes downregulated by RpoS are differentially expressed based on the LRT analysis (supplementary data Table S37). To gain information about morphotype-dependent expression patterns of genes upregulated or downregulated by RpoS, I visualized their standardized gene expressions (Fig. 14). Standardized expression of genes upregulated by RpoS is higher in blebs than in any other morphotype for the majority of RpoS upregulated genes (Fig. 14a). On the other hand, standardized expression of genes downregulated by RpoS does not show a morphotype-dependent expression pattern (Fig. 14b).

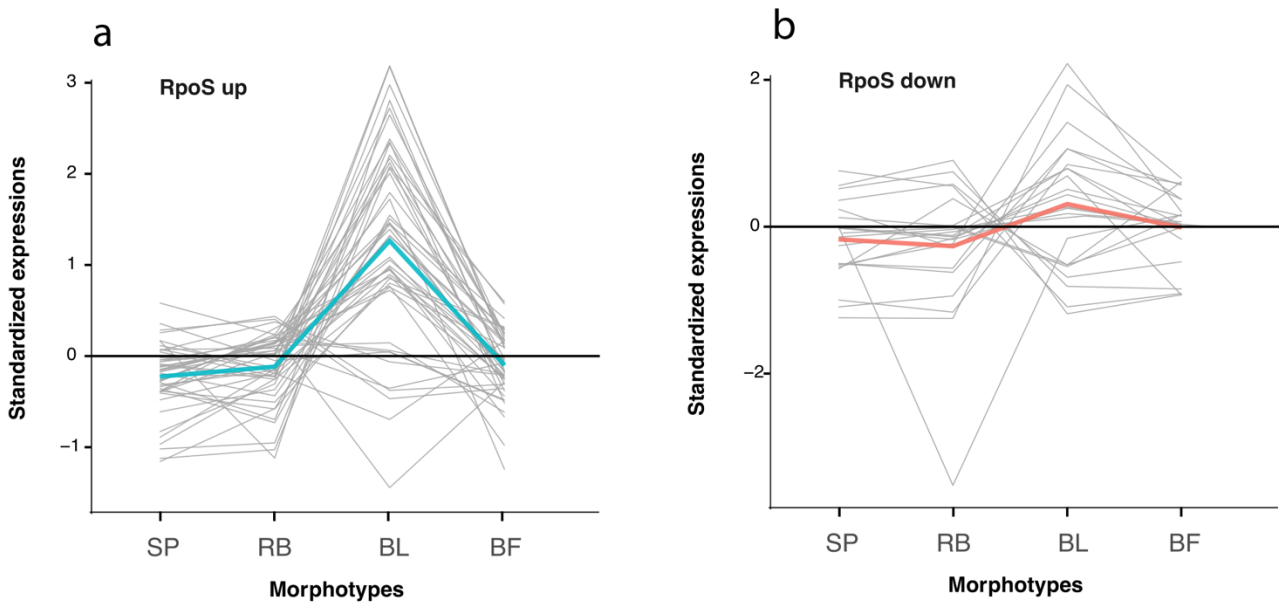



**Figure 14. Standardized transcription profiles genes regulated by RpoS.** Standardized expression profiles of (a) RpoS upregulated (RpoS up) or (b) RpoS downregulated (RpoS down) genes are shown as grey lines, while the arithmetic mean of their expression is shown in blue (RpoS up) and orange (RpoS down). Genes were categorized as upregulated or downregulated by RpoS based on the paper published by Caimano et al. (2019).

To confirm the observations made on the basis of standard transcription profiles of RpoS regulated genes, the distribution of all *B. burgdroferi* genes across clusters and the difference in the distribution of genes determined by Caimano et al. (2019) to be upregulated by RNA polymerase sigma factor RpoS across clusters were tested for statistical significance. It is revealed that RpoS-upregulated genes (supplementary data Table S36) predominantly correspond with expression values of genes grouped in the 15th cluster (Fig. 15, supplementary data Table S37), while genes downregulated by RpoS (supplementary data Table S37) are not corresponding with any of 22 gene clusters (Fig. 15, supplementary data Table S41). Furthermore, while RpoS-upregulated genes are absent from 10 clusters, the greatest amount of RpoS-upregulated genes (35%) is placed in the 15th cluster (supplementary data Table S3).

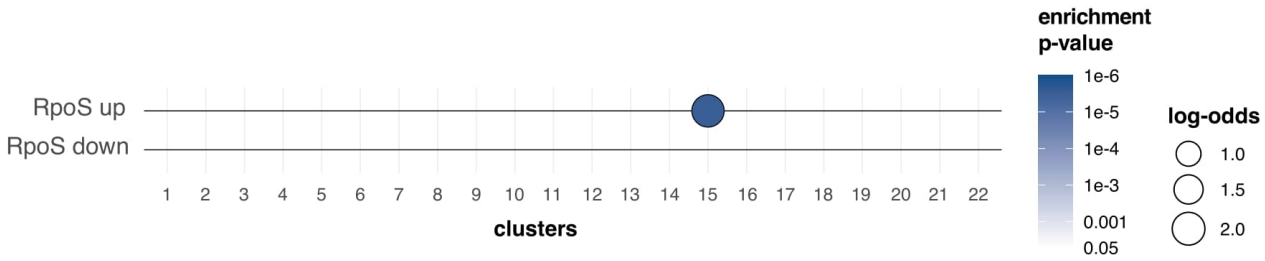



**Figure 15. The enrichment of RpoS modulated genes in gene expression clusters.** Clusters were generated with the DP_GP_cluster algorithm at maximum Gibbs sampling iterations set to 500 based on standardized expression across *B. burgdroferi* morphotypes. Enrichment of cluster-specific genes among genes upregulated (RpoS up) or downregulated (RpoS down) by RpoS was tested by two-tailed hypergeometric test corrected for multiple comparisons at 0.05 level. Enrichment of cluster-specific genes was shown by log-odds. Log-odds levels were shown by circles of different sizes and p-values by blue color shades.

Cluster 15 (Fig. 11) makes up almost 8% of protein-coding genes (supplementary data Table S3). More importantly, standardized expression profiles are showing that genes included in this cluster are showing a small difference in expression between spirochetes, round bodies, and biofilms, while they are significantly overexpressed in blebs (Fig. 11). Despite the insignificant variation in *rpoS* expression across morphotypes, its expression is increased in blebs compared to other morphotypes (Fig. 13). However, this slight increase remains meaningful since the expression of transcriptional factors and their target genes generally are not correlated (Zaborowski & Walther, 2020).

To gain insight into evolutionary signatures of genes upregulated by RpoS, enrichment analysis of genes upregulated by RpoS among phylostratum-specific genes was performed (Fig. 16). As a result, it was determined that most RpoS-overexpressed genes originate in the 6$^{th}$ phylostratum, corresponding to the origin of the *Borreliacea* taxon (supplementary data Table S38). This finding correlates with the results obtained by phylostratigraphic analysis of the *B. burgdorferi* B31 morphotypes, by which it is demonstrated that genes overexpressed in blebs contain a significant amount of genes originating in the 6$^{th}$ phylostratum (Fig. 7. a, b). On the other hand, RpoS-downregulated genes include a significant number of genes originating in the 7$^{th}$ phylostratum corresponding to the *Borrelia* genus (supplementary data Table S39).

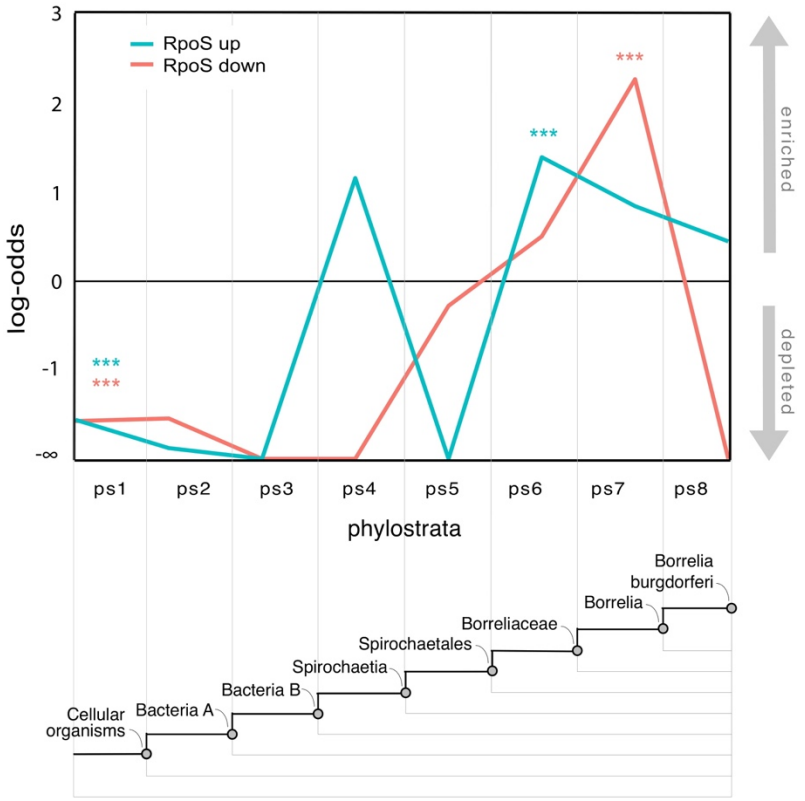



**Figure 16. Phylostratigraphic analysis of *B. burgdorferi* genes modulated by RpoS.** A horizontal grid depicts the eight phylostrata (ps1-ps8) assigned using the e-value $10^{-3}$ BLASTp cutoff value. The frequency of phylostratum-specific genes among upregulated (RpoS up, blue line) genes or downregulated (RpoS down, orange line) genes is compared to the frequency of phylostratum-specific genes in the complete genome and deviations are shown by log-odds (y-axis). The log-odds of zero marks that the frequency of genes upregulated or downregulated by RpoS in a phylostratum equals the expected frequency estimated from the distribution of all genes across phylostrata. Genes were categorized as upregulated or downregulated by RpoS based on the paper published by Caimano et al. (2019). Alterations from the expected frequencies were tested by a two-tailed hypergeometric test corrected for multiple comparisons at 0.05 level (*P < 0.05; **P < 0.01; ***P < 0.001).

# 5. Discussion

Since its initial discovery in 1975 (Steere et al., 1977), *Borrelia burgdorferi*, the causative agent of Lyme disease, has remained a relevant topic among scientists and medical professionals alike. Evolutionary innovations, such as exploitation of tick salivary protein for early host immune response delayment, usurpation of host's plasminogen activating system, and deceivement of alternative complement pathways by surface antigen masking (Berndtson, 2013), are being extensively researched since they represent potential drug targeting pathways. Nonetheless, prevention, diagnostics, and treatment strategies remain only partly effective, resulting in a relatively high number of patients suffering from Post-Treatment Lyme Disease Syndrome (PTLDS) and Chronic Lyme Disease (CLD) (Mead, 2015). Based on that, *B. burgdroferi* is recognized as an escalating public health problem that demands an improved understanding of this pathogen's sophisticated survival strategies, life cycle, and parasitic lifestyle adaptations (Berndtson, 2013).

An additional layer of complexity was introduced with the discovery of atypical borrelial morphologies found in various *in vitro* (Alban, Johnson & Nelson, 2010; Brorson & Brorson, 1998; Murgia et al., 2002; Feng, Shi, Zhang & Zhang, 2015; Meriläinen, Herranen, Schwarzbach & Gilbert, 2015; Vancová et al., 2017; Brorson et al., 2009; Sapi et al., 2011; Bamm, Ko, Mainprize, Sanderson & Wills, 2019; Brorson et al., 2009; Kersten, Poitschek, Rauch & Aberer, 1995; Barbour & Hayes, 1986; Sapi et al., 2012) and *in vivo* (Miklossy et al., 2008; Sapi et al., 2019; Aberer, Kersten, Klade, Poitschek & Jurecka, 1996; Sapi et al., 2016; Hulínská et al., 1994; Kersten, Poitschek, Rauch & Aberer, 1995) environmental conditions. Notably, those morphotypes were observed in tissue samples of patients suffering from Lyme disease manifestations such as neuroborreliosis and Lyme arthritis (Miklossy et al., 2008, Sapi et al., 2019). Round bodies, blebs, and biofilms were found to be present in small numbers in *B. burgdorferi* populations *in vitro* (Meriläinen, Herranen, Schwarzbach & Gilbert, 2015). In addition, they are showing typical persister characteristics such as high antibiotic tolerance, low metabolic activity, and resistance to environmental changes (Sapi et al., 2011, Meriläinen, Herranen, Schwarzbach & Gilbert, 2015, Bernatdtson 2013). In addition, round bodies, blebs and biofilms possess immune-modulating

activity in *in vitro* systems (Meriläinen, Brander, Herranen, Schwarzbach & Gilbert, 2016; Whitmire & Garon, 1993). Interestingly, although there is a plethora of research pointing toward the medical relevance of *B. burgdorferi* round body, bleb, and biofilm morphotypes (Bamm, Ko, Mainprize, Sanderson & Wills, 2019; Di Domenico et al., 2018; Margulis, Maniotis & MacAllister, 2009; Meriläinen, Herranen, Schwarzbach & Gilbert, 2015; Meriläinen, Brander, Herranen, Schwarzbach & Gilbert, 2016; Bernardtson, 2013; Vancová et al., 2017; Rudenko, Golovchenko, Kybicova & Vancova, 2019; Sapi et al., 2019), this thesis is the first to analyze their differential gene expression on a global scale.

## 5.1. *Borrelia burgdorferi* morphotypes show distinct expression profiles

When spirochete preculture was exposed to different growing environments, the gained populations resulted in morphologically distinct *B. burgdorferi* cell types. Round bodies were gained by exposure of spirochetes to distilled water, putting the bacteria under osmotic stress and nutrient deprivation. The transformation of spirochetes to round bodies was rapid, taking only 10 min. Blebs were induced with increased oxygen exposure, and biofilms were raised in high cell density populations. Unlike round bodies, the formation of blebs and biofilms was a gradual process that took several days. Transcription values of those morphotype-rich cultures were used for further bioinformatic analysis.

### 5.1.1. Genes upregulated in round bodies are coding for structural components of the bacterial ribosome

Based both on principal component analysis and pairwise gene expression analysis, a significant difference in gene expression between *B. burgdorferi* morphotypes was observed. In particular, transcription profiles of blebs and biofilms are distinctive from each other, as from spirochetes and round bodies whose gene expression is quite similar. The similarity in gene expression between spirochetes and round bodies is quite surprising, considering the major morphological differences these two morphotypes are showing (Alban, Johnson & Nelson, 2010; Meriläinen, Herranen, Schwarzbach & Gilbert, 2015). In addition, round bodies possess different

biochemical features (Meriläinen, Herranen, Schwarzbach & Gilbert, 2015) and effects on the immune cells of the vertebrate host (Meriläinen, Brander, Herranen, Schwarzbach & Gilbert, 2016). Taken all together, a significant difference in gene expression in round bodies when compared to spirochetes was expected. Surprisingly, out of all *B. burgdorferi* genes, 4.34% are differentially expressed. Moreover, only 2.85% are upregulated in round bodies compared to spirochetes. Based on the functional enrichment analysis, a significant amount of these genes are annotated as possessing "translational, ribosomal structure, and biogenesis" function. A more detailed view of those genes reveals that they are coding for structural components of the bacterial ribosome and not participating in translation and ribosome biogenesis in any other way. However, it is worth noticing that the expression of almost 46% of ribosomal proteins differs between round bodies and spirochetes and thus reveals there is a partial difference in the expression of ribosomal proteins between round bodies and spirochetes.

The partial difference in expression of ribosomal proteins between round bodies and spirochetes indicated that *B. burgdorferi* ribosomes are heterogeneous in their composition and have a degree of specialization in their function (Byrgazov, Vesper & Moll, 2013). Namely, despite the catalytic activity of the rRNA in protein synthesis, the lack of some ribosomal proteins could contribute to fine-tuning of ribosome function and, in particular, to its selectivity for distinct transcripts (Byrgazov, Vesper & Moll, 2013). Because of this heterogeneity, a minor difference in transcription between round bodies and spirochetes could still result in substantial physiological divergence if the expression of those transcripts is regulated on a translational level by differential ribosome composition. Comparable effects were found in *E. coli* (Deusser, 1972; Deusser & Wittmann, 1972), cultured in minimal media similar to the one previously used for *B. burgdroferi* round body induction (Meriläinen, Herranen, Schwarzbach & Gilbert, 2015; Alban, Johnson & Nelson, 2010; Murgia et al., 2002).

### 5.1.2. Functional characteristics of genes differentially expressed in blebs and biofilms

Based on a more stringent criteria, pairwise expression analysis of blebs and biofilms demonstrates they share approximately 70% of differentially expressed genes. Additionally, a considerable fraction of "replication, recombination and repair" and "cell wall/membrane/envelope biogenesis" genes downregulated in blebs are also downregulated in biofilms, pinpointing the similarities in the protoplasmic envelope, and replication, recombination, and repair mechanism in blebs and biofilms. Since blebs and biofilms are persisters (Bamm, Ko, Mainprize, Sanderson & Wills, 2019), and persisters are known to have low replication rates (Rudenko, Golovchenko, Kybicova & Vancova, 2019), downregulation of genes taking part in replication is anticipated. Furthermore, differential expression of genes involved in the cell wall, membrane, and envelope biogenesis in the bleb morphotype is also expected since alterations of the outer membrane define blebs (Vancová et al., 2017; Berndtson, 2013).

Although genes involved in the cell wall, membrane, and envelope biogenesis are important for biofilm formation in general (Jefferson, 2004), these genes are functionally enriched among genes downregulated in *B. burgdorferi* biofilms. Notably, this downregulation is accompanied by an increased expression of genes of unknown function. Based on that, it is highly probable that among genes of unknown function more expressed in biofilms than in spirochetes lies a considerable number of genes controlling cell wall, membrane, and envelope biogenesis in *B. burgdorferi* biofilms. The differential expression of genes taking part in "cell wall/membrane/envelope biogenesis" in blebs and biofilms when compared to spirochetes might have consequences regarding Lyme disease treatment. Namely, many antibiotics used for Lyme disease treatment primarily target the cell membrane (Bamm, Ko, Mainprize, Sanderson & Wills, 2019; Feng, Zhang, Shi & Zhang, 2016), making variation of these structures among morphotypes potentially medically problematic.

### 5.1.3. Functional characteristics of gene clusters

Functional enrichment analysis of clusters occupied by genes grouped on the basis of their expression across *B. burgdorferi* morphotypes has shown that in some cases, genes that share a similar expression profile also share a similar function. For example, clusters 3, 9, and 11 are occupied with genes more expressed in blebs and biofilms than in round bodies and spirochetes. Since functional enrichment analysis of morphotype-specific genes has shown that those two morphotypes are enriched with genes of unknown function, functional enrichment of clusters 3, 9, and 11 with genes of unknown functions is expected. Additionally, a significant number of genes populating cluster 1 are involved in "cell wall/membrane/envelope biogenesis." Since genes present in these clusters have a higher expression in spirochetes and round bodies than in blebs and biofilms, genes labeled as "cell wall/membrane/envelope biogenesis" placed in cluster 1 may be involved in these processes in both of those morphotypes. Similarly, another cluster of genes more expressed in spirochetes and round bodies than in blebs and biofilms is cluster 5. In this cluster, a significant amount of genes are taking part in "replication, recombination and repair" and "lipid transport and metabolism" processes. Spirochetes are the replicative form of *B. burgdorferi* (Meriläinen, Herranen, Schwarzbach & Gilbert, 2015), and because of that, enrichment of cluster 5 with genes involved in replication is expected. In contrast, enrichment of a cluster containing genes highly expressed in round bodies is unexpected since round bodies are known to be reproductively inactive (Meriläinen, Herranen, Schwarzbach & Gilbert, 2015).

Since it is shown that *B. burgdorferi* genes tend to cluster based on their expression in various morphotypes, it is possible to expect that genes of unknown functions placed in the same clusters as genes of known functions may share those functions. For example, cluster 16 is enriched with genes involved in "translation, ribosomal structure and biogenesis", and around 30% of genes occupying these clusters are functionally undescribed. Based on that, it is expected that a significant number of genes of unknown function are indeed taking part in "translation, ribosomal structure and biogenesis". However, because of the lacking annotation of *B. burgdorferi* genes in general, the vast majority of clusters are not enriched with any functional annotation. It is worth keeping in mind that a more detailed annotation of genes may potentially result in a significantly different distribution of gene annotations across clusters. Until that, cluster analysis of *B. burgdorferi* genes is a good starting point for further research of *B. bugdorferi* gene function.

## 5.2. Genes upregulated in blebs and biofilms are enriched with plasmid-encoded genes

Based on the distribution of differentially expressed genes across *B. burgdorferi* plasmids and the chromosome, it is possible to conclude that the transition from spirochetes into round bodies is primarily associated with expressional changes in bacterial chromosome genes, while the conversion of spirochetes into blebs or biofilms is accompanied with significantly different expression of plasmid genes. Bleb morphotype is characterized by the formation of outer membrane bulges that bud into small outer membrane vesicles (OMVs) (Vancová et al., 2017; Berndtson, 2013) and shed off the bacterial surface. OMVs released from the surface of *B. burgdorferi* blebs are enriched in plasmid-encoded mRNA transcripts, while the cell body is enriched with mRNA transcripts located on the chromosome (Malge et al., 2018). According to that, the presence of plasmid transcript enriched OMVs in the bleb cell culture could potentially explain the increased abundance of these transcripts in the culture as a whole. In addition, research of synovial fluid samples has shown that plasmid DNA, but not chromosomal DNA was present inside *B. burgdroferi* blebs shed into the joint space (Persing et al., 1994). Moreover, observed transcripts were lipoproteins or putative lipoproteins involved in modulating the host response to pathogenesis during the development of Lyme disease (Malge et al., 2018). Based on increased expression of plasmid-encoded genes in bleb cultures, the presence of *B. burgdroferi* plasmid-encoded genes in OMVs shed into the joint space (Persing et al., 1994), and the virulence potential of their cargo (Toledo, Coleman, Kuhlow, Crowley & Benach, 2012), it is possible to speculate that blebs may play a role in the progression of Lyme arthritis and Lyme disease in general. However, further research is needed.

The enrichment of plasmid-encoded genes among biofilm upregulated genes is in accordance with previously published research on *B. burgdorferi* biofilms. Namely, although the correlation between the high expression of plasmid-encoded genes and *B. burgdorferi* biofilm formation was not previously established, it is known that *B. burgdorferi* biofilms accommodate spirochetes, round bodies, and, most importantly, blebs (Meriläinen, Herranen, Schwarzbach & Gilbert, 2015). Therefore, it would not be surprising if *B. burgdorferi* biofilm cultures contained OMVs shed by blebs and, consequently, a meaningful number of plasmid-encoded transcripts. *B. burgdorferi* would not be the only bacteria showing this type of adjustment since budding of OMVs

under *in vitro* conditions was previously reported in biofilms of other bacterial species (Schooling & Beveridge, 2006; Klimentová & Stulík, 2015). OMVs are generally involved in stress response, quorums sensing, transfer of genetic materials by plasmid exchange, resistance against antibiotics, and modulation of host immune response. These processes commonly occur during biofilm formation (Molin & Tolker-Nielsen, 2003; Schwechheimer and Kuehn, 2015; Jan, 2017). Furthermore, bacterial OMVs regularly contain lipoproteins responsible for biofilm adhesion, growth factors, and extracellular matrix components such as exopolysaccharides, increasing co-aggregation of cells in the biofilms (Schooling & Beveridge, 2006; Klimentová & Stulík, 2015; Jan, 2017). According to the distribution of morphotype-specific genes across *B. burgdroferi* plasmids and *B. burgdroferi* chromosome, and previously published OMV research, it is possible to propose that increased expression of plasmid-encoded genes in blebs and biofilms is a consequence of OMV shedding happening in both of these morphotypes. Despite that, there is a possibility that the upregulation of plasmid transcripts in blebs and biofilms is not related to the production of OMVs, and based on that, this topic demands further research.

## 5.3. Evolutionary expression suggests a recent origin of biofilm and bleb morphotype

Phylostratigraphic analysis of *B. burgdorferi* morphotypes revealed that the origins of genes differentially expressed among *B. burgdorferi* spirochete, round body, bleb, and biofilm morphotypes can be traced to distinct periods in the evolutionary history of this bacterial species. Namely, genes differentially expressed between round bodies and spirochetes are enriched with genes developed early during the evolution of cellular organisms. Furthermore, this is true for both upregulated and downregulated genes, suggesting that basic genetic prerequisitions governing this transition have developed at the origin of cellular organisms. In contrast to round bodies, genes upregulated by blebs and biofilms are rich with genes corresponding to the emergence of *Borreliaceae*. Species belonging to this phylum have evolved a biphasic life cycle, parasiting both on arthropod and vertebrate hosts (Bamm, Ko, Mainprize, Sanderson & Wills, 2019). Based on that, it is possible to conclude that genes developed in the *Borreliaceae* phylum are involved in host-specific adaptation processes. Phylostratigraphic analysis revealed that these exact genes are

upregulated in blebs and biofilm, suggesting bleb and biofilm morphotype may function in regulating complex interactions between the bacteria and its hosts.

The expression of genes whose origin can be traced to the emergence of cellular life and Bacteria taxa are significantly downregulated in blebs and biofilms. Additionally, genes that originated during the development of Spirochaetales are significantly downregulated in the bleb morphotype. These older genes have developed in phyla whose species are not participating in the enzootic life cycle (Gupta, 2000), indicating they are not crucial in forming blebs and biofilms and their interactions with the host.

Although the reliability of the phylostratigraphic approach regarding the sensitivity of the BLAST algorithm was confirmed both experimentally (Shi et al., 2020) and *in silico* (Futo et al., 2021, Domazet-Lošo et al., 2017.), enrichments of phylostratum-specific genes among genes differentially expressed among morphotypes was tested by using a broad range of e-value thresholds while searching for *B. burgdorferi* homologs in the protein database. By applying e-value thresholds lower than $10^{-3}$, genes were pulled toward younger phylostrata, while e-value thresholds higher than $10^{-3}$ pushed the genes into older phylostrata. Different distributions of genes across phylostrata were tested by a hypergeometric test. Although shifts in gene distribution across phylostrata were notable, differentially expressed genes between round bodies and spirochetes, blebs and spirochetes, and biofilms and spirochetes were mostly enriched with genes placed in the exact same phylostratum as if it was the case when e-value threshold od $10^{-3}$ was used. This is especially true for enrichments of genes corresponding to the origin of life (ps1) and the emergence of the *Borreliacea* phylum (ps6). In contrast, enrichments of morphotype-specific genes with genes placed in the Bacteria A (ps2), Bacteria B (ps3), and *Spirochaetales* (ps5) are worthy of further investigation. Overall, it is safe to say that results gained by the phylostratigraphic approach indicate that genes overexpressed in blebs and biofilms have evolved in the *Borreliaceae* phylum, while genes significantly more expressed in round bodies than in spirochetes emerged at the origin of life.

## 5.4. *Borrelia burgdorferi* genes of unknown function

The function of a vast majority of *B. burgdroferi* genes is poorly understood. To be precise, only 47% are assigned with COG annotations, while the rest of genes remains functionally unannotated. Moreover, genes with known functions are mostly shared with Bacteria and cellular organisms in general, while genes developed in younger taxons lack functional annotations. This is especially evident in the case of genes developed within the *Borreliaceae* phylum, *Borrelia* genus, and *B. burgdroferi* species. Notably, those exact genes are particularly interesting since they have developed after the bacteria has evolved to take part in the biphasic life cycle. Furthermore, genes originating in the *Borreliacea* phylum are significantly more expressed in blebs and biofilms when compared to spirochetes, and those are the same genes that remain to be functionally annotated.

In addition to the lack of functionally annotated genes in younger phylostrata, the distribution of functionally annotated genes based on their expression in morphotypes is indeed morphotype-specific. Functional enrichment analysis of differentially expressed genes in *B. burgdorferi* morphotypes revealed that genes overexpressed in blebs and biofilms significantly lack functional annotations, reflecting a poor understanding of molecular processes governing the formation of these morphotypes. Additionally, a significant number of genes of unknown function are populating clusters 3, 9, and 11. Interestingly, the average standardized expression of genes occupying these clusters is higher in blebs and biofilms than in round bodies and spirochetes. Because of that, functionally unannotated genes significantly more expressed in blebs and biofilms than in spirochetes, along with genes placed in clusters 3, 9, and 11, represent interesting candidates for functional research. The knock-outs of these genes that show phenotypic changes could improve our understanding of bleb and biofilm morphotypes.

## 5.5. Genes involved in the zoonotic life cycle and persistence in the mammalian host are differentially expressed among *Borrelia burgdorferi* morphotypes

According to standardized expression profiles of *B. burgdorferi* genes taking part in the enzootic life cycle and bacterial persistence in the mammalian host, one of the most well-described alternative sigma factors, RpoS, is not differentially expressed in spirochetes, round bodies, blebs, and biofilms. Since this gene is differentially expressed during the stringent response (Caimano et al., 2019), and the stringent response in *B. burgdorferi* is accompanied by the formation of round bodies (Drecktrah et al., 2015), the lacking difference in expression of the *rpoS* gene revealed by my research is not in correlation with the research done by Caimano (2019) and Drecktrah (2015). Genes *rpoS, rpoN,* and *luxS* are proven to be crucial for the development of a true *B. burgdorferi* biofilm (Sapi, Theophilus, Pham, Burugu & Luecke, 2016). As such, an increased expression of these genes in the biofilm morphotype was expected. Surprisingly, neither of those genes have shown increased expression in the biofilm morphotype, and furthermore, *luxS*, a quorum sensing gene (Stevenson et al., 2003), was significantly downregulated in biofilms when compared to spirochetes, round bodies, and blebs.

Out of genes *ospA*, *ospB*, and *ospC*, which are all coding for outer surface lipoproteins (Ouyang, Blevins & Norgard, 2008), the only gene showing differences in gene expression between morphotypes is the gene *ospC*. Namely, this gene is more expressed in spirochetes and round bodies than in biofilms, and more expressed in biofilms than in blebs. Since spirochetes are the most motile *B. burgdorferi* morphotype (Meriläinen, Herranen, Schwarzbach & Gilbert, 2015) and the expression of ospC is at its peak in the early phases of mammalian infection during which *B. burgdorferi* disseminates throughout various tissues (Ouyang, Blevins & Norgard, 2008), increased expression of *ospC* in spirochetes is expected. Additionally, as the acute infection in mammal's transitions into its chronic phase, *B. burgdorferi* downregulates *ospC* expression (Ouyang, Blevins & Norgard, 2008). Since biofilms are known to be taking part in a broad spectrum of bacterial infections (Flemming & Wuertz, 2019), the decreased expression of *ospC* in *B. burgdorferi* biofilms was also expected.

Genes *dbpA*, *dbpB*, *vlsE*, *revA*, and *erpY* are all well-known virulence factors involved in the adhesion of *B. burgdorferi* to the vertebrate extracellular matrix and immune system evasion (Schwartz, Margos, Casjens, Qiu & Eggers, 2020; Verhey, Castellanos & Chaconas, 2019). Notably, all of them are more expressed in blebs than in any other morphotype, indicating that blebs may present at the later stages of chronic infection and take part in adhesion and immune avoidance processes. Interestingly, although it is more expressed in blebs, than in spirochetes and round bodies, the *erpM* gene, which codes for a plasminogen-binding protein, has the highest expression in biofilms. Furthermore, since the *erpY* gene, another plasminogen-binding protein, is more expressed in blebs than in biofilms, it is possible to conclude that different morphotypes adopt different plasminogen binding strategies. Overall, based on standardized expression profiles of *B. burgdorferi* genes taking part in the enzootic life cycle and bacterial persistence in the mammalian host, it is possible to conclude that these genes are differentially expressed among spirochetes, round bodies, blebs, and biofilms.

## 5.6. Genes upregulated by RpoS are showing a morphotype-dependent transcription profile

Since *B. burgdroferi* genes are poorly annotated, functional characteristics of blebs and biofilms remained ill-described. To fill this void, further enrichment analyses were performed, testing the significance of the distribution of RpoS-regulated genes. RNA polymerase sigma factor RpoS is the central regulator of general stress response in various bacterial species (Hengge-Aronis, 2002). In *B. burgdorferi*, it is required for tick to mammal transmission (Dunham-Ems, Caimano, Eggers & Radolf, 2012) and keeping maximum fitness through mammalian infection (Caimano et al., 2019). Using transcriptional reporters and mutagenesis, Caimano et al. (2019) acquired a list of genes upregulated or downregulated by RpoS within a dialysis membrane chamber (DMC) peritoneal cultivation system, an important *in vitro* proxy for infected mammalian tissue (Caimano, 2018). On the basis of standardized expression profiles and the distribution of RpoS-regulated genes across clusters of genes grouped based on their differential expression across morphotypes, I discovered that most RpoS-upregulated genes have a significantly higher expression in blebs than

other morphotypes. Since genes upregulated by RpoS are also upregulated in bleb morphotype, it is applied that bleb morphotype may play an important role during mammalian infection.

To gain insight into evolutionary signatures of genes whose expression is regulated by RpoS, distributions of RpoS-upregulated and RpoS-downregulated genes across phylostrata were tested for statistical significance. The analysis has shown that RpoS-overexpressed genes primarily originate in the *Borreliaceae* phylum (ps6), while RpoS-downregulated genes originate in the genus *Borrelia* (ps7) phylostratum. Since the *Borreliaceae* phylum (ps6) is characterized by the development of a biphasic life cycle, and RpoS-upregulates genes inside the mammalian host, it is possible that the bleb morphotype had emerged during the development of *Borreliaceae* as an adaptation for surviving the transition to the mammalian host. This notion goes hand in hand with previous research in which it has been shown that blebs form as a response to components of the complement system (Meriläinen, Herranen, Schwarzbach & Gilbert, 2015), induce B-cell response in mouse models (Whitmire & Garon, 1993), and most probably participate in the proteolytic degradation of the extracellular matrix (Toledo, Coleman, Kuhlow, Crowley & Benach, 2012). On the other hand, *Borrelia* species (ps7) are distinctive from other members of the *Borreliaceae* phylum based on their vector preference. Namely, while other *Borreliaceae* species parasite on Argasid ticks, human body louse, and order *Ixodida* which includes the *Ixodes* genus, *Borrelia* has adapted for survival in primarily genus *Ixodes* (Oppler, O'Keeffe, McCoy & Brisson, 2020). This means that the majority of genes acquired for survival in this specific tick genus are not required inside the mammalian host and are consequently downregulated by RpoS in DMC.

Comparable to bleb morphotype, phylostratigraphic analysis has shown that genes upregulated in biofilm morphotype predominantly originated in the *Borreliaceae* phylum, suggesting they developed during *B. burgdorferi* adaptation to the mammalian host. Additionally, the crucial role of biofilms in the colonization and persistence of pathogen bacteria is supported by numerous research papers (Bamm, Ko, Mainprize, Sanderson & Wills, 2019). Thus, it may come as a surprise that RpoS-induced genes generally have similar expression across spirochete, round body, and biofilm morphotypes. However, although the DMC peritoneal cultivation system effectively simulates the exposure of the bacteria to soluble elements of the mammalian environment, it does not accommodate for the contact of bacteria and the extracellular matrix of the vertebrate host (Caimano et al., 2019). Since the adhesion of bacteria to the extracellular matrix

strongly induces biofilm development (Caimano et al., 2019), genes participating in the mammalian host-specific biofilm formation may not have been induced in the experiment performed by Caimano et al. (2019).

## 5.7. Future directions

Based on functional enrichment analysis of differentially expressed genes in *B. burgdorferi* morphotypes, many genes highly expressed in blebs and biofilms remain functionally undescribed. Considering potential bleb and biofilm involvement in the progression of Lyme disease (Meriläinen, Herranen, Schwarzbach & Gilbert, 2015), the functional characterization of genes taking part in bleb and biofilm formation is of crucial importance. In this regard, stratification of bleb and biofilm-specific genes with respect to their evolutionary origin, similar to that performed by Shi et al. (2020), could provide a basis for the identification of new genes crucial for the development of respected morphotypes. Their research included phylostratigraphic analysis of *B. subtilis* genes and tested the distribution of sporulation genes for statistical significance. The analysis revealed that sporulation genes cluster at distinct evolutionary time points. Unknown genes placed in the same phylostrata as major clusters of sporulation genes were inactivated, and mutant phenotypes were observed. Shi et al. revealed that the vast majority of inactivated genes significantly affected sporulation in *B. burgdorferi* biofilms, confirming that genomic phylostratigraphy is a valuable tool for predicting functions of unknown genes. Similar to previously mentioned studies (Shi et al., 2020), future research on *B. burgdorferi* could aim to deactivate genes of unknown function highly expressed in round body, bleb, and biofilm morphotypes and observe if those knock-outs significantly affect morphotype formation. Genes found to have a significant role in the development of pleomorphic variants could represent a target for future drugs, vaccines, and diagnostic methods against *B. burgdorferi*.

# 6. Conclusion

This doctoral thesis provides the first analysis of differential expression and evolutionary signatures of *Borrelia burgdorferi* spirochete, round body, bleb, and biofilm morphotypes. It revealed that biofilm and bleb morphotypes have clearly distinct transcriptomes between each other, and compared to spirochete and round body morphotypes that cluster together. The distribution of morphotype-specific genes across *B. burgdorferi* plasmids and chromosome, displays that the transition from spirochetes into round bodies is primarily associated with expressional changes of bacterial chromosome genes, while plasmid genes are significantly differentially expressed during the conversion of spirochetes into blebs or biofilms. Additionally, phylostratigraphic analysis has shown that genes required for round body development emerged early during the evolution of cellular organisms, whereas genes upregulated by blebs and biofilms are rich with genes that emerged at the origin of *Borreliaceae*. A significant proportion of genes upregulated in round bodies are taking part in translational, ribosomal structure, and biogenesis function, while a huge number of genes highly expressed in blebs and biofilms remain to be functionally described. These results are a significant contributing to our understanding of the development of *B. burgdorferi* round body, bleb, and biofilm morphotypes and bacterial pleomorphism in general.

# 7. References


1. Rosenberg, R., Lindsey, N. P., Fischer, M., Gregory, C. J., Hinckley, A. F., Mead, P. S., Paz-Bailey, G., Waterman, S. H., Drexler, N. A., Kersh, G. J., Hooks, H., Partridge, S. K., Visser, S. N., Beard, C. B., & Petersen, L. R. (2018). Vital Signs: Trends in Reported Vectorborne Disease Cases - United States and Territories, 2004-2016. *MMWR. Morbidity and mortality weekly report*, *67*(17), 496–501. https://doi.org/10.15585/mmwr.mm6717e1
2. Barbour, A. G., & Hayes, S. F. (1986). Biology of Borrelia species. *Microbiological reviews*, *50*(4), 381–400. https://doi.org/10.1128/mr.50.4.381-400.1986
3. Radolf, J. D., Caimano, M. J., Stevenson, B., & Hu, L. T. (2012). Of ticks, mice and men: understanding the dual-host lifestyle of Lyme disease spirochaetes. *Nature reviews. Microbiology*, *10*(2), 87–99. https://doi.org/10.1038/nrmicro2714
4. Stafford, K. C., 3rd, Cartter, M. L., Magnarelli, L. A., Ertel, S. H., & Mshar, P. A. (1998). Temporal correlations between tick abundance and prevalence of ticks infected with Borrelia burgdorferi and increasing incidence of Lyme disease. *Journal of clinical microbiology*, *36*(5), 1240–1244. https://doi.org/10.1128/JCM.36.5.1240-1244.1998
5. Dumic, I., & Severnini, E. (2018). "Ticking Bomb": The Impact of Climate Change on the Incidence of Lyme Disease. *The Canadian journal of infectious diseases & medical microbiology = Journal canadien des maladies infectieuses et de la microbiologie medicale*, *2018*, 5719081. https://doi.org/10.1155
6. 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
7. Berndtson K. (2013). Review of evidence for immune evasion and persistent infection in Lyme disease. *International journal of general medicine*, *6*, 291–306. https://doi.org/10.2147/IJGM.S44114
8. Hündersen, F., Forst, S., & Kasten, E. (2021). Neuropsychiatric and Psychological Symptoms in Patients with Lyme Disease: A Study of 252 Patients. *Healthcare (Basel, Switzerland)*, *9*(6), 733. https://doi.org/10.3390/healthcare9060733
9. Lindgren, E., Jaenson T.G.T. & Menne B. (2006) Lyme Borreliosis in Europe : influences of climate and climate change, epidemiology, ecology and adaptation measures. *WHO.*

*2006*. https://apps.who.int/iris/bitstream/ handle/10665/107800/E89522.pdf?sequence=1&isAllowed=y (WHO, 2006).

10. Bamm, V. V., Ko, J. T., Mainprize, I. L., Sanderson, V. P., & Wills, M. (2019). Lyme Disease Frontiers: Reconciling *Borrelia* Biology and Clinical Conundrums. *Pathogens (Basel, Switzerland)*, *8*(4), 299. https://doi.org/10.3390/pathogens8040299
11. Drecktrah, D., Lybecker, M., Popitsch, N., Rescheneder, P., Hall, L. S., & Samuels, D. S. (2015). The Borrelia burgdorferi RelA/SpoT Homolog and Stringent Response Regulate Survival in the Tick Vector and Global Gene Expression during Starvation. *PLoS pathogens*, *11*(9), e1005160. https://doi.org/10.1371/journal.ppat.1005160
12. Kurokawa, C., Lynn, G. E., Pedra, J., Pal, U., Narasimhan, S., & Fikrig, E. (2020). Interactions between Borrelia burgdorferi and ticks. *Nature reviews. Microbiology*, *18*(10), 587–600. https://doi.org/10.1038/s41579-020-0400-5
13. Piesman, J., Schneider, B. S., & Zeidner, N. S. (2001). Use of quantitative PCR to measure density of Borrelia burgdorferi in the midgut and salivary glands of feeding tick vectors. *Journal of clinical microbiology*, *39*(11), 4145–4148. https://doi.org/10.1128/JCM.39.11.4145-4148.2001
14. Schwan, T. G., & Piesman, J. (2002). Vector interactions and molecular adaptations of lyme disease and relapsing fever spirochetes associated with transmission by ticks. *Emerging infectious diseases*, *8*(2), 115–121. https://doi.org/10.3201/eid0802.010198
15. Kung, F., Anguita, J., & Pal, U. (2013). Borrelia burgdorferi and tick proteins supporting pathogen persistence in the vector. *Future microbiology*, *8*(1), 41–56. https://doi.org/10.2217/fmb.12.121
16. Paster, B. J., & Dewhirst, F. E. (2000). Phylogenetic foundation of spirochetes. *Journal of molecular microbiology and biotechnology*, *2*(4), 341–344.
17. Biesiada, G., Czepiel, J., Leśniak, M. R., Garlicki, A., & Mach, T. (2012). Lyme disease: review. Archives of medical science: AMS, 8(6), 978–982. https://doi.org/10.5114/aoms.2012.30948
18. 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

19. Barbour, A. G., & Gupta, R. S. (2021). The Family Borreliaceae (Spirochaetales), a Diverse Group in Two Genera of Tick-Borne Spirochetes of Mammals, Birds, and Reptiles. *Journal of medical entomology*, *58*(4), 1513–1524. https://doi.org/10.1093/jme/tjab055
20. Assous, M. V., & Wilamowski, A. (2009). Relapsing fever borreliosis in Eurasia--forgotten, but certainly not gone!. *Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases*, *15*(5), 407–414. https://doi.org/10.1111/j.1469-0691.2009.02767.x
21. Steere, A. C., Coburn, J., & Glickstein, L. (2004). The emergence of Lyme disease. *The Journal of clinical investigation*, *113*(8), 1093–1101. https://doi.org/10.1172/JCI21681
22. Steere, A. C., Malawista, S. E., Snydman, D. R., Shope, R. E., Andiman, W. A., Ross, M. R., & Steele, F. M. (1977). Lyme arthritis: an epidemic of oligoarticular arthritis in children and adults in three connecticut communities. *Arthritis and rheumatism*, *20*(1), 7–17. https://doi.org/10.1002/art.1780200102
23. Burgdorfer, W., Barbour, A. G., Hayes, S. F., Benach, J. L., Grunwaldt, E., & Davis, J. P. (1982). Lyme disease-a tick-borne spirochetosis?. *Science (New York, N.Y.)*, *216*(4552), 1317–1319. https://doi.org/10.1126/science.7043737
24. Steere, A. C., Grodzicki, R. L., Kornblatt, A. N., Craft, J. E., Barbour, A. G., Burgdorfer, W., Schmid, G. P., Johnson, E., & Malawista, S. E. (1983). The spirochetal etiology of Lyme disease. *The New England journal of medicine*, *308*(13), 733–740. https://doi.org/10.1056/NEJM198303313081301
25. Coumou, J., van der Poll, T., Speelman, P., & Hovius, J. W. (2011). Tired of Lyme borreliosis. Lyme borreliosis in the Netherlands. *The Netherlands journal of medicine*, *69*(3), 101–111.
26. Steere, A. C., Strle, F., Wormser, G. P., Hu, L. T., Branda, J. A., Hovius, J. W., Li, X., & Mead, P. S. (2016). Lyme borreliosis. *Nature reviews. Disease primers*, *2*, 16090. https://doi.org/10.1038/nrdp.2016.90
27. Aguero-Rosenfeld, M. E., Wang, G., Schwartz, I., & Wormser, G. P. (2005). Diagnosis of lyme borreliosis. *Clinical microbiology reviews*, *18*(3), 484–509. https://doi.org/10.1128/CMR.18.3.484-509.2005

28. Steere, A. C., Gross, D., Meyer, A. L., & Huber, B. T. (2001). Autoimmune mechanisms in antibiotic treatment-resistant lyme arthritis. *Journal of autoimmunity*, *16*(3), 263–268. https://doi.org/10.1006/jaut.2000.0495
29. Hyde J. A. (2017). *Borrelia burgdorferi* Keeps Moving and Carries on: A Review of Borrelial Dissemination and Invasion. *Frontiers in immunology*, *8*, 114. https://doi.org/10.3389/fimmu.2017.00114
30. Peters, D. J., & Benach, J. L. (1997). Borrelia burgdorferi adherence and injury to undifferentiated and differentiated neural cells in vitro. *The Journal of infectious diseases*, *176*(2), 470–477. https://doi.org/10.1086/514066
31. Koedel, U., Fingerle, V., & Pfister, H. W. (2015). Lyme neuroborreliosis-epidemiology, diagnosis and management. *Nature reviews. Neurology*, *11*(8), 446–456. https://doi.org/10.1038/nrneurol.2015.121
32. Ogrinc, K., Lusa, L., Lotrič-Furlan, S., Bogovič, P., Stupica, D., Cerar, T., Ružić-Sabljić, E., & Strle, F. (2016). Course and Outcome of Early European Lyme Neuroborreliosis (Bannwarth Syndrome): Clinical and Laboratory Findings. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America*, *63*(3), 346–353. https://doi.org/10.1093/cid/ciw299
33. Uldry, P. A., Regli, F., & Bogousslavsky, J. (1987). Cerebral angiopathy and recurrent strokes following Borrelia burgdorferi infection. *Journal of neurology, neurosurgery, and psychiatry*, *50*(12), 1703–1704. https://doi.org/10.1136/jnnp.50.12.1703
34. Stanek, G., Fingerle, V., Hunfeld, K. P., Jaulhac, B., Kaiser, R., Krause, A., Kristoferitsch, W., O'Connell, S., Ornstein, K., Strle, F., & Gray, J. (2011). Lyme borreliosis: clinical case definitions for diagnosis and management in Europe. *Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases*, *17*(1), 69–79. https://doi.org/10.1111/j.1469-0691.2010.03175.x
35. Puius, Y. A., & Kalish, R. A. (2008). Lyme arthritis: pathogenesis, clinical presentation, and management. *Infectious disease clinics of North America*, *22*(2), 289–vii. https://doi.org/10.1016/j.idc.2007.12.014
36. Swanson, S. J., Neitzel, D., Reed, K. D., & Belongia, E. A. (2006). Coinfections acquired from ixodes ticks. *Clinical microbiology reviews*, *19*(4), 708–727. https://doi.org/10.1128/CMR.00011-06

37. Bockenstedt, L. K., Gonzalez, D. G., Haberman, A. M., & Belperron, A. A. (2012). Spirochete antigens persist near cartilage after murine Lyme borreliosis therapy. *The Journal of clinical investigation*, *122*(7), 2652–2660. https://doi.org/10.1172/JCI58813

38. Hodzic, E., Feng, S., Holden, K., Freet, K. J., & Barthold, S. W. (2008). Persistence of Borrelia burgdorferi following antibiotic treatment in mice. *Antimicrobial agents and chemotherapy*, *52*(5), 1728–1736. https://doi.org/10.1128/AAC.01050-07

39. Hodzic, E., Imai, D., Feng, S., & Barthold, S. W. (2014). Resurgence of persisting non-cultivable Borrelia burgdorferi following antibiotic treatment in mice. *PloS one*, *9*(1), e86907. https://doi.org/10.1371/journal.pone.0086907

40. Embers, M. E., Barthold, S. W., Borda, J. T., Bowers, L., Doyle, L., Hodzic, E., Jacobs, M. B., Hasenkampf, N. R., Martin, D. S., Narasimhan, S., Phillippi-Falkenstein, K. M., Purcell, J. E., Ratterree, M. S., & Philipp, M. T. (2012). Persistence of Borrelia burgdorferi in rhesus macaques following antibiotic treatment of disseminated infection. *PloS one*, *7*(1), e29914. https://doi.org/10.1371/journal.pone.0029914

41. Vancová, M., Rudenko, N., Vaněček, J., Golovchenko, M., Strnad, M., Rego, R., Tichá, L., Grubhoffer, L., & Nebesářová, J. (2017). Pleomorphism and Viability of the Lyme Disease Pathogen *Borrelia burgdorferi* Exposed to Physiological Stress Conditions: A Correlative Cryo-Fluorescence and Cryo-Scanning Electron Microscopy Study. *Frontiers in microbiology*, *8*, 596. https://doi.org/10.3389/fmicb.2017.00596

42. Motaleb, M. A., Corum, L., Bono, J. L., Elias, A. F., Rosa, P., Samuels, D. S., & Charon, N. W. (2000). Borrelia burgdorferi periplasmic flagella have both skeletal and motility functions. *Proceedings of the National Academy of Sciences of the United States of America*, *97*(20), 10899–10904. https://doi.org/10.1073/pnas.200221797

43. Yang, D. C., Blair, K. M., & Salama, N. R. (2016). Staying in Shape: the Impact of Cell Shape on Bacterial Survival in Diverse Environments. *Microbiology and molecular biology reviews : MMBR*, *80*(1), 187–203. https://doi.org/10.1128/MMBR.00031-15

44. Harman, M. W., Dunham-Ems, S. M., Caimano, M. J., Belperron, A. A., Bockenstedt, L. K., Fu, H. C., Radolf, J. D., & Wolgemuth, C. W. (2012). The heterogeneous motility of the Lyme disease spirochete in gelatin mimics dissemination through tissue. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(8), 3059–3064. https://doi.org/10.1073/pnas.1114362109

45. Charon, N. W., Cockburn, A., Li, C., Liu, J., Miller, K. A., Miller, M. R., Motaleb, M. A., & Wolgemuth, C. W. (2012). The unique paradigm of spirochete motility and chemotaxis. *Annual review of microbiology*, *66*, 349–370. https://doi.org/10.1146/annurev-micro-092611-150145

46. Meriläinen, L., Herranen, A., Schwarzbach, A., & Gilbert, L. (2015). Morphological and biochemical features of Borrelia burgdorferi pleomorphic forms. *Microbiology (Reading, England)*, *161*(Pt 3), 516–527. https://doi.org/10.1099/mic.0.000027

47. Groshong, A.M., Blevins, J.S. (2014). Insights into the biology of Borrelia burgdorferi gained through the application of molecular genetics. *Adv Appl Microbiol*. 41-143. doi: 10.1016/B978-0-12-800262-9.00002-0.

48. Hoon-Hanks, L. L., Morton, E. A., Lybecker, M. C., Battisti, J. M., Samuels, D. S., & Drecktrah, D. (2012). Borrelia burgdorferi malQ mutants utilize disaccharides and traverse the enzootic cycle. *FEMS immunology and medical microbiology*, *66*(2), 157–165. https://doi.org/10.1111/j.1574-695X.2012.00996.x

49. Pappas, C. J., Iyer, R., Petzke, M. M., Caimano, M. J., Radolf, J. D., & Schwartz, I. (2011). Borrelia burgdorferi requires glycerol for maximum fitness during the tick phase of the enzootic cycle. *PLoS pathogens*, *7*(7), e1002102. https://doi.org/10.1371/journal.ppat.1002102

50. Tilly, K., Grimm, D., Bueschel, D. M., Krum, J. G., & Rosa, P. (2004). Infectious cycle analysis of a Borrelia burgdorferi mutant defective in transport of chitobiose, a tick cuticle component. *Vector borne and zoonotic diseases (Larchmont, N.Y.)*, *4*(2), 159–168. https://doi.org/10.1089/1530366041210738

51. Kazmierczak, M. J., Wiedmann, M., & Boor, K. J. (2005). Alternative sigma factors and their roles in bacterial virulence. *Microbiology and molecular biology reviews : MMBR*, *69*(4), 527–543. https://doi.org/10.1128/MMBR.69.4.527-543.2005

52. Ouyang, Z., Blevins, G.S. & Norgard, M.V. (2008). Transcriptional inter- play among the regulators Rrp2, RpoN and RpoS in *Borrelia burgdorferi. Microbiol*. 154, 2641–2658 doi: 10.1099/mic.0.2008/019992-0

53. Ouyang, Z., Narasimhan, S., Neelakanta, G., Kumar, M., Pal, U., Fikrig, E., & Norgard, M. V. (2012). Activation of the RpoN-RpoS regulatory pathway during the enzootic life cycle

of Borrelia burgdorferi. *BMC microbiology*, *12*, 44. https://doi.org/10.1186/1471-2180-12-44

54. Cabello, F. C., Godfrey, H. P., Bugrysheva, J. V., & Newman, S. A. (2017). Sleeper cells: the stringent response and persistence in the Borreliella (Borrelia) burgdorferi enzootic cycle. *Environmental microbiology*, *19*(10), 3846–3862. https://doi.org/10.1111/1462-2920.13897
55. Haseltine, W. A., & Block, R. (1973). Synthesis of guanosine tetra- and pentaphosphate requires the presence of a codon-specific, uncharged transfer ribonucleic acid in the acceptor site of ribosomes. *Proceedings of the National Academy of Sciences of the United States of America*, *70*(5), 1564–1568. https://doi.org/10.1073/pnas.70.5.1564
56. De Keersmaecker, S.C., Sonck, K. & Vanderleyden J. (2006). Let LuxS speak up in AI-2 signaling. *Trends Microbiol*. 14(3):114-9. doi: 10.1016/j.tim.2006.01.003. Epub 2006 Feb 3. PMID: 16459080.
57. Stevenson, B., von Lackum, K., Wattier, R. L., McAlister, J. D., Miller, J. C., & Babb, K. (2003). Quorum sensing by the Lyme disease spirochete. *Microbes and infection*, *5*(11), 991–997. https://doi.org/10.1016/s1286-4579(03)00184-9
58. Fraser, C. M., Casjens, S., Huang, W. M., Sutton, G. G., Clayton, R., Lathigra, R., White, O., Ketchum, K. A., Dodson, R., Hickey, E. K., Gwinn, M., Dougherty, B., Tomb, J. F., Fleischmann, R. D., Richardson, D., Peterson, J., Kerlavage, A. R., Quackenbush, J., Salzberg, S., Hanson, M., … Venter, J. C. (1997). Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi. *Nature*, *390*(6660), 580–586. https://doi.org/10.1038/37551
59. Casjens, S., Palmer, N., van Vugt, R., Huang, W. M., Stevenson, B., Rosa, P., Lathigra, R., Sutton, G., Peterson, J., Dodson, R. J., Haft, D., Hickey, E., Gwinn, M., White, O., & Fraser, C. M. (2000). A bacterial genome in flux: the twelve linear and nine circular extrachromosomal DNAs in an infectious isolate of the Lyme disease spirochete Borrelia burgdorferi. *Molecular microbiology*, *35*(3), 490–516. https://doi.org/10.1046/j.1365-2958.2000.01698.x
60. Schwartz, I., Margos, G., Casjens, S. R., Qiu, W. G., & Eggers, C. H. (2021). Multipartite Genome of Lyme Disease *Borrelia*: Structure, Variation and Prophages. *Current issues in molecular biology*, *42*, 409–454. https://doi.org/10.21775/cimb.042.409

61. 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

62. Verhey, T. B., Castellanos, M., & Chaconas, G. (2019). Antigenic variation in the Lyme spirochete: detailed functional assessment of recombinational switching at vlsE in the JD1 strain of Borrelia burgdorferi. *Molecular microbiology*, *111*(3), 750–763. https://doi.org/10.1111/mmi.14189

63. Caimano, M. J., Iyer, R., Eggers, C. H., Gonzalez, C., Morton, E. A., Gilbert, M. A., Schwartz, I., & Radolf, J. D. (2007). Analysis of the RpoS regulon in Borrelia burgdorferi in response to mammalian host signals provides insight into RpoS function during the enzootic cycle. *Molecular microbiology*, *65*(5), 1193–1217. https://doi.org/10.1111/j.1365-2958.2007.05860.x

64. Tilly, K., Checroun, C., & Rosa, P. A. (2012). Requirements for Borrelia burgdorferi plasmid maintenance. *Plasmid*, *68*(1), 1–12. https://doi.org/10.1016/j.plasmid.2012.01.009

65. Caimano, M. J., Groshong, A. M., Belperron, A., Mao, J., Hawley, K. L., Luthra, A., Graham, D. E., Earnhart, C. G., Marconi, R. T., Bockenstedt, L. K., Blevins, J. S., & Radolf, J. D. (2019). The RpoS Gatekeeper in *Borrelia burgdorferi*: An Invariant Regulatory Scheme That Promotes Spirochete Persistence in Reservoir Hosts and Niche Diversity. *Frontiers in microbiology*, *10*, 1923. https://doi.org/10.3389/fmicb.2019.01923

66. Brüggemann, H., Hagman, A., Jules, M., Sismeiro, O., Dillies, M. A., Gouyette, C., Kunst, F., Steinert, M., Heuner, K., Coppée, J. Y., & Buchrieser, C. (2006). Virulence strategies for infecting phagocytes deduced from the in vivo transcriptional program of Legionella pneumophila. *Cellular microbiology*, *8*(8), 1228–1240. https://doi.org/10.1111/j.1462-5822.2006.00703.x

67. Hindré, T., Brüggemann, H., Buchrieser, C., & Héchard, Y. (2008). Transcriptional profiling of Legionella pneumophila biofilm cells and the influence of iron on biofilm formation. *Microbiology (Reading, England)*, *154*(Pt 1), 30–41. https://doi.org/10.1099/mic.0.2007/008698-0

68. Oliva, G., Sahr, T., & Buchrieser, C. (2018). The Life Cycle of *L. pneumophila*: Cellular Differentiation Is Linked to Virulence and Metabolism. *Frontiers in cellular and infection microbiology*, *8*, 3. https://doi.org/10.3389/fcimb.2018.00003
69. Gober, J. W., & Marques, M. V. (1995). Regulation of cellular differentiation in Caulobacter crescentus. *Microbiological reviews*, *59*(1), 31–47. https://doi.org/10.1128/mr.59.1.31-47.1995
70. Sézary, J.L. (1907). Des formes atypiques et dégénératives du tréponéme pâle [The atypical and degenerative forms of Treponema pallidum]. *Bull Mem Soc Med Hop Paris*. 24:114. French.
71. Herxheimer, K. (1905). Zur kenntnis der Spirochaeta pallida [Knowledge advances on pallidum spirochetes]. *München Med Wochshr* (journal out of print). 53:310–312.
72. Umemoto, T., Namikawa, I., Yoshii, Z., & Konishi, H. (1982). An internal view of the spherical body of Treponema macrodentium as revealed by scanning electron microscopy. *Microbiology and immunology*, *26*(3), 191–198. https://doi.org/10.1111/j.1348-0421.1982.tb00171.x
73. Umemoto, T., Namikawa, I., & Yamamoto, M. (1984). Colonial morphology of treponemes observed by electron microscopy. *Microbiology and immunology*, *28*(1), 11–22. https://doi.org/10.1111/j.1348-0421.1984.tb02943.x
74. Helaine, S., & Kugelberg, E. (2014). Bacterial persisters: formation, eradication, and experimental systems. *Trends in microbiology*, *22*(7), 417–424. https://doi.org/10.1016/j.tim.2014.03.008
75. Harms, A., Maisonneuve, E., & Gerdes, K. (2016). Mechanisms of bacterial persistence during stress and antibiotic exposure. *Science (New York, N.Y.)*, *354*(6318), aaf4268. https://doi.org/10.1126/science.aaf4268
76. Michiels, J. E., Van den Bergh, B., Verstraeten, N., & Michiels, J. (2016). Molecular mechanisms and clinical implications of bacterial persistence. *Drug resistance updates : reviews and commentaries in antimicrobial and anticancer chemotherapy*, *29*, 76–89. https://doi.org/10.1016/j.drup.2016.10.002
77. Ayrapetyan, M., Williams, T., & Oliver, J. D. (2018). Relationship between the Viable but Nonculturable State and Antibiotic Persister Cells. *Journal of bacteriology*, *200*(20), e00249-18. https://doi.org/10.1128/JB.00249-18

78. Rudenko, N., Golovchenko, M., Kybicova, K., & Vancova, M. (2019). Metamorphoses of Lyme disease spirochetes: phenomenon of Borrelia persisters. *Parasites & vectors*, *12*(1), 237. https://doi.org/10.1186/s13071-019-3495-7

79. Xu, H. S., Roberts, N., Singleton, F. L., Attwell, R. W., Grimes, D. J., & Colwell, R. R. (1982). Survival and viability of nonculturableEscherichia coli andVibrio cholerae in the estuarine and marine environment. *Microbial ecology*, *8*(4), 313–323. https://doi.org/10.1007/BF02010671

80. Mali, S., Mitchell, M., Havis, S., Bodunrin, A., Rangel, J., Olson, G., Widger, W. R., & Bark, S. J. (2017). A Proteomic Signature of Dormancy in the Actinobacterium Micrococcus luteus. *Journal of bacteriology*, *199*(14), e00206-17. https://doi.org/10.1128/JB.00206-17

81. Oliver J. D. (2010). Recent findings on the viable but nonculturable state in pathogenic bacteria. *FEMS microbiology reviews*, *34*(4), 415–425. https://doi.org/10.1111/j.1574-6976.2009.00200.x

82. Balaban, N. Q., Merrin, J., Chait, R., Kowalik, L., & Leibler, S. (2004). Bacterial persistence as a phenotypic switch. *Science (New York, N.Y.)*, *305*(5690), 1622–1625. https://doi.org/10.1126/science.1099390

83. Fisher, R. A., Gollan, B., & Helaine, S. (2017). Persistent bacterial infections and persister cells. *Nature reviews. Microbiology*, *15*(8), 453–464. https://doi.org/10.1038/nrmicro.2017.42

84. Brorson, Ø., Brorson, S. H., Scythes, J., MacAllister, J., Wier, A., & Margulis, L. (2009). Destruction of spirochete Borrelia burgdorferi round-body propagules (RBs) by the antibiotic tigecycline. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(44), 18656–18661. https://doi.org/10.1073/pnas.0908236106

85. Timmaraju, V. A., Theophilus, P. A., Balasubramanian, K., Shakih, S., Luecke, D. F., & Sapi, E. (2015). Biofilm formation by Borrelia burgdorferi sensu lato. *FEMS microbiology letters*, *362*(15), fnv120. https://doi.org/10.1093/femsle/fnv120

86. Feng, J., Zhang, S., Shi, W., & Zhang, Y. (2016). Ceftriaxone Pulse Dosing Fails to Eradicate Biofilm-Like Microcolony *B. burgdorferi* Persisters Which Are Sterilized by Daptomycin/ Doxycycline/Cefuroxime without Pulse Dosing. *Frontiers in microbiology*, *7*, 1744. https://doi.org/10.3389/fmicb.2016.01744

87. Nocton, J. J., Dressler, F., Rutledge, B. J., Rys, P. N., Persing, D. H., & Steere, A. C. (1994). Detection of Borrelia burgdorferi DNA by polymerase chain reaction in synovial fluid from patients with Lyme arthritis. *The New England journal of medicine*, *330*(4), 229–234. https://doi.org/10.1056/NEJM199401273300401

88. Li, X., McHugh, G. A., Damle, N., Sikand, V. K., Glickstein, L., & Steere, A. C. (2011). Burden and viability of Borrelia burgdorferi in skin and joints of patients with erythema migrans or lyme arthritis. *Arthritis and rheumatism*, *63*(8), 2238–2247. https://doi.org/10.1002/art.30384

89. Zhang Y. (2014). Persisters, persistent infections and the Yin-Yang model. *Emerging microbes & infections*, *3*(1), e3. https://doi.org/10.1038/emi.2014.3

90. Barthold, S. W., Hodzic, E., Imai, D. M., Feng, S., Yang, X., & Luft, B. J. (2010). Ineffectiveness of tigecycline against persistent Borrelia burgdorferi. *Antimicrobial agents and chemotherapy*, *54*(2), 643–651. https://doi.org/10.1128/AAC.00788-09

91. Feng, J., Shi, W., Zhang, S., & Zhang, Y. (2015). Persister mechanisms in Borrelia burgdorferi: implications for improved intervention. *Emerging microbes & infections*, *4*(8), e51. https://doi.org/10.1038/emi.2015.51

92. Vesey, P. M., & Kuramitsu, H. K. (2004). Genetic analysis of Treponema denticola ATCC 35405 biofilm formation. *Microbiology (Reading, England)*, *150*(Pt 7), 2401–2407. https://doi.org/10.1099/mic.0.26816-0

93. Ristow, P., Bourhy, P., Kerneis, S., Schmitt, C., Prevost, M. C., Lilenbaum, W., & Picardeau, M. (2008). Biofilm formation by saprophytic and pathogenic leptospires. *Microbiology (Reading, England)*, *154*(Pt 5), 1309–1317. https://doi.org/10.1099/mic.0.2007/014746-0

94. Barbour A. G. (1984). Isolation and cultivation of Lyme disease spirochetes. *The Yale journal of biology and medicine*, *57*(4), 521–525.

95. Domingue, G. J., Sr, & Woody, H. B. (1997). Bacterial persistence and expression of disease. *Clinical microbiology reviews*, *10*(2), 320–344. https://doi.org/10.1128/CMR.10.2.320

96. Stricker, R. B., & Johnson, L. (2011). Lyme disease: the next decade. *Infection and drug resistance*, *4*, 1–9. https://doi.org/10.2147/IDR.S15653

97. Margulis, L., Maniotis, A. & MacAllister, J. (2009). Spirochete round bodies Syphilis, Lyme disease & AIDS: Resurgence of "the great imitator"?. *Symbiosis*. 47, 51–58 https://doi.org/10.1007/BF03179970

98. Alban, P. S., Johnson, P. W., & Nelson, D. R. (2000). Serum-starvation-induced changes in protein synthesis and morphology of Borrelia burgdorferi. *Microbiology (Reading, England)*, *146 ( Pt 1)*, 119–127. https://doi.org/10.1099/00221287-146-1-119

99. Hulínská, D., Barták, P., Hercogová, J., Hancil, J., Basta, J., & Schramlová, J. (1994). Electron microscopy of Langerhans cells and Borrelia burgdorferi in Lyme disease patients. *Zentralblatt fur Bakteriologie : international journal of medical microbiology*, *280*(3), 348–359. https://doi.org/10.1016/s0934-8840(11)80597-9

100. Brorson, O., & Brorson, S. H. (1998). In vitro conversion of Borrelia burgdorferi to cystic forms in spinal fluid, and transformation to mobile spirochetes by incubation in BSK-H medium. *Infection*, *26*(3), 144–150. https://doi.org/10.1007/BF02771839

101. Murgia, R., Piazzetta, C., & Cinco, M. (2002). Cystic forms of Borrelia burgdorferi sensu lato: induction, development, and the role of RpoS. *Wiener klinische Wochenschrift*, *114*(13-14), 574–579.

102. Sapi, E., Kaur, N., Anyanwu, S., Luecke, D. F., Datar, A., Patel, S., Rossi, M., & Stricker, R. B. (2011). Evaluation of in-vitro antibiotic susceptibility of different morphological forms of Borrelia burgdorferi. *Infection and drug resistance*, *4*, 97–113. https://doi.org/10.2147/IDR.S19201

103. Miklossy, J., Kasas, S., Zurn, A. D., McCall, S., Yu, S., & McGeer, P. L. (2008). Persisting atypical and cystic forms of Borrelia burgdorferi and local inflammation in Lyme neuroborreliosis. *Journal of neuroinflammation*, *5*, 40. https://doi.org/10.1186/1742-2094-5-40

104. Groshong, A. M., Dey, A., Bezsonova, I., Caimano, M. J., & Radolf, J. D. (2017). Peptide Uptake Is Essential for *Borrelia burgdorferi* Viability and Involves Structural and Regulatory Complexity of its Oligopeptide Transporter. *mBio*, *8*(6), e02047-17. https://doi.org/10.1128/mBio.02047-17

105. Kersten, A., Poitschek, C., Rauch, S., & Aberer, E. (1995). Effects of penicillin, ceftriaxone, and doxycycline on morphology of Borrelia burgdorferi. *Antimicrobial agents and chemotherapy*, *39*(5), 1127–1133. https://doi.org/10.1128/AAC.39.5.1127

106. Duray, P. H., Yin, S. R., Ito, Y., Bezrukov, L., Cox, C., Cho, M. S., Fitzgerald, W., Dorward, D., Zimmerberg, J., & Margolis, L. (2005). Invasion of human tissue ex vivo by Borrelia burgdorferi. *The Journal of infectious diseases*, *191*(10), 1747–1754. https://doi.org/10.1086/429632

107. Dunham-Ems, S. M., Caimano, M. J., Eggers, C. H., & Radolf, J. D. (2012). Borrelia burgdorferi requires the alternative sigma factor RpoS for dissemination within the vector during tick-to-mammal transmission. *PLoS pathogens*, *8*(2), e1002532. https://doi.org/10.1371/journal.ppat.1002532

108. Janus, I., Noszczyk-Nowak, A., Nowak, M., Cepiel, A., Ciaputa, R., Pasławska, U., Dzięgiel, P., & Jabłońska, K. (2014). Myocarditis in dogs: etiology, clinical and histopathological features (11 cases: 2007-2013). *Irish veterinary journal*, *67*(1), 28. https://doi.org/10.1186/s13620-014-0028-8

109. Aberer, E., Kersten, A., Klade, H., Poitschek, C., & Jurecka, W. (1996). Heterogeneity of Borrelia burgdorferi in the skin. *The American Journal of dermatopathology*, *18*(6), 571–579. https://doi.org/10.1097/00000372-199612000-00004

110. Strnad, M., Grubhoffer, L., & Rego, R. (2020). Novel targets and strategies to combat borreliosis. *Applied microbiology and biotechnology*, *104*(5), 1915–1925. https://doi.org/10.1007/s00253-020-10375-8

111. Al-Robaiy, S., Dihazi, H., Kacza, J., Seeger, J., Schiller, J., Huster, D., Knauer, J., Straubinger, R.K. (2010). Metamorphosis of Borrelia burgdorferi organisms--RNA, lipid and protein composition in context with the spirochetes' shape. *J Basic Microbiol*. 50 Suppl 1:S5-17. doi: 10.1002/jobm.201000074. PMID: 20967786.

112. Lawrence, C., Lipton, R. B., Lowy, F. D., & Coyle, P. K. (1995). Seronegative chronic relapsing neuroborreliosis. *European neurology*, *35*(2), 113–117. https://doi.org/10.1159/000117104

113. Meriläinen, L., Brander, H., Herranen, A., Schwarzbach, A., & Gilbert, L. (2016). Pleomorphic forms of Borrelia burgdorferi induce distinct immune responses. *Microbes and infection*, *18*(7-8), 484–495. https://doi.org/10.1016/j.micinf.2016.04.002

114. Gu, L., Tseng, S., Horner, R. M., Tam, C., Loda, M., & Rollins, B. J. (2000). Control of TH2 polarization by the chemokine monocyte chemoattractant protein-1. *Nature*, *404*(6776), 407–411. https://doi.org/10.1038/35006097

115. Brown, C. R., Blaho, V. A., & Loiacono, C. M. (2003). Susceptibility to experimental Lyme arthritis correlates with KC and monocyte chemoattractant protein-1 production in joints and requires neutrophil recruitment via CXCR2. *Journal of immunology (Baltimore, Md. : 1950)*, *171*(2), 893–901. https://doi.org/10.4049/jimmunol.171.2.893

116. Deshmane, S. L., Kremlev, S., Amini, S., & Sawaya, B. E. (2009). Monocyte chemoattractant protein-1 (MCP-1): an overview. *Journal of interferon & cytokine research : the official journal of the International Society for Interferon and Cytokine Research*, *29*(6), 313–326. https://doi.org/10.1089/jir.2008.0027

117. Jan A. T. (2017). Outer Membrane Vesicles (OMVs) of Gram-negative Bacteria: A Perspective Update. *Frontiers in microbiology*, *8*, 1053. https://doi.org/10.3389/fmicb.2017.01053

118. Toledo, A., Coleman, J. L., Kuhlow, C. J., Crowley, J. T., & Benach, J. L. (2012). The enolase of Borrelia burgdorferi is a plasminogen receptor released in outer membrane vesicles. *Infection and immunity*, *80*(1), 359–368. https://doi.org/10.1128/IAI.05836-11

119. Li, Z., Clarke, A. J., & Beveridge, T. J. (1998). Gram-negative bacteria produce membrane vesicles which are capable of killing other bacteria. *Journal of bacteriology*, *180*(20), 5478–5483. https://doi.org/10.1128/JB.180.20.5478-5483.1998

120. Malge, A., Ghai, V., Reddy, P. J., Baxter, D., Kim, T. K., Moritz, R. L., & Wang, K. (2018). mRNA transcript distribution bias between Borrelia burgdorferi bacteria and their outer membrane vesicles. *FEMS microbiology letters*, *365*(13), fny135. https://doi.org/10.1093/femsle/fny135

121. Shoberg, R. J., & Thomas, D. D. (1993). Specific adherence of Borrelia burgdorferi extracellular vesicles to human endothelial cells in culture. *Infection and immunity*, *61*(9), 3892–3900. https://doi.org/10.1128/iai.61.9.3892-3900.1993

122. Whitmire, W. M., & Garon, C. F. (1993). Specific and nonspecific responses of murine B cells to membrane blebs of Borrelia burgdorferi. *Infection and immunity*, *61*(4), 1460–1467. https://doi.org/10.1128/iai.61.4.1460-1467.1993

123. Song, T., Duperthuy, M., & Wai, S. N. (2016). Sub-Optimal Treatment of Bacterial Biofilms. *Antibiotics (Basel, Switzerland)*, *5*(2), 23. https://doi.org/10.3390/antibiotics5020023

124. Vlamakis, H., Aguilar, C., Losick, R., & Kolter, R. (2008). Control of cell fate by the formation of an architecturally complex bacterial community. *Genes & development*, *22*(7), 945–953. https://doi.org/10.1101/gad.1645008

125. Vlamakis, H., Chai, Y., Beauregard, P., Losick, R., & Kolter, R. (2013). Sticking together: building a biofilm the Bacillus subtilis way. *Nature reviews. Microbiology*, *11*(3), 157–168. https://doi.org/10.1038/nrmicro2960

126. Donlan R. M. (2002). Biofilms: microbial life on surfaces. *Emerging infectious diseases*, *8*(9), 881–890. https://doi.org/10.3201/eid0809.020063

127. O'Toole, G., Kaplan, H. B., & Kolter, R. (2000). Biofilm formation as microbial development. *Annual review of microbiology*, *54*, 49–79. https://doi.org/10.1146/annurev.micro.54.1.49

128. Flemming, H. C., & Wuertz, S. (2019). Bacteria and archaea on Earth and their abundance in biofilms. *Nature reviews. Microbiology*, *17*(4), 247–260. https://doi.org/10.1038/s41579-019-0158-9

129. Sapi, E., Bastian, S. L., Mpoy, C. M., Scott, S., Rattelle, A., Pabbati, N., Poruri, A., Burugu, D., Theophilus, P. A., Pham, T. V., Datar, A., Dhaliwal, N. K., MacDonald, A., Rossi, M. J., Sinha, S. K., & Luecke, D. F. (2012). Characterization of biofilm formation by Borrelia burgdorferi in vitro. *PloS one*, *7*(10), e48277. https://doi.org/10.1371/journal.pone.0048277

130. Futo, M., Opašić, L., Koska, S., Čorak, N., Široki, T., Ravikumar, V., Thorsell, A., Lenuzzi, M., Kifer, D., Domazet-Lošo, M., Vlahoviček, K., Mijakovic, I., & Domazet-Lošo, T. (2021). Embryo-Like Features in Developing Bacillus subtilis Biofilms. *Molecular biology and evolution*, *38*(1), 31–47. https://doi.org/10.1093/molbev/msaa217

131. Jefferson, K.K. (2004). What drives bacteria to produce biofilm? *FEMS Microbiol Lett*. 236:163–73. doi: 1 0.1111/j.1574-6968.2004.tb09643.

132. Sapi, E., Theophilus, P. A., Pham, T. V., Burugu, D., & Luecke, D. F. (2016). Effect of RpoN, RpoS and LuxS Pathways on the Biofilm Formation and Antibiotic Sensitivity of *Borrelia Burgdorferi*. *European journal of microbiology & immunology*, *6*(4), 272–286. https://doi.org/10.1556/1886.2016.00026

133. Okshevsky, M., & Meyer, R. L. (2015). The role of extracellular DNA in the establishment, maintenance and perpetuation of bacterial biofilms. *Critical reviews in microbiology*, *41*(3), 341–352. https://doi.org/10.3109/1040841X.2013.841639

134. Srivastava, S. Y., & de Silva, A. M. (2009). Characterization of Borrelia burgdorferi aggregates. *Vector borne and zoonotic diseases (Larchmont, N.Y.)*, *9*(3), 323–329. https://doi.org/10.1089/vbz.2008.0148

135. Sapi, E., Kasliwala, R. S., Ismail, H., Torres, J. P., Oldakowski, M., Markland, S., Gaur, G., Melillo, A., Eisendle, K., Liegner, K. B., Libien, J., & Goldman, J. E. (2019). The Long-Term Persistence of *Borrelia burgdorferi* Antigens and DNA in the Tissues of a Patient with Lyme Disease. *Antibiotics (Basel, Switzerland)*, *8*(4), 183. https://doi.org/10.3390/antibiotics8040183

136. Di Domenico, E. G., Cavallo, I., Bordignon, V., D'Agosto, G., Pontone, M., Trento, E., Gallo, M. T., Prignano, G., Pimpinelli, F., Toma, L., & Ensoli, F. (2018). The Emerging Role of Microbial Biofilm in Lyme Neuroborreliosis. *Frontiers in neurology*, *9*, 1048. https://doi.org/10.3389/fneur.2018.01048

137. Sapi, E., Balasubramanian, K., Poruri, A., Maghsoudlou, J. S., Socarras, K. M., Timmaraju, A. V., Filush, K. R., Gupta, K., Shaikh, S., Theophilus, P. A., Luecke, D. F., MacDonald, A., & Zelger, B. (2016). Evidence of In Vivo Existence of Borrelia Biofilm in Borrelial Lymphocytomas. *European journal of microbiology & immunology*, *6*(1), 9–24. https://doi.org/10.1556/1886.2015.00049

138. Feng, J., Shi, W., Miklossy, J., Tauxe, G. M., McMeniman, C. J., & Zhang, Y. (2018). Identification of Essential Oils with Strong Activity against Stationary Phase *Borrelia burgdorferi*. *Antibiotics (Basel, Switzerland)*, *7*(4), 89. https://doi.org/10.3390/antibiotics7040089

139. Grothe, A.N. (2019). Identifying factors controlling cell shape and virulence gene expression in *Borrelia burgdorferi* [dissertation]. Department of Microbiology and Immunology: Indiana Univesity.

140. Lantos, P. M., Auwaerter, P. G., & Wormser, G. P. (2014). A systematic review of Borrelia burgdorferi morphologic variants does not support a role in chronic Lyme disease. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America*, *58*(5), 663–671. https://doi.org/10.1093/cid/cit810

141. Onwuamaegbu, M. E., Belcher, R. A., & Soare, C. (2005). Cell wall-deficient bacteria as a cause of infections: a review of the clinical significance. *The Journal of international medical research*, *33*(1), 1–20. https://doi.org/10.1177/147323000503300101

142. Schnell, B., Staubli, T., Harris, N. L., Rogler, G., Kopf, M., Loessner, M. J., & Schuppler, M. (2014). Cell-wall deficient L. monocytogenes L-forms feature abrogated pathogenicity. *Frontiers in cellular and infection microbiology*, *4*, 60. https://doi.org/10.3389/fcimb.2014.00060

143. Bushnell, B. (2014). BBMap: A Fast, Accurate, Splice-Aware Aligner. https://www.osti.gov/servlets/purl/1241166.

144. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., & 1000 Genome Project Data Processing Subgroup (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics (Oxford, England)*, *25*(16), 2078–2079. https://doi.org/10.1093/bioinformatics/btp352

145. R Development Core Team (2019). R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing.

146. Morgan, M., Pagès, H., Obenchain, V. & Hayden, N. (2021). Rsamtools: Binary alignment (BAM), FASTA, variant call (BCF), and tabix file import. R package version 2.10.0, https://bioconductor.org/packages/Rsamtools.

147. Lawrence, M., Huber, W., Pagès, H., Aboyoun, P., Carlson, M., Gentleman, R., Morgan, M. T., & Carey, V. J. (2013). Software for computing and annotating genomic ranges. *PLoS computational biology*, *9*(8), e1003118. https://doi.org/10.1371/journal.pcbi.1003118

148. Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome biology*, *15*(12), 550. https://doi.org/10.1186/s13059-014-0550-8

149. Wickham, H. (2016). ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York. ISBN 978-3-319-24277-4, https://ggplot2.tidyverse.org.

150. Li, B., Ruotti, V., Stewart, R. M., Thomson, J. A., & Dewey, C. N. (2010). RNA-Seq gene expression estimation with read mapping uncertainty. *Bioinformatics (Oxford, England)*, *26*(4), 493–500. https://doi.org/10.1093/bioinformatics/btp692

151. McDowell, I. C., Manandhar, D., Vockley, C. M., Schmid, A. K., Reddy, T. E., & Engelhardt, B. E. (2018). Clustering gene expression time series data using an infinite

Gaussian process mixture model. *PLoS computational biology*, *14*(1), e1005896. https://doi.org/10.1371/journal.pcbi.1005896

152. Domazet-Loso, T., Brajković, J., & Tautz, D. (2007). A phylostratigraphy approach to uncover the genomic history of major adaptations in metazoan lineages. *Trends in genetics : TIG*, *23*(11), 533–539. https://doi.org/10.1016/j.tig.2007.08.014
153. Margos, G., Gofton, A., Wibberg, D., Dangel, A., Marosevic, D., Loh, S. M., Oskam, C., & Fingerle, V. (2018). The genus Borrelia reloaded. *PloS one*, *13*(12), e0208432. https://doi.org/10.1371/journal.pone.0208432
154. Parks, D. H., Chuvochina, M., Waite, D. W., Rinke, C., Skarshewski, A., Chaumeil, P. A., & Hugenholtz, P. (2018). A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. *Nature biotechnology*, *36*(10), 996–1004. https://doi.org/10.1038/nbt.4229
155. Mukherjee, S., Seshadri, R., Varghese, N. J., Eloe-Fadrosh, E. A., Meier-Kolthoff, J. P., Göker, M., Coates, R. C., Hadjithomas, M., Pavlopoulos, G. A., Paez-Espino, D., Yoshikuni, Y., Visel, A., Whitman, W. B., Garrity, G. M., Eisen, J. A., Hugenholtz, P., Pati, A., Ivanova, N. N., Woyke, T., Klenk, H. P., … Kyrpides, N. C. (2017). 1,003 reference genomes of bacterial and archaeal isolates expand coverage of the tree of life. *Nature biotechnology*, *35*(7), 676–683. https://doi.org/10.1038/nbt.3886
156. Hug, L. A., Baker, B. J., Anantharaman, K., Brown, C. T., Probst, A. J., Castelle, C. J., Butterfield, C. N., Hernsdorf, A. W., Amano, Y., Ise, K., Suzuki, Y., Dudek, N., Relman, D. A., Finstad, K. M., Amundson, R., Thomas, B. C., & Banfield, J. F. (2016). A new view of the tree of life. *Nature microbiology*, *1*, 16048. https://doi.org/10.1038/nmicrobiol.2016.48
157. Raymann, K., Brochier-Armanet, C., & Gribaldo, S. (2015). The two-domain tree of life is linked to a new root for the Archaea. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(21), 6670–6675. https://doi.org/10.1073/pnas.1420858112
158. Di, L., Pagan, P. E., Packer, D., Martin, C. L., Akther, S., Ramrattan, G., Mongodin, E. F., Fraser, C. M., Schutzer, S. E., Luft, B. J., Casjens, S. R., & Qiu, W. G. (2014). BorreliaBase: a phylogeny-centered browser of Borrelia genomes. *BMC bioinformatics*, *15*, 233. https://doi.org/10.1186/1471-2105-15-233

159. Gupta, R.S., Mahmood, S. & Adeolu, M. (2013). A phylogenomic and molecular signature based approach for characterization of the phylum Spirochaetes and its major clades: proposal for a taxonomic revision of the phylum. *Frontiers in microbiology*. 217. https://doi.org/10.3389/fmicb.2013.00217.

160. Rinke, C., Schwientek, P., Sczyrba, A., Ivanova, N. N., Anderson, I. J., Cheng, J. F., Darling, A., Malfatti, S., Swan, B. K., Gies, E. A., Dodsworth, J. A., Hedlund, B. P., Tsiamis, G., Sievert, S. M., Liu, W. T., Eisen, J. A., Hallam, S. J., Kyrpides, N. C., Stepanauskas, R., Rubin, E. M., … Woyke, T. (2013). Insights into the phylogeny and coding potential of microbial dark matter. *Nature*, *499*(7459), 431–437. https://doi.org/10.1038/nature12352

161. Wu, D., Hugenholtz, P., Mavromatis, K., Pukall, R., Dalin, E., Ivanova, N. N., Kunin, V., Goodwin, L., Wu, M., Tindall, B. J., Hooper, S. D., Pati, A., Lykidis, A., Spring, S., Anderson, I. J., D'haeseleer, P., Zemla, A., Singer, M., Lapidus, A., Nolan, M., … Eisen, J. A. (2009). A phylogeny-driven genomic encyclopaedia of Bacteria and Archaea. *Nature*, *462*(7276), 1056–1060. https://doi.org/10.1038/nature08656

162. Richter, D., Postic, D., Sertour, N., Livey, I., Matuschka, F. R., & Baranton, G. (2006). Delineation of Borrelia burgdorferi sensu lato species by multilocus sequence analysis and confirmation of the delineation of Borrelia spielmanii sp. nov. *International journal of systematic and evolutionary microbiology*, *56*(Pt 4), 873–881. https://doi.org/10.1099/ijs.0.64050-0

163. Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V., & Zdobnov, E. M. (2015). BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics (Oxford, England)*, *31*(19), 3210–3212. https://doi.org/10.1093/bioinformatics/btv351

164. Yates, A. D., Achuthan, P., Akanni, W., Allen, J., Allen, J., Alvarez-Jarreta, J., Amode, M. R., Armean, I. M., Azov, A. G., Bennett, R., Bhai, J., Billis, K., Boddu, S., Marugán, J. C., Cummins, C., Davidson, C., Dodiya, K., Fatima, R., Gall, A., Giron, C. G., … Flicek, P. (2020). Ensembl 2020. *Nucleic acids research*, *48*(D1), D682–D688. https://doi.org/10.1093/nar/gkz966

165. Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of molecular biology*, *215*(3), 403–410. https://doi.org/10.1016/S0022-2836(05)80360-2

166. Tatusov, R. L., Galperin, M. Y., Natale, D. A., & Koonin, E. V. (2000). The COG database: a tool for genome-scale analysis of protein functions and evolution. *Nucleic acids research*, *28*(1), 33–36. https://doi.org/10.1093/nar/28.1.33

167. Huerta-Cepas, J., Szklarczyk, D., Heller, D., Hernández-Plaza, A., Forslund, S. K., Cook, H., Mende, D. R., Letunic, I., Rattei, T., Jensen, L. J., von Mering, C., & Bork, P. (2019). eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. *Nucleic acids research*, *47*(D1), D309–D314. https://doi.org/10.1093/nar/gky1085

168. Huerta-Cepas, J., Forslund, K., Coelho, L. P., Szklarczyk, D., Jensen, L. J., von Mering, C., & Bork, P. (2017). Fast Genome-Wide Functional Annotation through Orthology Assignment by eggNOG-Mapper. *Molecular biology and evolution*, *34*(8), 2115–2122. https://doi.org/10.1093/molbev/msx148

169. Benjamini, Y. & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society Series B (Methodological), 57, 289-300. doi: http://dx.doi.org/10.2307/2346101

170. Arnold, W. K., Savage, C. R., Brissette, C. A., Seshu, J., Livny, J., & Stevenson, B. (2016). RNA-Seq of Borrelia burgdorferi in Multiple Phases of Growth Reveals Insights into the Dynamics of Gene Expression, Transcriptome Architecture, and Noncoding RNAs. *PloS one*, *11*(10), e0164165. https://doi.org/10.1371/journal.pone.0164165

171. Wu, Q., Guan, G., Liu, Z., Li, Y., Luo, J., & Yin, H. (2015). RNA-Seq-based analysis of changes in Borrelia burgdorferi gene expression linked to pathogenicity. *Parasites & vectors*, *8*, 155. https://doi.org/10.1186/s13071-014-0623-2

172. Schuwirth, B. S., Borovinskaya, M. A., Hau, C. W., Zhang, W., Vila-Sanjurjo, A., Holton, J. M., & Cate, J. H. (2005). Structures of the bacterial ribosome at 3.5 A resolution. *Science (New York, N.Y.)*, *310*(5749), 827–834. https://doi.org/10.1126/science.1117230

173. Zaborowski, A. B., & Walther, D. (2020). Determinants of correlated expression of transcription factors and their target genes. *Nucleic acids research*, *48*(20), 11347–11369. https://doi.org/10.1093/nar/gkaa927

174. Byrgazov, K., Vesper, O., & Moll, I. (2013). Ribosome heterogeneity: another level of complexity in bacterial translation regulation. *Current opinion in microbiology*, *16*(2), 133–139. https://doi.org/10.1016/j.mib.2013.01.009
175. Deusser E. (1972). Heterogeneity of ribosomal populations in Escherichia coli cells grown in different media. *Molecular & general genetics : MGG*, *119*(3), 249–258. https://doi.org/10.1007/BF00333862
176. Deusser, E., & Wittmann, H. G. (1972). Ribosomal proteins: variation of the protein composition in Escherichia coli ribosomes as function of growth rate. *Nature*, *238*(5362), 269–270. https://doi.org/10.1038/238269a0
177. Persing, D. H., Rutledge, B. J., Rys, P. N., Podzorski, D. S., Mitchell, P. D., Reed, K. D., Liu, B., Fikrig, E., & Malawista, S. E. (1994). Target imbalance: disparity of Borrelia burgdorferi genetic material in synovial fluid from Lyme arthritis patients. *The Journal of infectious diseases*, *169*(3), 668–672. https://doi.org/10.1093/infdis/169.3.668
178. Schooling, S. R., & Beveridge, T. J. (2006). Membrane vesicles: an overlooked component of the matrices of biofilms. *Journal of bacteriology*, *188*(16), 5945–5957. https://doi.org/10.1128/JB.00257-06
179. Klimentová, J., & Stulík, J. (2015). Methods of isolation and purification of outer membrane vesicles from gram-negative bacteria. *Microbiological research*, *170*, 1–9. https://doi.org/10.1016/j.micres.2014.09.006
180. Molin, S., & Tolker-Nielsen, T. (2003). Gene transfer occurs with enhanced efficiency in biofilms and induces enhanced stabilisation of the biofilm structure. *Current opinion in biotechnology*, *14*(3), 255–261. https://doi.org/10.1016/s0958-1669(03)00036-3
181. Gupta R. S. (2019). Distinction between Borrelia and Borreliella is more robustly supported by molecular and phenotypic characteristics than all other neighbouring prokaryotic genera: Response to Margos' et al. "The genus Borrelia reloaded" (PLoS ONE 13(12): e0208432). *PloS one*, *14*(8), e0221397. https://doi.org/10.1371/journal.pone.0221397
182. Domazet-Lošo, T., Carvunis, A. R., Albà, M. M., Šestak, M. S., Bakaric, R., Neme, R., & Tautz, D. (2017). No Evidence for Phylostratigraphic Bias Impacting Inferences on Patterns of Gene Emergence and Evolution. Molecular biology and evolution, 34(4), 843–856. https://doi.org/10.1093/molbev/msw284

183. Shi, L., Derouiche, A., Pandit, S., Rahimi, S., Kalantari, A., Futo, M., Ravikumar, V., Jers, C., Mokkapati, V., Vlahoviček, K., & Mijakovic, I. (2020). Evolutionary Analysis of the Bacillus subtilis Genome Reveals New Genes Involved in Sporulation. *Molecular biology and evolution*, *37*(6), 1667–1678. https://doi.org/10.1093/molbev/msaa035
184. Hengge-Aronis R. (2002). Signal transduction and regulatory mechanisms involved in control of the sigma(S) (RpoS) subunit of RNA polymerase. *Microbiology and molecular biology reviews : MMBR*, *66*(3), 373–395. https://doi.org/10.1128/MMBR.66.3.373-395.2002
185. Caimano M. J. (2018). Generation of Mammalian Host-Adapted Borrelia burgdorferi by Cultivation in Peritoneal Dialysis Membrane Chamber Implantation in Rats. *Methods in molecular biology (Clifton, N.J.)*, *1690*, 35–45. https://doi.org/10.1007/978-1-4939-7383-5_3
186. Oppler, Z. J., O'Keeffe, K. R., McCoy, K. D., & Brisson, D. (2021). Evolutionary Genetics of *Borrelia*. *Current issues in molecular biology*, *42*, 97–112. https://doi.org/10.21775/cimb.042.097

# 8. Summary

*Borrelia burgdorferi* is a spirochete bacterium that causes tick-borne Lyme disease. In laboratory cultures, *B. burgdorferi* develops several pleomorphic forms (morphotypes). Functional and structural differences between some of the morphotypes have been studied before, but expression changes at the transcriptome level have never been investigated. To address this problem, spirochete, round body, bleb, and biofilm cultures were grown, their RNA was harvested, and transcriptomes were recovered by RNAseq profiling. The results have shown that spirochetes and round bodies, despite their morphological differences, share similar expression profiles. In contrast, blebs and biofilms showed a significant difference in expression patterns in comparison to spirochetes and round bodies. Regardless of the overall transcriptional similarity to spirochetes, the genes upregulated in round bodies are enriched with translation functions. Although the total number of upregulated genes is much higher in blebs and biofilms compared to round bodies, their function is mainly unknown. Interestingly, the genes that are upregulated in round bodies tend to be localized on the chromosome, while the genes that are upregulated in blebs and biofilms primarily derive from *B. burgdorferi* plasmids. To discern evolutionary imprints of differentially expressed genes, evolutionary age was assigned to *B. burgdorferi* genes by phylostratigraphic approach. The results demonstrated that round body upregulated genes are enriched for evolutionary old genes common to all life, while the genes upregulated in blebs and biofilms are evolutionary young and specific for *Borreliaceae*. It is possible to conclude that spirochete to round body transition relies on the delicate regulation of a relatively small number of highly evolutionary conserved genes involved in translation, while spirochete to bleb and biofilm transition includes substantial reshaping of transcription profiles towards evolutionary young genes of yet unknown function.

# 9. Sažetak

Bakterija *Borrelia burgdorferi* primarni je uzročnik Lajamske bolesti koji se vektorski se prenosi putem krpelja. Genom ove bakterije podijeljen je u bakterijski kromosom i 21 cirkularni ili linearni plazmid. Važna značajka životnog ciklusa ove bakterije je da ovisno o fazi ciklusa obitava u krpelju ili u nekoj vrsti sisavaca. Jedan od ključnih transkripcijskih faktora koji reguliraju ovaj proces je sigma faktor RpoS. Ekspresija proteina RpoS u bakteriji *B. burgdorferi* pojačana je u tijeku infekcije u sisavcu gdje RpoS pojačava transkripciju gena bitnih za virulenciju. S obzirom na to da *B. burgdorferi* predstavlja rastući zdravstveni problem, javlja se potreba za boljim razumijevanjem kompleksnih svojstava ove bakterije. Morfološka plastičnost značajka je individualnih bakterijskih stanica da mijenjaju morfologiju ovisno o okolišnim čimbenicima. Kao i mnoge druge bakterije, *B. burgdorferi* također posjeduje ovo svojstvo. U staničnim kulturama ove bakterije često je prisutno više morfoloških oblika (morfotipova), a njihov omjer se mijenja ovisno o uzgojnim uvjetima. U standardnim uzgojnim uvjetima većina stanica *B. burgdorferi* se poprima oblik spiroheta. Ovaj morfotip stoga je najbolje istražen i opisan. Za razliku od spiroheta, stanična svojstva i molekularne osobine ostalih morfotipova kao što su okrugla tjelešca, mjehuraste forme i biofilmovi većinski su neistražena. Poznato je da okrugla tjelešca imaju fleksibilnu vanjsku ovojnicu i smanjenu metaboličku aktivnost u odnosu na spirohete. Također, u staničnim kulturama humanih imunoloških stanica, okrugla tjelešca i spirohete uzrokuju različit imunološki odgovor. Karakteristično svojstvo mjehurastih formi je formiranje vezikula koje se odvajaju od površine vanjske ovojnice, a važno je naglasiti da opisane vezikule sadrže velik udio genskih produkata važnih za virulenciju. I naposljetku, biofilmovi bakterije *B. burgdorferi* sastoje se of skupine spiroheta, okruglih tjelešaca i mjehurastih formi koji su uklopljeni u ekstracelularni matriks bogat alginatom, kalcijem i ekstracelularnom DNA.

Iako mnoga istraživanja upućuju na biološku važnost ovih morfotipova, njihova uloga u životnom ciklusu *B. burgdorferi* i razvoju Lajmske bolesti nije poznata. Dosadašnja istraživanja bila su usmjerena prema razvoju metoda pogodnih za indukciju ciljanih morfotipova, određivanju njihovog udjela u heterogenim kulturama, utvrđivanju vijabilnosti, osjetljivosti na antibiotike te pronalaženja razlika u sastavu proteoma putem dvodimenzionalne gel-elektroforeze. Iako su mnoge funkcionalne i strukturne karakteristike nekih od morfotipova već prethodno opisane,

razlike u ekspresiji na razini transkriptoma do sada nisu istražene. S obzirom na to, primarni cilj ovog rada bilo je detektirati razlike u transkripciji između morfotipova bakterije *B. burgdorferi*.

S tim ciljem uzgojene su laboratorijske kulture različitih morfotipova (spiroheta, okruglih tjelešaca, mjehurastih formi i biofilmova), izolirana je ukupna RNA, te je ista sekvencirana na osnovu čega su skupljeni podaci o razlici u ekspresiji na razini transkriptoma. Kako bi se definiralo evolucijsko podrijetlo diferencijalno eksprimiranih gena bakterije *B. burgdorferi* korištena je genomska filostratigrafija. Uzgoj kultura obogaćenih morfotipovima bakterije *B. burgdorferi*, mikroskopsku analizu i izolaciju RNA proveli su moji suradnici iz BCA klinike locirane u Augsburgu u Njemačkoj. Ja osobno sam organizirala sekvenciranje RNA, provela mapiranje RNA sekvenci na referentni genom, kvantificirala mapirana očitanja, provela analizu transkriptomskih podataka, filostratigrafsku analizu i ostale oblike bioinformatičke analize.

Analiza transkripcijskih podataka podrazumijeva usporedbu ekspresijskih vrijednosti gena u morfotipovima putem analize glavnih komponenti, te usporednu analizu transkripcijskih razlika između okruglih tjelešaca i spiroheta, mjehurastih formi i spiroheta, te biofilmova i spiroheta. Osim toga, provedeno je i grupiranje gena bakterije *B. burgdorferi* na osnovu njihove ekspresije u morfotipovima. Također je provedena vizualizacija standardiziranih ekspresijskih vrijednosti pojedinačnih gena, kao i grupa u kojima se nalaze. U svrhu filostratigrafske analize izrađena je i referentna filogenija s bakterijom *B. burgdorferi* kao fokalnom vrstom. Čvorovi koji se nalaze u sklopu te filogenije odabrani su s obzirom na relevantnu literaturu, važnost evolucijskih prijelaza i dovršenost funkcionalnih anotacija. Na osnovu referentne filogenije primjenom filostratigrafskog pristupa, geni bakterije *B. burgdorferi* podijeljeni su u 8 starosnih kategorija (filostratuma). Distribucija funkcionalnih anotacija gena koji su diferencijalno eksprimirani u okruglim tjelešcima, mjehurastim formama i biofilmovima uspoređena je s distribucijom funkcionalnih anotaciju u genomu bakterije *B. burgdorferi*, te je statistička značajnost uočenih razlika procijenjena putem hipergeometrijskog testa. Istom metodom provjerena je statistička značajnost uočenih razlika u lokaciji gena koji su pojačano eksprimirani u okruglim tjelešcima, mjehurastim formama i biofilmovima u odnosu na lokaciju svih gena u genomu *B. burgdorferi*. Na isti način je testirana i statistička značajnost filogenetske pripadnosti gena koji su pojačano eksprimirani u okruglim tjelešcima, mjehurastim formama i biofilmovima u odnosu na filogenetsku pripadnost svih gena u genomu *B. burgdorferi*. Dodatno, vizualizirane su standardizirane ekspresijske vrijednosti gena

reguliranih putem transkripcijskog regulatora RpoS, te je izračunata statistička značajnost filogenetske pripadnosti opisanih gena u odnosu na filogenetsku pripadnost svih gena u genomu *B. burgdorferi*.

Rezultati ovog istraživanja otkrivaju da unatoč morfološkim razlikama, spirohete i okrugla tjelešca imaju sličan ekspresijski profil, dok mjehuraste forme i biofilmovi pokazuju značajnu razliku u ekspresiji kako međusobno tako i u odnosu na spirohete i okrugla tjelešca. Unatoč tome što okrugla tjelešca imaju transkripciju sličnu spirohetama, geni koji su pojačano eksprimirani u okruglim tjelešcima su obogaćeni genima uključenima u procese translacije. Iako mjehuraste forme i biofilmovi imaju veći broj pojačano eksprimiranih gena o odnosu na okrugla tjelešca, funkcija tih gena je pretežito nepoznata. Također, geni koji su pojačano eksprimirani u okruglim tjelešcima su pretežito locirani na bakterijskom kromosomu, dok se pojačano eksprimirani geni u mjehurastim formama i biofilmovima uglavnom nalaze na plazmidima. Primjenom genomske filostratigrafije otkriveno je da su geni koji su pojačano eksprimirani u okruglim tjelešcima obogaćeni evolucijski starim genima karakterističnim za sve stanične organizme, dok su geni pojačano eksprimirani u mjehurastim formama i biofilmovima evolucijski mlađi i specifični za porodicu Borreliaceae. Kako bi se provjerila značajnost ovih rezultata, ista analiza je provedena uz korištenje različitih e-vrijednosti za traženje homologa u korištenoj proteomskoj bazi. Navedena analiza je potvrdila značajnost rezultata dobivenih filostratigrafskom analizom. Distribucija funkcionalno anotiranih gena po filostratumima pokazala je da su evolucijski stariji geni pretežito dobro anotirani, dok broj funkcionalno anotiranih gena značajno pada kako prilazimo mlađim filostratumima, broj funkcionalno anotiranih gena značajno opada.

Geni koji su diferencijalno eksprimirani među morfotipovima *B. burgdorferi* grupirani su s obzirom na razlike u njihovoj ekspresiji. Tom metodom geni su grupirani u 22 grupe. Među njima, grupa 1 je obogaćena genima koji sudjeluju u biogenezi stanične stijenke, membrane ili ovojnice. Grupa 5 sadrži značajan udio gena koji sudjeluju u replikaciji, rekombinaciji i popravku nukleinskih kiselina, kao i gena koji su uključeni u transport i metabolizam lipida. Grupa 16 je obogaćena genima uključenim u translaciju i biogenezu ribosoma ili kodiraju za strukturne komponente ribosoma. Naposljetku, grupe 3, 9 i 11 obogaćene su genima kojima funkcija nije poznata. Prilikom analize transkripcijskih razlika između morfotipova, utvrđeno je da se protein koji se nalazi na vanjskoj ovojnici OspC, transkripcijski regulator RpoN, varijabilni površinski

antigen VlsE, proteini koji vežu dekorin DbpA i DbpB, S-ribozilhomocitein liaza LuxS, protien koji veže fibronektin RevA i proteini koji vežu plazminogen ErpM i ErpY značajno razlikuju u ekspresiji. Također, proteini koji se nalaze na vanjskoj ovojnici OspA i OspB, te transkripcijski regulator RpoS nisu diferencijalno eksprimirani među morfotipovima. Analizom ekspresije gena koji su pojačano eksprimirani djelovanjem transkripcijskog regulatora RpoS, utvrđeno je da je većina takvih gena grupirana u grupu 15. Grupa 15 karakteristična je po tome što sadrži gene čija ekspresija je veća u mjehurastim formama nego u ostalim morfotipovima. Dodatno, filostratigrafskom analizom gena koji su pojačano eksprimirani putem transkripcijskog regulatora RpoS, utvrđeno je da većina gena koji su regulirani na opisani nastala prilikom razvoja Borreliacea.

S obzirom na distribuciju funkcionalnih anotacija među genima koji su diferencijalno eksprimirani između okruglih tjelešaca i spiroheta, zaključila sam da okrugla tjelešca i spirohete imaju različite proteine u sastavu svojih ribosoma. S obzirom na to da varijacije u proteinskom sastavu ribosoma reguliraju ekspresiju gena na razini translacije, razlike u ekspresiji gena koji kodiraju za strukturne komponente ribosoma potencijalno mogu objasniti male razlike u transkriptomima spiroheta i okruglih tjelešaca. Kako je pokazano da vezikule koje se odvajaju od vanjske stanične ovojnice sadrže produkte kodirane genima koji se nalaze na plazmidima, a geni koji su pojačano eksprimirani u mjehurastim tvorbama i biofilmovima sadrže značajan broj gena koji se nalaze na plazmidima, postoji mogućnost da su vezikule koje se odvajaju od vanjske ovojnice uzrok pojačane ekspresije gena kodiranih na plazmidima kod mjehurastih tvorbi i biofilmova. Dodatno, na osnovu filostratigrafske analize, kao i činjenice da su Borreliaceae u tijeku evolucijskog razvoja razvile prilagodbe na bifazni životni ciklus, moguće je pretpostaviti da su mnogi geni koji su pojačano eksprimirani u mjehurastim tvorbama i biofilmovima vjerojatno ključni za adaptaciju *B. burgdorferi* na bifazni životni ciklus. Analizom ekspresije gena koji su regulirani transkripcijskim faktorom RpoS, pokazano je da je značajan broj gena koji su regulirani tim transkripcijskom faktorom grupiran u grupi gena koji imaju veću ekspresiju u mjehurastim tvorbama nego li i u jednom drugom morfotipu. S obzirom na to da transkripcijski faktor RpoS pojačava ekspresiju gena koji su nužni za infekciju sisavaca bakterijom *B. burgdorferi*, moguće je pretpostaviti da su mjehuraste forme važne za virulenciju bakterije *B. burgdorferi*.

Ova doktorska disertacija istražuje diferencijalnu ekspresiju i evolucijske značajke morfotipova bakterije *Borrelia burgdorferi*. Proučavani morfotipovi uključuju spirohete, okrugla tjelešca, mjehuraste tvorbe i biofilmove. Rezultati istraživanja upućuju na to da se transkriptom mjehurastih tvorbi i biofilmova vidno razlikuje od transkriptoma spiroheta i okruglih tjelešca, dok spirohete i okrugla tjelešca imaju sličan transkriptom. Također, pokazano je da je tranzicija spiroheta u okrugla tjelešca primarno popraćena promjenom ekspresije gena koji su kodirani na bakterijskom kromosomu, dok je tranzicija spiroheta u mjehuraste tvorbe i tranzicija spiroheta u biofilmove popraćena promjenama u ekspresiji gena kodiranih na plazmidima. Također, filostratigrafska analiza je pokazala kako su geni čija se ekspresija mijenja tijekom tranzicije spiroheta u okrugla tjelešca evolucijski mlađeg podrijetla, dok je većina gena čija je ekspresija pojačana u mjehurastim tvorbama i biofilmovima u odnosu na spirohete nastala prilikom razvoja Borreliaceae. Značajan broj gena koji su pojačano eksprimirani u okruglim tjelešcima u odnosu na spirohete sudjeluju u translaciji, kodiraju proteine koji su strukturne komponente ribosoma ili sudjeluju u biogenezi ribosoma. S druge strane, funkcija gena koji su pojačano eksprimirani u mjehurastim tvorbama i biofilmovima u odnosu na spirohete nije poznata. Ovi rezultati značajno doprinose razumijevanju morfotipova bakterije *B. burgdorferi*, kao i bakterijskog pleomorfizma općenito.

# 10. Abbreviations

| | |
|---|---|
| BF | Biofilm |
| BL | Bleb |
| BSK | Barbour-Stoenner-Kelly |
| BUSCO | Benchmarking Universal Single-Copy Orthologs |
| CLD | Chronic Lyme Disease |
| COG | Clusters of Orthologous Genes |
| DMC | Dialysis membrane chamber |
| eDNA | Extracellular DNA |
| FC | Fold change |
| LRT | Likelihood ration test |
| mRNA | Messenger RNA |
| OMV | Outer membrane vesicles |
| PCA | Principal component analysis |
| PH | Phase-contrast |
| PTLDS | Post-Treatment Lyme Disease Syndrome |
| RB | Round body |
| SP | Spirochete |
| VBNC | Viable but nonculturable |

# 11. Supplements

**Table S1.** Transcriptome read counts used for mapping on the *Borellia burgdorferi* B31 genome sequence. Abbreviations SP1, SP2, SP3 stand for spirochete replicates, abbreviations RB1, RB2, RB3 stand for round body replicates, abbreviations BL1, BL2 stand for bleb replicates, and abbreviations BF1, BF2, and BF3 stand for biofilm replicates.

| Morphotype replicate | Reads used for mapping | Total number of mapped reads | Percentage of mapped reads (%) | Number of bases used for mapping | Total number of mapped bases | Protein coding nucleotides in the *B. burgdorferi* genome | Coverage |
|---|---|---|---|---|---|---|---|
| **SP1** | 80628390 | 76424906 | 94.79 | 6450271200 | 5811813280 | 1260954 | 4609.060505 |
| **SP2** | 89764630 | 84958944 | 94.65 | 7181170400 | 6447147040 | 1260954 | 5112.91216 |
| **SP3** | 66473270 | 63960466 | 96.22 | 5317861600 | 4979484160 | 1260954 | 3948.981612 |
| **RB1** | 88408228 | 82495969 | 93.31 | 7072658240 | 6365817600 | 1260954 | 5048.41382 |
| **RB2** | 10211821 0 | 99790148 | 97.72 | 8169456800 | 7906756640 | 1260954 | 6270.456052 |
| **RB3** | 78169592 | 72485324 | 92.73 | 6253567360 | 5759493600 | 1260954 | 4567.568365 |
| **BL1** | 77770810 | 71676484 | 92.16 | 6221664800 | 5421871360 | 1260954 | 4299.816932 |
| **BL2** | 85133958 | 76646629 | 90.03 | 6810716640 | 5891320160 | 1260954 | 4672.113463 |
| **BF1** | 11794577 2 | 11273882 7 | 95.59 | 9435661760 | 8724904640 | 1260954 | 6919.288602 |
| **BF2** | 11139754 2 | 10606354 0 | 95.21 | 8911803360 | 8171744800 | 1260954 | 6480.605002 |
| **BF3** | 83869240 | 79806479 | 95.16 | 6709539200 | 6159292640 | 1260954 | 4884.62913 |
| **AVERAGE** | **89243603.82** | **84277065.09** | **94.32373411** | **7139488305** | **6512695084** | **1260954** | **5164.895059** |
| **SUM** | **981679642** | **927047716** | | | | | |

**Table S3.** Distribution of *Borrelia burgdorferi B31* genes across clusters based on their normalized expression. Percentage of cluster-specific genes among protein-coding genes and percentage of cluster-specific genes among RpoS-upregulated genes. The clustering was performed by DP_GP_cluster algorithm (McDowell et al., 2018) with the maximum Gibbs sampling iterations set to 500.

| Cluster number | Number of genes | Percentage cluster genes among protein coding genes | Percentage cluster genes among RpoS-upregulated genes | Percentage of COG annotated genes |
|---|---|---|---|---|
| 1 | 102 | 7.6% | 4% | 72.5% |
| 2 | 45 | 3.3% | 4% | 84.4% |
| 3 | 169 | 12.5% | 27% | 40.2% |
| 4 | 124 | 9.2% | 0% | 57.3% |
| 5 | 47 | 3.5% | 0% | 89.4% |
| 6 | 86 | 6.4% | 4% | 86.0% |
| 7 | 69 | 5.1% | 0% | 62.3% |
| 8 | 62 | 4.6% | 0% | 59.7% |
| 9 | 177 | 13.1% | 2% | 40.0% |
| 10 | 9 | 0.7% | 0% | 88.9% |
| 11 | 27 | 2% | 8% | 29.6% |
| 12 | 14 | 1% | 0% | 71.4% |
| 13 | 55 | 4.1% | 2% | 58.2% |
| 14 | 21 | 1.6% | 0% | 33.3% |
| 15 | 107 | 7.9% | 35% | 44.9% |
| 16 | 49 | 3.6% | 0% | 71.4% |
| 17 | 20 | 1.5% | 2% | 90% |
| 18 | 22 | 1.6% | 2% | 77.3% |
| 19 | 6 | 0.4% | 2% | 50.0% |
| 20 | 15 | 1.1% | 0% | 80.0% |
| 21 | 42 | 3.1% | 6% | 50.0% |
| 22 | 8 | 0.6% | 0% | 87.5% |

**Table S3.** Distribution of *Borrelia burgdorferi* B31 genes across phylostrata (ps) made by BLASTp with different e-value cutoffs.

| **ps** | **1E-30** | **1E-20** | **1E-15** | **1E-10** | **1E-5** | **1E-3** | **1E-2** | **1E-1** | **1** | **10** |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 286 | 354 | 406 | 456 | 535 | 557 | 582 | 615 | 760 | 1144 |
| 2 | 122 | 155 | 152 | 158 | 138 | 148 | 158 | 177 | 193 | 100 |
| 3 | 31 | 25 | 26 | 18 | 22 | 24 | 19 | 25 | 64 | 32 |
| 4 | 8 | 10 | 16 | 12 | 18 | 18 | 19 | 19 | 14 | 5 |
| 5 | 72 | 63 | 66 | 64 | 53 | 46 | 42 | 35 | 25 | 7 |
| 6 | 585 | 551 | 522 | 504 | 464 | 446 | 423 | 387 | 235 | 38 |
| 7 | 174 | 135 | 116 | 102 | 92 | 88 | 86 | 74 | 46 | 18 |
| 8 | 42 | 44 | 37 | 27 | 22 | 17 | 16 | 13 | 8 | 1 |
| **total** | 1320 | 1337 | 1341 | 1341 | 1344 | 1344 | 1345 | 1345 | 1345 | 1345 |

**Table S4.** Functional enrichment of COG functional annotations in genes more expressed in round bodies than spirochetes. For a gene to be categorized as differentially expressed, the adjusted p-value ($p_{ajd}$) calculated by DeSeq2 pairwise comparison analysis had to be below 0.05. q = the number of annotations in round body upregulated genes with a particular COG term, s = the total number of annotations in round body upregulated genes for all COG terms, h = the number of annotations in *B. burgdorferi* genome with a particular COG term, t = the total number of annotations in *B. burgdorferi* genome for all COG terms. The abbreviation –inf marks it was not possible to calculate the log-odds.

| COG annotation | Annotation description | q | s | h | t | $p_{adj}$ | log-odds |
|---|---|---|---|---|---|---|---|
| C | Energy production and conversion | 2 | 46 | 26 | 1387 | 0.902 | 0.914 |
| D | Cell cycle control, cell division, chromosome partitioning | 0 | 46 | 42 | 1387 | 0.902 | -inf |
| E | Amino acid transport and metabolism | 2 | 46 | 27 | 1387 | 0.902 | 0.872 |
| F | Nucleotide transport and metabolism | 0 | 46 | 45 | 1387 | 0.902 | -inf |
| G | Carbohydrate transport and metabolism | 0 | 46 | 47 | 1387 | 0.902 | -inf |
| H | Coenzyme transport and metabolism | 0 | 46 | 23 | 1387 | 1 | -inf |
| I | Lipid transport and metabolism | 1 | 46 | 18 | 1387 | 1 | 0.549 |
| J | Translation, ribosomal structure and biogenesis | 22 | 46 | 121 | 1387 | 4.053E-11 | 2.442 |
| K | Transcription | 4 | 46 | 21 | 1387 | 0.053 | 2.004 |
| L | Replication, recombination and repair | 4 | 46 | 64 | 1387 | 0.902 | 0.710 |
| M | Cell wall/membrane/envelope biogenesis | 1 | 46 | 53 | 1387 | 1 | -0.596 |
| N | Cell motility | 1 | 46 | 53 | 1387 | 1 | -0.596 |
| O | Post-translational modification, protein turnover, and chaperones | 0 | 46 | 28 | 1387 | 1 | -inf |
| P | Inorganic ion transport and metabolism | 0 | 46 | 35 | 1387 | 1 | -inf |
| Q | Secondary metabolites biosynthesis, transport, and catabolism | 0 | 46 | 5 | 1387 | 1 | -inf |
| S | Function unknown | 5 | 46 | 716 | 1387 | 7.539 E-08 | -2.225 |
| T | Signal transduction mechanisms | 1 | 46 | 35 | 1387 | 1 | -0.156 |
| U | Intracellular trafficking, secretion, and vesicular transport | 3 | 46 | 20 | 1387 | 0.249 | 1.693 |
| V | Defense mechanisms | 0 | 46 | 8 | 1387 | 1 | -inf |

**Table S5.** Functional enrichment of COG functional annotations in genes more expressed in blebs than spirochetes. For a gene to be categorized as differentially expressed, the adjusted p-value ($p_{ajd}$) calculated by DeSeq2 pairwise comparison analysis had to be below 0.05. q = the number of annotations in bleb upregulated genes with a particular COG term, s = the total number of annotations in bleb upregulated genes for all COG terms, h = the number of annotations in *B. burgdorferi* genome with a particular COG term, t = the total number of annotations in *B. burgdorferi* genome for all COG terms. The abbreviation –inf marks it was not possible to calculate the log-odds.

| COG annotation | Annotation description | q | s | h | t | $p_{adj}$ | log-odds |
|---|---|---|---|---|---|---|---|
| C | Energy production and conversion | 4 | 484 | 26 | 1387 | 0.151 | -1.097 |
| D | Cell cycle control, cell division, chromosome partitioning | 22 | 484 | 42 | 1387 | 0.125 | 0.743 |
| E | Amino acid transport and metabolism | 9 | 484 | 27 | 1387 | 1 | -0.071 |
| F | Nucleotide transport and metabolism | 16 | 484 | 45 | 1387 | 1 | 0.030 |
| G | Carbohydrate transport and metabolism | 12 | 484 | 47 | 1387 | 0.420 | -0.461 |
| H | Coenzyme transport and metabolism | 7 | 484 | 23 | 1387 | 1 | -0.206 |
| I | Lipid transport and metabolism | 3 | 484 | 18 | 1387 | 0.372 | -0.996 |
| J | Translation, ribosomal structure and biogenesis | 37 | 484 | 121 | 1387 | 0.598 | -0.214 |
| K | Transcription | 5 | 484 | 21 | 1387 | 0.639 | -0.547 |
| L | Replication, recombination and repair | 15 | 484 | 64 | 1387 | 0.166 | -0.584 |
| M | Cell wall/membrane/envelope biogenesis | 9 | 484 | 53 | 1387 | 0.036 | -0.995 |
| N | Cell motility | 3 | 484 | 53 | 1387 | 6.8753 E-06 | -2.241 |
| O | Post-translational modification, protein turnover, and chaperones | 6 | 484 | 28 | 1387 | 0.388 | -0.688 |
| P | Inorganic ion transport and metabolism | 12 | 484 | 35 | 1387 | 1 | -0.028 |
| Q | Secondary metabolites biosynthesis, transport, and catabolism | 1 | 484 | 5 | 1387 | 1 | -0.7658 |
| S | Function unknown | 309 | 484 | 716 | 1387 | 5.805 E-10 | 0.766 |
| T | Signal transduction mechanisms | 6 | 484 | 35 | 1387 | 0.125 | -0.972 |
| U | Intracellular trafficking, secretion, and vesicular transport | 6 | 484 | 20 | 1387 | 1 | -0.227 |
| V | Defense mechanisms | 2 | 484 | 8 | 1387 | 1 | -0.478 |

**Table S6.** Functional enrichment of COG functional annotations in genes more expressed in biofilms than spirochetes. For a gene to be categorized as differentially expressed, the adjusted p-value ($p_{ajd}$) calculated by DeSeq2 pairwise comparison analysis had to be below 0,05. q = the number of annotations in biofilm upregulated genes with a particular COG term, s = the total number of annotations in biofilm upregulated genes for all COG terms, h = the number of annotations in *B. burgdorferi* genome with a particular COG term, t = the total number of annotations in *B. burgdorferi* genome for all COG terms. The abbreviation –inf marks it was not possible to calculate the log-odds.

| COG annotation | Annotation description | q | s | h | t | $p_{adj}$ | log-odds |
|---|---|---|---|---|---|---|---|
| C | Energy production and conversion | 6 | 445 | 26 | 1387 | 0.885 | -0.462 |
| D | Cell cycle control, cell division, chromosome partitioning | 17 | 445 | 42 | 1387 | 0.752 | 0.376 |
| E | Amino acid transport and metabolism | 10 | 445 | 27 | 1387 | 0.967 | 0.224 |
| F | Nucleotide transport and metabolism | 13 | 445 | 45 | 1387 | 0.980 | -0.156 |
| G | Carbohydrate transport and metabolism | 15 | 445 | 47 | 1387 | 1 | -0.008 |
| H | Coenzyme transport and metabolism | 6 | 445 | 23 | 1387 | 0.966 | -0.296 |
| I | Lipid transport and metabolism | 2 | 445 | 18 | 1387 | 0.536 | -1.342 |
| J | Translation, ribosomal structure and biogenesis | 47 | 445 | 121 | 1387 | 0.536 | 0.326 |
| K | Transcription | 6 | 445 | 21 | 1387 | 1 | -0.169 |
| L | Replication, recombination and repair | 13 | 445 | 64 | 1387 | 0.536 | -0.643 |
| M | Cell wall/membrane/envelope biogenesis | 11 | 445 | 53 | 1387 | 0.536 | -0.610 |
| N | Cell motility | 13 | 445 | 53 | 1387 | 0.752 | -0.388 |
| O | Post-translational modification, protein turnover, and chaperones | 11 | 445 | 28 | 1387 | 0.909 | 0.321 |
| P | Inorganic ion transport and metabolism | 13 | 445 | 35 | 1387 | 0.966 | 0.230 |
| Q | Secondary metabolites biosynthesis, transport, and catabolism | 1 | 445 | 5 | 1387 | 1 | -0.638 |
| S | Function unknown | 243 | 445 | 716 | 1387 | 0.536 | 0.176 |
| T | Signal transduction mechanisms | 8 | 445 | 35 | 1387 | 0.752 | -0.477 |
| U | Intracellular trafficking, secretion, and vesicular transport | 6 | 445 | 20 | 1387 | 1 | -0.099 |
| V | Defense mechanisms | 4 | 445 | 8 | 1387 | 0.885 | 0.755 |

**Table S7.** Functional enrichment of COG functional annotations in genes more expressed in blebs than spirochetes. For a gene to be categorized as differentially expressed, the adjusted p-value ($p_{ajd}$) and absolute value of log2FC (|log2 FC|) calculated by DeSeq2 pairwise comparison analysis had to be below 0.05. q = the number of annotations in bleb upregulated genes with a particular COG term, s = the total number of annotations in bleb upregulated genes for all COG terms, h = the number of annotations in *B. burgdorferi* genome with a particular COG term, t = the total number of annotations in *B. burgdorferi* genome for all COG terms. The abbreviation –inf marks it was not possible to calculate the log-odds.

| COG annotation | Annotation description | q | s | h | t | $p_{adj}$ | log-odds |
|---|---|---|---|---|---|---|---|
| C | Energy production and conversion | 0 | 240 | 26 | 1387 | 0.038 | -inf |
| D | Cell cycle control, cell division, chromosome partitioning | 9 | 240 | 42 | 1387 | 0.727 | 0.274 |
| E | Amino acid transport and metabolism | 2 | 240 | 27 | 1387 | 0.441 | -0.975 |
| F | Nucleotide transport and metabolism | 8 | 240 | 45 | 1387 | 1 | 0.034 |
| G | Carbohydrate transport and metabolism | 2 | 240 | 47 | 1387 | 0.038 | -1.581 |
| H | Coenzyme transport and metabolism | 2 | 240 | 23 | 1387 | 0.591 | -0.797 |
| I | Lipid transport and metabolism | 1 | 240 | 18 | 1387 | 0.488 | -1.280 |
| J | Translation, ribosomal structure and biogenesis | 2 | 240 | 121 | 1387 | 2.590 E-07 | -2.623 |
| K | Transcription | 0 | 240 | 21 | 1387 | 0.085 | -inf |
| L | Replication, recombination and repair | 9 | 240 | 64 | 1387 | 0.727 | -0.257 |
| M | Cell wall/membrane/envelope biogenesis | 1 | 240 | 53 | 1387 | 0.004 | -2.429 |
| N | Cell motility | 0 | 240 | 53 | 1387 | 0.0004 | -inf |
| O | Post-translational modification, protein turnover, and chaperones | 4 | 240 | 28 | 1387 | 1 | -0.232 |
| P | Inorganic ion transport and metabolism | 2 | 240 | 35 | 1387 | 0.1619 | -1.260 |
| Q | Secondary metabolites biosynthesis, transport, and catabolism | 1 | 240 | 5 | 1387 | 1 | 0.179 |
| S | Function unknown | 197 | 240 | 716 | 1387 | 1.853 E-25 | 1.713 |
| T | Signal transduction mechanisms | 0 | 240 | 35 | 1387 | 0.009 | -inf |
| U | Intracellular trafficking, secretion, and vesicular transport | 0 | 240 | 20 | 1387 | 0.092 | -inf |
| V | Defense mechanisms | 0 | 240 | 8 | 1387 | 0.436 | -inf |

**Table S8.** Functional enrichment of COG functional annotations in genes more expressed in biofilms than spirochetes. For a gene to be categorized as differentially expressed, the adjusted p-value ($p_{ajd}$) and absolute value of log2FC (|log2 FC|) calculated by DeSeq2 pairwise comparison analysis had to be below 0.05. q = the number of annotations in biofilm upregulated genes with a particular COG term, s = the total number of annotations in biofilm upregulated genes for all COG terms, h = the number of annotations in *B. burgdorferi* genome with a particular COG term, t = the total number of annotations in *B. burgdorferi* genome for all COG terms. The abbreviation –inf marks it was not possible to calculate the log-odds.

| COG annotation | Annotation description | q | s | h | t | $p_{adj}$ | log-odds |
|---|---|---|---|---|---|---|---|
| C | Energy production and conversion | 0 | 137 | 26 | 1387 | 0.288 | -inf |
| D | Cell cycle control, cell division, chromosome partitioning | 6 | 137 | 42 | 1387 | 0.690 | 0.435 |
| E | Amino acid transport and metabolism | 1 | 137 | 27 | 1387 | 0.690 | -1.061 |
| F | Nucleotide transport and metabolism | 8 | 137 | 45 | 1387 | 0.288 | 0.710 |
| G | Carbohydrate transport and metabolism | 1 | 137 | 47 | 1387 | 0.238 | -1.648 |
| H | Coenzyme transport and metabolism | 0 | 137 | 23 | 1387 | 0.341 | -inf |
| I | Lipid transport and metabolism | 1 | 137 | 18 | 1387 | 0.972 | -0.629 |
| J | Translation, ribosomal structure and biogenesis | 5 | 137 | 121 | 1387 | 0.127 | -0.993 |
| K | Transcription | 0 | 137 | 21 | 1387 | 0.382 | -inf |
| L | Replication, recombination and repair | 2 | 137 | 64 | 1387 | 0.238 | -1.259 |
| M | Cell wall/membrane/envelope biogenesis | 0 | 137 | 53 | 1387 | 0.046 | -inf |
| N | Cell motility | 0 | 137 | 53 | 1387 | 0.046 | -inf |
| O | Post-translational modification, protein turnover, and chaperones | 4 | 137 | 28 | 1387 | 0.805 | 0.429 |
| P | Inorganic ion transport and metabolism | 4 | 137 | 35 | 1387 | 0.972 | 0.168 |
| Q | Secondary metabolites biosynthesis, transport, and catabolism | 0 | 137 | 5 | 1387 | 1 | -inf |
| S | Function unknown | 104 | 137 | 716 | 1387 | 2.895 E-08 | 1.189 |
| T | Signal transduction mechanisms | 0 | 137 | 35 | 1387 | 0.190 | -inf |
| U | Intracellular trafficking, secretion, and vesicular transport | 1 | 137 | 20 | 1387 | 0.972 | -0.742 |
| V | Defense mechanisms | 0 | 137 | 8 | 1387 | 0.972 | -inf |

**Table S9.** Functional enrichment of COG functional annotations in genes less expressed in round bodies than spirochetes. For a gene to be categorized as differentially expressed, the adjusted p-value ($p_{ajd}$) calculated by DeSeq2 pairwise comparison analysis had to be below 0.05. q = the number of annotations in round body downregulated genes with a particular COG term, s = the total number of annotations in round body downregulated genes for all COG terms, h = the number of annotations in *B. burgdorferi* genome with a particular COG term, t = the total number of annotations in *B. burgdorferi* genome for all COG terms. The abbreviation –inf marks it was not possible to calculate the log-odds.

| COG annotation | Annotation description | q | s | h | t | $p_{adj}$ | log-odds |
|---|---|---|---|---|---|---|---|
| C | Energy production and conversion | 0 | 23 | 26 | 1387 | 1 | -inf |
| D | Cell cycle control, cell division, chromosome partitioning | 1 | 23 | 42 | 1387 | 1 | 0.383 |
| E | Amino acid transport and metabolism | 1 | 23 | 27 | 1387 | 1 | 0.850 |
| F | Nucleotide transport and metabolism | 0 | 23 | 45 | 1387 | 1 | -inf |
| G | Carbohydrate transport and metabolism | 2 | 23 | 47 | 1387 | 1 | 1.027 |
| H | Coenzyme transport and metabolism | 0 | 23 | 23 | 1387 | 1 | -inf |
| I | Lipid transport and metabolism | 0 | 23 | 18 | 1387 | 1 | -inf |
| J | Translation, ribosomal structure and biogenesis | 1 | 23 | 121 | 1387 | 1 | -0.752 |
| K | Transcription | 0 | 23 | 21 | 1387 | 1 | -inf |
| L | Replication, recombination and repair | 0 | 23 | 64 | 1387 | 1 | -inf |
| M | Cell wall/membrane/envelope biogenesis | 0 | 23 | 53 | 1387 | 1 | -inf |
| N | Cell motility | 0 | 23 | 53 | 1387 | 1 | -inf |
| O | Post-translational modification, protein turnover, and chaperones | 0 | 23 | 28 | 1387 | 1 | -inf |
| P | Inorganic ion transport and metabolism | 1 | 23 | 35 | 1387 | 1 | 0.576 |
| Q | Secondary metabolites biosynthesis, transport, and catabolism | 1 | 23 | 5 | 1387 | 1 | 2.738 |
| S | Function unknown | 16 | 23 | 716 | 1387 | 1 | 0.774 |
| T | Signal transduction mechanisms | 0 | 23 | 35 | 1387 | 1 | -inf |
| U | Intracellular trafficking, secretion, and vesicular transport | 0 | 23 | 20 | 1387 | 1 | -inf |
| V | Defense mechanisms | 0 | 23 | 8 | 1387 | 1 | -inf |

**Table S10.** Functional enrichment of COG functional annotations in genes more less in blebs than spirochetes. For a gene to be categorized as differentially expressed, the adjusted p-value ($p_{ajd}$) calculated by DeSeq2 pairwise comparison analysis had to be below 0.05. q = the number of annotations in bleb downregulated genes with a particular COG term, s = the total number of annotations in bleb downregulated genes for all COG terms, h = the number of annotations in *B. burgdorferi* genome with a particular COG term, t = the total number of annotations in *B. burgdorferi* genome for all COG terms. The abbreviation –inf marks it was not possible to calculate the log-odds.

| COG annotation | Annotation description | q | s | h | t | $p_{adj}$ | log-odds |
|---|---|---|---|---|---|---|---|
| C | Energy production and conversion | 16 | 532 | 26 | 1387 | 0.072 | 1.0 |
| D | Cell cycle control, cell division, chromosome partitioning | 8 | 532 | 42 | 1387 | 0.041 | -1.0 |
| E | Amino acid transport and metabolism | 14 | 532 | 27 | 1387 | 0.267 | 0.6 |
| F | Nucleotide transport and metabolism | 19 | 532 | 45 | 1387 | 0.739 | 0.2 |
| G | Carbohydrate transport and metabolism | 23 | 532 | 47 | 1387 | 0.241 | 0.448 |
| H | Coenzyme transport and metabolism | 10 | 532 | 23 | 1387 | 0.759 | 0.216 |
| I | Lipid transport and metabolism | 11 | 532 | 18 | 1387 | 0.158 | 0.9 |
| J | Translation, ribosomal structure and biogenesis | 54 | 532 | 121 | 1387 | 0.241 | 0.3 |
| K | Transcription | 12 | 532 | 21 | 1387 | 0.212 | 0.8 |
| L | Replication, recombination and repair | 35 | 532 | 64 | 1387 | 0.041 | 0.7 |
| M | Cell wall/membrane/envelope biogenesis | 39 | 532 | 53 | 1387 | 2.27 E-06 | 1.559 |
| N | Cell motility | 36 | 532 | 53 | 1387 | 0.0001 | 1.3 |
| O | Post-translational modification, protein turnover, and chaperones | 16 | 532 | 28 | 1387 | 0.137 | 0.8 |
| P | Inorganic ion transport and metabolism | 20 | 532 | 35 | 1387 | 0.083 | 0.8 |
| Q | Secondary metabolites biosynthesis, transport, and catabolism | 0 | 532 | 5 | 1387 | 0.241 | -inf |
| S | Function unknown | 186 | 532 | 716 | 1387 | 2.262 E-21 | -1.1 |
| T | Signal transduction mechanisms | 21 | 532 | 35 | 1387 | 0.045 | 0.9 |
| U | Intracellular trafficking, secretion, and vesicular transport | 10 | 532 | 20 | 1387 | 0.469 | 0.5 |
| V | Defense mechanisms | 2 | 532 | 8 | 1387 | 0.739 | -0.6 |

**Table S11.** Functional enrichment of COG functional annotations in genes less expressed in biofilms than spirochetes. For a gene to be categorized as differentially expressed, the adjusted p-value ($p_{ajd}$) calculated by DeSeq2 pairwise comparison analysis had to be below 0.05. q = the number of annotations in biofilm downregulated genes with a particular COG term, s = the total number of annotations in biofilm downregulated genes for all COG terms, h = the number of annotations in *B. burgdorferi* genome with a particular COG term, t = the total number of annotations in *B. burgdorferi* genome for all COG terms. The abbreviation –inf marks it was not possible to calculate the log-odds.

| COG annotation | Annotation description | q | s | h | t | $p_{adj}$ | log-odds |
|---|---|---|---|---|---|---|---|
| C | Energy production and conversion | 18 | 463 | 26 | 1387 | 0.001 | 1.5 |
| D | Cell cycle control, cell division, chromosome partitioning | 4 | 463 | 42 | 1387 | 0.002 | -1.6 |
| E | Amino acid transport and metabolism | 12 | 463 | 27 | 1387 | 0.446 | 0.5 |
| F | Nucleotide transport and metabolism | 18 | 463 | 45 | 1387 | 0.573 | 0.3 |
| G | Carbohydrate transport and metabolism | 18 | 463 | 47 | 1387 | 0.712 | 0.2 |
| H | Coenzyme transport and metabolism | 13 | 463 | 23 | 1387 | 0.086 | 1.0 |
| I | Lipid transport and metabolism | 14 | 463 | 18 | 1387 | 0.001 | 2.0 |
| J | Translation, ribosomal structure and biogenesis | 42 | 463 | 121 | 1387 | 0.912 | 0.1 |
| K | Transcription | 10 | 463 | 21 | 1387 | 0.393 | 0.6 |
| L | Replication, recombination and repair | 39 | 463 | 64 | 1387 | 4.395 E-05 | 1.2 |
| M | Cell wall/membrane/envelope biogenesis | 35 | 463 | 53 | 1387 | 1.388 E-05 | 1.415 |
| N | Cell motility | 25 | 463 | 53 | 1387 | 0.099 | 0.6 |
| O | Post-translational modification, protein turnover, and chaperones | 14 | 463 | 28 | 1387 | 0.170 | 0.7 |
| P | Inorganic ion transport and metabolism | 13 | 463 | 35 | 1387 | 0.895 | 0.2 |
| Q | Secondary metabolites biosynthesis, transport, and catabolism | 1 | 463 | 5 | 1387 | 0.971 | -0.7 |
| S | Function unknown | 155 | 463 | 716 | 1387 | 1.926 E-20 | -1.1 |
| T | Signal transduction mechanisms | 17 | 463 | 35 | 1387 | 0.162 | 0.7 |
| U | Intracellular trafficking, secretion, and vesicular transport | 12 | 463 | 20 | 1387 | 0.068 | 1.1 |
| V | Defense mechanisms | 3 | 463 | 8 | 1387 | 1 | 0.2 |

**Table S12.** Functional enrichment of COG functional annotations in genes less expressed in blebs than spirochetes. For a gene to be categorized as differentially expressed, the adjusted p-value ($p_{ajd}$) and absolute value of log2FC (|log2 FC|) calculated by DeSeq2 pairwise comparison analysis had to be below 0.05 and more that 1, respectively. q = the number of annotations in bleb downregulated genes with a particular COG term, s = the total number of annotations in bleb downregulated genes for all COG terms, h = the number of annotations in *B. burgdorferi* genome with a particular COG term, t = the total number of annotations in *B. burgdorferi* genome for all COG terms. The abbreviation –inf marks it was not possible to calculate the log-odds.

| COG annotation | Annotation description | q | s | h | t | $p_{adj}$ | log-odds |
|---|---|---|---|---|---|---|---|
| C | Energy production and conversion | 5 | 143 | 26 | 1387 | 0.747 | 0.7 |
| D | Cell cycle control, cell division, chromosome partitioning | 1 | 143 | 42 | 1387 | 0.731 | -1.6 |
| E | Amino acid transport and metabolism | 2 | 143 | 27 | 1387 | 1 | -0.4 |
| F | Nucleotide transport and metabolism | 7 | 143 | 45 | 1387 | 0.747 | 0.5 |
| G | Carbohydrate transport and metabolism | 5 | 143 | 47 | 1387 | 1 | 0.04 |
| H | Coenzyme transport and metabolism | 1 | 143 | 23 | 1387 | 0.951 | -1.0 |
| I | Lipid transport and metabolism | 2 | 143 | 18 | 1387 | 1 | 0.1 |
| J | Translation, ribosomal structure and biogenesis | 9 | 143 | 121 | 1387 | 0.747 | -0.4 |
| K | Transcription | 2 | 143 | 21 | 1387 | 1 | -0.1 |
| L | Replication, recombination and repair | 10 | 143 | 64 | 1387 | 0.747 | 0.5 |
| M | Cell wall/membrane/envelope biogenesis | 11 | 143 | 53 | 1387 | 0.432 | 0.9 |
| N | Cell motility | 7 | 143 | 53 | 1387 | 0.951 | 0.3 |
| O | Post-translational modification, protein turnover, and chaperones | 4 | 143 | 28 | 1387 | 0.951 | 0.4 |
| P | Inorganic ion transport and metabolism | 6 | 143 | 35 | 1387 | 0.747 | 0.6 |
| Q | Secondary metabolites biosynthesis, transport, and catabolism | 0 | 143 | 5 | 1387 | 1 | -inf |
| S | Function unknown | 62 | 143 | 716 | 1387 | 0.432 | -0.4 |
| T | Signal transduction mechanisms | 5 | 143 | 35 | 1387 | 0.951 | 0.4 |
| U | Intracellular trafficking, secretion, and vesicular transport | 4 | 143 | 20 | 1387 | 0.747 | 0.8 |
| V | Defense mechanisms | 0 | 143 | 8 | 1387 | 1 | -inf |

**Table S13.** Functional enrichment of COG functional annotations in genes less expressed in biofilms than spirochetes. For a gene to be categorized as differentially expressed, the adjusted p-value ($p_{ajd}$) and absolute value of log2FC (|log2 FC|) calculated by DeSeq2 pairwise comparison analysis had to be below 0.05 and more that 1, respectively. q = the number of annotations in biofilm downregulated genes with a particular COG term, s = the total number of annotations in biofilm downregulated genes for all COG terms, h = the number of annotations in *B. burgdorferi* genome with a particular COG term, t = the total number of annotations in *B. burgdorferi* genome for all COG terms. The abbreviation –inf marks it was not possible to calculate the log-odds.

| COG annotation | Annotation description | q | s | h | t | $p_{adj}$ | log-odds |
|---|---|---|---|---|---|---|---|
| C | Energy production and conversion | 2 | 52 | 26 | 1387 | 0.981 | 0.8 |
| D | Cell cycle control, cell division, chromosome partitioning | 0 | 52 | 42 | 1387 | 0.931 | -inf |
| E | Amino acid transport and metabolism | 1 | 52 | 27 | 1387 | 1 | -0.01 |
| F | Nucleotide transport and metabolism | 2 | 52 | 45 | 1387 | 1 | 0.2 |
| G | Carbohydrate transport and metabolism | 2 | 52 | 47 | 1387 | 1 | 0.1 |
| H | Coenzyme transport and metabolism | 3 | 52 | 23 | 1387 | 0.657 | 1.4 |
| I | Lipid transport and metabolism | 0 | 52 | 18 | 1387 | 1 | -inf |
| J | Translation, ribosomal structure and biogenesis | 1 | 52 | 121 | 1387 | 0.657 | -1.6 |
| K | Transcription | 2 | 52 | 21 | 1387 | 0.931 | 1.02 |
| L | Replication, recombination and repair | 6 | 52 | 64 | 1387 | 0.657 | 1.06 |
| M | Cell wall/membrane/envelope biogenesis | 4 | 52 | 53 | 1387 | 0.844 | 0.8 |
| N | Cell motility | 3 | 52 | 53 | 1387 | 1 | 0.5 |
| O | Post-translational modification, protein turnover, and chaperones | 3 | 52 | 28 | 1387 | 0.800 | 1.2 |
| P | Inorganic ion transport and metabolism | 1 | 52 | 35 | 1387 | 1 | -0.3 |
| Q | Secondary metabolites biosynthesis, transport, and catabolism | 0 | 52 | 5 | 1387 | 1 | -inf |
| S | Function unknown | 22 | 52 | 716 | 1387 | 0.833 | -0.4 |
| T | Signal transduction mechanisms | 0 | 52 | 35 | 1387 | 0.981 | -inf |
| U | Intracellular trafficking, secretion, and vesicular transport | 0 | 52 | 20 | 1387 | 1 | -inf |
| V | Defense mechanisms | 0 | 52 | 8 | 1387 | 1 | -inf |

**Table S14.** Enrichment of location-specific genes among genes more expressed in round bodies than spirochetes. For a gene to be categorized as differentially expressed, the adjusted p-value ($p_{ajd}$) calculated by DeSeq2 pairwise comparison analysis had to be below 0.05. q = the number of location-specific genes in round body upregulated genes, s = the total number of location-specific genes in round body upregulated genes, h = the number of location-specific genes in *B. burgdorferi* genome, t = the total number of location-specific in *B. burgdorferi* genome. The abbreviation –inf marks it was not possible to calculate the log-odds.

| location | q | s | h | t | $p_{adj}$ | log-odds |
|---|---|---|---|---|---|---|
| chromosome | 43 | 44 | 856 | 1544 | 5.428 E-09 | 3.6 |
| cp26 | 0 | 44 | 26 | 1544 | 1 | -inf |
| cp32-1 | 1 | 44 | 42 | 1544 | 1 | -0.2 |
| cp32-3 | 0 | 44 | 43 | 1544 | 1 | -inf |
| cp32-4 | 0 | 44 | 43 | 1544 | 1 | -inf |
| cp32-6 | 0 | 44 | 41 | 1544 | 1 | -inf |
| cp32-7 | 0 | 44 | 42 | 1544 | 1 | -inf |
| cp32-8 | 0 | 44 | 42 | 1544 | 1 | -inf |
| cp32-9 | 0 | 44 | 43 | 1544 | 1 | -inf |
| cp9 | 0 | 44 | 9 | 1544 | 1 | -inf |
| lp17 | 0 | 44 | 17 | 1544 | 1 | -inf |
| lp21 | 0 | 44 | 11 | 1544 | 1 | -inf |
| lp25 | 0 | 44 | 16 | 1544 | 1 | -inf |
| lp28-1 | 0 | 44 | 30 | 1544 | 1 | -inf |
| lp28-2 | 0 | 44 | 34 | 1544 | 1 | -inf |
| lp28-3 | 0 | 44 | 23 | 1544 | 1 | -inf |
| lp28-4 | 0 | 44 | 25 | 1544 | 1 | -inf |
| lp36 | 0 | 44 | 33 | 1544 | 1 | -inf |
| lp38 | 0 | 44 | 33 | 1544 | 1 | -inf |
| lp5 | 0 | 44 | 6 | 1544 | 1 | -inf |
| lp54 | 0 | 44 | 64 | 1544 | 1 | -inf |
| lp56 | 0 | 44 | 65 | 1544 | 1 | -inf |

**Table S15.** Enrichment of location-specific genes among genes more expressed in blebs than spirochetes. For a gene to be categorized as differentially expressed, the the adjusted p-value ($p_{ajd}$) calculated by DeSeq2 pairwise comparison analysis had to be below 0,05. q = the number of location-specific genes in bleb upregulated genes, s = the total number of location-specific genes in bleb upregulated genes, h = the number of location-specific genes in *B. burgdorferi* genome, t = the total number of location-specific in *B. burgdorferi* genome. The abbreviation –inf marks it was not possible to calculate the log-odds.

| location | q | s | h | t | $p_{adj}$ | log-odds |
|---|---|---|---|---|---|---|
| chromosome | 160 | 529 | 856 | 1544 | 6.9E-46 | -1.6 |
| cp26 | 10 | 529 | 26 | 1544 | 0.828 | 0.2 |
| cp32-1 | 33 | 529 | 42 | 1544 | 2.3E-08 | 2 |
| cp32-3 | 27 | 529 | 43 | 1544 | 0.0005 | 1.2 |
| cp32-4 | 37 | 529 | 43 | 1544 | 2.3E-11 | 2.5 |
| cp32-6 | 21 | 529 | 41 | 1544 | 0.051 | 0.7 |
| cp32-7 | 16 | 529 | 42 | 1544 | 0.775 | 0.2 |
| cp32-8 | 14 | 529 | 42 | 1544 | 1 | -0.04 |
| cp32-9 | 26 | 529 | 43 | 1544 | 0.001 | 1.1 |
| cp9 | 1 | 529 | 9 | 1544 | 0.3 | -1.4 |
| lp17 | 10 | 529 | 17 | 1544 | 0.089 | 1 |
| lp21 | 0 | 529 | 11 | 1544 | 0.031 | -inf |
| lp25 | 2 | 529 | 16 | 1544 | 0.13 | -1.3 |
| lp28-1 | 23 | 529 | 30 | 1544 | 1.1E-05 | 1.9 |
| lp28-2 | 29 | 529 | 34 | 1544 | 6.9E-09 | 2.5 |
| lp28-3 | 17 | 529 | 23 | 1544 | 0.0005 | 1.7 |
| lp28-4 | 1 | 529 | 25 | 1544 | 0.001 | -2.6 |
| lp36 | 3 | 529 | 33 | 1544 | 0.003 | -1.7 |
| lp38 | 0 | 529 | 33 | 1544 | 5.1E-06 | -inf |
| lp5 | 0 | 529 | 6 | 1544 | 0.196 | -inf |
| lp54 | 48 | 529 | 64 | 1544 | 1.3E-10 | 1.8 |
| lp56 | 51 | 529 | 65 | 1544 | 2.4E-12 | 2 |

**Table S16.** Enrichment of location-specific genes among genes more expressed in biofilms than spirochetes. For a gene to be categorized as differentially expressed, the adjusted p-value ($p_{ajd}$) calculated by DeSeq2 pairwise comparison analysis had to be below 0,05. q = the number of location-specific genes in biofilm upregulated genes, s = the total number of location-specific genes in biofilm upregulated genes, h = the number of location-specific genes in *B. burgdorferi* genome, t = the total number of location-specific in *B. burgdorferi* genome. The abbreviation –inf marks it was not possible to calculate the log-odds.

| location | q | s | h | t | padj | log-odds |
|---|---|---|---|---|---|---|
| chromosome | 219 | 467 | 856 | 1544 | 5.2E-05 | -0.5 |
| cp26 | 8 | 467 | 26 | 1544 | 1 | 0.03 |
| cp32-1 | 33 | 467 | 42 | 1544 | 3.3E-09 | 2.2 |
| cp32-3 | 23 | 467 | 43 | 1544 | 0.006 | 1.007 |
| cp32-4 | 28 | 467 | 43 | 1544 | 3.9E-05 | 1.5 |
| cp32-6 | 18 | 467 | 41 | 1544 | 0.144 | 0.6 |
| cp32-7 | 13 | 467 | 42 | 1544 | 1 | 0.03 |
| cp32-8 | 11 | 467 | 42 | 1544 | 0.9 | -0.2 |
| cp32-9 | 20 | 467 | 43 | 1544 | 0.073 | 0.7 |
| cp9 | 0 | 467 | 9 | 1544 | 0.142 | -inf |
| lp17 | 6 | 467 | 17 | 1544 | 0.954 | 0.2 |
| lp21 | 0 | 467 | 11 | 1544 | 0.075 | -inf |
| lp25 | 2 | 467 | 16 | 1544 | 0.278 | -1.1 |
| lp28-1 | 5 | 467 | 30 | 1544 | 0.222 | -0.8 |
| lp28-2 | 11 | 467 | 34 | 1544 | 1 | 0.1 |
| lp28-3 | 8 | 467 | 23 | 1544 | 0.954 | 0.2 |
| lp28-4 | 0 | 467 | 25 | 1544 | 0.0008 | -inf |
| lp36 | 0 | 467 | 33 | 1544 | 5.2E-05 | -inf |
| lp38 | 0 | 467 | 33 | 1544 | 5.2E-05 | -inf |
| lp5 | 0 | 467 | 6 | 1544 | 0.315 | -inf |
| lp54 | 28 | 467 | 64 | 1544 | 0.066 | 0.6 |
| lp56 | 34 | 467 | 65 | 1544 | 0.0008 | 1 |

**Table S17.** Enrichment of location-specific genes among genes more expressed in blebs than spirochetes. For a gene to be categorized as differentially expressed, the adjusted p-value ($p_{ajd}$) and absolute value of log2FC (|log2 FC|) calculated by DeSeq2 pairwise comparison analysis had to be below 0,05 and more that 1, respectively. q = the number of location-specific genes in bleb upregulated genes, s = the total number of location-specific genes in bleb upregulated genes, h = the number of location-specific genes in *B. burgdorferi* genome, t = the total number of location-specific in *B. burgdorferi* genome. The abbreviation –inf marks it was not possible to calculate the log-odds.

| location | q | s | h | t | $p_{adj}$ | log-odds |
|---|---|---|---|---|---|---|
| chromosome | 19 | 274 | 856 | 1544 | 6E-77 | -3.3 |
| cp26 | 2 | 274 | 26 | 1544 | 0.374 | -1 |
| cp32-1 | 20 | 274 | 42 | 1544 | 3.3E-05 | 1.5 |
| cp32-3 | 17 | 274 | 43 | 1544 | 0.003 | 1.2 |
| cp32-4 | 22 | 274 | 43 | 1544 | 3E-06 | 1.6 |
| cp32-6 | 14 | 274 | 41 | 1544 | 0.03 | 0.9 |
| cp32-7 | 10 | 274 | 42 | 1544 | 0.5 | 0.4 |
| cp32-8 | 8 | 274 | 42 | 1544 | 0.951 | 0.1 |
| cp32-9 | 13 | 274 | 43 | 1544 | 0.11 | 0.7 |
| cp9 | 0 | 274 | 9 | 1544 | 0.419 | -inf |
| lp17 | 5 | 274 | 17 | 1544 | 0.419 | 0.7 |
| lp21 | 0 | 274 | 11 | 1544 | 0.364 | -inf |
| lp25 | 2 | 274 | 16 | 1544 | 0.921 | -0.4 |
| lp28-1 | 19 | 274 | 30 | 1544 | 2.3E-07 | 2.1 |
| lp28-2 | 23 | 274 | 34 | 1544 | 1.2E-09 | 2.4 |
| lp28-3 | 14 | 274 | 23 | 1544 | 2.5E-05 | 2 |
| lp28-4 | 1 | 274 | 25 | 1544 | 0.159 | -1.7 |
| lp36 | 3 | 274 | 33 | 1544 | 0.374 | -0.8 |
| lp38 | 0 | 274 | 33 | 1544 | 0.006 | -inf |
| lp5 | 0 | 274 | 6 | 1544 | 0.68 | -inf |
| lp54 | 37 | 274 | 64 | 1544 | 2.7E-12 | 2 |
| lp56 | 45 | 274 | 65 | 1544 | 1.4E-20 | 1.5E-19 |

**Table S18.** Enrichment of location-specific genes among genes more expressed in biofilms than spirochetes. For a gene to be categorized as differentially expressed, the adjusted p-value ($p_{ajd}$) and absolute value of log2FC (|log2 FC|) calculated by DeSeq2 pairwise comparison analysis had to be below 0.05 and more that 1, respectively. q = the number of location-specific genes in biofilm upregulated genes, s = the total number of location-specific genes in biofilm upregulated genes, h = the number of location-specific genes in *B. burgdorferi* genome, t = the total number of location-specific in *B. burgdorferi* genome. The abbreviation –inf marks it was not possible to calculate the log-odds.

| location | q | s | h | t | $p_{adj}$ | log-odds |
|---|---|---|---|---|---|---|
| chromosome | 26 | 156 | 856 | 1544 | 3.6E-24 | -2 |
| cp26 | 2 | 156 | 26 | 1544 | 1 | -0.3 |
| cp32-1 | 19 | 156 | 42 | 1544 | 5.4E-08 | 2.1 |
| cp32-3 | 13 | 156 | 43 | 1544 | 0.001 | 1.4 |
| cp32-4 | 15 | 156 | 43 | 1544 | 6.6E-05 | 1.6 |
| cp32-6 | 15 | 156 | 41 | 1544 | 4.1E-05 | 1.7 |
| cp32-7 | 8 | 156 | 42 | 1544 | 0.217 | 0.8 |
| cp32-8 | 9 | 156 | 42 | 1544 | 0.131 | 0.9 |
| cp32-9 | 6 | 156 | 43 | 1544 | 0.722 | 0.4 |
| cp9 | 0 | 156 | 9 | 1544 | 0.935 | -inf |
| lp17 | 1 | 156 | 17 | 1544 | 1 | -0.6 |
| lp21 | 0 | 156 | 11 | 1544 | 0.799 | -inf |
| lp25 | 2 | 156 | 16 | 1544 | 1 | 0.2 |
| lp28-1 | 0 | 156 | 30 | 1544 | 0.174 | -inf |
| lp28-2 | 6 | 156 | 34 | 1544 | 0.379 | 0.7 |
| lp28-3 | 5 | 156 | 23 | 1544 | 0.251 | 0.9 |
| lp28-4 | 0 | 156 | 25 | 1544 | 0.25 | -inf |
| lp36 | 0 | 156 | 33 | 1544 | 0.14 | -inf |
| lp38 | 0 | 156 | 33 | 1544 | 0.14 | -inf |
| lp5 | 0 | 156 | 6 | 1544 | 1 | -inf |
| lp54 | 9 | 156 | 64 | 1544 | 0.558 | 0.4 |
| lp56 | 20 | 156 | 65 | 1544 | 2.9E-05 | 1.5 |

**Table S19.** Enrichment of location-specific genes among genes less expressed in round bodies than spirochetes. For a gene to be categorized as differentially expressed, the adjusted p-value ($p_{ajd}$) calculated by DeSeq2 pairwise comparison analysis had to be below 0.05. q = the total number of location-specific genes in round body downregulated genes, s = the number of location-specific genes in round body downregulated genes, h = the number of location-specific genes in *B. burgdorferi* genome, t = the total number of location-specific in *B. burgdorferi* genome. The abbreviation –inf marks it was not possible to calculate the log-odds.

| location | q | s | h | t | $p_{adj}$ | log-odds |
|---|---|---|---|---|---|---|
| chromosome | 19 | 23 | 856 | 1544 | 0.257 | 1.4 |
| cp26 | 0 | 23 | 26 | 1544 | 1 | -inf |
| cp32-1 | 0 | 23 | 42 | 1544 | 1 | -inf |
| cp32-3 | 0 | 23 | 43 | 1544 | 1 | -inf |
| cp32-4 | 0 | 23 | 43 | 1544 | 1 | -inf |
| cp32-6 | 0 | 23 | 41 | 1544 | 1 | -inf |
| cp32-7 | 1 | 23 | 42 | 1544 | 1 | 0.5 |
| cp32-8 | 0 | 23 | 42 | 1544 | 1 | -inf |
| cp32-9 | 0 | 23 | 43 | 1544 | 1 | -inf |
| cp9 | 0 | 23 | 9 | 1544 | 1 | -inf |
| lp17 | 0 | 23 | 17 | 1544 | 1 | -inf |
| lp21 | 0 | 23 | 11 | 1544 | 1 | -inf |
| lp25 | 0 | 23 | 16 | 1544 | 1 | -inf |
| lp28-1 | 0 | 23 | 30 | 1544 | 1 | -inf |
| lp28-2 | 0 | 23 | 34 | 1544 | 1 | -inf |
| lp28-3 | 0 | 23 | 23 | 1544 | 1 | -inf |
| lp28-4 | 0 | 23 | 25 | 1544 | 1 | -inf |
| lp36 | 0 | 23 | 33 | 1544 | 1 | -inf |
| lp38 | 0 | 23 | 33 | 1544 | 1 | -inf |
| lp5 | 0 | 23 | 6 | 1544 | 1 | -inf |
| lp54 | 3 | 23 | 64 | 1544 | 1 | 1.3 |
| lp56 | 0 | 23 | 65 | 1544 | 1 | -inf |

**Table S20.** Enrichment of location-specific genes among genes less expressed in blebs than spirochetes. For a gene to be categorized as differentially expressed, the adjusted p-value ($p_{ajd}$) calculated by DeSeq2 pairwise comparison analysis had to be below 0.05. q = the number of location-specific genes in bleb downregulated genes, s = the total number of location-specific genes in bleb downregulated genes, h = the number of location-specific genes in *B. burgdorferi* genome, t = the total number of location-specific in *B. burgdorferi* genome. The abbreviation –inf marks it was not possible to calculate the log-odds.

| location | q | s | h | t | $p_{adj}$ | log-odds |
|---|---|---|---|---|---|---|
| chromosome | 486 | 522 | 856 | 1544 | 1.6E-113 | 3.2 |
| cp26 | 10 | 522 | 26 | 1544 | 0.752 | 0.2 |
| cp32-1 | 1 | 522 | 42 | 1544 | 3.8E-06 | -3.1 |
| cp32-3 | 2 | 522 | 43 | 1544 | 1.8E-05 | -2.4 |
| cp32-4 | 1 | 522 | 43 | 1544 | 3.8E-06 | -3.1 |
| cp32-6 | 1 | 522 | 41 | 1544 | 4.9E-06 | -3.1 |
| cp32-7 | 1 | 522 | 42 | 1544 | 3.8E-06 | -3.1 |
| cp32-8 | 0 | 522 | 42 | 1544 | 3.3E-07 | -inf |
| cp32-9 | 6 | 522 | 43 | 1544 | 0.007 | -1.2 |
| cp9 | 0 | 522 | 9 | 1544 | 0.053 | -inf |
| lp17 | 1 | 522 | 17 | 1544 | 0.021 | -2.1 |
| lp21 | 0 | 522 | 11 | 1544 | 0.024 | -inf |
| lp25 | 0 | 522 | 16 | 1544 | 0.004 | -inf |
| lp28-1 | 1 | 522 | 30 | 1544 | 0.0002 | -2.7 |
| lp28-2 | 3 | 522 | 34 | 1544 | 0.002 | -1.7 |
| lp28-3 | 2 | 522 | 23 | 1544 | 0.015 | -1.7 |
| lp28-4 | 0 | 522 | 25 | 1544 | 0.0001 | -inf |
| lp36 | 1 | 522 | 33 | 1544 | 7.5E-05 | -2.8 |
| lp38 | 0 | 522 | 33 | 1544 | 5.6E-06 | -inf |
| lp5 | 0 | 522 | 6 | 1544 | 0.175 | -inf |
| lp54 | 6 | 522 | 64 | 1544 | 1.8E-05 | -1.6 |
| lp56 | 0 | 522 | 65 | 1544 | 2.4E-11 | -inf |

**Table S21.** Enrichment of location-specific genes among genes less expressed in biofilms than spirochetes. For a gene to be categorized as differentially expressed, the adjusted p-value ($p_{ajd}$) calculated by DeSeq2 pairwise comparison analysis had to be below 0.05. q = the number of location-specific genes in biofilm downregulated genes, s = the total number of location-specific genes in biofilm downregulated genes, h = the number of location-specific genes in *B. burgdorferi* genome, t = the total number of location-specific in *B. burgdorferi* genome. The abbreviation –inf marks it was not possible to calculate the log-odds.

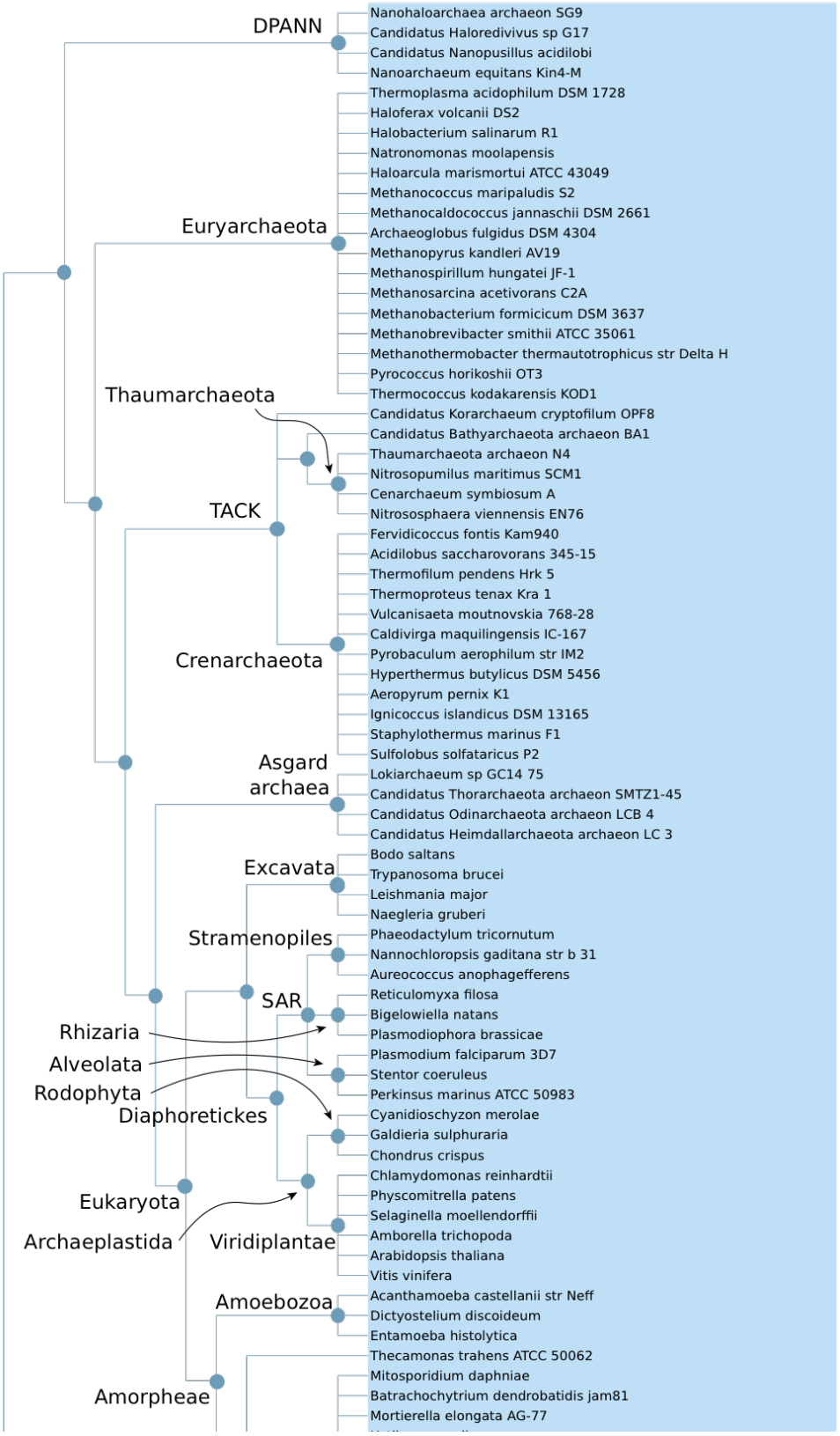
| location | q | s | h | t | $p_{adj}$ | log-odds |
|---|---|---|---|---|---|---|
| chromosome | 406 | 464 | 856 | 1544 | 1.4E-66 | 2.3 |
| cp26 | 10 | 464 | 26 | 1544 | 0.481 | 0.4 |
| cp32-1 | 0 | 464 | 42 | 1544 | 2.6E-06 | -inf |
| cp32-3 | 2 | 464 | 43 | 1544 | 0.0001 | -2.2 |
| cp32-4 | 0 | 464 | 43 | 1544 | 2.6E-06 | -inf |
| cp32-6 | 1 | 464 | 41 | 1544 | 3.2E-05 | -2.9 |
| cp32-7 | 1 | 464 | 42 | 1544 | 3.2E-05 | -2.9 |
| cp32-8 | 1 | 464 | 42 | 1544 | 3.2E-05 | -2.9 |
| cp32-9 | 6 | 464 | 43 | 1544 | 0.036 | -1 |
| cp9 | 0 | 464 | 9 | 1544 | 0.103 | -inf |
| lp17 | 4 | 464 | 17 | 1544 | 0.772 | -0.3 |
| lp21 | 0 | 464 | 11 | 1544 | 0.0575 | -inf |
| lp25 | 0 | 464 | 16 | 1544 | 0.011 | -inf |
| lp28-1 | 12 | 464 | 30 | 1544 | 0.349 | 0.4 |
| lp28-2 | 5 | 464 | 34 | 1544 | 0.087 | -0.9 |
| lp28-3 | 4 | 464 | 23 | 1544 | 0.307 | -0.7 |
| lp28-4 | 0 | 464 | 25 | 1544 | 0.0005 | -inf |
| lp36 | 0 | 464 | 33 | 1544 | 3.2E-05 | -inf |
| lp38 | 0 | 464 | 33 | 1544 | 3.2E-05 | -inf |
| lp5 | 0 | 464 | 6 | 1544 | 0.285 | -inf |
| lp54 | 9 | 464 | 64 | 1544 | 0.008 | -1 |
| lp56 | 3 | 464 | 65 | 1544 | 2.6E-06 | -2.2 |

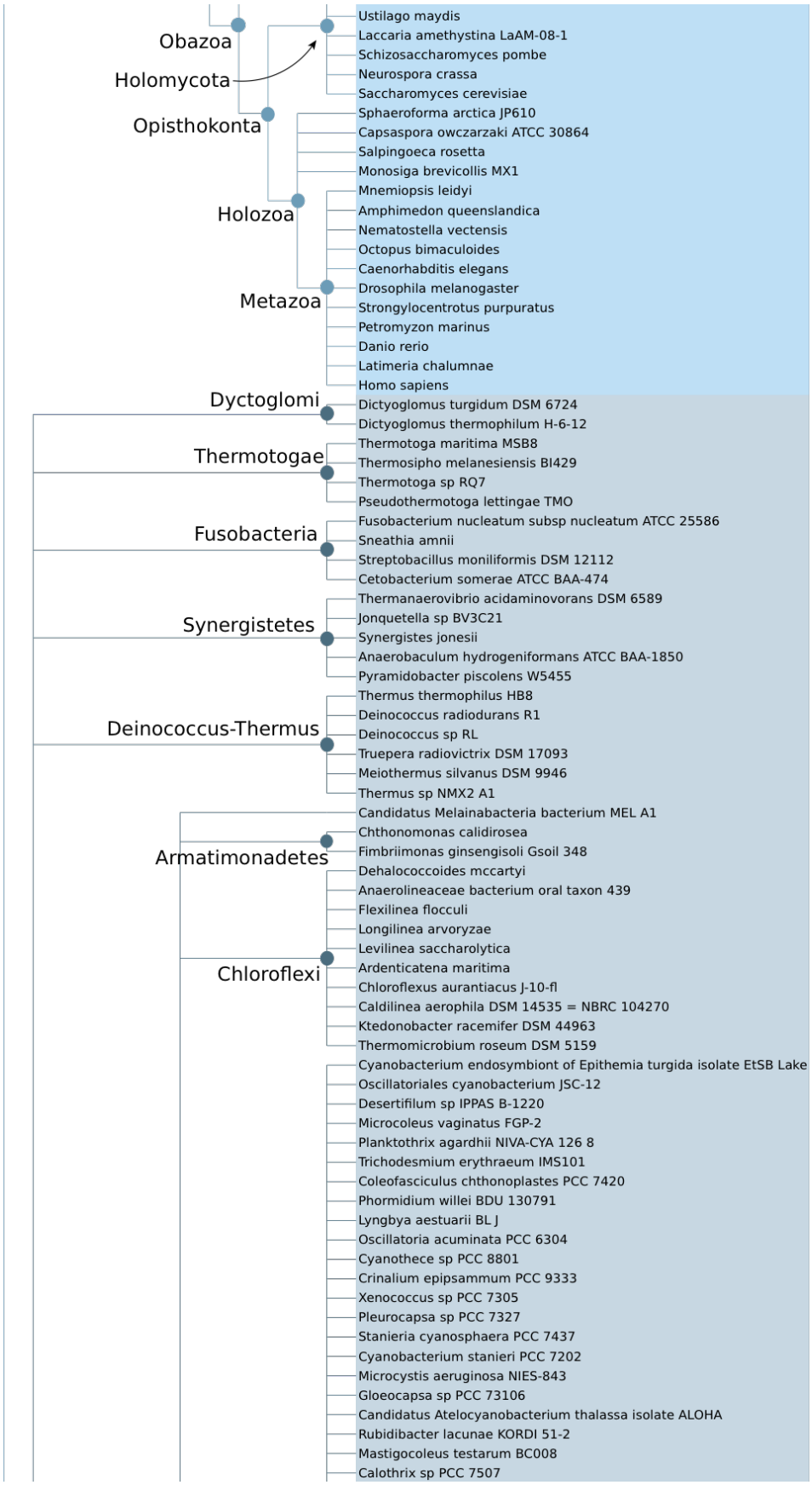
**Table S22.** Enrichment of location-specific genes among genes less expressed in blebs than spirochetes. For a gene to be categorized as differentially expressed, the adjusted p-value ($p_{ajd}$) and absolute value of log2FC (|log2 FC|) calculated by DeSeq2 pairwise comparison analysis had to be below 0.05 and more that 1, respectively. q = the number of location-specific genes in bleb downregulated genes, s = the total number of location-specific genes in bleb downregulated genes, h = the number of location-specific genes in *B. burgdorferi* genome, t = the total number of location-specific in *B. burgdorferi* genome. The abbreviation –inf marks it was not possible to calculate the log-odds.

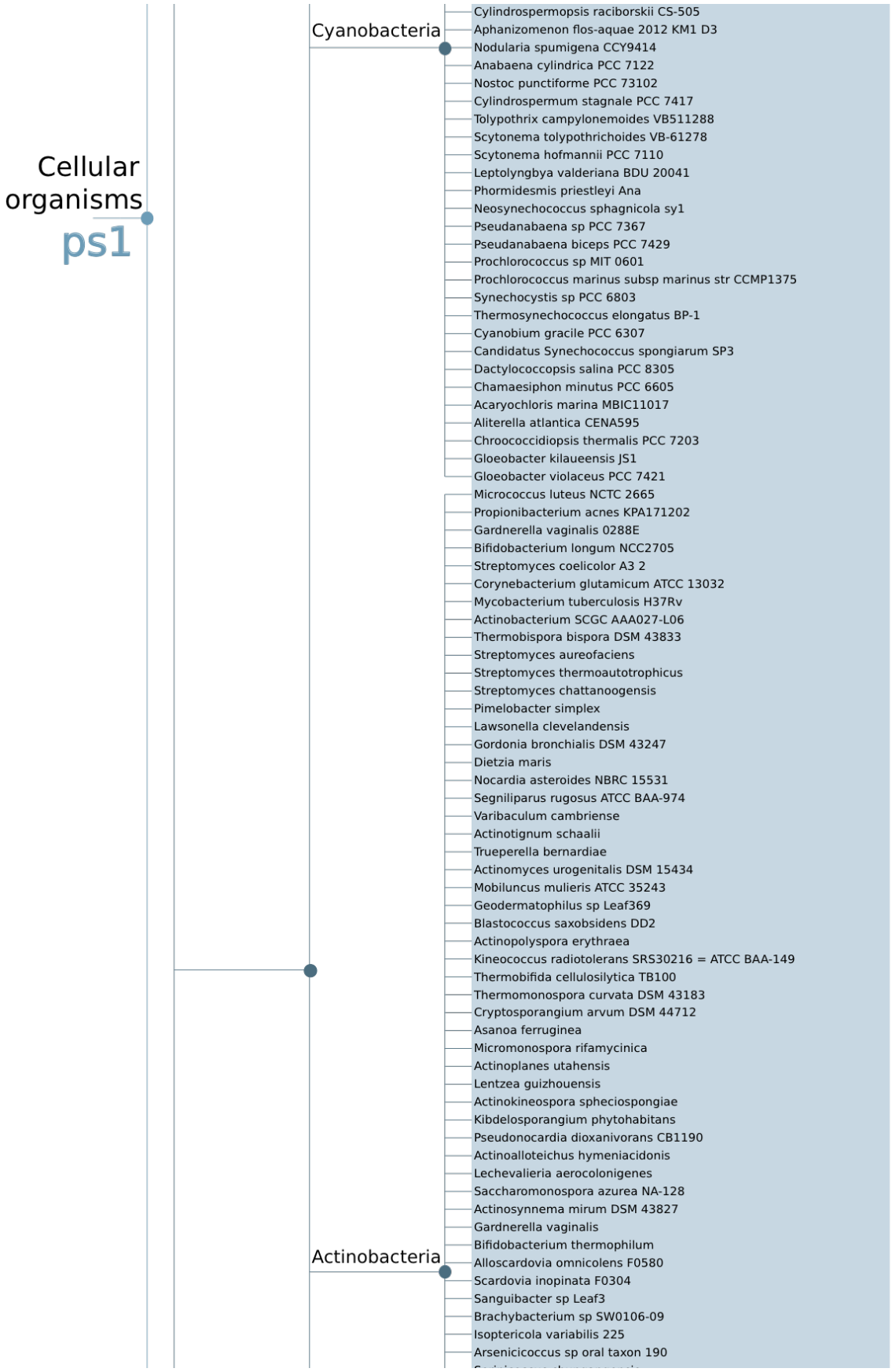
| location | q | s | h | t | $p_{adj}$ | log-odds |
|---|---|---|---|---|---|---|
| chromosome | 134 | 142 | 856 | 1544 | 1.4E-25 | 2.8 |
| cp26 | 2 | 142 | 26 | 1544 | 1 | -0.2 |
| cp32-1 | 0 | 142 | 42 | 1544 | 0.114 | -inf |
| cp32-3 | 1 | 142 | 43 | 1544 | 0.278 | -1.5 |
| cp32-4 | 0 | 142 | 43 | 1544 | 0.114 | -inf |
| cp32-6 | 0 | 142 | 41 | 1544 | 0.114 | -inf |
| cp32-7 | 1 | 142 | 42 | 1544 | 0.278 | -1.4 |
| cp32-8 | 0 | 142 | 42 | 1544 | 0.114 | -inf |
| cp32-9 | 0 | 142 | 43 | 1544 | 0.114 | -inf |
| cp9 | 0 | 142 | 9 | 1544 | 0.921 | -inf |
| lp17 | 0 | 142 | 17 | 1544 | 0.529 | -inf |
| lp21 | 0 | 142 | 11 | 1544 | 0.834 | -inf |
| lp25 | 0 | 142 | 16 | 1544 | 0.548 | -inf |
| lp28-1 | 0 | 142 | 30 | 1544 | 0.22 | -inf |
| lp28-2 | 0 | 142 | 34 | 1544 | 0.2 | -inf |
| lp28-3 | 1 | 142 | 23 | 1544 | 0.834 | -0.8 |
| lp28-4 | 0 | 142 | 25 | 1544 | 0.278 | -inf |
| lp36 | 1 | 142 | 33 | 1544 | 0.519 | -1.2 |
| lp38 | 0 | 142 | 33 | 1544 | 0.2 | -inf |
| lp5 | 0 | 142 | 6 | 1544 | 1 | -inf |
| lp54 | 2 | 142 | 64 | 1544 | 0.22 | -1.2 |
| lp56 | 0 | 142 | 65 | 1544 | 0.036 | -inf |

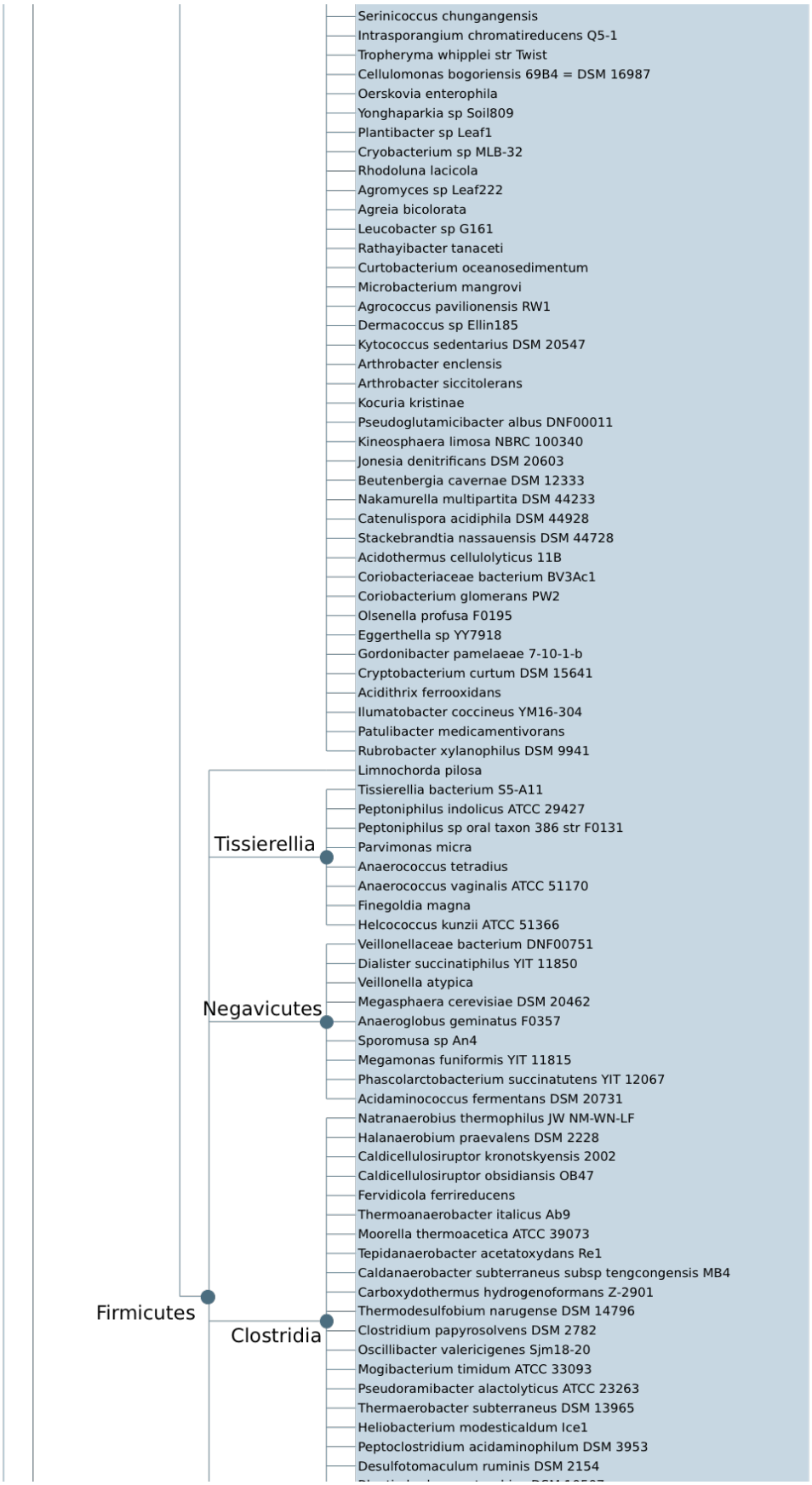
**Table S23.** Enrichment of location-specific genes among genes less expressed in biofilms than spirochetes. For a gene to be categorized as differentially expressed, the adjusted p-value ($p_{ajd}$) and absolute value of log2FC (|log2 FC|) calculated by DeSeq2 pairwise comparison analysis had to be below 0.05 and more that 1, respectively. q = the number of location-specific genes in biofilm downregulated genes, s = the total number of location-specific genes in biofilm downregulated genes, h = the number of location-specific genes in *B. burgdorferi* genome, t = the total number of location-specific in *B. burgdorferi* genome. The abbreviation –inf marks it was not possible to calculate the log-odds.

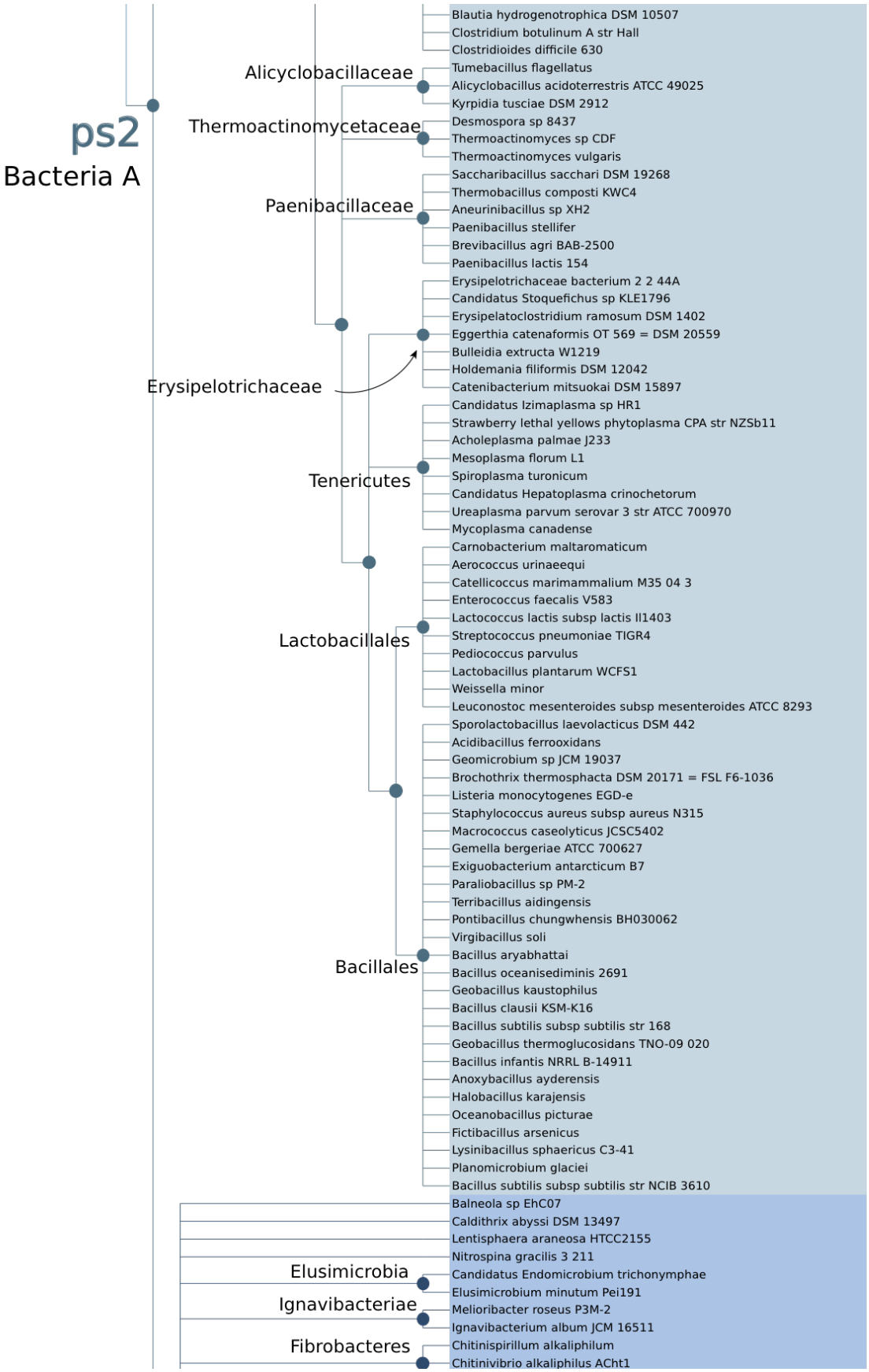
| location | q | s | h | t | $p_{adj}$ | log-odds |
|---|---|---|---|---|---|---|
| chromosome | 45 | 60 | 856 | 1544 | 0.025 | 0.9 |
| cp26 | 2 | 60 | 26 | 1544 | 0.842 | 0.7 |
| cp32-1 | 0 | 60 | 42 | 1544 | 0.842 | -inf |
| cp32-3 | 0 | 60 | 43 | 1544 | 0.842 | -inf |
| cp32-4 | 0 | 60 | 43 | 1544 | 0.842 | -inf |
| cp32-6 | 0 | 60 | 41 | 1544 | 0.842 | -inf |
| cp32-7 | 0 | 60 | 42 | 1544 | 0.842 | -inf |
| cp32-8 | 0 | 60 | 42 | 1544 | 0.842 | -inf |
| cp32-9 | 1 | 60 | 43 | 1544 | 1 | -0.5 |
| cp9 | 0 | 60 | 9 | 1544 | 1 | -inf |
| lp17 | 2 | 60 | 17 | 1544 | 0.842 | 1.2 |
| lp21 | 0 | 60 | 11 | 1544 | 1 | -inf |
| lp25 | 0 | 60 | 16 | 1544 | 1 | -inf |
| lp28-1 | 8 | 60 | 30 | 1544 | 0.0004 | 2.3 |
| lp28-2 | 0 | 60 | 34 | 1544 | 0.842 | -inf |
| lp28-3 | 0 | 60 | 23 | 1544 | 1 | -inf |
| lp28-4 | 0 | 60 | 25 | 1544 | 1 | -inf |
| lp36 | 0 | 60 | 33 | 1544 | 0.842 | -inf |
| lp38 | 0 | 60 | 33 | 1544 | 0.842 | -inf |
| lp5 | 0 | 60 | 6 | 1544 | 1 | -inf |
| lp54 | 0 | 60 | 64 | 1544 | 0.842 | -inf |
| lp56 | 2 | 60 | 65 | 1544 | 1 | -0.3 |

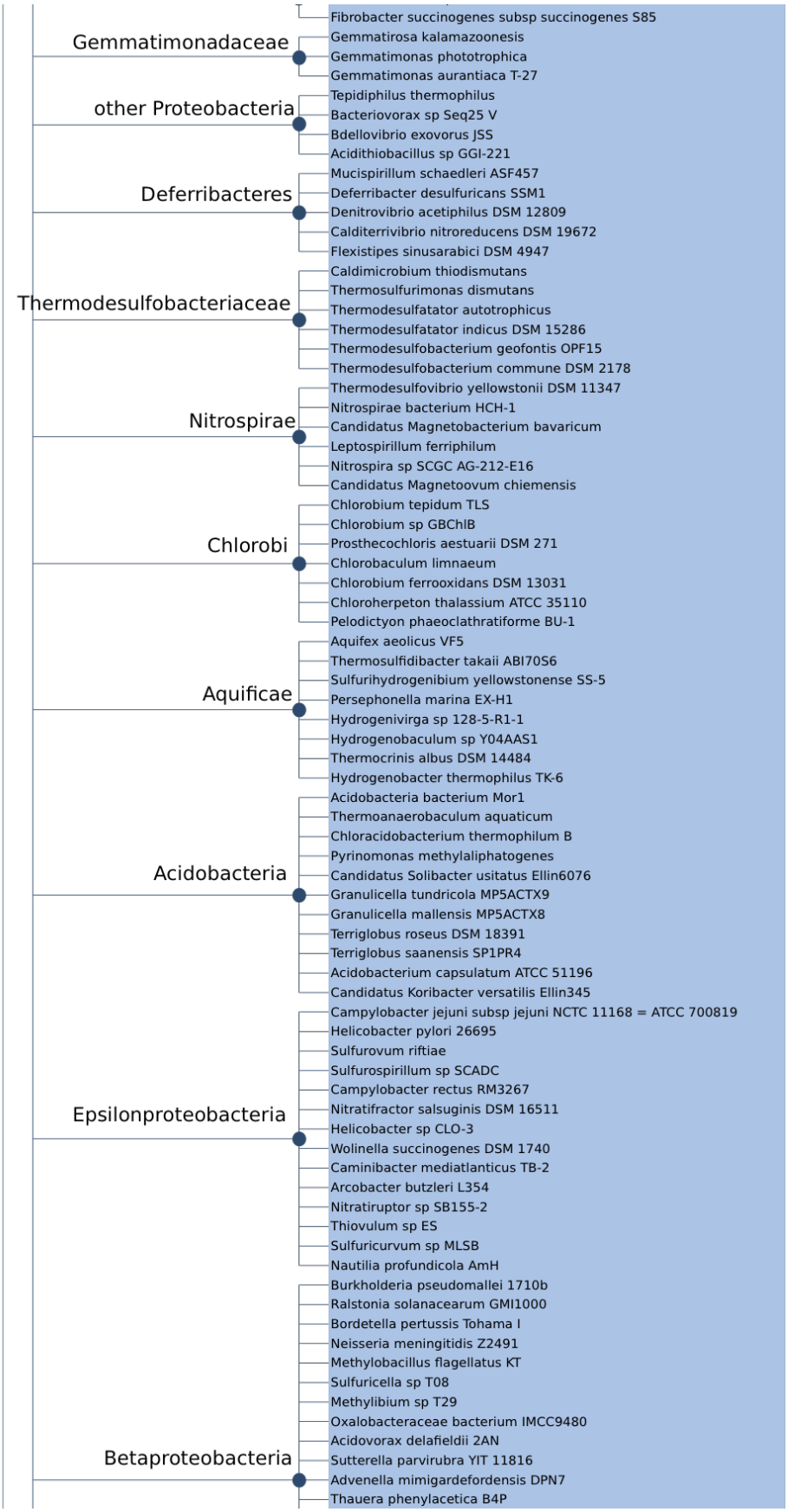

DPANN
Euryarchaeota
Thaumarchaeota
TACK
Crenarchaeota
Asgard archaea
Excavata
Stramenopiles
SAR
Rhizaria
Alveolata
Rodophyta
Diaphoretickes
Eukaryota
Archaeplastida
Viridiplantae
Amoebozoa
Amorpheae
Nanohaloarchaea archaeon SG9
Candidatus Haloredivivus sp G17
Candidatus Nanopusillus acidilobi
Nanoarchaeum equitans Kin4-M
Thermoplasma acidophilum DSM 1728
Haloferax volcanii DS2
Halobacterium salinarum R1
Natronomonas moolapensis
Haloarcula marismortui ATCC 43049
Methanococcus maripaludis S2
Methanocaldococcus jannaschii DSM 2661
Archaeoglobus fulgidus DSM 4304
Methanopyrus kandleri AV19
Methanospirillum hungatei JF-1
Methanosarcina acetivorans C2A
Methanobacterium formicicum DSM 3637
Methanobrevibacter smithii ATCC 35061
Methanothermobacter thermautotrophicus str Delta H
Pyrococcus horikoshii OT3
Thermococcus kodakarensis KOD1
Candidatus Korarchaeum cryptofilum OPF8
Candidatus Bathyarchaeota archaeon BA1
Thaumarchaeota archaeon N4
Nitrosopumilus maritimus SCM1
Cenarchaeum symbiosum A
Nitrososphaera viennensis EN76
Fervidicoccus fontis Kam940
Acidilobus saccharovorans 345-15
Thermofilum pendens Hrk 5
Thermoproteus tenax Kra 1
Vulcanisaeta moutnovskia 768-28
Caldivirga maquilingensis IC-167
Pyrobaculum aerophilum str IM2
Hyperthermus butylicus DSM 5456
Aeropyrum pernix K1
Ignicoccus islandicus DSM 13165
Staphylothermus marinus F1
Sulfolobus solfataricus P2
Lokiarchaeum sp GC14 75
Candidatus Thorarchaeota archaeon SMTZ1-45
Candidatus Odinarchaeota archaeon LCB 4
Candidatus Heimdallarchaeota archaeon LC 3
Bodo saltans
Trypanosoma brucei
Leishmania major
Naegleria gruberi
Phaeodactylum tricornutum
Nannochloropsis gaditana str b 31
Aureococcus anophagefferens
Reticulomyxa filosa
Bigelowiella natans
Plasmodiophora brassicae
Plasmodium falciparum 3D7
Stentor coeruleus
Perkinsus marinus ATCC 50983
Cyanidioschyzon merolae
Galdieria sulphuraria
Chondrus crispus
Chlamydomonas reinhardtii
Physcomitrella patens
Selaginella moellendorffii
Amborella trichopoda
Arabidopsis thaliana
Vitis vinifera
Acanthamoeba castellanii str Neff
Dictyostelium discoideum
Entamoeba histolytica
Thecamonas trahens ATCC 50062
Mitosporidium daphniae
Batrachochytrium dendrobatidis jam81
Mortierella elongata AG-77

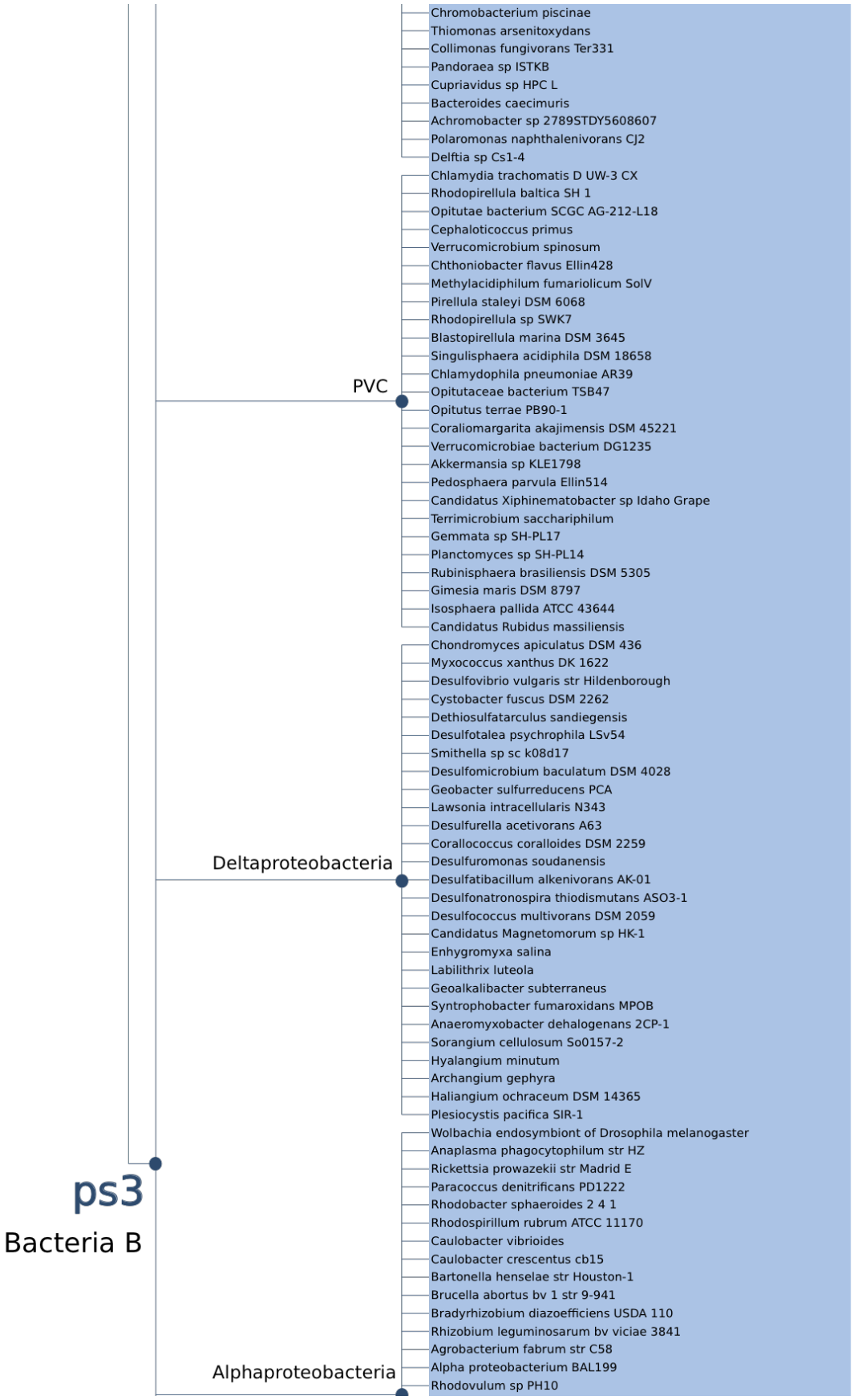

Obazoa
Holomycota
Opisthokonta
Holozoa
Metazoa
Dyctoglomi
Thermotogae
Fusobacteria
Synergistetes
Deinococcus-Thermus
Armatimonadetes
Chloroflexi
Ustilago maydis
Laccaria amethystina LaAM-08-1
Schizosaccharomyces pombe
Neurospora crassa
Saccharomyces cerevisiae
Sphaeroforma arctica JP610
Capsaspora owczarzaki ATCC 30864
Salpingoeca rosetta
Monosiga brevicollis MX1
Mnemiopsis leidyi
Amphimedon queenslandica
Nematostella vectensis
Octopus bimaculoides
Caenorhabditis elegans
Drosophila melanogaster
Strongylocentrotus purpuratus
Petromyzon marinus
Danio rerio
Latimeria chalumnae
Homo sapiens
Dictyoglomus turgidum DSM 6724
Dictyoglomus thermophilum H-6-12
Thermotoga maritima MSB8
Thermosipho melanesiensis BI429
Thermotoga sp RQ7
Pseudothermotoga lettingae TMO
Fusobacterium nucleatum subsp nucleatum ATCC 25586
Sneathia amnii
Streptobacillus moniliformis DSM 12112
Cetobacterium somerae ATCC BAA-474
Thermanaerovibrio acidaminovorans DSM 6589
Jonquetella sp BV3C21
Synergistes jonesii
Anaerobaculum hydrogeniformans ATCC BAA-1850
Pyramidobacter piscolens W5455
Thermus thermophilus HB8
Deinococcus radiodurans R1
Deinococcus sp RL
Truepera radiovictrix DSM 17093
Meiothermus silvanus DSM 9946
Thermus sp NMX2 A1
Candidatus Melainabacteria bacterium MEL A1
Chthonomonas calidirosea
Fimbriimonas ginsengisoli Gsoil 348
Dehalococcoides mccartyi
Anaerolineaceae bacterium oral taxon 439
Flexilinea flocculi
Longilinea arvoryzae
Levilinea saccharolytica
Ardenticatena maritima
Chloroflexus aurantiacus J-10-fl
Caldilinea aerophila DSM 14535 = NBRC 104270
Ktedonobacter racemifer DSM 44963
Thermomicrobium roseum DSM 5159
Cyanobacterium endosymbiont of Epithemia turgida isolate EtSB Lake
Oscillatoriales cyanobacterium JSC-12
Desertifilum sp IPPAS B-1220
Microcoleus vaginatus FGP-2
Planktothrix agardhii NIVA-CYA 126 8
Trichodesmium erythraeum IMS101
Coleofasciculus chthonoplastes PCC 7420
Phormidium willei BDU 130791
Lyngbya aestuarii BL J
Oscillatoria acuminata PCC 6304
Cyanothece sp PCC 8801
Crinalium epipsammum PCC 9333
Xenococcus sp PCC 7305
Pleurocapsa sp PCC 7327
Stanieria cyanosphaera PCC 7437
Cyanobacterium stanieri PCC 7202
Microcystis aeruginosa NIES-843
Gloeocapsa sp PCC 73106
Candidatus Atelocyanobacterium thalassa isolate ALOHA
Rubidibacter lacunae KORDI 51-2
Mastigocoleus testarum BC008
Calothrix sp PCC 7507

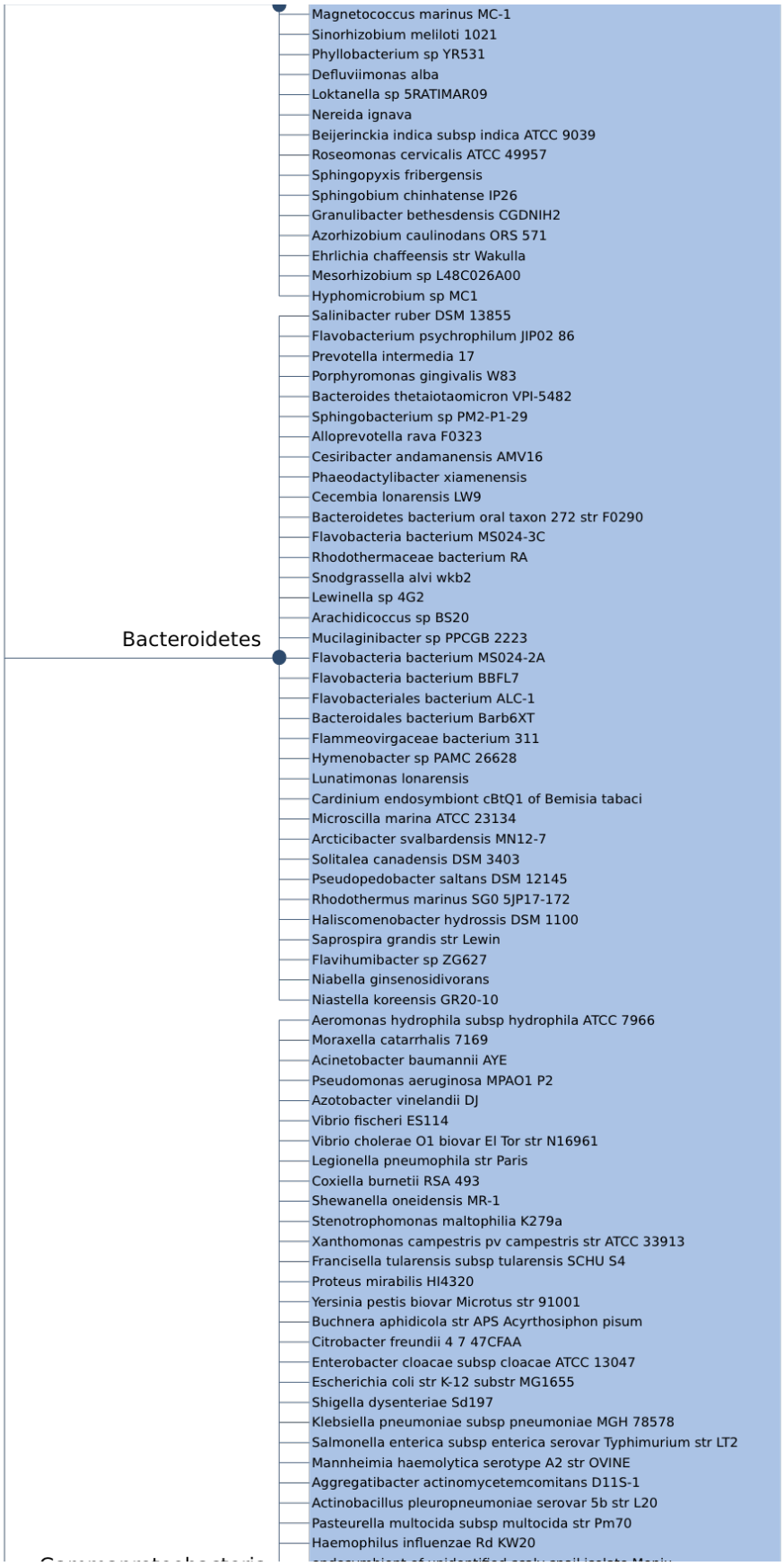

Cellular organisms
ps1
Cyanobacteria
Cylindrospermopsis raciborskii CS-505
Aphanizomenon flos-aquae 2012 KM1 D3
Nodularia spumigena CCY9414
Anabaena cylindrica PCC 7122
Nostoc punctiforme PCC 73102
Cylindrospermum stagnale PCC 7417
Tolypothrix campylonemoides VB511288
Scytonema tolypothrichoides VB-61278
Scytonema hofmannii PCC 7110
Leptolyngbya valderiana BDU 20041
Phormidesmis priestleyi Ana
Neosynechococcus sphagnicola sy1
Pseudanabaena sp PCC 7367
Pseudanabaena biceps PCC 7429
Prochlorococcus sp MIT 0601
Prochlorococcus marinus subsp marinus str CCMP1375
Synechocystis sp PCC 6803
Thermosynechococcus elongatus BP-1
Cyanobium gracile PCC 6307
Candidatus Synechococcus spongiarum SP3
Dactylococcopsis salina PCC 8305
Chamaesiphon minutus PCC 6605
Acaryochloris marina MBIC11017
Aliterella atlantica CENA595
Chroococcidiopsis thermalis PCC 7203
Gloeobacter kilaueensis JS1
Gloeobacter violaceus PCC 7421
Micrococcus luteus NCTC 2665
Propionibacterium acnes KPA171202
Gardnerella vaginalis 0288E
Bifidobacterium longum NCC2705
Streptomyces coelicolor A3 2
Corynebacterium glutamicum ATCC 13032
Mycobacterium tuberculosis H37Rv
Actinobacterium SCGC AAA027-L06
Thermobispora bispora DSM 43833
Streptomyces aureofaciens
Streptomyces thermoautotrophicus
Streptomyces chattanoogensis
Pimelobacter simplex
Lawsonella clevelandensis
Gordonia bronchialis DSM 43247
Dietzia maris
Nocardia asteroides NBRC 15531
Segniliparus rugosus ATCC BAA-974
Varibaculum cambriense
Actinotignum schaalii
Trueperella bernardiae
Actinomyces urogenitalis DSM 15434
Mobiluncus mulieris ATCC 35243
Geodermatophilus sp Leaf369
Blastococcus saxobsidens DD2
Actinopolyspora erythraea
Kineococcus radiotolerans SRS30216 = ATCC BAA-149
Thermobifida cellulosilytica TB100
Thermomonospora curvata DSM 43183
Cryptosporangium arvum DSM 44712
Asanoa ferruginea
Micromonospora rifamycinica
Actinoplanes utahensis
Lentzea guizhouensis
Actinokineospora spheciospongiae
Kibdelosporangium phytohabitans
Pseudonocardia dioxanivorans CB1190
Actinoalloteichus hymeniacidonis
Lechevalieria aerocolonigenes
Saccharomonospora azurea NA-128
Actinosynnema mirum DSM 43827
Gardnerella vaginalis
Bifidobacterium thermophilum
Actinobacteria
Alloscardovia omnicolens F0580
Scardovia inopinata F0304
Sanguibacter sp Leaf3
Brachybacterium sp SW0106-09
Isoptericola variabilis 225
Arsenicicoccus sp oral taxon 190

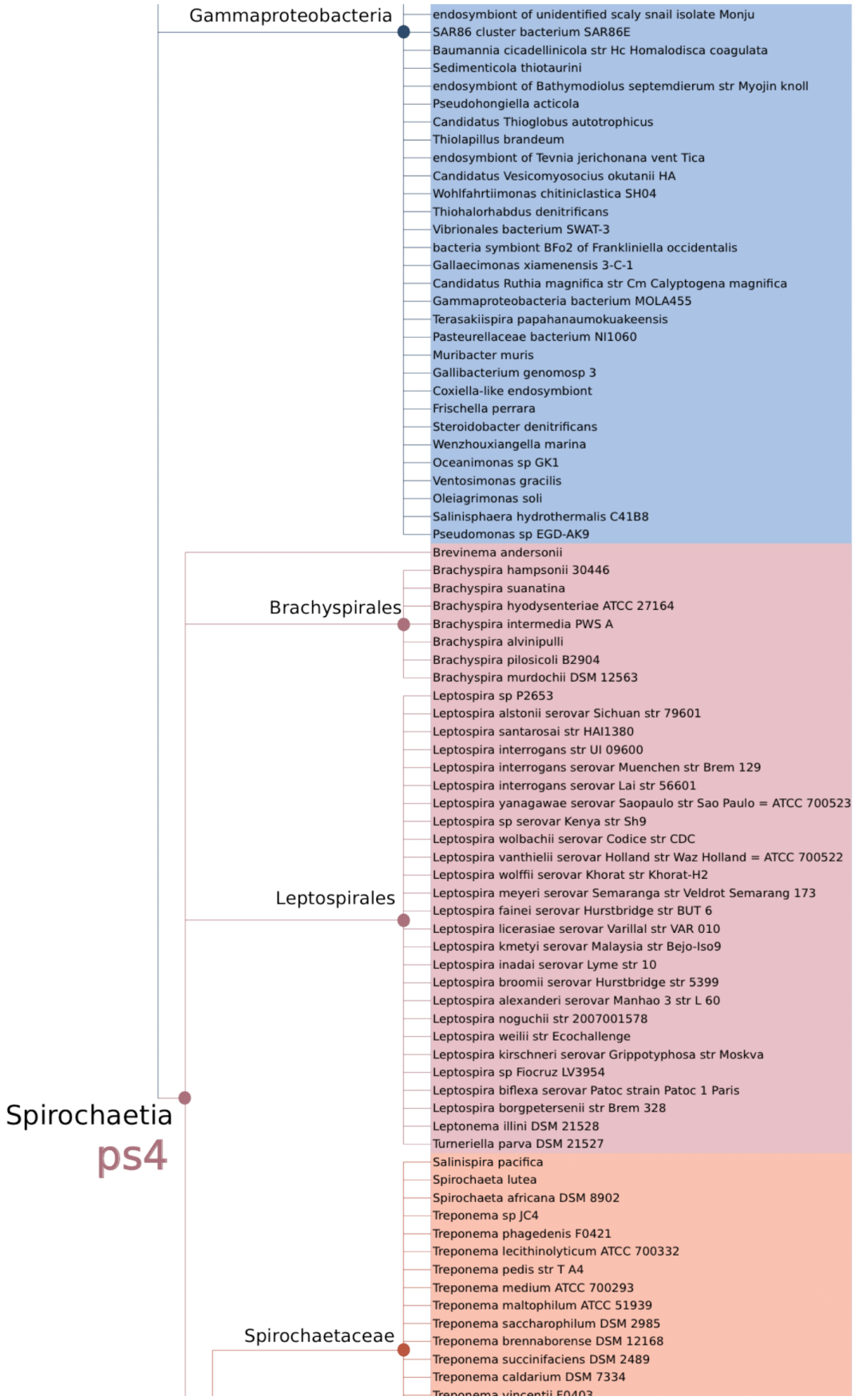

Serinicoccus chungangensis
Intrasporangium chromatireducens Q5-1
Tropheryma whipplei str Twist
Cellulomonas bogoriensis 69B4 = DSM 16987
Oerskovia enterophila
Yonghaparkia sp Soil809
Plantibacter sp Leaf1
Cryobacterium sp MLB-32
Rhodoluna lacicola
Agromyces sp Leaf222
Agreia bicolorata
Leucobacter sp G161
Rathayibacter tanaceti
Curtobacterium oceanosedimentum
Microbacterium mangrovi
Agrococcus pavilionensis RW1
Dermacoccus sp Ellin185
Kytococcus sedentarius DSM 20547
Arthrobacter enclensis
Arthrobacter siccitolerans
Kocuria kristinae
Pseudoglutamicibacter albus DNF00011
Kineosphaera limosa NBRC 100340
Jonesia denitrificans DSM 20603
Beutenbergia cavernae DSM 12333
Nakamurella multipartita DSM 44233
Catenulispora acidiphila DSM 44928
Stackebrandtia nassauensis DSM 44728
Acidothermus cellulolyticus 11B
Coriobacteriaceae bacterium BV3Ac1
Coriobacterium glomerans PW2
Olsenella profusa F0195
Eggerthella sp YY7918
Gordonibacter pamelaeae 7-10-1-b
Cryptobacterium curtum DSM 15641
Acidithrix ferrooxidans
Ilumatobacter coccineus YM16-304
Patulibacter medicamentivorans
Rubrobacter xylanophilus DSM 9941
Limnochorda pilosa
Tissierellia bacterium S5-A11
Peptoniphilus indolicus ATCC 29427
Peptoniphilus sp oral taxon 386 str F0131
Tissierellia
Parvimonas micra
Anaerococcus tetradius
Anaerococcus vaginalis ATCC 51170
Finegoldia magna
Helcococcus kunzii ATCC 51366
Veillonellaceae bacterium DNF00751
Dialister succinatiphilus YIT 11850
Veillonella atypica
Megasphaera cerevisiae DSM 20462
Negavicutes
Anaeroglobus geminatus F0357
Sporomusa sp An4
Megamonas funiformis YIT 11815
Phascolarctobacterium succinatutens YIT 12067
Acidaminococcus fermentans DSM 20731
Natranaerobius thermophilus JW NM-WN-LF
Halanaerobium praevalens DSM 2228
Caldicellulosiruptor kronotskyensis 2002
Caldicellulosiruptor obsidiansis OB47
Fervidicola ferrireducens
Thermoanaerobacter italicus Ab9
Moorella thermoacetica ATCC 39073
Tepidanaerobacter acetatoxydans Re1
Caldanaerobacter subterraneus subsp tengcongensis MB4
Carboxydothermus hydrogenoformans Z-2901
Firmicutes
Thermodesulfobium narugense DSM 14796
Clostridia
Clostridium papyrosolvens DSM 2782
Oscillibacter valericigenes Sjm18-20
Mogibacterium timidum ATCC 33093
Pseudoramibacter alactolyticus ATCC 23263
Thermaerobacter subterraneus DSM 13965
Heliobacterium modesticaldum Ice1
Peptoclostridium acidaminophilum DSM 3953
Desulfotomaculum ruminis DSM 2154


ps2
Bacteria A
Alicyclobacillaceae
Thermoactinomycetaceae
Paenibacillaceae
Erysipelotrichaceae
Tenericutes
Lactobacillales
Bacillales
Elusimicrobia
Ignavibacteriae
Fibrobacteres
Blautia hydrogenotrophica DSM 10507
Clostridium botulinum A str Hall
Clostridioides difficile 630
Tumebacillus flagellatus
Alicyclobacillus acidoterrestris ATCC 49025
Kyrpidia tusciae DSM 2912
Desmospora sp 8437
Thermoactinomyces sp CDF
Thermoactinomyces vulgaris
Saccharibacillus sacchari DSM 19268
Thermobacillus composti KWC4
Aneurinibacillus sp XH2
Paenibacillus stellifer
Brevibacillus agri BAB-2500
Paenibacillus lactis 154
Erysipelotrichaceae bacterium 2 2 44A
Candidatus Stoquefichus sp KLE1796
Erysipelatoclostridium ramosum DSM 1402
Eggerthia catenaformis OT 569 = DSM 20559
Bulleidia extructa W1219
Holdemania filiformis DSM 12042
Catenibacterium mitsuokai DSM 15897
Candidatus Izimaplasma sp HR1
Strawberry lethal yellows phytoplasma CPA str NZSb11
Acholeplasma palmae J233
Mesoplasma florum L1
Spiroplasma turonicum
Candidatus Hepatoplasma crinochetorum
Ureaplasma parvum serovar 3 str ATCC 700970
Mycoplasma canadense
Carnobacterium maltaromaticum
Aerococcus urinaeequi
Catellicoccus marimammalium M35 04 3
Enterococcus faecalis V583
Lactococcus lactis subsp lactis Il1403
Streptococcus pneumoniae TIGR4
Pediococcus parvulus
Lactobacillus plantarum WCFS1
Weissella minor
Leuconostoc mesenteroides subsp mesenteroides ATCC 8293
Sporolactobacillus laevolacticus DSM 442
Acidibacillus ferrooxidans
Geomicrobium sp JCM 19037
Brochothrix thermosphacta DSM 20171 = FSL F6-1036
Listeria monocytogenes EGD-e
Staphylococcus aureus subsp aureus N315
Macrococcus caseolyticus JCSC5402
Gemella bergeriae ATCC 700627
Exiguobacterium antarcticum B7
Paraliobacillus sp PM-2
Terribacillus aidingensis
Pontibacillus chungwhensis BH030062
Virgibacillus soli
Bacillus aryabhattai
Bacillus oceanisediminis 2691
Geobacillus kaustophilus
Bacillus clausii KSM-K16
Bacillus subtilis subsp subtilis str 168
Geobacillus thermoglucosidans TNO-09 020
Bacillus infantis NRRL B-14911
Anoxybacillus ayderensis
Halobacillus karajensis
Oceanobacillus picturae
Fictibacillus arsenicus
Lysinibacillus sphaericus C3-41
Planomicrobium glaciei
Bacillus subtilis subsp subtilis str NCIB 3610
Balneola sp EhC07
Caldithrix abyssi DSM 13497
Lentisphaera araneosa HTCC2155
Nitrospina gracilis 3 211
Candidatus Endomicrobium trichonymphae
Elusimicrobium minutum Pei191
Melioribacter roseus P3M-2
Ignavibacterium album JCM 16511
Chitinispirillum alkaliphilum
Chitinivibrio alkaliphilus ACht1


Fibrobacter succinogenes subsp succinogenes S85
Gemmatimonadaceae
Gemmatirosa kalamazoonesis
Gemmatimonas phototrophica
Gemmatimonas aurantiaca T-27
other Proteobacteria
Tepidiphilus thermophilus
Bacteriovorax sp Seq25 V
Bdellovibrio exovorus JSS
Acidithiobacillus sp GGI-221
Deferribacteres
Mucispirillum schaedleri ASF457
Deferribacter desulfuricans SSM1
Denitrovibrio acetiphilus DSM 12809
Calditerrivibrio nitroreducens DSM 19672
Flexistipes sinusarabici DSM 4947
Thermodesulfobacteriaceae
Caldimicrobium thiodismutans
Thermosulfurimonas dismutans
Thermodesulfatator autotrophicus
Thermodesulfatator indicus DSM 15286
Thermodesulfobacterium geofontis OPF15
Thermodesulfobacterium commune DSM 2178
Nitrospirae
Thermodesulfovibrio yellowstonii DSM 11347
Nitrospirae bacterium HCH-1
Candidatus Magnetobacterium bavaricum
Leptospirillum ferriphilum
Nitrospira sp SCGC AG-212-E16
Candidatus Magnetoovum chiemensis
Chlorobi
Chlorobium tepidum TLS
Chlorobium sp GBChlB
Prosthecochloris aestuarii DSM 271
Chlorobaculum limnaeum
Chlorobium ferrooxidans DSM 13031
Chloroherpeton thalassium ATCC 35110
Pelodictyon phaeoclathratiforme BU-1
Aquificae
Aquifex aeolicus VF5
Thermosulfidibacter takaii ABI70S6
Sulfurihydrogenibium yellowstonense SS-5
Persephonella marina EX-H1
Hydrogenivirga sp 128-5-R1-1
Hydrogenobaculum sp Y04AAS1
Thermocrinis albus DSM 14484
Hydrogenobacter thermophilus TK-6
Acidobacteria
Acidobacteria bacterium Mor1
Thermoanaerobaculum aquaticum
Chloracidobacterium thermophilum B
Pyrinomonas methylaliphatogenes
Candidatus Solibacter usitatus Ellin6076
Granulicella tundricola MP5ACTX9
Granulicella mallensis MP5ACTX8
Terriglobus roseus DSM 18391
Terriglobus saanensis SP1PR4
Acidobacterium capsulatum ATCC 51196
Candidatus Koribacter versatilis Ellin345
Epsilonproteobacteria
Campylobacter jejuni subsp jejuni NCTC 11168 = ATCC 700819
Helicobacter pylori 26695
Sulfurovum riftiae
Sulfurospirillum sp SCADC
Campylobacter rectus RM3267
Nitratifractor salsuginis DSM 16511
Helicobacter sp CLO-3
Wolinella succinogenes DSM 1740
Caminibacter mediatlanticus TB-2
Arcobacter butzleri L354
Nitratiruptor sp SB155-2
Thiovulum sp ES
Sulfuricurvum sp MLSB
Nautilia profundicola AmH
Betaproteobacteria
Burkholderia pseudomallei 1710b
Ralstonia solanacearum GMI1000
Bordetella pertussis Tohama I
Neisseria meningitidis Z2491
Methylobacillus flagellatus KT
Sulfuricella sp T08
Methylibium sp T29
Oxalobacteraceae bacterium IMCC9480
Acidovorax delafieldii 2AN
Sutterella parvirubra YIT 11816
Advenella mimigardefordensis DPN7
Thauera phenylacetica B4P


ps3
Bacteria B
PVC
Deltaproteobacteria
Alphaproteobacteria
Chromobacterium piscinae
Thiomonas arsenitoxydans
Collimonas fungivorans Ter331
Pandoraea sp ISTKB
Cupriavidus sp HPC L
Bacteroides caecimuris
Achromobacter sp 2789STDY5608607
Polaromonas naphthalenivorans CJ2
Delftia sp Cs1-4
Chlamydia trachomatis D UW-3 CX
Rhodopirellula baltica SH 1
Opitutae bacterium SCGC AG-212-L18
Cephaloticoccus primus
Verrucomicrobium spinosum
Chthoniobacter flavus Ellin428
Methylacidiphilum fumariolicum SolV
Pirellula staleyi DSM 6068
Rhodopirellula sp SWK7
Blastopirellula marina DSM 3645
Singulisphaera acidiphila DSM 18658
Chlamydophila pneumoniae AR39
Opitutaceae bacterium TSB47
Opitutus terrae PB90-1
Coraliomargarita akajimensis DSM 45221
Verrucomicrobiae bacterium DG1235
Akkermansia sp KLE1798
Pedosphaera parvula Ellin514
Candidatus Xiphinematobacter sp Idaho Grape
Terrimicrobium sacchariphilum
Gemmata sp SH-PL17
Planctomyces sp SH-PL14
Rubinisphaera brasiliensis DSM 5305
Gimesia maris DSM 8797
Isosphaera pallida ATCC 43644
Candidatus Rubidus massiliensis
Chondromyces apiculatus DSM 436
Myxococcus xanthus DK 1622
Desulfovibrio vulgaris str Hildenborough
Cystobacter fuscus DSM 2262
Dethiosulfatarculus sandiegensis
Desulfotalea psychrophila LSv54
Smithella sp sc k08d17
Desulfomicrobium baculatum DSM 4028
Geobacter sulfurreducens PCA
Lawsonia intracellularis N343
Desulfurella acetivorans A63
Corallococcus coralloides DSM 2259
Desulfuromonas soudanensis
Desulfatibacillum alkenivorans AK-01
Desulfonatronospira thiodismutans ASO3-1
Desulfococcus multivorans DSM 2059
Candidatus Magnetomorum sp HK-1
Enhygromyxa salina
Labilithrix luteola
Geoalkalibacter subterraneus
Syntrophobacter fumaroxidans MPOB
Anaeromyxobacter dehalogenans 2CP-1
Sorangium cellulosum So0157-2
Hyalangium minutum
Archangium gephyra
Haliangium ochraceum DSM 14365
Plesiocystis pacifica SIR-1
Wolbachia endosymbiont of Drosophila melanogaster
Anaplasma phagocytophilum str HZ
Rickettsia prowazekii str Madrid E
Paracoccus denitrificans PD1222
Rhodobacter sphaeroides 2 4 1
Rhodospirillum rubrum ATCC 11170
Caulobacter vibrioides
Caulobacter crescentus cb15
Bartonella henselae str Houston-1
Brucella abortus bv 1 str 9-941
Bradyrhizobium diazoefficiens USDA 110
Rhizobium leguminosarum bv viciae 3841
Agrobacterium fabrum str C58
Alpha proteobacterium BAL199
Rhodovulum sp PH10


Magnetococcus marinus MC-1
Sinorhizobium meliloti 1021
Phyllobacterium sp YR531
Defluviimonas alba
Loktanella sp 5RATIMAR09
Nereida ignava
Beijerinckia indica subsp indica ATCC 9039
Roseomonas cervicalis ATCC 49957
Sphingopyxis fribergensis
Sphingobium chinhatense IP26
Granulibacter bethesdensis CGDNIH2
Azorhizobium caulinodans ORS 571
Ehrlichia chaffeensis str Wakulla
Mesorhizobium sp L48C026A00
Hyphomicrobium sp MC1
Salinibacter ruber DSM 13855
Flavobacterium psychrophilum JIP02 86
Prevotella intermedia 17
Porphyromonas gingivalis W83
Bacteroides thetaiotaomicron VPI-5482
Sphingobacterium sp PM2-P1-29
Alloprevotella rava F0323
Cesiribacter andamanensis AMV16
Phaeodactylibacter xiamenensis
Cecembia lonarensis LW9
Bacteroidetes bacterium oral taxon 272 str F0290
Flavobacteria bacterium MS024-3C
Rhodothermaceae bacterium RA
Snodgrassella alvi wkb2
Lewinella sp 4G2
Arachidicoccus sp BS20
Bacteroidetes
Mucilaginibacter sp PPCGB 2223
Flavobacteria bacterium MS024-2A
Flavobacteria bacterium BBFL7
Flavobacteriales bacterium ALC-1
Bacteroidales bacterium Barb6XT
Flammeovirgaceae bacterium 311
Hymenobacter sp PAMC 26628
Lunatimonas lonarensis
Cardinium endosymbiont cBtQ1 of Bemisia tabaci
Microscilla marina ATCC 23134
Arcticibacter svalbardensis MN12-7
Solitalea canadensis DSM 3403
Pseudopedobacter saltans DSM 12145
Rhodothermus marinus SG0 5JP17-172
Haliscomenobacter hydrossis DSM 1100
Saprospira grandis str Lewin
Flavihumibacter sp ZG627
Niabella ginsenosidivorans
Niastella koreensis GR20-10
Aeromonas hydrophila subsp hydrophila ATCC 7966
Moraxella catarrhalis 7169
Acinetobacter baumannii AYE
Pseudomonas aeruginosa MPAO1 P2
Azotobacter vinelandii DJ
Vibrio fischeri ES114
Vibrio cholerae O1 biovar El Tor str N16961
Legionella pneumophila str Paris
Coxiella burnetii RSA 493
Shewanella oneidensis MR-1
Stenotrophomonas maltophilia K279a
Xanthomonas campestris pv campestris str ATCC 33913
Francisella tularensis subsp tularensis SCHU S4
Proteus mirabilis HI4320
Yersinia pestis biovar Microtus str 91001
Buchnera aphidicola str APS Acyrthosiphon pisum
Citrobacter freundii 4 7 47CFAA
Enterobacter cloacae subsp cloacae ATCC 13047
Escherichia coli str K-12 substr MG1655
Shigella dysenteriae Sd197
Klebsiella pneumoniae subsp pneumoniae MGH 78578
Salmonella enterica subsp enterica serovar Typhimurium str LT2
Mannheimia haemolytica serotype A2 str OVINE
Aggregatibacter actinomycetemcomitans D11S-1
Actinobacillus pleuropneumoniae serovar 5b str L20
Pasteurella multocida subsp multocida str Pm70
Haemophilus influenzae Rd KW20


Gammaproteobacteria
endosymbiont of unidentified scaly snail isolate Monju
SAR86 cluster bacterium SAR86E
Baumannia cicadellinicola str Hc Homalodisca coagulata
Sedimenticola thiotaurini
endosymbiont of Bathymodiolus septemdierum str Myojin knoll
Pseudohongiella acticola
Candidatus Thioglobus autotrophicus
Thiolapillus brandeum
endosymbiont of Tevnia jerichonana vent Tica
Candidatus Vesicomyosocius okutanii HA
Wohlfahrtiimonas chitiniclastica SH04
Thiohalorhabdus denitrificans
Vibrionales bacterium SWAT-3
bacteria symbiont BFo2 of Frankliniella occidentalis
Gallaecimonas xiamenensis 3-C-1
Candidatus Ruthia magnifica str Cm Calyptogena magnifica
Gammaproteobacteria bacterium MOLA455
Terasakiispira papahanaumokuakeensis
Pasteurellaceae bacterium NI1060
Muribacter muris
Gallibacterium genomosp 3
Coxiella-like endosymbiont
Frischella perrara
Steroidobacter denitrificans
Wenzhouxiangella marina
Oceanimonas sp GK1
Ventosimonas gracilis
Oleiagrimonas soli
Salinisphaera hydrothermalis C41B8
Pseudomonas sp EGD-AK9
Brachyspirales
Brevinema andersonii
Brachyspira hampsonii 30446
Brachyspira suanatina
Brachyspira hyodysenteriae ATCC 27164
Brachyspira intermedia PWS A
Brachyspira alvinipulli
Brachyspira pilosicoli B2904
Brachyspira murdochii DSM 12563
Leptospirales
Leptospira sp P2653
Leptospira alstonii serovar Sichuan str 79601
Leptospira santarosai str HAI1380
Leptospira interrogans str UI 09600
Leptospira interrogans serovar Muenchen str Brem 129
Leptospira interrogans serovar Lai str 56601
Leptospira yanagawae serovar Saopaulo str Sao Paulo = ATCC 700523
Leptospira sp serovar Kenya str Sh9
Leptospira wolbachii serovar Codice str CDC
Leptospira vanthielii serovar Holland str Waz Holland = ATCC 700522
Leptospira wolffii serovar Khorat str Khorat-H2
Leptospira meyeri serovar Semaranga str Veldrot Semarang 173
Leptospira fainei serovar Hurstbridge str BUT 6
Leptospira licerasiae serovar Varillal str VAR 010
Leptospira kmetyi serovar Malaysia str Bejo-Iso9
Leptospira inadai serovar Lyme str 10
Leptospira broomii serovar Hurstbridge str 5399
Leptospira alexanderi serovar Manhao 3 str L 60
Leptospira noguchii str 2007001578
Leptospira weilii str Ecochallenge
Leptospira kirschneri serovar Grippotyphosa str Moskva
Leptospira sp Fiocruz LV3954
Leptospira biflexa serovar Patoc strain Patoc 1 Paris
Leptospira borgpetersenii str Brem 328
Leptonema illini DSM 21528
Turneriella parva DSM 21527
Spirochaetia
ps4
Salinispira pacifica
Spirochaeta lutea
Spirochaeta africana DSM 8902
Treponema sp JC4
Treponema phagedenis F0421
Treponema lecithinolyticum ATCC 700332
Treponema pedis str T A4
Treponema medium ATCC 700293
Treponema maltophilum ATCC 51939
Treponema saccharophilum DSM 2985
Spirochaetaceae
Treponema brennaborense DSM 12168
Treponema succinifaciens DSM 2489
Treponema caldarium DSM 7334

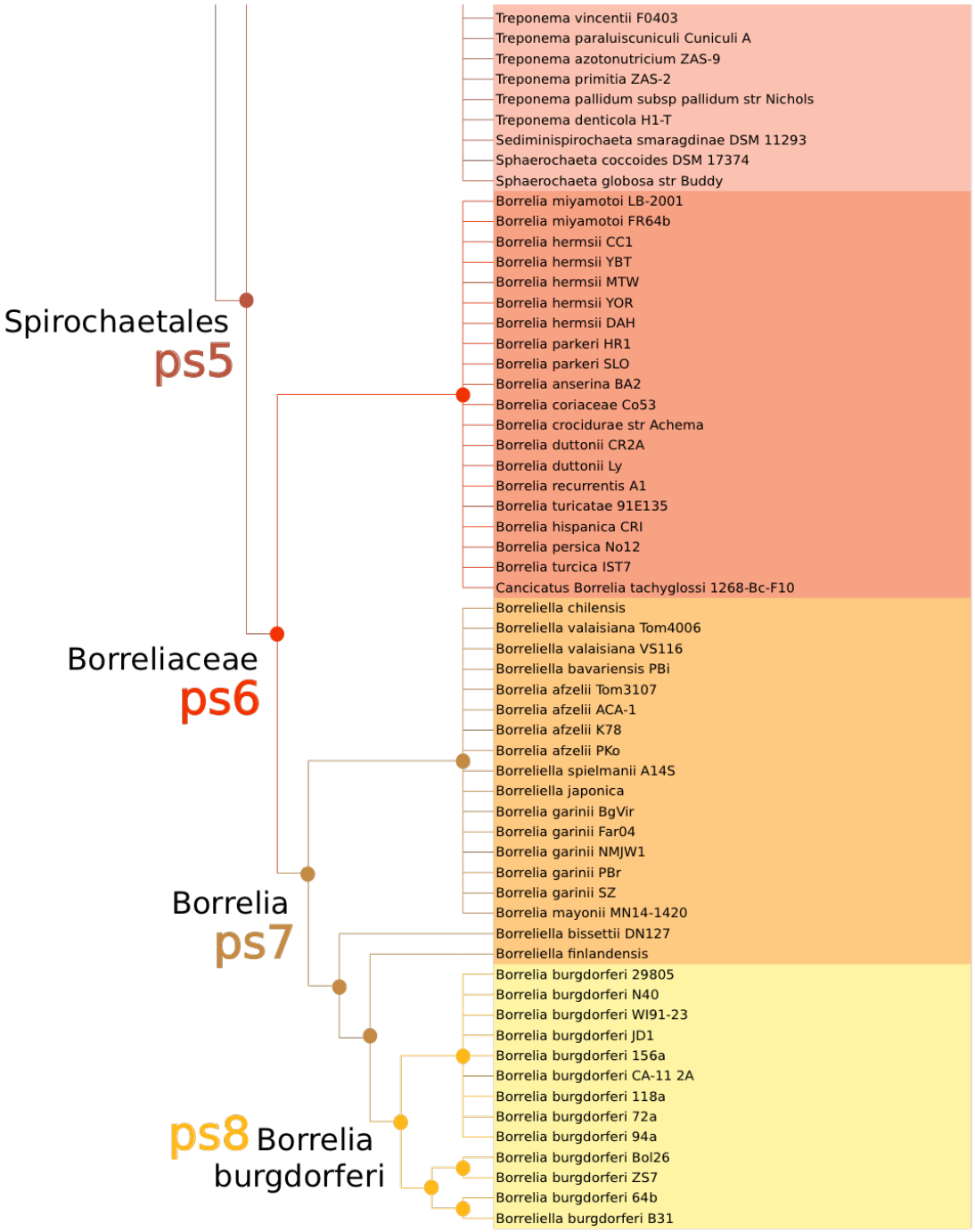



**Figure S24.** The consensus phylogeny used in the phylostratigraphic analysis. The consensus tree covers divergence from the last common ancestor of cellular organisms to *B. burgdorferi* B31 as a focal organism. It is constructed based on the importance of evolutionary transitions, availability of reference genomes and thier completeness estimated using BUSCO scores. Eight phylostrata defined in the phylostratigraphic analysis are marked by ps1-ps8.

**Table S25.** Enrichment of phylostratum (ps) specific genes in *B. burgdorferi* round body upregulated genes. For a gene to be categorized as differentially expressed, the adjusted p-value ($p_{ajd}$) calculated by DeSeq2 pairwise comparison analysis had to be below 0.05. P values were adjusted for multiple comparisons using the Benjamini and Hochberg procedure ($p_{adj}$). q = the number of phylostratum genes in round body upregulated genes, s = the total number of phylostratum genes in round body upregulated genes, h = the number of phylostratum genes in *B. burgdorferi* genome, t = the total number of genes in *B. burgdorferi* genome which passed the phylostratigraphic procedure. The abbreviation –inf marks it was not possible to calculate the log-odds.

| ps | ps_name | q | s | h | t | $p_{adj}$ | log-odds |
|---|---|---|---|---|---|---|---|
| 1 | Cellular organisms | 39 | 44 | 557 | 1344 | 8.078E-10 | 2.5 |
| 2 | Bacteria A | 4 | 44 | 148 | 1344 | 1 | -0.2 |
| 3 | Bacteria B | 0 | 44 | 24 | 1344 | 1 | -inf |
| 4 | Spirochaetia | 0 | 44 | 18 | 1344 | 1 | -inf |
| 5 | Spirochaetales | 0 | 44 | 46 | 1344 | 0.842 | -inf |
| 6 | Borreliaceae | 0 | 44 | 446 | 1344 | 1.1E-07 | -inf |
| 7 | Borrelia | 1 | 44 | 88 | 1344 | 0.842 | -1.1 |
| 8 | Borrelia burgdorferi | 0 | 44 | 17 | 1344 | 1 | -inf |

**Table S26.** Enrichment of phylostratum (ps) specific genes in *B. burgdorferi* bleb upregulated genes. For a gene to be categorized as differentially expressed, the adjusted p-value ($p_{ajd}$) calculated by DeSeq2 pairwise comparison analysis had to be below 0.05. P values were adjusted for multiple comparisons using the Benjamini and Hochberg procedure ($p_{adj}$). q = the number of phylostratum genes in bleb upregulated genes, s = the total number of all phylostratum genes in bleb upregulated genes, h = the number of phylostratum genes in *B. burgdorferi* genome, t = the total number of all genes in *B. burgdorferi* genome which passed the phylostratigraphic procedure. The abbreviation –inf marks it was not possible to calculate the log-odds.

| ps | ps_name | q | s | h | t | $p_{adj}$ | log-odds |
|---|---|---|---|---|---|---|---|
| 1 | Cellular organisms | 158 | 478 | 557 | 1344 | 1.6E-05 | -0.5 |
| 2 | Bacteria A | 39 | 478 | 148 | 1344 | 0.024 | -0.5 |
| 3 | Bacteria B | 3 | 478 | 24 | 1344 | 0.029 | -1.4 |
| 4 | Spirochaetia | 1 | 478 | 18 | 1344 | 0.015 | -2.3 |
| 5 | Spirochaetales | 5 | 478 | 46 | 1344 | 0.0006 | -1.5 |
| 6 | Borreliaceae | 227 | 478 | 446 | 1344 | 1.9E-15 | 1 |
| 7 | Borrelia | 39 | 478 | 88 | 1344 | 0.114 | 0.4 |
| 8 | Borrelia burgdorferi | 6 | 478 | 17 | 1344 | 1 | -0.01 |

**Table S27.** Enrichment of phylostratum (ps) specific genes in *B. burgdorferi* biofilm upregulated genes. For a gene to be categorized as differentially expressed, the adjusted p-value ($p_{ajd}$) calculated by DeSeq2 pairwise comparison analysis had to be below 0.05. P values were adjusted for multiple comparisons using the Benjamini and Hochberg procedure ($p_{adj}$). q = the number of phylostratum genes in biofilm upregulated genes, s = the total number of phylostratum genes in biofilm upregulated genes, h = the number of phylostratum genes in *B. burgdorferi* genome, t = the total number of genes in *B. burgdorferi* genome which passed the phylostratigraphic procedure. The abbreviation –inf marks it was not possible to calculate the log-odds.

| ps | ps_name | q | s | h | t | $p_{adj}$ | log-odds |
|---|---|---|---|---|---|---|---|
| 1 | Cellular organisms | 181 | 435 | 557 | 1344 | 0.978 | 0.01 |
| 2 | Bacteria A | 38 | 435 | 148 | 1344 | 0.164 | -0.4 |
| 3 | Bacteria B | 3 | 435 | 24 | 1344 | 0.164 | -1.2 |
| 4 | Spirochaetia | 3 | 435 | 18 | 1344 | 0.309 | -0.9 |
| 5 | Spirochaetales | 11 | 435 | 46 | 1344 | 0.315 | -0.4 |
| 6 | Borreliaceae | 176 | 435 | 446 | 1344 | 0.001 | 0.5 |
| 7 | Borrelia | 21 | 435 | 88 | 1344 | 0.164 | -0.5 |
| 8 | Borrelia burgdorferi | 2 | 435 | 17 | 1344 | 0.164 | -1.3 |

**Table S28.** Enrichment of phylostratum (ps) specific genes in *B. burgdorferi* bleb upregulated genes. For a gene to be categorized as differentially expressed, the adjusted p-value ($p_{ajd}$) and the absolute value of log2FC (|log2 FC|) calculated by DeSeq2 pairwise comparison analysis had to be below 0.05 and more than 1, respectively. P values were adjusted for multiple comparisons using the Benjamini and Hochberg procedure ($p_{adj}$). q = the number of phylostratum genes in bleb upregulated genes, s = the total number of phylostratum genes in bleb upregulated genes, h = the number of phylostratum genes in *B. burgdorferi* genome, t = the total number of genes in *B. burgdorferi* genome which passed the phylostratigraphic procedure. The abbreviation –inf marks it was not possible to calculate the log-odds.

| ps | ps_name | q | s | h | t | $p_{adj}$ | log-odds |
|---|---|---|---|---|---|---|---|
| 1 | Cellular organisms | 37 | 238 | 557 | 1344 | 4.4E-20 | -1.6 |
| 2 | Bacteria A | 9 | 238 | 148 | 1344 | 6.9E-05 | -1.3 |
| 3 | Bacteria B | 0 | 238 | 24 | 1344 | 0.034 | -inf |
| 4 | Spirochaetia | 0 | 238 | 18 | 1344 | 0.067 | -inf |
| 5 | Spirochaetales | 3 | 238 | 46 | 1344 | 0.067 | -1.2 |
| 6 | Borreliaceae | 161 | 238 | 446 | 1344 | 1.8E-32 | 1.8 |
| 7 | Borrelia | 23 | 238 | 88 | 1344 | 0.067 | 0.5 |
| 8 | Borrelia burgdorferi | 5 | 238 | 17 | 1344 | 0.335 | 0.7 |

**Table S29.** Enrichment of phylostratum (ps) specific genes in *B. burgdorferi* biofilm upregulated genes. For a gene to be categorized as differentially expressed, the adjusted p-value ($p_{ajd}$) and the absolute value of log2FC (|log2 FC|) calculated by DeSeq2 pairwise comparison analysis had to be below 0.05 and more than 1, respectively. P values were adjusted for multiple comparisons using the Benjamini and Hochberg procedure ($p_{adj}$). q = the number of phylostratum genes in biofilm upregulated genes, s = the total number of phylostratum genes in biofilm upregulated genes, h = the number of phylostratum genes in *B. burgdorferi* genome, t = the total number of genes in *B. burgdorferi* genome which passed the phylostratigraphic procedure. The abbreviation –inf marks it was not possible to calculate the log-odds.

| ps | ps_name | q | s | h | t | $p_{adj}$ | log-odds |
|---|---|---|---|---|---|---|---|
| 1 | Cellular organisms | 29 | 136 | 557 | 1344 | 1.3E-06 | -1.1 |
| 2 | Bacteria A | 6 | 136 | 148 | 1344 | 0.021 | -1.1 |
| 3 | Bacteria B | 1 | 136 | 24 | 1344 | 0.907 | -1 |
| 4 | Spirochaetia | 0 | 136 | 18 | 1344 | 0.579 | -inf |
| 5 | Spirochaetales | 4 | 136 | 46 | 1344 | 1 | -0.2 |
| 6 | Borreliaceae | 86 | 136 | 446 | 1344 | 5.3E-13 | 1.4 |
| 7 | Borrelia | 9 | 136 | 88 | 1344 | 1 | 0.01 |
| 8 | Borrelia burgdorferi | 1 | 136 | 17 | 1344 | 1 | -0.6 |

**Table S30.** Enrichment of phylostratum (ps) specific genes in *B. burgdorferi* round body downregulated genes. For a gene to be categorized as differentially expressed, the adjusted p-value ($p_{ajd}$) calculated by DeSeq2 pairwise comparison analysis had to be below 0.05. P values were adjusted for multiple comparisons using the Benjamini and Hochberg procedure ($p_{adj}$). q = the number of phylostratum genes in round body downregulated genes, s = the total number of phylostratum genes in round body downregulated genes, h = the number of phylostratum genes in *B. burgdorferi* genome, t = the total number of genes in *B. burgdorferi* genome which passed the phylostratigraphic procedure. The abbreviation –inf marks it was not possible to calculate the log-odds.

| ps | ps_name | q | s | h | t | $p_{adj}$ | log-odds |
|---|---|---|---|---|---|---|---|
| 1 | Cellular organisms | 8 | 23 | 557 | 1344 | 1 | -0.3 |
| 2 | Bacteria A | 2 | 23 | 148 | 1344 | 1 | -0.3 |
| 3 | Bacteria B | 1 | 23 | 24 | 1344 | 1 | 0.9 |
| 4 | Spirochaetia | 1 | 23 | 18 | 1344 | 1 | 1.3 |
| 5 | Spirochaetales | 4 | 23 | 46 | 1344 | 0.106 | 1.9 |
| 6 | Borreliaceae | 5 | 23 | 446 | 1344 | 1 | -0.6 |
| 7 | Borrelia | 2 | 23 | 88 | 1344 | 1 | 0.3 |
| 8 | Borrelia burgdorferi | 0 | 23 | 17 | 1344 | 1 | -inf |

**Table S31.** Enrichment of phylostratum (ps) specific genes in *B. burgdorferi* bleb downregulated genes. For a gene to be categorized as differentially expressed, the adjusted p-value ($p_{ajd}$) calculated by DeSeq2 pairwise comparison analysis had to be below 0.05. P values were adjusted for multiple comparisons using the Benjamini and Hochberg procedure ($p_{adj}$). q = the number of phylostratum genes in bleb downregulated genes, s = the total number of phylostratum genes in bleb downregulated genes, h = the number of phylostratum genes in *B. burgdorferi* genome, t = the total number of genes in *B. burgdorferi* genome which passed the phylostratigraphic procedure. The abbreviation –inf marks it was not possible to calculate the log-odds.

| ps | ps_name | q | s | h | t | $p_{adj}$ | log-odds |
|---|---|---|---|---|---|---|---|
| 1 | Cellular organisms | 272 | 508 | 557 | 1344 | 1.4E-11 | 0.8 |
| 2 | Bacteria A | 80 | 508 | 148 | 1344 | 4.8E-05 | 0.7 |
| 3 | Bacteria B | 16 | 508 | 24 | 1344 | 0.008 | 1.2 |
| 4 | Spirochaetia | 9 | 508 | 18 | 1344 | 0.403 | 0.5 |
| 5 | Spirochaetales | 35 | 508 | 46 | 1344 | 3.5E-07 | 1.7 |
| 6 | Borreliaceae | 85 | 508 | 446 | 1344 | 1.2E-23 | -1.3 |
| 7 | Borrelia | 11 | 508 | 88 | 1344 | 2.9E-07 | -1.5 |
| 8 | Borrelia burgdorferi | 0 | 508 | 17 | 1344 | 0.0008 | -inf |

**Table S32.** Enrichment of phylostratum (ps) specific genes in *B. burgdorferi* biofilm downregulated genes. For a gene to be categorized as differentially expressed, the adjusted p-value ($p_{ajd}$) calculated by DeSeq2 pairwise comparison analysis had to be below 0.05. P values were adjusted for multiple comparisons using the Benjamini and Hochberg procedure ($p_{adj}$). q = the number of phylostratum genes in biofilm downregulated genes, s = the total number of phylostratum genes in biofilm downregulated genes, h = the number of phylostratum genes in *B. burgdorferi* genome, t = the total number of genes in *B. burgdorferi* genome which passed the phylostratigraphic procedure. The abbreviation –inf marks it was not possible to calculate the log-odds.

| ps | ps_name | q | s | h | t | $p_{adj}$ | log-odds |
|---|---|---|---|---|---|---|---|
| 1 | Cellular organisms | 248 | 441 | 557 | 1344 | 1.1E-13 | 0.9 |
| 2 | Bacteria A | 65 | 441 | 148 | 1344 | 0.007 | 0.5 |
| 3 | Bacteria B | 14 | 441 | 24 | 1344 | 0.027 | 1.1 |
| 4 | Spirochaetia | 8 | 441 | 18 | 1344 | 0.415 | 0.5 |
| 5 | Spirochaetales | 22 | 441 | 46 | 1344 | 0.051 | 0.7 |
| 6 | Borreliaceae | 71 | 441 | 446 | 1344 | 7.8E-21 | -1.3 |
| 7 | Borrelia | 12 | 441 | 88 | 1344 | 0.0001 | -1.2 |
| 8 | Borrelia burgdorferi | 1 | 441 | 17 | 1344 | 0.028 | -2.1 |

**Table S33.** Enrichment of phylostratum (ps) specific genes in *B. burgdorferi* bleb downregulated genes. For a gene to be categorized as differentially expressed, the adjusted p-value ($p_{ajd}$) and the absolute value of log2FC (|log2 FC|) calculated by DeSeq2 pairwise comparison analysis had to be below 0.05 and more than 1, respectively. P values were adjusted for multiple comparisons using the Benjamini and Hochberg procedure ($p_{adj}$). q = the number of phylostratum genes in bleb downregulated genes, s = the total number of phylostratum genes in bleb downregulated genes, h = the number of phylostratum genes in *B. burgdorferi* genome, t = the total number of genes in *B. burgdorferi* genome which passed the phylostratigraphic procedure. The abbreviation –inf marks it was not possible to calculate the log-odds.

| ps | ps_name | q | s | h | t | $p_{adj}$ | log-odds |
|---|---|---|---|---|---|---|---|
| 1 | Cellular organisms | 53 | 138 | 557 | 1344 | 0.559 | -0.1 |
| 2 | Bacteria A | 21 | 138 | 148 | 1344 | 0.272 | 0.4 |
| 3 | Bacteria B | 4 | 138 | 24 | 1344 | 0.559 | 0.6 |
| 4 | Spirochaetia | 3 | 138 | 18 | 1344 | 0.559 | 0.6 |
| 5 | Spirochaetales | 18 | 138 | 46 | 1344 | 2.04E-06 | 1.8 |
| 6 | Borreliaceae | 37 | 138 | 446 | 1344 | 0.272 | -0.3 |
| 7 | Borrelia | 2 | 138 | 88 | 1344 | 0.029 | -1.7 |
| 8 | Borrelia burgdorferi | 0 | 138 | 17 | 1344 | 0.501 | -inf |

**Table S34.** Enrichment of phylostratum (ps) specific genes in *B. burgdorferi* biofilm downregulated genes. For a gene to be categorized as differentially expressed, the adjusted p-value ($p_{ajd}$) and the absolute value of log2FC (|log2 FC|) calculated by DeSeq2 pairwise comparison analysis had to be below 0.05 and more than 1, respectively. P values were adjusted for multiple comparisons using the Benjamini and Hochberg procedure ($p_{adj}$). q = the number of phylostratum genes in biofilm downregulated genes, s = the total number of phylostratum genes in biofilm downregulated genes, h = the number of phylostratum genes in *B. burgdorferi* genome, t = the total number of genes in *B. burgdorferi* genome which passed the phylostratigraphic procedure. The abbreviation –inf marks it was not possible to calculate the log-odds.

| ps | ps_name | q | s | h | t | $p_{adj}$ | log-odds |
|---|---|---|---|---|---|---|---|
| 1 | Cellular organisms | 22 | 50 | 557 | 1344 | 0.931 | 0.1 |
| 2 | Bacteria A | 8 | 50 | 148 | 1344 | 0.601 | 0.5 |
| 3 | Bacteria B | 2 | 50 | 24 | 1344 | 0.601 | 0.9 |
| 4 | Spirochaetia | 2 | 50 | 18 | 1344 | 0.601 | 1.2 |
| 5 | Spirochaetales | 4 | 50 | 46 | 1344 | 0.601 | 1 |
| 6 | Borreliaceae | 7 | 50 | 446 | 1344 | 0.025 | -1.1 |
| 7 | Borrelia | 5 | 50 | 88 | 1344 | 0.601 | 0.5 |
| 8 | Borrelia burgdorferi | 0 | 50 | 17 | 1344 | 1 | -inf |

**Table S35.** Enrichment of phylostratum-specific COG annotated genes compared to phylostratum-specific genes in *B. burgdorferi* genome. P values were adjusted for multiple comparisons using the Benjamini and Hochberg procedure ($p_{adj}$). q = the number of phylostratum-specific genes in COG annotated genes, s = the total number of phylostratum-specific genes in COG annotated genes, h = the number of phylostratum-specific genes in *B. burgdorferi* genome, t = the total number of phylostratum-specific genes in *B. burgdorferi* genome. The abbreviation –inf marks it was not possible to calculate the log-odds.

| phylostratum | phylostratum_name | q | s | h | t | $p_{adj}$ | log-odds |
|---|---|---|---|---|---|---|---|
| 1 | Cellular organisms | 475 | 631 | 557 | 1344 | 7.3E-133 | 3.2 |
| 2 | Bacteria A | 103 | 631 | 148 | 1344 | 1.3E-08 | 1.1 |
| 3 | Bacteria B | 10 | 631 | 24 | 1344 | 0.755 | -0.2 |
| 4 | Spirochaetia | 7 | 631 | 18 | 1344 | 0.749 | -0.3 |
| 5 | Spirochaetales | 13 | 631 | 46 | 1344 | 0.018 | -0.8 |
| 6 | Borreliaceae | 22 | 631 | 446 | 1344 | 3.5E-121 | -3.7 |
| 7 | Borrelia | 1 | 631 | 88 | 1344 | 1.9E-23 | -4.5 |
| 8 | Borrelia burgdorferi | 0 | 631 | 17 | 1344 | 6.1E-05 | -inf |

**Table S36.** The list of RpoS-upregulated genes acquired from Caimano et al. (2019). Associated p-value is calculated by LRT analysis (see Methods).

| Locus tag | Name | COG annotation | COG description | Seq location | Phylostratum (ps) | $p_{adj}$ |
|---|---|---|---|---|---|---|
| BB_RS02825 | DUF3996 domain-containing protein | S | Function unknown | chromosome | 1 | 9.15E-26 |
| BB_RS02835 | chemotaxis protein CheW | N; T | Cell motility; Signal transduction mechanisms | chromosome | 1 | 8.13E-43 |
| BB_RS02840 | STAS domain-containing protein | P | Inorganic ion transport and metabolism | chromosome | 1 | 2.39E-08 |
| BB_RS02845 | chemotaxis protein CheA | N; T | Cell motility; Signal transduction mechanisms | chromosome | 1 | 6.56E-23 |
| BB_RS03445 | methyl-accepting chemotaxis protein | N; T | Cell motility; Signal transduction mechanisms | chromosome | 1 | 1.24E-26 |
| BB_RS03450 | methyl-accepting chemotaxis protein | N; T | Cell motility; Signal transduction mechanisms | chromosome | 1 | 1.07E-30 |
| BB_RS04280 | P12 family lipoprotein | S | Function unknown | chromosome | 6 | 9.73E-157 |
| BB_RS04335 | class I SAM-dependent DNA methyltransferase | L | Replication, recombination and repair | lp28-3 | 6 | 3.85E-124 |
| BB_RS04410 | helix-turn-helix domain-containing protein | L | Replication, recombination and repair | lp28-3 | 6 | 0.000273 |
| BB_RS04415 | P13 family porin | S | Function unknown | lp28-3 | 6 | 0.000967 |
| BB_RS04550 | hypothetical protein | S | Function unknown | lp28-2 | 6 | 1.19E-11 |
| BB_RS04555 | hypothetical protein | S | Function unknown | lp28-2 | 6 | 7.17E-16 |
| BB_RS04560 | hypothetical protein | S | Function unknown | lp28-2 | 6 | 2.54E-05 |
| BB_RS04565 | hypothetical protein | S | Function unknown | lp28-2 | 6 | 2.20E-08 |
| BB_RS04570 | hypothetical protein | S | Function unknown | lp28-2 | 6 | 0.000689 |

| | | | | | | |
|---|---|---|---|---|---|---|
| BB_RS04575 | hypothetical protein | S | Function unknown | lp28-2 | 6 | 5.14E-02 |
| BB_RS04580 | hypothetical protein | S | Function unknown | lp28-2 | 6 | 1.41E-08 |
| BB_RS04585 | hypothetical protein | S | Function unknown | lp28-2 | 6 | 0.000012 |
| BB_RS04590 | DUF1073 domain-containing protein | F | Nucleotide transport and metabolism | lp28-2 | 6 | 2.31E-03 |
| BB_RS04595 | PBSX family phage terminase large subunit | S | Function unknown | lp28-2 | 6 | 4.10E-14 |
| BB_RS04600 | hypothetical protein | S | Function unknown | lp28-2 | 6 | 8.68E-19 |
| BB_RS04605 | hypothetical protein | S | Function unknown | lp28-2 | 6 | 5.27E-48 |
| BB_RS04610 | hypothetical protein | S | Function unknown | lp28-2 | 6 | 7.44E-73 |
| BB_RS04615 | right-handed parallel beta-helix repeat-containing protein | S | Function unknown | lp28-2 | 6 | 1.99E-43 |
| BB_RS04620 | hypothetical protein | S | Function unknown | lp28-2 | 6 | 6.40E-01 |
| BB_RS04625 | hypothetical protein | S | Function unknown | lp28-2 | 6 | 6.81E-01 |
| BB_RS04630 | DUF261 family protein | S | Function unknown | lp28-2 | 6 | 3.93E-18 |
| BB_RS04635 | hypothetical protein | S | Function unknown | lp28-2 | 6 | 4.33E-12 |
| BB_RS04640 | hypothetical protein | S | Function unknown | lp28-2 | 6 | 3.66E-38 |
| BB_RS04645 | hypothetical protein | S | Function unknown | lp28-2 | 6 | 1.08E-07 |
| BB_RS04730 | hypothetical protein | S | Function unknown | lp38 | 6 | 0.091 |
| BB_RS04845 | lipoprotein | S | Function unknown | lp36 | 6 | 0.112 |
| BB_RS04915 | fibronectin-binding protein | S | Function unknown | lp36 | 6 | 0.485 |
| BB_RS05135 | hypothetical protein | S | Function unknown | lp54 | 6 | 3.63E-73 |
| BB_RS05140 | BBA07 family lipoprotein | S | Function unknown | lp54 | 6 | 8.15E-05 |
| BB_RS05215 | decorin-binding | S | Function unknown | lp54 | 6 | 3.08E-109 |

| | | | | | | |
|---|---|---|---|---|---|---|
| | protein DbpA | | | | | |
| BB_RS05220 | decorin-binding protein DbpB | S | Function unknown | lp54 | 6 | 3.02E-112 |
| BB_RS05240 | lipoprotein | S | Function unknown | lp54 | 6 | 5.58E-36 |
| BB_RS05245 | peptide ABC transporter substrate-binding protein | E | Amino acid transport and metabolism | lp54 | 6 | 3.36E-110 |
| BB_RS05250 | Mlp family lipoprotein | S | Function unknown | lp54 | 6 | 3.38E-106 |
| BB_RS05255 | hypothetical protein | S | Function unknown | lp54 | 6 | 1.73E-189 |
| BB_RS05390 | complement regulator-acquiring protein | S | Function unknown | lp54 | 6 | 1.94E-65 |
| BB_RS05420 | complement regulator-acquiring protein | S | Function unknown | lp54 | 6 | 3.25E-28 |
| BB_RS05565 | outer surface protein OspC | S | Function unknown | cp26 | 6 | 1.61E-169 |
| BB_RS05695 | ErpD protein | S | Function unknown | lp28-1 | 6 | 9.68E-40 |
| BB_RS05720 | hypothetical protein | S | Function unknown | lp28-1 | 6 | 0 |
| BB_RS05840 | variable surface antigen VlsE | S | Function unknown | lp28-1 | 6 | 3.13E-33 |
| BB_RS05975 | fibronectin-binding protein RevA | S | Function unknown | cp32-1 | 6 | 4.39E-14 |
| BB_RS06615 | fibronectin-binding protein RevA | S | Function unknown | cp32-6 | 7 | 5.13E-07 |
| BB_RS06620 | Mlp family lipoprotein | S | Function unknown | cp32-6 | 7 | 1.94E-02 |
| BB_RS06670 | ErpK protein | S | Function unknown | cp32-6 | 7 | 0.563 |
| BB_RS06875 | hypothetical protein | S | Function unknown | cp32-7 | 8 | 0.083 |

**Table S37.** The list of RpoS-downregulated genes acquired from Caimano et al. (2019). Associated p-value is calculated by LRT analysis (see Methods).

| Locus tag | Name | COG annotation | COG description | Seq location | Phylostratum (ps) | $P_{adj}$ |
|---|---|---|---|---|---|---|
| BB_RS00170 | P13 family porin | S | Function unknown | chromosome | 5 | 7.05E-24 |
| BB_RS00410 | cysteine desulfurase | E | Amino acid transport and metabolism | chromosome | 1 | 0.318 |
| BB_RS01190 | aquaporin family protein | P | Inorganic ion transport and metabolism | chromosome | 1 | 2.12E-41 |
| BB_RS01195 | glycerol kinase GlpK | F | Nucleotide transport and metabolism | chromosome | 1 | 1.35E-41 |
| BB_RS01200 | glycerol-3-phosphate dehydrogenase/oxidase | C | Energy production and conversion | chromosome | 1 | 1.47E-71 |
| BB_RS01795 | hypothetical protein | S | Function unknown | chromosome | 6 | 1.19E-16 |
| BB_RS03195 | hypothetical protein | S | Function unknown | chromosome | 7 | 4.81E-180 |
| BB_RS04690 | P12 family lipoprotein | S | Function unknown | lp38 | 6 | 0.999 |
| BB_RS04695 | outer surface protein OspD | S | Function unknown | lp38 | 6 | 0. 438 |
| BB_RS04785 | complement regulator-acquiring protein | S | Function unknown | lp38 | 7 | 0. 499 |
| BB_RS04835 | P12 family lipoprotein | S | Function unknown | lp36 | 6 | 0.989 |
| BB_RS04865 | SIMPL domain-containing protein | S | Function unknown | lp36 | 2 | 0. 699 |
| BB_RS04870 | complement regulator-acquiring protein | S | Function unknown | lp36 | 6 | 0. 638 |
| BB_RS04935 | hypothetical protein | S | Function unknown | lp36 | 1 | 0. 704 |
| BB_RS04950 | immunogenic protein P37 | S | Function unknown | lp36 | 6 | 0.6 |
| BB_RS05085 | complement regulator-acquiring protein | S | Function unknown | lp28-4 | 7 | 0. 515 |
| BB_RS05090 | complement regulator- | S | Function unknown | lp28-4 | 7 | 0. 541 |

| | | | | | | |
|---|---|---|---|---|---|---|
| | acquiring protein | | | | | |
| BB_RS05095 | complement regulator-acquiring protein | S | Function unknown | lp28-4 | 7 | 0. 195 |
| BB_RS05125 | outer membrane protein | S | Function unknown | lp54 | 6 | 4.47E-93 |
| BB_RS05180 | outer surface lipoprotein OspA | S | Function unknown | lp54 | 7 | 4.33E-09 |
| BB_RS05185 | outer surface lipoprotein OspB | S | Function unknown | lp54 | 7 | 3.06E-08 |
| BB_RS05210 | DUF261 domain-containing protein | S | Function unknown | lp54 | 6 | 3.34E-65 |
| BB_RS05360 | hypothetical protein | S | Function unknown | lp54 | 7 | 2.35E-03 |
| BB_RS05370 | hypothetical protein | S | Function unknown | lp54 | 6 | 0.757 |
| BB_RS05375 | Lp6.6 family lipoprotein | S | Function unknown | lp54 | 7 | 2.80E-21 |
| BB_RS05395 | complement regulator-acquiring protein | S | Function unknown | lp54 | 7 | 3.50E-103 |
| BB_RS05400 | complement regulator-acquiring protein | S | Function unknown | lp54 | 6 | 3.58E-202 |
| BB_RS05425 | porin Osm28 | S | Function unknown | lp54 | 7 | 9.18E-251 |
| BB_RS05435 | hypothetical protein | S | Function unknown | cp9 | 6 | 0. 052 |
| BB_RS05440 | DUF226 domain-containing protein | S | Function unknown | cp9 | 6 | 0. 094 |
| BB_RS05455 | exported protein A EppA | S | Function unknown | cp9 | 6 | 0. 812 |
| BB_RS05460 | hypothetical protein | S | Function unknown | cp9 | 7 | 0. 435 |
| BB_RS05465 | fibronectin-binding protein RevA | S | Function unknown | cp9 | 7 | 0. 526 |

| | | | | | | |
|---|---|---|---|---|---|---|
| BB_RS05470 | site-specific integrase | L | Replication, recombination and repair | cp9 | 6 | 0. 582 |
| BB_RS05475 | DUF244 domain-containing protein | S | Function unknown | cp9 | 6 | 5.65E-07 |
| BB_RS05670 | hypothetical protein | S | Function unknown | lp17 | 6 | 3.50E-160 |
| BB_RS05765 | DUF5425 domain-containing protein | S | Function unknown | lp28-1 | 7 | 1.89E-08 |
| BB_RS07750 | hypothetical protein | S | Function unknown | lp36 | 6 | 0.325 |

**Table S38.** Enrichment of phylostratum (ps) specific genes in *B. burgdorferi* genes upregulated by RpoS. The list of RpoS-upregulated genes was acquired from Caimano et al. (2019). P values were adjusted for multiple comparisons using the Benjamini and Hochberg procedure ($p_{adj}$). q = the number of phylostratum genes in RpoS-upregulated genes, s = the total number of RpoS-upregulated genes, h = the number of phylostratum genes in *B. burgdorferi* genome, t = the total number of genes in *B. burgdorferi* genome which passed the phylostratigraphic procedure. The abbreviation –inf marks it was not possible to calculate the log-odds.

| ps | ps_name | q | s | h | t | $p_{adj}$ | log-odds |
|---|---|---|---|---|---|---|---|
| 1 | Cellular organisms | 7 | 52 | 557 | 1344 | 7.1E-05 | -1.6 |
| 2 | Bacteria A | 1 | 52 | 148 | 1344 | 0.084 | -1.9 |
| 3 | Bacteria B | 0 | 52 | 24 | 1344 | 0.879 | -inf |
| 4 | Spirochaetia | 2 | 52 | 18 | 1344 | 0.421 | 1.2 |
| 5 | Spirochaetales | 0 | 52 | 46 | 1344 | 0.421 | -inf |
| 6 | Borreliaceae | 34 | 52 | 446 | 1344 | 2E-05 | 1.4 |
| 7 | Borrelia | 7 | 52 | 88 | 1344 | 0.194 | 0.8 |
| 8 | Borrelia burgdorferi | 1 | 52 | 17 | 1344 | 0.982 | 0.4 |

**Table S39.** Enrichment of phylostratum (ps) specific genes in *B. burgdorferi* genes upregulated by RpoS. The list of RpoS-upregulated genes was acquired from Caimano et al. (2019). P values were adjusted for multiple comparisons using the Benjamini and Hochberg ($p_{adj}$). q = the number of phylostratum genes in RpoS-upregulated genes, s = the total number of RpoS-upregulated genes, h = the number of phylostratum genes in *B. burgdorferi* genome, t = the total number of genes in *B. burgdorferi* genome which passed the phylostratigraphic procedure. The abbreviation –inf marks it was not possible to calculate the log-odds.

| ps | ps_name | q | s | h | t | $p_{adj}$ | log-odds |
|---|---|---|---|---|---|---|---|
| 1 | Cellular organisms | 5 | 38 | 557 | 1344 | 0.001 | -1.6 |
| 2 | Bacteria A | 1 | 38 | 148 | 1344 | 0.347 | -1.6 |
| 3 | Bacteria B | 0 | 38 | 24 | 1344 | 1 | -inf |
| 4 | Spirochaetia | 0 | 38 | 18 | 1344 | 1 | -inf |
| 5 | Spirochaetales | 1 | 38 | 46 | 1344 | 1 | -0.3 |
| 6 | Borreliaceae | 17 | 38 | 446 | 1344 | 0.356 | 0.5 |
| 7 | Borrelia | 14 | 38 | 88 | 1344 | 4.1E-07 | 2.3 |
| 8 | Borrelia burgdorferi | 0 | 38 | 17 | 1344 | 1 | -inf |

**Table S40.** Enrichment of cluster-specific genes in *B. burgdorferi* RpoS-upregulated genes. The list of RpoS-upregulated genes was acquired from Caimano et al. (2019). Gene clustering was based on normalized gene expression among morphotypes and performed by the DP_GP_cluster algorithm (McDowell et al., 2018), with the maximum Gibbs sampling iterations set to 500. P values were adjusted for multiple comparisons using the Benjamini and Hochberg procedure ($p_{adj}$). q = the number of cluster-specific genes upregulated by RpoS, s = the total number of clustered genes upregulated by RpoS, h = the number of cluster-specific genes in *B. burgdorferi* genome, t = the total number of clustered genes in *B. burgdorferi* genome. The abbreviation –inf marks it was not possible to calculate the log-odds.

| cluster | q | s | h | t | $p_{adj}$ | log-odds |
|---|---|---|---|---|---|---|
| 1 | 2 | 48 | 102 | 1276 | 0.9 | -0.7 |
| 2 | 2 | 48 | 45 | 1276 | 1 | 0.2 |
| 3 | 13 | 48 | 169 | 1276 | 0.077 | 0.9 |
| 4 | 0 | 48 | 124 | 1276 | 0.077 | -inf |
| 5 | 0 | 48 | 47 | 1276 | 0.779 | -inf |
| 6 | 2 | 48 | 86 | 1276 | 1 | -0.5 |
| 7 | 0 | 48 | 69 | 1276 | 0.483 | -inf |
| 8 | 0 | 48 | 62 | 1276 | 0.549 | -inf |
| 9 | 1 | 48 | 177 | 1276 | 0.077 | -2.1 |
| 10 | 0 | 48 | 9 | 1276 | 1 | -inf |
| 11 | 4 | 48 | 27 | 1276 | 0.144 | 1.6 |
| 12 | 0 | 48 | 14 | 1276 | 1 | -inf |
| 13 | 1 | 48 | 55 | 1276 | 1 | -0.8 |
| 14 | 0 | 48 | 21 | 1276 | 1 | -inf |
| 15 | 17 | 48 | 107 | 1276 | 3.1E-06 | 1.9 |
| 16 | 0 | 48 | 49 | 1276 | 0.779 | -inf |
| 17 | 1 | 48 | 20 | 1276 | 1 | 0.3 |
| 18 | 1 | 48 | 22 | 1276 | 1 | 0.2 |
| 19 | 1 | 48 | 6 | 1276 | 0.829 | 1.6 |
| 20 | 0 | 48 | 15 | 1276 | 1 | -inf |
| 21 | 3 | 48 | 42 | 1276 | 0.829 | 0.7 |
| 22 | 0 | 48 | 8 | 1276 | 1 | -inf |

**Table S41.** Enrichment of cluster-specific genes in *B. burgdorferi* RpoS-downregulated genes. The list of RpoS-downregulated genes was acquired from Caimano et al. (2019). Gene clustering was based on normalized gene expression among morphotypes and performed by the DP_GP_cluster algorithm (McDowell et al., 2018), with the maximum Gibbs sampling iterations set to 500. P values were adjusted for multiple comparisons using the Benjamini and Hochberg procedure ($p_{adj}$). q = the number of cluster-specific genes downregulated by RpoS, s = the total number of clustered genes downregulated by RpoS, h = the number of cluster-specific genes in *B. burgdorferi* genome, t = the total number of clustered genes in *B. burgdorferi* genome. The abbreviation –inf marks it was not possible to calculate the log-odds.

| cluster | q | s | h | t | $p_{adj}$ | log-odds |
|---|---|---|---|---|---|---|
| 1 | 1 | 38 | 102 | 1276 | 1 | -1.2 |
| 2 | 0 | 38 | 45 | 1276 | 1 | -inf |
| 3 | 4 | 38 | 169 | 1276 | 1 | -0.3 |
| 4 | 0 | 38 | 124 | 1276 | 0.851 | -inf |
| 5 | 0 | 38 | 47 | 1276 | 1 | -inf |
| 6 | 3 | 38 | 86 | 1276 | 1 | 0.2 |
| 7 | 0 | 38 | 69 | 1276 | 1 | -inf |
| 8 | 1 | 38 | 62 | 1276 | 1 | -0.7 |
| 9 | 4 | 38 | 177 | 1276 | 1 | -0.3 |
| 10 | 0 | 38 | 9 | 1276 | 1 | -inf |
| 11 | 3 | 38 | 27 | 1276 | 0.958 | 1.5 |
| 12 | 1 | 38 | 14 | 1276 | 1 | 0.9 |
| 13 | 1 | 38 | 55 | 1276 | 1 | -0.5 |
| 14 | 1 | 38 | 21 | 1276 | 1 | 0.5 |
| 15 | 1 | 38 | 107 | 1276 | 1 | -1.2 |
| 16 | 2 | 38 | 49 | 1276 | 1 | 0.3 |
| 17 | 0 | 38 | 20 | 1276 | 1 | -inf |
| 18 | 0 | 38 | 22 | 1276 | 1 | -inf |
| 19 | 0 | 38 | 6 | 1276 | 1 | -inf |
| 20 | 1 | 38 | 15 | 1276 | 1 | 0.9 |
| 21 | 0 | 38 | 42 | 1276 | 1 | -inf |
| 22 | 0 | 38 | 8 | 1276 | 1 | -inf |

**Table S42.** Genes differentially expressed in round when compared to spirochetes. For a gene to be categorized as differentially expressed, adjusted p-value (p) calculated DeSeq2 pairwise comparison analysis had to be below 0.05.

| Locus tag | Name | Regulation compared to SP | COG annotation | COG description | Seq location | Phylostratum (ps) |
|---|---|---|---|---|---|---|
| BB_RS00135 | hypothetical protein | downregulated | S | Function unknown | chromosome | 5 |
| BB_RS00245 | queuosine precursor transporter | downregulated | S | Function unknown | chromosome | 1 |
| BB_RS00585 | phosphatidate cytidylyltransferase | downregulated | S | Function unknown | chromosome | 1 |
| BB_RS00755 | lipoprotein | downregulated | S | Function unknown | chromosome | 5 |
| BB_RS00800 | hypothetical protein | downregulated | S | Function unknown | chromosome | 1 |
| BB_RS01070 | phosphate ABC transporter permease subunit PstC | downregulated | P | Inorganic ion transport and metabolism | chromosome | 1 |
| BB_RS01180 | hypothetical protein | downregulated | S | Function unknown | chromosome | 6 |
| BB_RS01475 | DNA-protecting protein DprA | upregulated | L; U | Replication, recombination and repair; Intracellular trafficking, secretion, and vesicular transport | chromosome | 1 |
| BB_RS01480 | hypothetical protein | upregulated | S | Function unknown | chromosome | 1 |
| BB_RS01600 | outer membrane protein assembly factor BamD | downregulated | S | Function unknown | chromosome | 3 |
| BB_RS01655 | hypothetical protein | downregulated | G | Carbohydrate transport and metabolism | chromosome | 4 |
| BB_RS01740 | hypothetical protein | downregulated | S | Function unknown | chromosome | 5 |
| BB_RS01910 | DNA-directed RNA polymerase subunit beta' | upregulated | K | Transcription | chromosome | 1 |
| BB_RS01915 | DNA-directed RNA polymerase subunit beta | upregulated | K | Transcription | chromosome | 1 |
| BB_RS01925 | 50S ribosomal protein L10 | upregulated | J | Translation, ribosomal structure and biogenesis | chromosome | 1 |

| | | | | | | |
|---|---|---|---|---|---|---|
| BB_RS01930 | 50S ribosomal protein L1 | upregulated | J | Translation, ribosomal structure and biogenesis | chromosome | 1 |
| BB_RS01945 | preprotein translocase subunit SecE | upregulated | U | Intracellular trafficking, secretion, and vesicular transport | chromosome | 2 |
| BB_RS02145 | hypothetical protein | downregulated | S | Function unknown | chromosome | 6 |
| BB_RS02180 | DNA topoisomerase (ATP-hydrolyzing) subunit B | upregulated | L | Replication, recombination and repair | chromosome | 1 |
| BB_RS02215 | protein jag | upregulated | S | Function unknown | chromosome | 2 |
| BB_RS02405 | elongation factor Tu | upregulated | J | Translation, ribosomal structure and biogenesis | chromosome | 1 |
| BB_RS02415 | 50S ribosomal protein L3 | upregulated | J | Translation, ribosomal structure and biogenesis | chromosome | 1 |
| BB_RS02420 | 50S ribosomal protein L4 | upregulated | J | Translation, ribosomal structure and biogenesis | chromosome | 1 |
| BB_RS02425 | 50S ribosomal protein L23 | upregulated | J | Translation, ribosomal structure and biogenesis | chromosome | 1 |
| BB_RS02430 | 50S ribosomal protein L2 | upregulated | J | Translation, ribosomal structure and biogenesis | chromosome | 1 |
| BB_RS02440 | 50S ribosomal protein L22 | upregulated | J | Translation, ribosomal structure and biogenesis | chromosome | 1 |
| BB_RS02445 | 30S ribosomal protein S3 | upregulated | J | Translation, ribosomal structure and biogenesis | chromosome | 1 |
| BB_RS02450 | 50S ribosomal protein L16 | upregulated | J | Translation, ribosomal structure and biogenesis | chromosome | 1 |
| BB_RS02460 | 30S ribosomal protein S17 | upregulated | J | Translation, ribosomal structure and biogenesis | chromosome | 1 |

| | | | | | | |
|---|---|---|---|---|---|---|
| BB_RS02465 | 50S ribosomal protein L14 | upregulated | J | Translation, ribosomal structure and biogenesis | chromosome | 1 |
| BB_RS02470 | 50S ribosomal protein L24 | upregulated | J | Translation, ribosomal structure and biogenesis | chromosome | 1 |
| BB_RS02475 | 50S ribosomal protein L5 | upregulated | J | Translation, ribosomal structure and biogenesis | chromosome | 1 |
| BB_RS02480 | type Z 30S ribosomal protein S14 | upregulated | J | Translation, ribosomal structure and biogenesis | chromosome | 1 |
| BB_RS02485 | 30S ribosomal protein S8 | upregulated | J | Translation, ribosomal structure and biogenesis | chromosome | 1 |
| BB_RS02490 | 50S ribosomal protein L6 | upregulated | J | Translation, ribosomal structure and biogenesis | chromosome | 1 |
| BB_RS02495 | 50S ribosomal protein L18 | upregulated | J | Translation, ribosomal structure and biogenesis | chromosome | 1 |
| BB_RS02510 | 50S ribosomal protein L15 | upregulated | J | Translation, ribosomal structure and biogenesis | chromosome | 1 |
| BB_RS02515 | preprotein translocase subunit SecY | upregulated | U | Intracellular trafficking, secretion, and vesicular transport | chromosome | 1 |
| BB_RS02525 | 30S ribosomal protein S13 | upregulated | J | Translation, ribosomal structure and biogenesis | chromosome | 1 |
| BB_RS02535 | DNA-directed RNA polymerase subunit alpha | upregulated | K | Transcription | chromosome | 1 |
| BB_RS02570 | hypothetical protein | downregulated | S | Function unknown | chromosome | 5 |
| BB_RS02700 | polymer-forming cytoskeletal protein | upregulated | M | Cell wall/membrane/ envelope biogenesis | chromosome | 2 |
| BB_RS02710 | elongation factor G | upregulated | J | Translation, ribosomal structure and biogenesis | chromosome | 1 |

| | | | | | | |
|---|---|---|---|---|---|---|
| BB_RS02775 | hypothetical protein | downregulated | Q | Secondary metabolites biosynthesis, transport, and catabolism | chromosome | 6 |
| BB_RS02835 | chemotaxis protein CheW | upregulated | N; T | Cell motility; Signal transduction mechanisms | chromosome | 1 |
| BB_RS02950 | methionine--tRNA ligase | downregulated | J | Translation, ribosomal structure and biogenesis | chromosome | 1 |
| BB_RS03195 | hypothetical protein | downregulated | S | Function unknown | chromosome | 7 |
| BB_RS03330 | ribose 5-phosphate isomerase A | downregulated | G | Carbohydrate transport and metabolism | chromosome | 1 |
| BB_RS03365 | lipoprotein | downregulated | S | Function unknown | chromosome | 2 |
| BB_RS03425 | HAD family hydrolase | downregulated | S | Function unknown | chromosome | 1 |
| BB_RS03430 | ABC transporter ATP-binding protein | upregulated | S | Function unknown | chromosome | 1 |
| BB_RS03465 | type 2 isopentenyl-diphosphate Delta-isomerase | upregulated | C | Energy production and conversion | chromosome | 1 |
| BB_RS03470 | hydroxymethylglutaryl-CoA reductase, degradative | upregulated | I | Lipid transport and metabolism | chromosome | 1 |
| BB_RS03745 | tRNA (5-methylaminomethyl-2-thiouridine)(34)-methyltransferase MnmD | downregulated | E | Amino acid transport and metabolism | chromosome | 2 |
| BB_RS03955 | 50S ribosomal protein L27 | upregulated | J | Translation, ribosomal structure and biogenesis | chromosome | 1 |
| BB_RS03960 | Obg family GTPase CgtA | upregulated | C | Energy production and conversion | chromosome | 1 |
| BB_RS04060 | transcription termination/antitermination protein NusA | upregulated | K | Transcription | chromosome | 1 |
| BB_RS04235 | excinuclease ABC subunit B | upregulated | L | Replication, recombination and repair | chromosome | 1 |

| | | | | | | |
|---|---|---|---|---|---|---|
| BB_RS04240 | excinuclease ABC subunit UvrA | upregulated | L | Replication, recombination and repair | chromosome | 1 |
| BB_RS04265 | arginine deiminase | upregulated | E | Amino acid transport and metabolism | chromosome | 1 |
| BB_RS04270 | ornithine carbamoyltransferase | upregulated | E | Amino acid transport and metabolism | chromosome | 1 |
| BB_RS04275 | YfcC family protein | upregulated | S | Function unknown | chromosome | 2 |
| BB_RS05145 | DUF2634 domain-containing protein | downregulated | S | Function unknown | lp54 | 6 |
| BB_RS05420 | complement regulator-acquiring protein | downregulated | S | Function unknown | lp54 | 6 |
| BB_RS06035 | ErpB | upregulated | S | Function unknown | cp32-1 | 7 |
| BB_RS06845 | ParA family protein | downregulated | D | Cell cycle control, cell division, chromosome partitioning | cp32-7 | 1 |
| BB_RS07755 | hypothetical protein | downregulated | S | Function unknown | lp54 | 7 |

# 12. Curriculum vitae

**First name:** Nina

**Family name:** Čorak

**Date of birth:** April 5th 1994

**Place of birth:** Karlovac, Croatia

**Present position:** Research assistent, Laboratory for Evolutionary Genetics, Department of Molecular Biology, Ruđer Bošković Institute, Zagreb

**Education:**

- Research assistent – Postgraduate Interdisciplinary Doctoral Study of Molecular Biosciences, joit study of the J.J. Strossmayer University of osijek, Ruđer Bošković Institute, Zagreb and University of Dubrovnik – in progress
- Master's degree in molecular biology, Department of Biology, Faculty of Science, University of Zagreb, 2018
- Bachelors's degree in molecular biology, Department of Biology, Faculty of Science, University of Zagreb, 2016

**Projects:**

- *Phylostratigraphy of gene gain and loss* (HRZZ-IP-2016-06-5925) - Croatian National Science Foundation, Project leader Tomislav Domazet-Lošo, dr. rer. nat., Associate Professor and Senior Research Associate

**Presentations at scientific conferences:**

1. Nina Čorak, Domagoj Kifer, Tomislav Domazet-Lošo. Towards phylostratigraphy of bacteria: A Bacillus case. 52nd Population Genetics Group Meeting. 3. siječanj 2019. – 6. siječanj 2019., Brookes University, Oxford, United Kingdom
2. Nina Čorak, Christina Daschkin, Viktoria Krey, Sara Koska, Momir Futo, Tin Široki, Sirli Anniko, Innokenty Woichansky, Luka Opašić, Domagoj Kifer, Kristian Vlahoviček, Horst-Günter Maxeiner, Mirjana Domazet-Lošo, Carsten Nicolaus, Tomislav Domazet-Lošo. Transcriptomes of Borrelia burgdorferi pleomorphic variants reveal distinct expression programs and evolutionary signatures. PhD Student Symposium, April 24- 25, 2021., Faculty of Science, University of Zagreb, Zagreb.
3. Postersko konferencijko izlaganje: Nina Čorak, Christina Daschkin, Viktoria Krey, Sara Koska, Momir Futo, Tin Široki, Sirli Anniko, Innokenty Woichansky, Luka Opašić, Domagoj Kifer, Kristian Vlahoviček, Horst-Günter Maxeiner, Mirjana Domazet-Lošo, Carsten Nicolaus, Tomislav Domazet-Lošo. Transcriptomes of Borrelia burgdorferi pleomorphic variants reveal distinct expression programs and evolutionary signatures. Predicting Evolution EMBO workshop, June 14-16, 2021. EMBL, Heidelberg, Germany
4. Nina Čorak, Christina Daschkin, Viktoria Krey, Sara Koska, Momir Futo, Tin Široki, Sirli Anniko, Innokenty Woichansky, Luka Opašić, Domagoj Kifer, Kristian Vlahoviček, Horst-Günter Maxeiner, Mirjana Domazet-Lošo, Carsten Nicolaus, Tomislav Domazet-Lošo. Transcriptomes of Borrelia burgdorferi pleomorphic variants reveal distinct expression programs and evolutionary signatures. Virtual Evolution, Society of Naturalists, the Society for the Study of Evolution, and the Society of Systematic Biologists, June 21-26, 2021, USA
5. Nina Čorak, Christina Daschkin, Viktoria Krey, Sara Koska, Momir Futo, Tin Široki, Sirli Anniko, Innokenty Woichansky, Luka Opašić, Domagoj Kifer, Kristian Vlahoviček, Horst-Günter Maxeiner, Mirjana Domazet-Lošo, Carsten Nicolaus, Tomislav Domazet-Lošo. Transcriptomes of Borrelia burgdorferi pleomorphic variants reveal distinct expression programs and evolutionary signatures. New Approaches and Concepts in Microbiology, July 7-9, 2021, EMBL, Heidelberg, Germany

**Publications:**


1. Futo, M., Opašić, L., Koska, S., Čorak, N., Široki, T., Ravikumar, V., Thorsell, A., Lenuzzi, M., Kifer, D., Domazet-Lošo, M., Vlahoviček, K., Mijakovic, I., & Domazet-Lošo, T. (2021). Embryo-Like Features in Developing Bacillus subtilis Biofilms. Molecular biology and evolution, 38(1), 31–47. https://doi.org/10.1093/molbev/msaa217
2. Široki, T., Koska, S., Čorak, N., Futo, M., Domazet-Lošo, T. & Domazet-Lošo, M. (2019) Correspondence Analysis Applied to Large Scale Evo-Devo Data. U: MIPRO 2019 Proceedings. https://doi.org/10.23919/MIPRO.2019.8756945