Přírodovědecká fakulta
Faculty of Science

Jihočeská univerzita v Českých Budějovicích
University of South Bohemia in České Budějovice

Understanding the role of *bb0454* during infection by *Borrelia afzelii*

Laboratory of Molecular Ecology of Vectors and Pathogens

INSTITUTE OF PARASITOLOGY

Biology Center, ASCR

Lisa Hain

Supervisor: Ryan O.M. Rego, PhD

Annotation: The aim of this thesis was to further our understanding of the importance of *bb0454*, a gene located at the linear chromosome of the *Borrelia* genome, which encodes a glycosyltransferase catalyzing the synthesis of one of the major glycolipids found in the outer membrane of this pathogenic bacteria.

Affirmation: I hereby declare that I have worked on the submitted bachelor thesis independently. All additional sources are listed in the bibliography section. I hereby declare that, in accordance with Article 47b of Act No. 111/1998 in the valid wording, I agree with the publication of my bachelor thesis, in full form to be kept in the Faculty of Science archive, in electronic form in publicly accessible part of the STAG database operated by the University of South Bohemia in České Budějovice accessible through its web pages. Further, I agree to the electronic publication of the comments of my supervisor and thesis opponents and the record of the proceedings and results of the thesis defense in accordance with aforementioned Act No. 111/1998. I also agree to the comparison of the text of my thesis with the Theses.cz thesis database operated by the National Registry of University Theses and a plagiarism detection system.

Ceske Budejovice, Date Lisa Hain

Table of Content

1 Literature Research

1.1 *Borrelia*

Borrelia is a bacterium belonging to the phylum of spirochetes. All spirochetes share three structural characteristics, including helically shaped cells and three forms of movement. Furthermore, the cytoplasm, the inner cell membrane, and the peptidoglycan of spirochetes form the so-called protoplasmic cylinder complex, which is surrounded by an outer cell membrane. Additionally, all spirochetes have a flagellum, which is located in the periplasmic space between the outer cell membrane and the protoplasmic cylinder complex[1].

Genetic analysis of different *Borrelia* species allows the division of *Borrelia* into different groups. One of these groups contains more than 20 species, including the causative agents of Lyme disease and it is referred to as *B. burgdorferi* sensu lato (s.l.). Another group, outside of this is formed by spirochetes responsible for relapsing fever[2].

With *B. afzelii*, *B. garinii*, *B. burgdorferi* sensu stricto (s.s.), *B. spielmanii* and *B. bavariensis*, there are at least five genospecies of *B. burgdorferi* s.l. that are pathogenic to humans. *B. burgdorferi* s.s. is the only *Borrelia* species infectious to humans in Northern America, whereas all five genospecies are found in Europe[3,4].

Hematophagous arthropods are responsible for the transmission of all *Borrelia* species to vertebrates[1]. *B. burgdorferi* s.l. is transmitted by vectors belonging to the *Ixodes* genus. Each one of the different feeding stages of these ticks (larvae, nymphs and adult) uses a different individual host. The bacteria is taken up by the vector during a blood meal at one of these stages and transmitted to another host during the next blood meal in the following feeding stage, making the life cycle of *B.* burgdorferi s.l. a dynamic interplay between bacteria, reservoir hosts and vectors[5].

In Europe the primary vector of *Borrelia burgdorferi* s.l., which causes Lyme disease in humans, is *Ixodes ricinus*, causing over 60,000 cases of this inflammatory disorder per year, while in the US it is vectored by the tick *Ixodes scapularis* with over 300,000 cases a year. Lyme disease is thus the most common tick-borne disease in the northern hemisphere[3].

1.2 Genetic Composition of *Borrelia burgdorferi* s.l.

B. burgdorferi s.l.'s genome is unique among bacteria in being segmented and actively evolving. In total the genome is about 1.5 megabase big, consisting of a small linear

chromosome (0.9 megabase), and a varying amount of linear and circular plasmids ranging in size from 5 to 56 kilobase[6].

The linear chromosome contains housekeeping genes, and is constant in organization, and content throughout the genus. The variability of content and number of plasmids is much greater and they are not equally represented in all strains[2]. Studies have found that the genetic information present on some of these plasmids play a key role in infection of mice and ticks in laboratories[7,8].

Passaging of isolates within the *B. burgdorferi s*.l. family leads to loss of plasmids in vitro, consequently affecting the life cycle of the bacteria[9].

1.3 Membrane Architecture of *Borrelia burgdorferi* s.l.

Borrelia has two membranes, which are organized in a similar manner as those in Gram-negative bacteria. However, studies have shown that *Borrelia* lack genes for the synthesis of the Gram-negative cell wall typical lipopolysaccharide (LPS). Also *Trepnoema*, which belongs like *Borrelia* to the spirochetes class, does not have these genes, whereas *Leptospira* another spirochete does contain a LPS with a unique chemical composition[10].

The peptidoglycan layer is not attached to the outer membrane, therefore a periplasmic space is formed containing the flagellum. The flagellum is located between the outer cell wall and the protoplasmic cylinder complex causing the *Borrelia burgdorferi* s.l.'s spherical shape and its motility. The position of this flagellum is unique for spirochete species[10].

The differences between the cell membrane of Gram-negative bacteria and spirochetes can be seen in Figure 1.

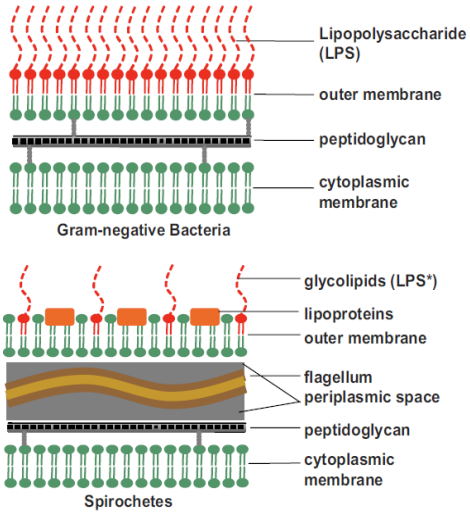

Figure 1 - Comparison of the structure of the spirochete cell wall and the one of Gram-negative bacteria, LPS* is only present in Leptospira (adapted from Schröder, 2008)

1.4 Glycosylation in Bacteria

During glycosylation glycans (polysaccharides) are attached to proteins, lipids, or other organic molecules. Glycosylation of proteins is present in all domains of life and is found in eukaryotes, prokaryotes, and viruses[11]. It is the most common post-translational modification of synthesized proteins controlling various cellular processes, including signal transduction, protein folding, targeting, stability, cell-cell and virus-cell interactions, and host immune responses[12,13].

In bacteria, these modifications are responsible for the glycosylation of enzymes, antigens, cell-envelope components, and most important surface layer (S-layer) proteins. Compared to eukaryotes, glycosylation in bacteria does not only lead to the formation of N-glycosylated linkages at asparagine, or O-glycosylated linkage at threonine, serine, hydroxylproline or hydroxyllysine, additionally, also tyrosine was identified as a possible glycosylation site[12].

N-glycosylation has only been proven in representatives of Gram-negative bacteria containing two membranes. One example is the glycosylation system of *Campylobacter jejuni,* a gastrointestinal bacterium. It was the first one described showing N-glycosylation of a great number of periplasmic and membrane-bound proteins with a conserved heptasaccharide[14]. Studies showed that the disruption of this N-glycosylation pathway reduces the bacteria's colonization of mice and chicken in-vivo, as well as leading to a reduced adhesion of the bacteria on intestinal epithelial cells, proving the importance of such glycosylation pathways for the life cycle of this kind of bacteria[13,15].

The position of most glycoproteins found in pathogenic bacteria suggests a role for the glycoproteins in the interaction with the host. They are either located on the surface of the organism as in pili or flagella or they secrete into the environment, resulting in the direct exposure to the host cells and its defence system[13].

The flagellin and outer surface proteins from *Borrelia burgdorferi* s.l. were suspected to be glycosylated as well, however, studies showed that *B. burgdorferi* s.l.'s outer membrane proteins OspA, and OspB, as well as the structural flagellar proteins FlaB, and FlaA are not N-glycosylated. Furthermore, bioinformatics analysis of *B. burgdorferi* s.l representatives provided evidence that *B. burgdorferi* s.l. lack the enzyme required for the transfer of N-linked glycans from a lipid carrier to an asparagine residue of a protein[16].

1.5 Glycosyltransferases

Glycosyltransferases form a very large enzyme family and are responsible for the synthesis of glycans by assembling monosaccharides into linear and branched glycan chains. Their common property is the catalysis of group-transfer reactions, in which a monosaccharide moiety of a simple nucleotide sugar donor substrate, like UDP-Gal or GDP-Fuc, is transferred to an acceptor substrate. Most of these transferases are acting sequentially, meaning that one enzyme yields a preferred acceptor substrate for the subsequent action of another, resulting in a linear and/or branched polymer composed of monosaccharides linked to one another. Aside from elongating glycan chains, glycosyltransferases are also synthesising glycoproteins and glycolipids by attaching saccharides to a polypeptide side chain or a sphingolipid base[17].

Most glycosyltransferases exhibit a generally strict donor, acceptor, and linkage specificity, which defines and limits the number and type of glycan structures observed in an organism[18]s.

However, there are a few exceptions to this general rule. Some glycosyltransferases, such as the human B blood group α1-3 galactosyltransferase, are highly specific in respect to donor and acceptor. This enzyme only catalyses the attachment of galactose in a α-linkage to the C-3 hydroxyl group of a galactose residue on the acceptor substrate if the substrate contains fucose in a α1-2 linkage. In other cases, including the human fucosyltransferases III-VII, more than one enzyme can use the same acceptor to make the same linkage. These transferases are all attaching fucose in a α1-3 linkage to *N*-acetylglucosamine moieties on glycans. Furthermore, some glycosyltransferases have two separate active sites. Transferases that synthesize the backbones of glycosaminoglycans have one side for the attachment of *N*-acetylglucsoamine to glucuronic acid and another that attaches glucuronic acid to *N*-acetylglucosamine[17].

The importance of glycosyltransferases in bacteria is seen, when looking at the glycosyltransferase encoded by *waaG* of *Escherichia coli*. *E. coli,* like all other gram-negative bacteria, has an outer membrane with the main structural feature LPS, functioning as a barrier to antibiotics and other host defence factors[19]. The LPS of *E. coli* is divided into three parts. The first part is formed by Lipid A, a hydrophobic membrane anchor, which is responsible for endotoxic properties. Then, there is the core region with phosphorylated oligosaccharides as well as the O antigen which is a structurally variable polysaccharide made up of repeating oligomeric units. The phosphate-containing substituents in the core region allow cross-linkage of neighbouring LPS molecules by divalent cations and are therefore essential for the membrane stability[19,20]. Mutations of the glycosyltransferase waaG results in a decrease of phosphorylation in the inner core of the LPS and therefore reduce the stability of LPS, leading to an increase of susceptibility to seven classes of antibiotics[21].

1.6 Glycolipids in *Borrelia burgdorferi* s.l.

Approximately 36% of the total lipid mass of *B. burgdorferi* s.l. is comprised by two major glycolipids, an acylated cholesteryl galactoside, cholesteryl 6-*O*-acyl-β-D-galactopyranoside (ACG, BbGL-I) and a monogalactosyl diacylglycerol, 1,2-di-*O*-acyl-3-O-α-D-galactopyranosyl-*sn*-glycerol (MGalDAG, BbGL-II). Both show several properties suggesting they may function as some kind of LPS. These glycolipids are the main components of the bacterial membrane, surface exposed and have a three-domain structure[22].

Additionally, they elicit antibodies in mice and rabbits. The antibodies elicited by BbGL-I are specific to this one protein, whereas the ones elicited by BbGL-II recognize both glycolipids[22].

Especially in the late stages of Lyme disease a strong immunoglobulin (IgG) reactivity with both glycolipids was recorded, while in early stages the IgG antibody responses to MGalDAG were much more frequent compared with those to ACG. This reactivity declined gradually in antibiotic-treated patients after the therapy[23,24].

Antibody responses to spirochetal lipoproteins are directed against protein epitopes and not lipid moieties together. Furthermore, antibody responses were only seen with α-MGalDAG and not with the commercially β-version of the protein. This suggests that the antibodies may bind the alpha configuration of the galactose moiety of α-MgalDAG[23,24].

In patients suffering from late stage infection the antibody responses to *B. burgdorferi* lipoproteins last for years after spirochetal killing and no re-infection occurs in these patients. The antibody reactivity with *B. burgdorferi* s.l. glycolipids appear to be long lasting as well, therefore suggesting immunization with *B. burgdorferi* s.l. glycolipids as one possibility for vaccination development[23].

1.7 Monogalactosyl-1,2-Diacylglycerol Synthase (bbMGS)

Monogalactosyl-1,2-diacylglycerol synthase (bbMGS) from *B. burgdorferi* s.l. is encoded by the gene *bb0454*, which is located on the linear chromosome of *B. burgdorferi* s.l.. According to the protein domain database at NCBI, bbMGS belongs to pfam00534, the glycosyltransferase group 1 that transfers nucleoside diphosphate (NDP)-linked sugars, like glucose and galactose to a variety of acceptor substrates.

Homologs of this gene were found in various Gram-positive bacteria, thermophiles, *Archaea*, photosynthetic bacteria and other spirochetes, namely *Treponema denticola* and *Leptospira interrogans*[25].

Expression of bbMGS in *E. coli* showed that this protein catalyses the glycosylation of 1,2-diacylglycerol with specificity for the donor substrate UDP-Galactose yielding the major membrane lipid of *B.* burgdorferi s.l. α-MGalDAG[25].

Therefore, bbMGS plays an important role in the glycolipid biosynthetic pathway of *B. burgdorferi* s.l., which is shown in Figure 2. UDP-Galactose, which is essential for this reaction, is believed to be derived from UDP-Glucose, however, this pathway is still unclear and has not been proven yet. The mechanism of the formation of the acceptor involved in the glycosylation reaction catalysed by bbMGS 1,2-diacylglycerol from CDP-1,2-diacylglycerol is also still unknown[26].

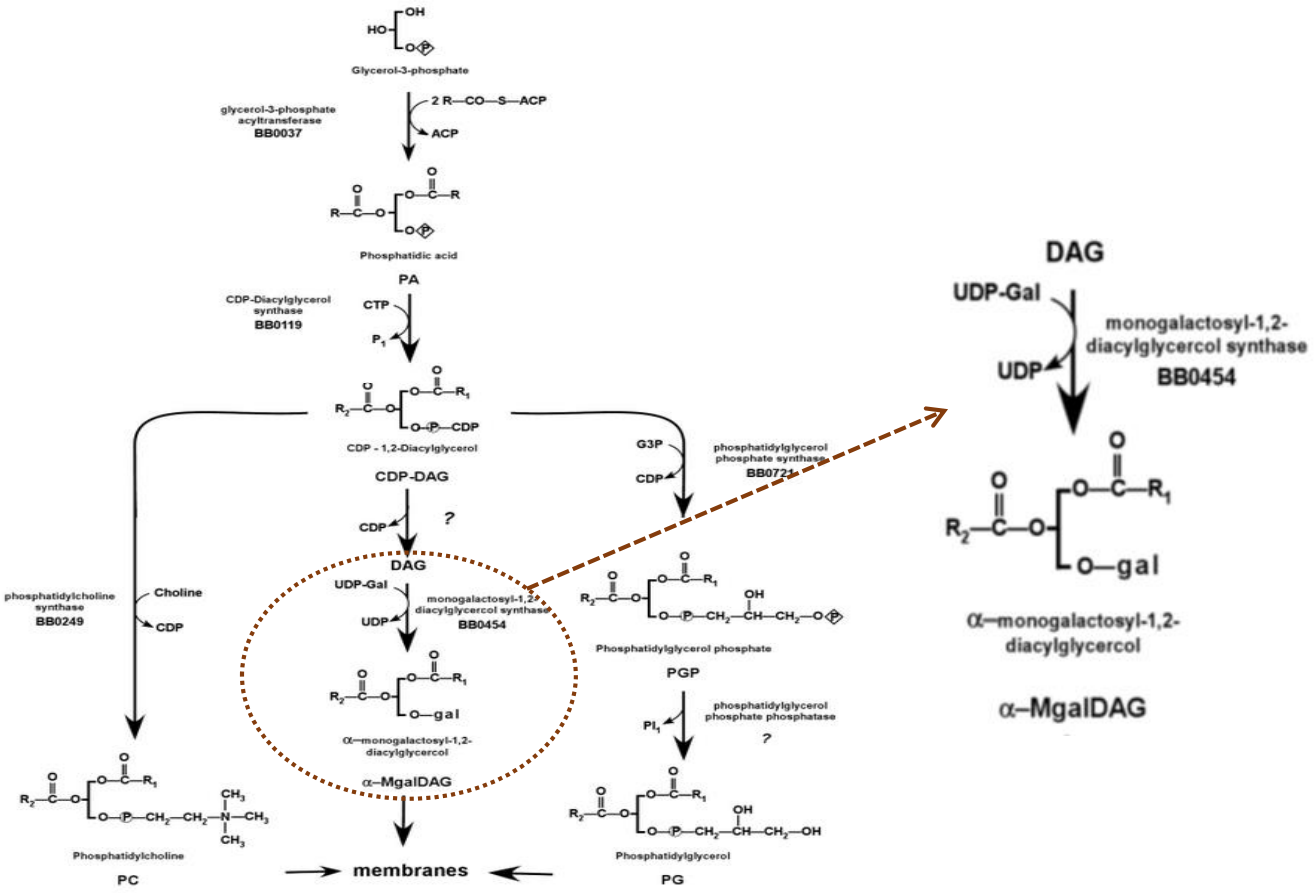

Figure 2 - Putative lipid and glycolipid biosynthetic pathway in *B. burgdorferi* (adapted from Samuels, 2010)

2 Aims of the work

To understand the role of *bb0454* during the infection by *B. afzelii* the following aims were attempted to be reached:

1. Sequencing of *bb0454* from a variety of *Borrelia* strains covering the different Lyme disease genospecies and a phylogenetic analysis as well as a structural analysis of selected examples.
2. Cloning of the gene of interest along with overlapping portions on either side.
3. Creating a knock-out construct using this cloned DNA for the infectious strains B31 (*B. burgdorferi* s.s.) and CB43 (*B. afzelii*).
4. Constructing a shuttle vector carrying the full gene, to be introduced into the parental strain, to observe a phenotype if there is overproduction of bb0454.

3 Material and Methods

During the experimental work of this thesis various techniques of molecular biology were employed to grow different Borrelia strains, isolate their genomic DNA, amplify the gene *bb0454*, followed by cloning and transformation in *E. coli* and sequencing. Furthermore, an attempt was made to construct shuttle vectors carrying the gene from representatives of *B. afzelii* (CB43) and *B. burgdorferi* (B31) as well as knock-out constructs of *bb0454* in these two strains.

3.1 Bacterial Strains

To cover different *Borrelia* genospecies, the following strains, were used: *B. afzelii* (CB43, PKo, ACA1, HL5, RU1, RU2), *B. americana*, *B. carolinensis*, *B. garinii* (PBr, ATCC13), *B. bavariensis* (PBi), *B. burgdorferi s.s.* (B31, ZS75, 297) and *B. finlandensis* (SV1).

E. coli DH5α was used in the cloning and transformation step.

3.2 Bacterial Growth Conditions

Borrelia from glycerol-stocks were grown in BSKII-media at 34°C in an incubator till the mid-log phase (10^7 *Borrelia*/mL) had been reached. The cultures were subsequently used for genomic DNA isolation.

E. coli was grown in LB media at 37°C in an orbital shaker.

3.3 Genomic DNA isolation

Genomic DNA (gDNA) samples had either already been prepared or were isolated from *Borrelia* cultures using Promega's kit for the isolation of Genomic DNA from Gram Positive and Gram Negative Bacteria and an adaptation of the manufacturer's manual.

First of all, the culture was transferred to a centrifugation tube and centrifuged at 8000 rpm for 10 minutes at 20°C. The supernatant was poured off and residues were removed with a pipette. Then, the pellet was resuspended in 600 µL of Nucleic Lysis Solution and transferred to a 1.5 mL microtube. This step was followed by 1 minute of vortexing and incubation at 80°C for 5 minutes.

After incubation the lysate was cooled to room temperature, 3 µL of RNAse was added and the tube was inverted 2 times for mixing, followed by15 minutes of incubation at 37°C. As soon as the solution had cooled down to room temperature, 200 µL of Protein Precipitation solution was added to the tube, which was then incubated on ice for 5 minutes.

The sample was centrifuged at 13000 rpm for 10 minutes and the supernatant was pipetted to a fresh 1.5mL tube already containing 600 µL isopropanol. After gentle mixing, the sample was centrifuged again (13000rpm, 10°C, 10 minutes). The supernatant was carefully poured off and the tube was turned on a paper towel four times. Then, 600 µL of room temperature 70% Ethanol was added and the sample was centrifuged once more (13000rpm, 10°C, 10 minutes).

After centrifugation, the supernatant was carefully removed with a pipette and small residual ethanol drops were removed with a clean absorbent paper. Furthermore, the DNA sample was air dried for 5-7 minutes.

In the end, 100 µL of DNA Rehydration Solution were added and the sample was incubated in the fridge overnight.

3.4 Polymerase Chain Reaction (PCR)

PCR was used for the amplification of the gene *bb0454* in various strains of *Borrelia* as well as the amplification of bb0454 plus extensions on both sides in representatives of *B. afzelii* (CB43) and *B. burgdorferi s.s* (B31). The gene plus extensions were used for the construction of a knock-out construct of this gene (chapter 3.9).

Furthermore, PCRs were employed in the amplification of DNA parts used in the construction of a shuttle vector carrying the gene for overexpression (chapter 3.10).

Table 1 lists the primers employed during these reactions and the ones used for sequencing.

Table 1 - List of Primers

Task	Primer	Sequence
Amplification of bb0454	BBGL-1F	5'-ATG AAA GTC GCA ATA TTT ACA G-3'
	BBGL-1R	5'-TTA ATG ATT TTT TCT TGC AAT AAT TTC TGA GTA-3'
Amplification of bb0454 + extensions	BBGL-1FL	5'-AAT AAA TTC TC AAA TAT AAA TCA AAC-3'
	BBGL-1RL	5'-TAA TTC AAT TAA TTT ATG TAA AGC CCG G-3'
Construction of knock-out construct	flgKan BsrGI F	5'-GAT GTA CAT ACC CGA GCT TCA AGG AAG-3'
	flgKan BsrGI R	5'-CCA ACA TTG TGA CCG TCT CAC ATG TAC-3'
	flgAclI F	5'-TCT AAC GTT TAA TAC CCG AGC TTC AAG GAA GAT-3'
	KanBmtI R	5'-GAA GCT AGC GCC GTC CGT CAA GTC AGC G-3'
Construction shuttle vector	Flg SbfI F	5'-CCT GCA GGT AAT ACC CGA GCT TCA AGG-3'
	Flg XhoI R	5'-CTC GAG ATG GAA ACC TCC CTC ATT TAA AAT TGC-3'
	454-XhoI F	5'-CTC GAG ATG AAA GTC GCA ATA TTT A-3'
	B31/454-SbfI R	5'-CCT GCA GGT TAA TGA TTT TTC TTG C-3'
	CB43/454-SbfI R	5'-CCT GCA GGT TAA TGA TTT TTT CT-3'
Sequencing	T7	5'-TAA TAC GAC TCA CTA TAG GG-3'
	M13R	5'-CAG GAA ACA GCT ATG AC-3'

For the amplification of the gene and the gene plus extensions various PCR programs and polymerases were tested, however only One Taq Hot Start DNA polymerase with the corresponding buffer from New England BioLabs® as well as the combination of this

polymerase with the buffers E, H, or G of the Fail Safe PCR Kit from Epicenter lead to successful amplification, transformation and cloning. In the following the PCR programs used to amplify both *bb0454* and *bb0454* plus extensions are listed.

Table 2 – PCR master mix and settings for the amplification of *bb0454* and *bb0454* plus extensions 1

Master Mix 1	**20 µL reaction**
10xOne Taq Standard Reaction Buffer	4 µL
dNTP's	0.4 µL
One Taq Hot Start Polymerase	0.1 µL
Primer F	1 µL
Primer R	1 µL
MiliQ H_2O	11.5 µL
DNA	2 µL

PCR settings 1	
Initial denaturation:	94°C for 30 sec.
30 cycles:	94°C for 30 sec.
	54°C for 45 sec.
	68°C for 1 min.
Final extension:	68°C for 5 min.
Hold:	16°C

Table 3 - PCR master mix and settings for the amplification of *bb0454* and *bb0454* plus extensions 2

Master Mix 2 – Fail Safe	**20 µL reaction**
Fail Safe Buffer	10 µL
One Taq Hot Start Polymerase	0.2 µL
Primer F	1 µL
Primer R	1 µL
MiliQ H_2O	5.8 µL
DNA	2 µL

PCR settings 2 – Fail Safe	
Initial denaturation:	95°C for 2 min.
30 cycles:	95°C for 45 sec.
	50°C for 45 sec.
	68°C for 90 sec.
Final extension:	68°C for 10 min.
Hold:	12°C

The success in amplifying the desired PCR product was visualized using gel electrophoresis. Part of the PCR samples were visualized using 10xSYBR and run together with 1kb+ DNA ladder from Invitrogen in a 5% agarose gel. Under UV light the size of the products was determined. The expected size for the amplification of *bb0454* was 1149 base pairs and the one for the gene plus extensions was 2149.

3.5 Cloning and Transformation in *E. coli*

After successful PCR, the products were cloned into a vector and then transformed into *E. coli*. Two different kinds of vectors were used in this step:

- Invitrogen's pCR®2.1-TOPO®
- Promega's pGEM®-T easy

The cloning was done by setting up one of the following reaction mixtures:

Table 4 - Reaction Mixture Cloning in pCR®2.1-TOPO® vector

pCR®2.1-TOPO® vector	
Reagent	**Amount [µL]**
DNA insert	3
Salt Solution	0.5
TOPO vector	0.5
MiliQ H_2O	2

Table 5 - Reaction Mixture Cloning in pGEM®-T easy vector

pGEM®-T easy vector	
Reagent	**Amount [µL]**
2x Rapid Ligation Buffer	2.5
pGEM®-T easy vector	0.5
DNA insert	1.5
T4 DNA Ligase	0.5

The TOPO reaction mixture was incubated for 5-10 minutes at room temperature before proceeding with the transformation in *E. coli.*. The pGEM-T easy reaction mixture had to be incubated overnight in the fridge before continuing.

For the transformation, NEB 5-alpha competent *E. coli* cells were thawed on ice for 5 minutes. To 25 µL of cells, 2 µL of Ligation reaction was added and the sample was mixed by flicking. Then, the mixture was incubated on ice for at least 30 minutes, followed by a heat shock for 30 seconds at 42°C. The cells were kept on ice for another 2 minutes. Afterwards, 250 µL of S.O.C.-Outgrowth media were added to the reactions.

After shaking at 37°C for 60-90 minutes, 75 µL and 150 µL cells were plated on prewarmed agar plates containing the antibiotic Carbenicillin and a layer of X-Gal.

The plates were kept at room temperature for 5 minutes and incubated overnight at 37°C.

3.6 Plasmid DNA isolation

To isolate plasmid DNA, white *E. coli* colonies were picked from incubated plates and shaken overnight in 6 mL of LB media at 37°C.

Plasmid DNA was isolated using QIAprep Spin Miniprep kit and NucleoSpin® Plasmid DNA Purification kit following the protocols provided by the manufacturers.

Before sequencing, the DNA isolates were tested for the insert using restriction digestion with the enzyme EcoR I. If an insert was present, 5 µL of the plasmid DNA sample containing an insert was mixed with 5 µL of either T7 or MT13R (concentration 5 pmol/µL) and then sent to GATC Biotech Lightrun sequencing service to sequence from both directions.

3.7 Restriction Digestion

Restriction Digestion uses enzymes recognizing and cutting certain nucleotide sequences. To check for inserts present in the isolated plasmid DNA samples NEB's EcoRI enzyme was used. Therefore, the following reaction was set up and incubated at 37°C for at least 30 minutes. Using gel electrophoresis the samples where checked for the presence of an insert.

Table 6 - Reaction Mixture Restriction Digestion EcoRI

Restriction Digestion	**Amount [µL]**
EcoRI Buffer	2 µL
EcoRI Enzyme	0.5 µL
MiliQ H_2O	14.5 µL
DNA	3 µL

3.8 Bioinformatic Analysis of *bb0454*

For each one of the analysed genospecies three different plasmids containing *bb0454* were sent for sequencing. The multiple alignment tool of the program Geneious V8.1.7 was used to obtain the nucleotide sequence of *bb0454* for each one of the genospecies analysed. Furthermore, the nucleotide sequences were translated in Geneious to amino acids[27].

Additional *bb0454* amino acid sequences of other *Borrelia* species, including some causing relapsing fever, were downloaded from the NCBI databank, as well as the sequence of an 1,2-diacylglycerol 3-glucostransferase from *Listeria booriae* to root the phylogenetic tree. All these sequences were used to create a phylogenetic tree in Geneious showing the relationship between the different species. The accession numbers of the protein sequences used in the analysis are listed in Table 7.

Furthermore, 3D models of *bb0454's* structure of representatives of *B. afzelii, B. burgdorferi, B. bavariensis, B. garinii* and *B. miyamotoi* were computed using the Phyre2 server, followed by analysis using Chimera structure modelling program[28,29].

Table 7 - List of Amino Acid accession numbers

Species	Amino Acid accession number
B. afzelii HLJ01	AFU74749
B. afzelii Tom3107	AIK18752
B. bissettii DN127	AEL18628
B. burgdorferi 156A	EEC21530
B. burgdorferi N40	ADQ29635
B. garinii BgVir	AEW68792
B. garinii NMJW1	AFT83778
B. garinii SZ	AHZ74095
B. miyamotoi FR64b	AHH04954
B. spielmanii A14S	EEF84283
B. valaisiana Tom4006	AIJ29811
B. hermsii MTW	AHH13956
B. parkeri HR1	AHE62775
B. persica	WP_024653575
B. turicatae	WP_011772401
B. recurrentis A1	ACH94692
B. duttonii Ly	ACH93398
Listeria booriae	WP_052167636

3.9 Construction of a knock-out construct

To observe the importance of *bb0454* during the infection by Borrelia a knock-out of the gene was constructed in representatives of *B. afzelii* (CB43) and *B. burgdorferi* (B31).

To create a knock-out of the gene, the nucleotide sequence of *bb0454* plus extensions on both sides were analysed using NEBcutter V2.0 to identify sites which would be cut by a restriction digestion enzyme and lead to the disruption of the gene[30].

After analysing CB43's nucleotide sequence, NEB's BsrGI enzyme was chosen to cut the gene open in the middle and insert a kanamycin antibiotic cassette (*Borrelia* promoter and kanamycin resistance gene). For B31's sequence AclI and BmtI were chosen to cut out a portion of the gene and replace it with the kanamycin cassette.

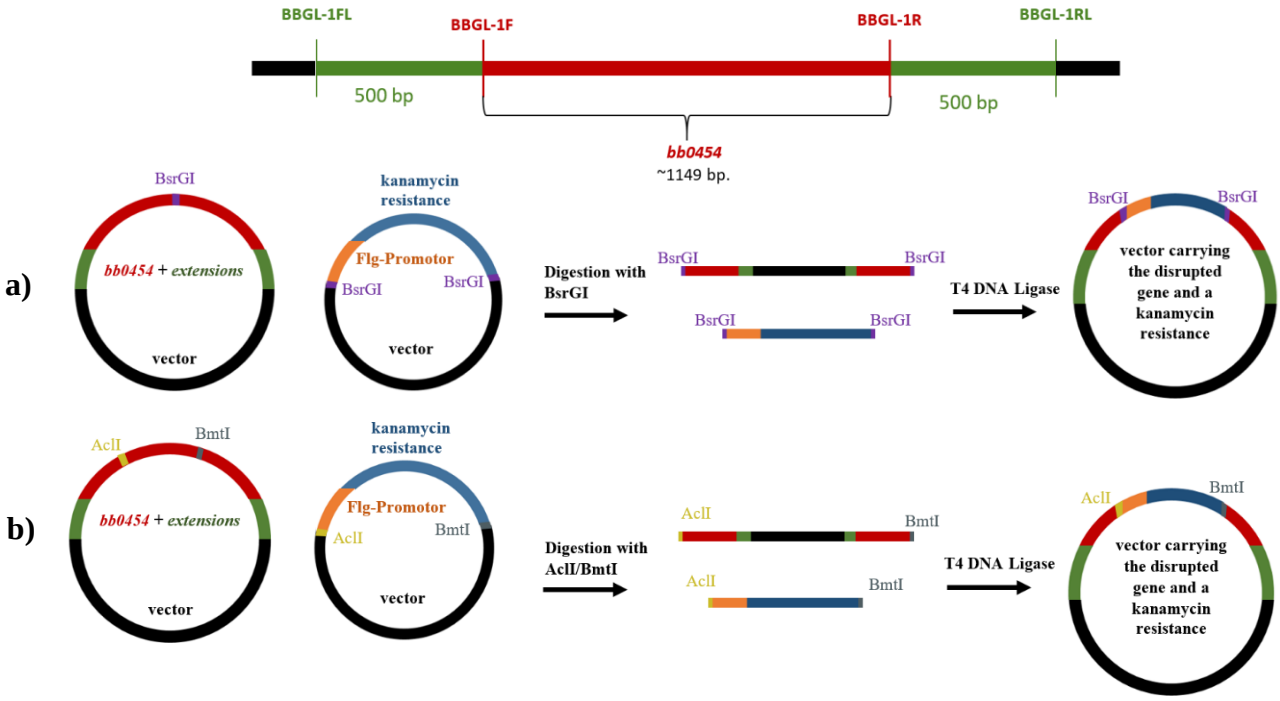

Figure 3 - Scheme of the construction of the knock-out construct of *bb0454;* a: CB43; b: B31

First of all, a vector carrying the Flg promotor and kanamycin antibiotic resistance with BsrGI cutting sites on both sides for CB43 and one with AclI cutting side in the beginning and BmtI in the end were created. Therefore, a PCR was run using the primer pair flgKan BsrGI F and flgKan BsrGI R for CB43 and flgAclI F and KanBmtI R for B31 with the following PCR settings.

Table 8 – PCR master mix and settings for the amplification of FlgKan – knock-out construct

Master Mix	**20 µL reaction**
10xOne Taq Standard Reaction Buffer	4 µL
dNTP's	0.4 µL
One Taq Hot Start Polymerase	0.1 µL
Primer F	1 µL
Primer R	1 µL
MiliQ H_2O	11.5 µL
DNA	2 µL

PCR settings	
Initial denaturation:	94°C for 30 sec.
30 cycles:	94°C for 30 sec.
	53°C for 45 sec.
	68°C for 1 min.
Final extension:	68°C for 5 min.
Hold:	16°C

The resulting PCR products were cloned into pGEM-T easy vector, transformed into *E. coli*, grown at carbenicillin plates and further processed until isolation of the plasmid DNA (3.4 –

3.6). Using Restriction Digestion with EcoRI and gel electrophoresis the DNA was checked for the presence of an insert with the right size (~12000 base pairs).

Using BsrGI, pGEM-T easy vector containing CB43's *bb0454* plus extensions was cut open. With the same enzyme the BsrGI-FlgKan was cut out of the pGEM-T easy vector. The DNA samples were run on a 5% agarose gel after colouring with 10xSYBR dye together with a 1kb+ DNA ladder from Invitrogen. After separation of the different parts of DNA, the FlgKan piece was cut out of the gel and the DNA was extracted using QIAquick Gel Extraction Kit from QIAGEN by following the manufacturer's instructions.

After DNA extraction, the FlgKan piece was inserted into the cut open CB43's *bb0454* gene by T4 DNA Ligase. The same was done with AclI and BmtI with B31 DNA and AclI/BmtI-FlgKan.

To ligate the two pieces of DNA together, a ligation reaction with T4 DNA Ligase was set up in micro centrifugation tubes on ice and gently mixed by pipetting. Then the reaction was incubated overnight at 16°C.

Table 9 - T4 Ligation Reaction

Substance	**Amount [µL]**
10x T4 DNA Ligase Buffer	2 µL
Vector DNA (*bb0454 + extension)*	1 µL
Insert DNA (FlgKan)	3 µL
MilliQ H_2O	13 µL
T4 DNA Ligase	1 µL

The ligation reactions were used for transformation in *E. coli* NEB 5 alpha cells by adding 3 µL of the reactions to 25 µL of cells followed by mixing by flicking. The cells were incubated for 30 minutes at ice and heat shocked at 42°C for 30 seconds. Then, they were kept on ice for another 2 minutes and 250 µL of S.O.C. media was added. The cells were shaken at 37°C for 60-90 minutes and then 150 µL and 75 µL of each sample were spread on agar plates containing kanamycin. The plates were incubated overnight at 37°C.

3.10 Construction of a shuttle vector carrying *bb0454*

pBSV2G was chosen as a shuttle vector for the gene *bb0454* to be introduced multiple times into the genome. The pBSV2G shuttle vector is carrying a gentamicin resistance gene and has a multiple cloning site (MCS), where multiple restriction enzyme are able to cut the vector

open. Into this MCS the *bb0454* of CB43 (*B. afzelii)* and B31 (*B. burgdorferi* s.s.) with the Flg promotor in front of it, was supposed to be inserted.

First of all, restriction enzymes were searched for, which do not cut the gene or the Flg promotor but are cutting the pBSV2G vector at its MCS. Using NEBcutter V2.0 SbfI and XohI from NEB were identified to not cut both sequence versions of *bb0454*. However, SbfI cuts in the MCS of the shuttle vector. The needed parts for the construction of the shuttle vector are shown in Figure 4.

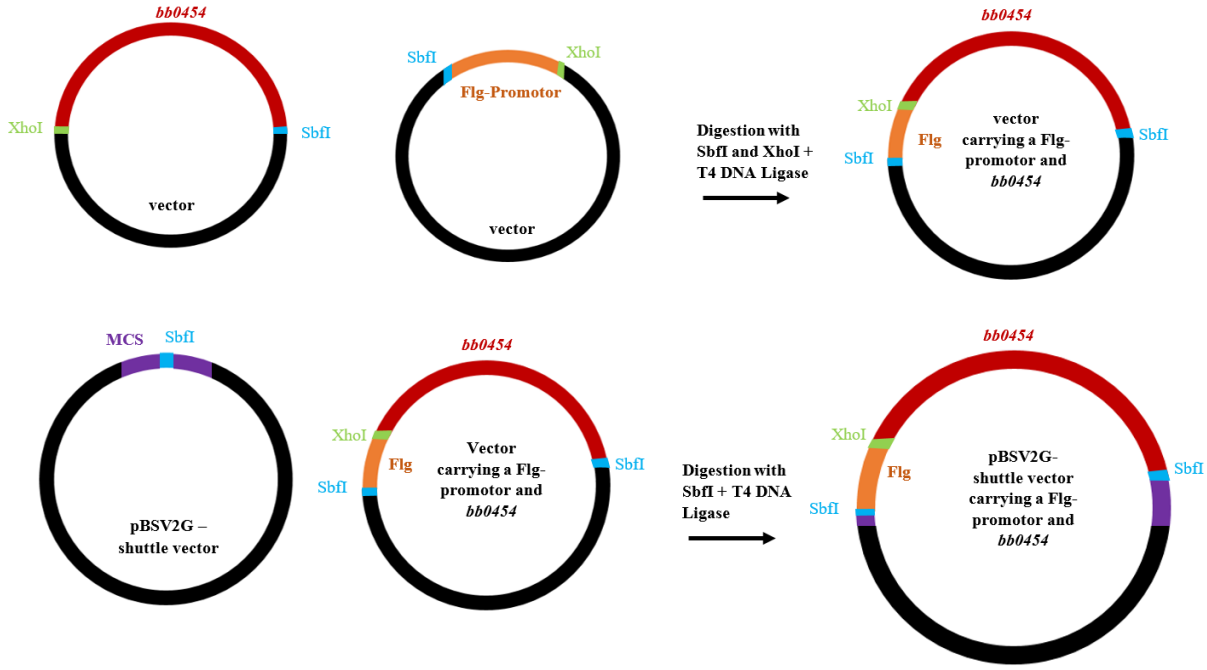

Figure 4 - Scheme of the construction of the shuttle vector carrying *bb0454*

PCR was performed with the primer pair Flg Sbf F/Flg XhoI R to create a Flg promotor with SbfI-restriction site in the beginning and XhoI-cutting site in the end. For this PCR an annealing temperature of 55°C was used and the expected size was 200 base pairs.

Furthermore, PCR was used to create *bb0454* with a XhoI-site in the beginning and a SbfI-site in the end. Due to the fact that the sequence of *bb0454* of the strain B31 and the one of CB43 differ in the end, different reverse primers were used. For both samples the forward primer 454-XhoI F was used. However, for B31 the reverse primer B31/454-SbfI R and for CB43 the primer CB43/454-SbfI R were used for the reaction. The annealing temperature was set to 50°C and the expected size of both products was approximately 1150 base pairs. In Table 10 the PCR settings for the construction of the shuttle vector are shown.

Table 10 - PCR master mix and settings for construction of the shuttle vector

Master Mix 3	**20 µL reaction**
10x Thermopol buffer	2 µL
dNTP's	2 µL
Taq DNA Polymerase	0.1 µL
Primer F	1 µL
Primer R	1 µL
MiliQ H_2O	11.9 µL
DNA	2 µL

PCR settings 3	
Initial denaturation:	95°C for 30 sec.
30 cycles:	95°C for 30 sec.
	50 (55)°C for 45 sec.
	68°C for 1 min.
Final extension:	68°C for 5 min.
Hold:	16°C

After confirming that the PCR products had the expected sizes by agarose gel electrophoresis, they were cloned into pCR®2.1-TOPO® vector and transformed into *E. coli*. Using blue-white screening colonies containing a vector where chosen, grown and the plasmid DNA was isolated. These steps were followed by restriction digestion to check if a vector carrying the insert was really present in the isolated DNA. The detailed procedures for these steps can be found in chapter 3.4 – 3.6.

4 Results

4.1 Nucleotide and Amino Acid sequences

The nucleotide sequences of *bb0454* obtained during the work differed in size from each other and were either 1149 or 1152 nucleotides long. Strains belonging to *B. burgdorferi s.s.*, *B. americana (B. am.)*, *B. carolinensis (B. caro.)*, and *B. finlandensis* are three nucleotides longer than the ones belonging to *B. afzelii*, *B. garinii*, and *B. bavariensis*. A multiple alignment of these sequences done by Geneious shows the additional TTC-triplet at the position 495.

Differences in the nucleotide and amino acid sequences become obvious when comparing them with each other and calculating identity scores. This was done for representatives of the different genospecies analysed. The calculated sequence identity scores of nucleotides and amino acids are shown in Table 11 and 12.

Table 10 - nucleotide sequence identity

	B31	SV1	*B. am.*	*B. caro.*	PBi	PBr	CB43	RU1
B31	-	-	-	-	-	-	-	-
SV1	98.177%	-	-	-	-	-	-	-
B. am.	94.705%	95.139%	-	-	-	-	-	-
B. caro.	94.878%	95.313%	99.826%	-	-	-	-	-
PBi	92.708%	92.622%	92.361%	92.535%	-	-	-	-
PBr	92.188%	92.708%	92.188%	92.361%	98.259%	-	-	-
CB43	94.010%	94.531%	92.882%	93.056%	94.691%	94.082%	-	-
RU1	94.079%	93.750%	93.056%	93.229%	95.126%	94.865%	94.952%	-

Table 11 – amino acid sequence identity

	B31	SV1	*B. am.*	*B. caro.*	PBi	PBr	CB43	RU1
B31	-	-	-	-	-	-	-	-
SV1	97.911%	-	-	-	-	-	-	-
B. am.	93.995%	93.734%	-	-	-	-	-	-
B. caro.	94.256%	93.995%	99.739%	-	-	-	-	-
PBi	92.689%	92.689%	92.689%	92.950%	-	-	-	-
PBr	92.689%	92.689%	92.689%	92.950%	98.953%	-	-	-
CB43	94.517%	93.995%	93.473%	93.734%	93.734%	95.288%	-	-
RU1	94.517%	93.473%	93.995%	94.256%	96.597%	96.073%	95.288%	-

A closer look on the amino acid sequences of *bb0454* revealed not only changes in amino acids but also the properties of these amino acids varied. Selected regions of the amino acids multiple alignment are shown in Figure 5. This alignment contains not only the sequences obtained during this work but also sequences of other *Borrelia* species, some of which cause relapsing fever.

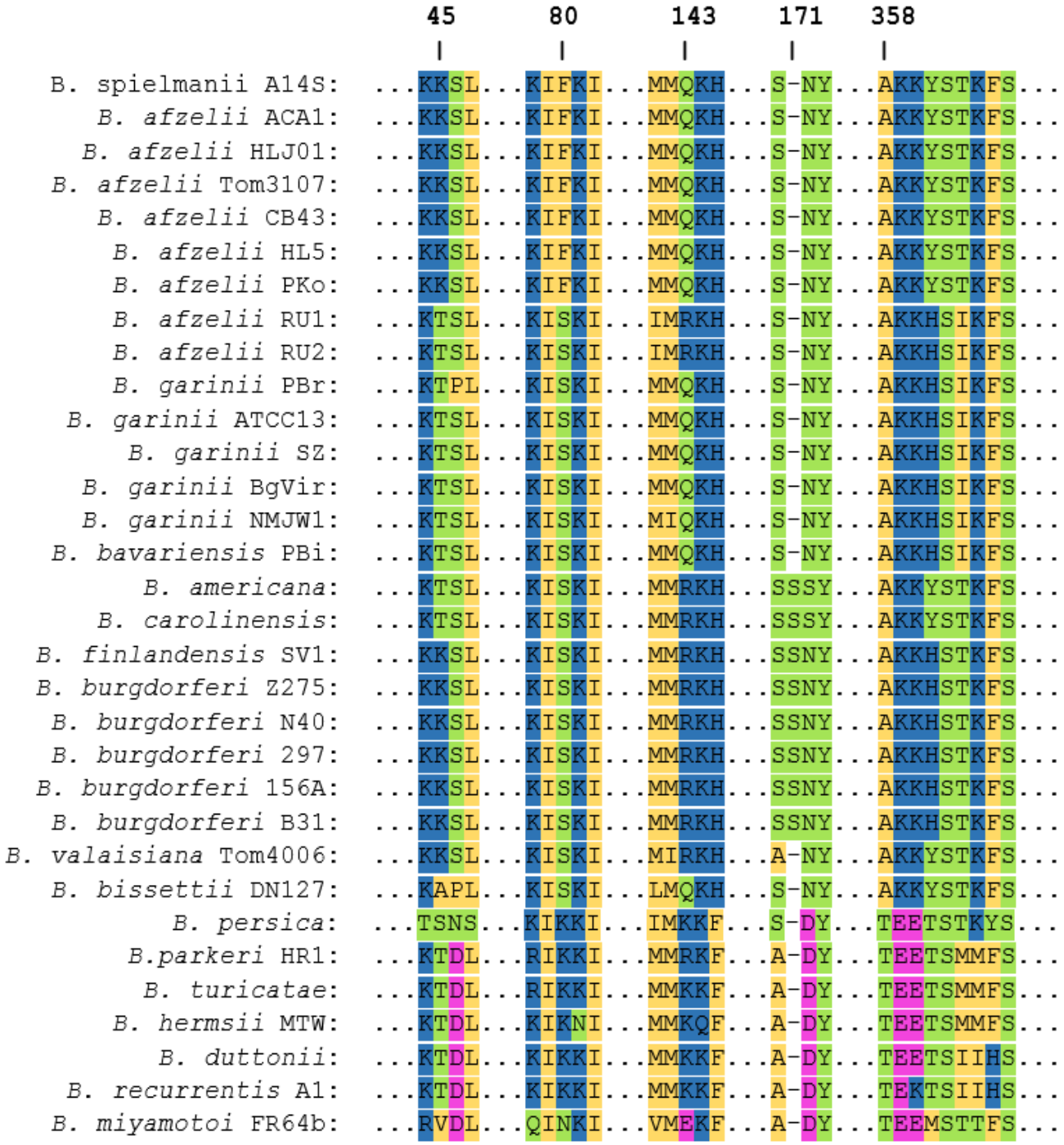

Figure 5 – Selected regions of the amino acid (AA) alignment showing the difference of *bb0454* between the genospecies with colours representing the difference in polarity of the amino acids; blue: polar, charged, basic AA; green: polar, neutral AA; yellow: neutral, hydrophobic AA; pink: polar, charged, acidic AA

The sequences, obtained during this work, were submitted to GenBank with a release date after publication of this work.

4.2 Phylogenetic Analysis of *bb0454*

Using the sequences obtained during the previous work, as well as additional sequences of *bb0454* of other representatives of the *Borrelia* genus downloaded from the NCBI database a phylogenetic analysis of *bb0454* was conducted to learn more about the relationship between the different *Borrelia* species.

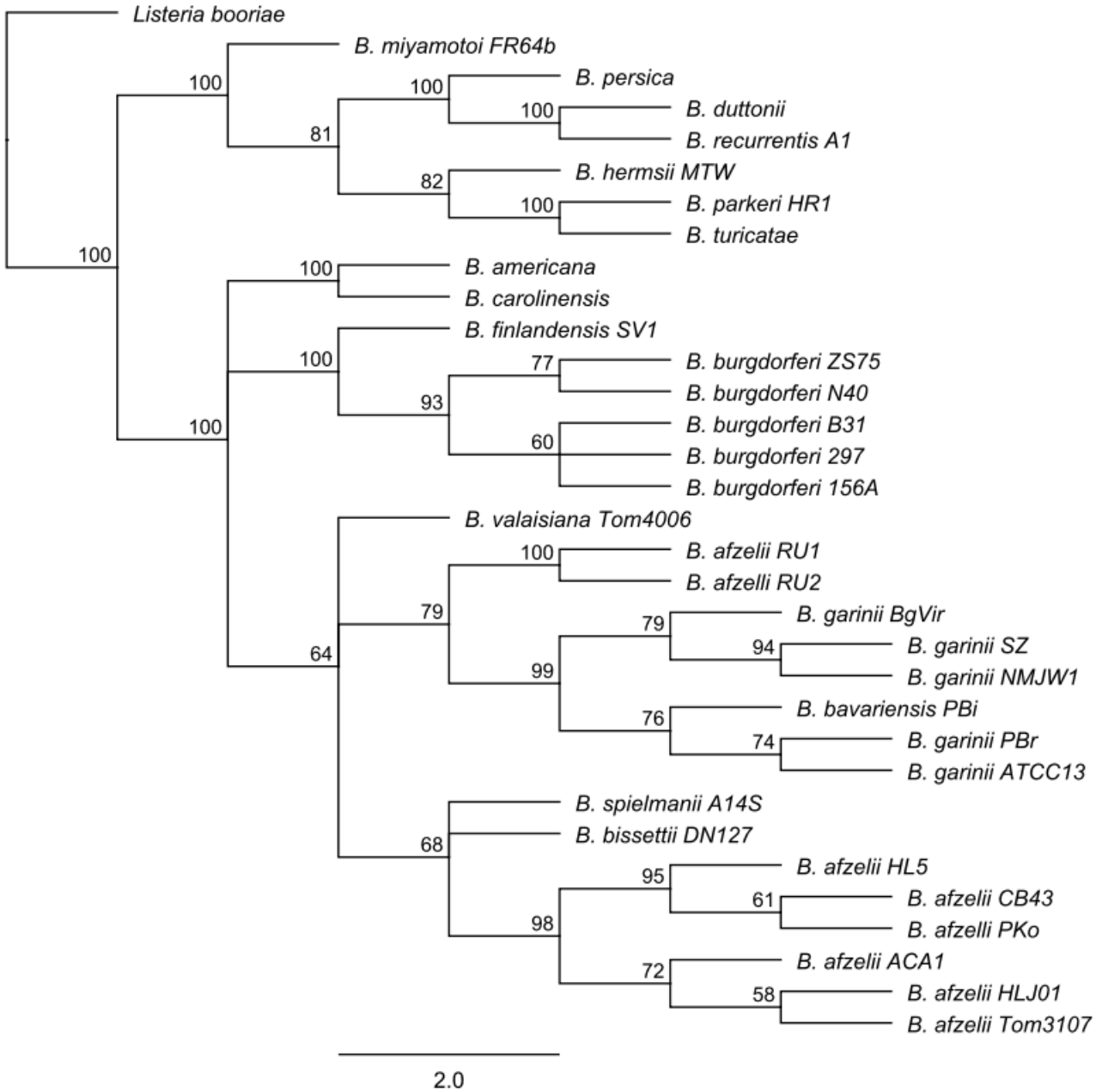

Figure 6 - Phylogenetic tree of *bb0454*

This phylogenetic tree shows the expected grouping of the different strains in the genospecies. The exception are RU1 and RU2 which do not sit with the other *B. afzelii* strains.

4.3 Comparison of 3D-structures

To compare structural elements of bb0454 3D models of the protein for representatives of *B. afzelii* (CB43), *B. burgdorferi* s.s. (B31), *B. bavariensis* (PBi), *B. garinii* (PBr) and of *B. miyamotoi* (FR64b) were computed. The resulting structures are shown in Figure 7.

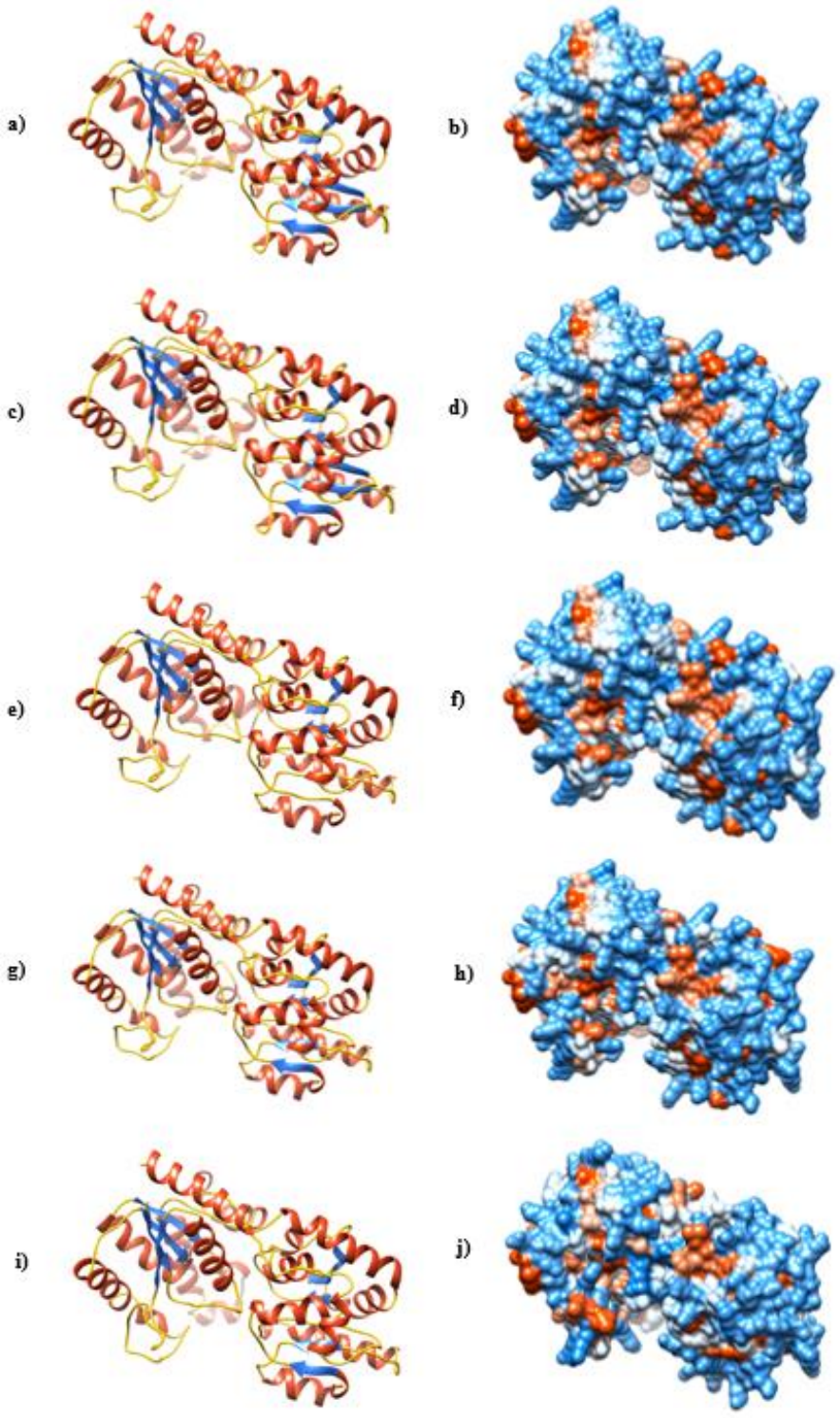

Figure 7 - 3D structure of *bb0454*; a, c, e, g, i: secondary structures (red: α-helix, blue: β-strand); b, d, f, h, j: surface structure displaying differences in hydrophobicity; a,b - B31; c,d - PBi; e,f - PBr, g,h – CB43; i,j – FR64b;

Using Chimera program the 3D structures were visualized and compared with each other. The analysis showed that the 3D structure of PBi is similar to the one of B31. Compared to B31 and PBi, the other strains have at least one β-strand less. *B. garinii* PBr has three β-strands less.

4.4 Knock-out Construct

The creation of the knock-out construct was unsuccessful. After restriction digestion of the vectors carrying the different pieces of DNA, and ligating them back together, the transformed *E. coli* cells were grown on plates containing kanamycin. Due to carryover of TOPO vector it was impossible to pick colonies containing the gene disrupted by the kanamycin cassette. The TOPO vector is already containing a kanamycin resistant gene and therefore it was impossible to distinguish between colonies with the TOPO vector carrying the full gene and the disrupted version.

Therefore, further work has to be done to complete the construction of the knock-out of *bb0454.* A possible solution is to transfer the gene *bb0454* from the TOPO vector to the pGEM-T easy one, which has no kanamycin resistance incorporated.

4.5 Shuttle vector carrying *bb0454*

Due to lack of time we had to stop the construction of the shuttle vector after obtaining the different required sequences within the plasmids.

Finish the construction of the shuttle vector and overexpression of *bb0454* in vivo are the next steps to further increase our knowledge about *bb0454* and its significance for *Borrelia burgdorferi* s.l..

5 Discussion

Together with other spirochetes, *Borrelia burgdorferi* s.l. has a specific structure of its two membranes, including the lack of LPS and missing linkages between the peptidoglycan layer and the outer membrane, therefore distinguishing it from Gram-negative bacteria[10].

Studies have shown the importance of glycosylation, the attachment of polysaccharides to proteins, lipids or other organic molecules, in bacteria and all other forms of life. In other pathogenic bacteria it was proven that the disruption of N-glycosylation pathways lead to a decrease in infectivity. However, earlier work demonstrated that *B. burgdorferi* s.l. outer membrane proteins and flagellin are not N-glycosylated. One possibility to synthesise glycosylated products is with the help of so called glycosyltransferases. These enzymes catalyse group-transfer reactions like the synthesis of glycolipids[13,15–17].

B. burgdorferi s.l. has two major glycolipids in its membrane, together making up more than one third of the total lipid mass of this kind of bacteria – ACG and MGalDAG. Previous studies showed that bbMGS encoded by *bb0454* is responsible for the synthesis of MGalDAG by glycosylating 1,2-diacylglycerol.[22–26]

Although *bb0454* encodes for an enzyme that we would expect to be highly homologous over all *Borrelia* genospecies, this does not seem to be true. The phylogenetic analysis clearly paints a picture of an ancestral form of this gene which diverged first into the relapsing fever *Borrelia* and Lyme disease *Borrelia* and then into 3 separate clades grouped under primarily *B. burgdorferi s.s.*, *B. afzelii* and *B. garinii*. It is important to note that the differences in the amino acid sequences and tertiary structures are significant enough to suggest that *bb0454* could have been under specific immunological pressure because of varied host specificity. Previous work has shown that the mutation of single genes encoding for functional glycosyltransferases has consequently affected the biosynthesis of different polysaccharides. For example, mutation in the *gpsX* gene which encodes a glycosyltransferase, in phytopathogenic *Xanthomonas*, lead to an altered LPS synthesis[31]. It also lead to a decrease in virulence which was associated with the altered LPS side chains as well as a reduction in extra polysaccharide (EPS) production[32] . This sort of evidence in gene mutations leading to a change in the LPS structure has also been shown in the genes *xanA* and *xanB*, involved in UDP-Glucose and GDP-Mannose biosynthesis in *Xanthomonas campestris*. The mutations also showed how a pathogen with a mutation in a glycosyltransferase gene within the same strain could lead to infection of a plant that was resistant to the parental strain[33]. We suggest that the differences we see of this gene within the different *Borrelia* clades are a result of trying to establish over time within the ecological niches which the *Borrelia* are circulating in based on the most abundant of reservoir hosts and the

immune system that they need to evade. *B. garinii* is primarily associated with birds as its reservoir host, whereas *B. afzelii*, *B. bavariensis* are associated with rodents, while *B. burgdorferi s.s.* is more of a generalist *Borrelia* species when looking at reservoir hosts that it can infect[34,35]. Future work in using a knock-out construct to inactivate the gene *bb0454* will help us determine the importance of the glycolipids in the *Borrelia* membrane for infection within a mammalian host and differences between various *Borrelia* genospecies for this gene.

An exception to this overall trend within the isolates used are RU1 and RU2. These strains were isolated from *I. ricinus* female ticks collected in Russia and then classified as *B. afzelii*. This was done using primers amplifying regions on the plasmids lp17 and lp25. These regions were sequenced and run against entries of the NCBI database. The best hits with highest identities classified them as *B. afzelii*. However, when looking at the multiple alignment of *bb0454* one can see that the sequences of these strains are at some positions similar to *B. garinii* strains (pos. 45, 80) and sometimes they correspond with *B. burgdorferi* s.s. and not with *B. garinii or B. afzelii* (pos.; 143). This suggest that RU1/RU2 might sit in between *B. garinii*, *B. afzelii*, and *B. burgdorferi* s.s. possibly hinting at another intermediate genospecies. In further work, this will need to be tested using multilocus sequence typing (MLST), the analysis of eight housekeeping genes on the chromosome[36]. This analysis could provide proof as to where exactly do these isolates fall under in the *B. burgdorferi* s.l. family.

The construction of the shuttle vector carrying the full gene was incomplete and is a task for future work. A study done with *Brucella*, a gram negative bacteria causing brucellosis in mammals, showed that overexpression of a putative glycosyltransferase wbkA, lead to the production of EPS and clumping of the bacteria. EPS has been shown to help bacteria to survive in hostile environment. Earlier work suggest that *B. burgdorferi* s.l. might be able to develop a protective EPS layer as well. So overexpression of *bb0454* might lead to the formation of such a protection layer of the bacteria[37,38].

For the first time an extensive phylogenetic study of *bb0454* of different isolates from European and North American genospecies has been done. Other studies showing changes in infectivity due to disruption of glycosylation in different bacteria suggest that a knock-out of this gene might influence the pathogenicity of *B. burgdorferi* s.l. as well[15]. Additional work with the knock-out might prove the importance of this transferase for *B. burgdorferi* s.l infectious cycle. Studying bacterial glycosyltransferases will help us provide an insight into how glycans or antigens are synthesized. This will help us in studying their functions and role in disease and the design of vaccines/diagnostic markers against bacterial pathogens like those belonging to the Lyme disease causing *Borrelia* group.

6 Acknowledgment of strains

For providing different *Borrelia* strains special thanks go to:

- *B. americana, B. carolinensis:* Natasha Rudenko and Maryna Golovchenko, Institute of Parasitology, Budweis, Czech Republic
- *B. bavariensis* (PBi): Peter Kraiczy, Institute of Medical Microbiology and Infection Control, Frankfurt/Main, Germany
- *B. garinii* (PBr), *B. afzelii* (PKo, ACA1): Ian Livey, Baxter, Vienna, Austria
- *B. burgdorferi* (297): Melissa Caimano, Department of Medicine, Connecticut, USA

7 References

1. Barbour, A. G. & Hayes, S. F. Biology of Borrelia species. *Microbiol. Rev.* **50,** 381–400 (1986).
2. Radolf, J. D., Caimano, M., Stevenson, B. & Hu, L. Of ticks, mice and men: understanding the dual-host lifestyle of Lyme disease spirochaetes. *Nat Rev Microbiol.* **10,** 87–99 (2012).
3. Rizzoli, A. *et al.* Lyme borreliosis in Europe. *Eurosurveillance* **16,** 1–8 (2011).
4. Coipan, E. C. *et al.* Imbalanced presence of Borrelia burgdorferi s.l. multilocus sequence types in clinical manifestations of Lyme borreliosis. *Infect. Genet. Evol.* 3–5 (2016). doi:10.1016/j.meegid.2016.04.019
5. Margos, G., Vollmer, S., Ogden, N. & Fish, D. Population genetics, taxonomy, phylogeny and evolution of Borrelia burgdorferi sensu lato. *... , Genet. Evol.* **11,** 1545–1563 (2011).
6. Norris, S. J. The dynamic proteome of Lyme disease Borrelia. *Genome Biol.* **7,** 209 (2006).
7. Purser, J. E. & Norris, S. J. Correlation between plasmid content and infectivity in Borrelia burgdorferi. *Proc. Natl. Acad. Sci. U. S. A.* **97,** 13865–70 (2000).
8. Xu, Y., Kodner, C., Coleman, L. & Johnson, R. C. Correlation of plasmids with infectivity of Borrelia burgdorferi sensu stricto type strain B31. *Infect. Immun.* **64,** 3870–3876 (1996).
9. Grimm, D., Elias, A. F., Tilly, K. & Rosa, P. A. Plasmid stability during in vitro propagation of Borrelia burgdorferi assessed at a clonal level. *Infect. Immun.* **71,** 3138–3145 (2003).
10. Schröder, N. W. J., Eckert, J., Stübs, G. & Schumann, R. R. Immune responses induced by spirochetal outer membrane lipoproteins and glycolipids. *Immunobiology* **213,** 329–340 (2008).
11. Nothaft, H. & Szymanski, C. M. Protein glycosylation in bacteria: sweeter than ever. *Nat. Rev. Microbiol.* **8,** 765–778 (2010).
12. Schäffer, C., Graninger, M. & Boltzmann-institut, L. Review Prokaryotic glycosylation. 248–261 (2001).
13. Nothaft, H. & Szymanski, C. M. Bacterial protein n-glycosylation: New perspectives and applications. *J. Biol. Chem.* **288,** 6912–6920 (2013).
14. Scott, N. E. *et al.* Simultaneous glycan-peptide characterization using hydrophilic interaction chromatography and parallel fragmentation by CID, higher energy collisional

dissociation, and electron transfer dissociation MS applied to the N-linked glycoproteome of Campylobact. *Mol. Cell. Proteomics* **10,** M000031–MCP201 (2011).

15. Karlyshev, A. V. *et al.* The Campylobacter jejuni general glycosylation system is important for attachment to human epithelial cells and in the colonization of chicks. *Microbiology* **150,** 1957–1964 (2004).
16. Štěrba, J. *et al.* Flagellin and outer surface proteins from Borrelia burgdorferi are not glycosylated. *J. Bacteriol.* **190,** 2619–2623 (2008).
17. Rini, J., Esko, J. & Varki, A. Chapter 5: Glycosyltransferases and Glycan-processing Enzymes. *Essentials Glycoboilogy. 2nd Ed.* 1–7 (2009). doi:NBK1921 [bookaccession]
18. Brockhausen, I. Crossroads between bacterial and mammalian glycosyltransferases. *Front. Immunol.* **5,** 1–21 (2014).
19. Vaara, M. Agents that increase the permeability of the outer membrane. *Microbiol. Rev.* **56,** 395–411 (1992).
20. Public, N. Molecular Basis of Bacterial Outer Membrane Permeability. **49,** 1–32 (1985).
21. Yethon, J. A., Vinogradov, E., Perry, M. B. & Whitfield, C. Mutation of the lipopolysaccharide core glycosyltransferase encoded by waaG destabilizes the outer membrane of Escherichia coli by interfering with core phosphorylation. *J. Bacteriol.* **182,** 5620–5623 (2000).
22. Ben-Menachem, G., Kubler-Kielb, J., Coxon, B., Yergey, A. & Schneerson, R. A newly discovered cholesteryl galactoside from Borrelia burgdorferi. *Proc. Natl. Acad. Sci.* **100,** 7913–7918 (2003).
23. Jones, K. L. *et al.* Strong IgG antibody responses to Borrelia burgdorferi glycolipids in patients with Lyme arthritis, a late manifestation of the infection. *Clin. Immunol.* **132,** 93–102 (2009).
24. Stübs, G. *et al.* Acylated cholesteryl galactosides are specific antigens of borrelia causing lyme disease and frequently induce antibodies in late stages of disease. *J. Biol. Chem.* **284,** 13326–13334 (2009).
25. Östberg, Y., Berg, S., Comstedt, P., Wieslander, Å.˚& Bergström, S. Functional analysis of a lipid galactosyltransferase synthesizing the major envelope lipid in the Lyme disease spirochete Borrelia burgdorferi. *FEMS Microbiol. Lett.* **272,** 22–29 (2007).
26. Samuels, D. & Radolf, J. D. Borrelia: Molecular Biology, Host Interaction and Pathogenesis. 124–126 (2010).
27. Kearse, M. *et al.* Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **28,** 1647–

1649 (2012).

28. Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N. & Sternberg, M. J. E. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* **10,** 845–858 (2015).
29. Pettersen, E. *et al.* UCSF Chimera - a visualization system for expolratory reserach and analysis. *J Comput Chem* **25,** 1605–12 (2004).
30. Vincze, T., Posfai, J. & Roberts, R. J. NEBcutter: A program to cleave DNA with restriction enzymes. *Nucleic Acids Res.* **31,** 3688–3691 (2003).
31. Li, J. & Wang, N. The gpsX gene encoding a glycosyltransferase is important for polysaccharide production and required for full virulence in Xanthomonas citri subsp. citri. *BMC Microbiol.* **12,** 31 (2012).
32. Kingsley, M. T., Gabriel, D. W., Marlow, G. C. & Roberts, P. D. The opsX locus of Xanthomonas campestris affects host range and biosynthesis of lipopolysaccharide and extracellular polysaccharide. *J. Bacteriol.* **175,** 5839–5850 (1993).
33. Koplin, R. *et al.* Genetics of xanthan production in Xanthomonas campestris: The xanA and xanB genes are involved in UDP-glucose and GDP-mannose biosynthesis. *J. Bacteriol.* **174,** 191–199 (1992).
34. Margos, G. *et al.* A new Borrelia species defined by multilocus sequence analysis of housekeeping genes. *Appl. Environ. Microbiol.* **75,** 5410–5416 (2009).
35. Mannelli, A., Bertolotti, L., Gern, L. & Gray, J. Ecology of Borrelia burgdorferi sensu lato in Europe: Transmission dynamics in multi-host systems, influence of molecular processes and effects of climate change. *FEMS Microbiol. Rev.* **36,** 837–861 (2012).
36. Margos, G. *et al.* MLST of housekeeping genes captures geographic population structure and suggests a European origin of Borrelia burgdorferi. *Proc. Natl. Acad. Sci. U. S. A.* **105,** 8730–5 (2008).
37. Dabral, N., Jain-Gupta, N., Seleem, M. N., Sriranganathan, N. & Vemulapalli, R. Overexpression of Brucella putative glycosyltransferase WbkA in B. abortus RB51 leads to production of exopolysaccharide. *Front. Cell. Infect. Microbiol.* **5,** 54 (2015).
38. Sapi, E. *et al.* Characterization of Biofilm Formation by Borrelia burgdorferi In Vitro. *PLoS One* **7,** 1–11 (2012).