Přírodovědecká fakulta
Faculty of Science

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

Further Delineation of *Borrelia burgdorferi* Restriction-Modification system and understanding antibiotic resistance in *Borrelia afzelii*

Laboratory of Molecular Ecology of Vectors and Pathogens

INSTITUTE OF PARASITOLOGY

Biology Center, ASCR

Stepan Demchyshyn

Supervisor: Ryan O.M. Rego, PhD

Annotation: The aim of this thesis was to further the understanding of restriction-modification and the homology within the restriction-modification genes in various *Borrelia* species including *Borrelia afzelii* as well as to investigate spontaneous antibiotic resistance within this particular *Borrelia* species.

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.

České Budějovice, Date
December 15, 2015

Stepan Demchyshyn

Contents

1 INTRODUCTION

1.1 LYME DISEASE

Lyme borreliosis, also known in the United States as Lyme disease is caused by the spirochete, *Borrelia burgdorferi*. This disease takes its name after a town in Connecticut, USA where an epidemic of the disease broke out during the 1970's[1]. Lyme disease is an issue in Europe, USA and Asia, though it is caused by different genospecies in different geographical regions. More than 20 species of *Borrelia* have been identified as Lyme disease spirochetes all together referred as *Borrelia burgdorferi* sensu lato. *Borrelia burgdorferi* sensu stricto (s.s.), *B. afzelii, B. bavariensis* and *B. garinii* have been identified as human pathogens[2]. *B. burgdorferi* s.s. is unique to the US[3] whereas the latter three are more common in Europe[4]. Even though the figures vary largely from one region to another, it has been reported that in Western Europe 22.4 persons per 100 000 each year are diagnosed with Lyme borreliosis[5]. This adds up to more than 200 000 cases per year[5] in the region and indicates slowly escalating new epidemic that is in need of closer attention and evaluation from local authorities.

Lyme disease usually manifests itself in three distinct stages[6]. The first stage is characterized by localized round skin lesions, which are called *erythema migrans*. It is also common for a patient to exhibit symptoms similar to influenza during this stage. After days or in some cases weeks the disease progresses to its second stage. In this stage infection proceeds to infect different systems of the human body. In 60% of the cases the infection is affecting joints but it also might affect nervous system and heart in 15% and 5% respectively[6]. This stage can be well treated with a few weeks course of antibiotics, though if not treated it can continue into the third stage of persistent infection. The treatment of Lyme disease is well defined and mostly consists of 2 to 4 weeks long course of various antibiotics, depending on the stage of the disease[7]. In addition, anti-inflammatory drugs can be applied to reduce the symptoms of arthritis or in case the symptoms persist (post-Lyme disease symptoms) they can be treated separately.

1.2 B. burgdorferi

B. burgdorferi sensu lato belongs to the phylum of *Spirochaetes*. This phylum is easily identified by an unusual spiral-shaped body of the bacteria and periplasmic flagella[8]. The genome of the various species that make up *B. burgdorferi* sensu lato is unique in having a segmented genome. It contains one large chromosome which is approximately 1 Mbp[9] in size and up to 22 smaller plasmids (linear and circular) ranging in size from 5 kbp to 56 kbp[10] (*Figure 1*). While the chromosomal genes are similar to genes of other bacteria, the majority of plasmid genes are unique to *Borrelia*[3]. It has been shown, that

passaging of *Borrelia* species in vitro leads to loss of some of the plasmids important for virulence, thus eventually leading to their inability to infect[11]. To this day 23 full or partial genomes have been sequenced that are associated with *Borrelia* genospecis/strains. There are 14 sequences available for *B. burgdorferi* sensu stricto, 5 for *B. garinii*, and one each for *B. bavariensis*, *B. finlandensis*, *B. spielmanii*, *B. valaisiana*[4].

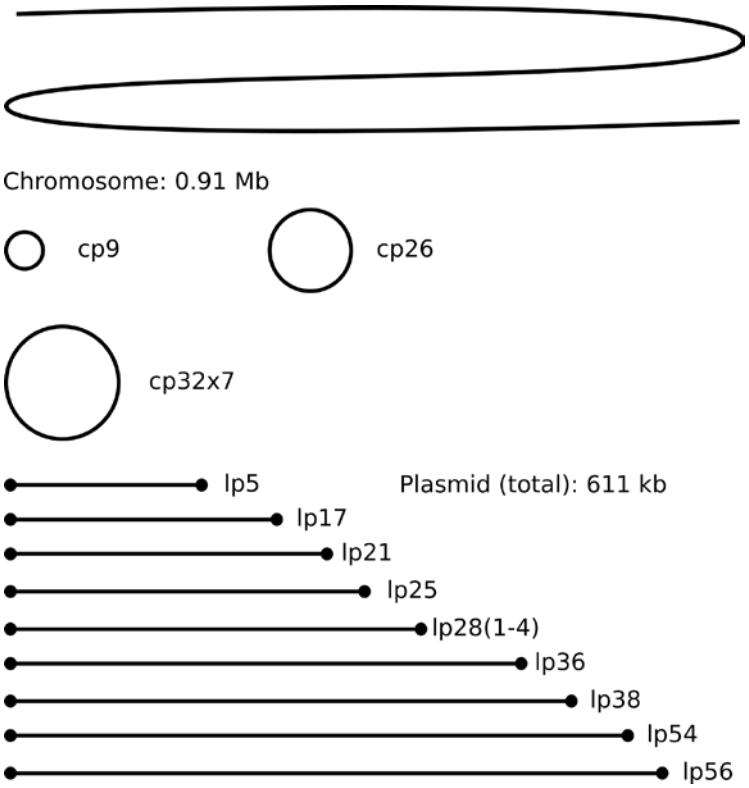

Figure 1. The illustration of genome of *B.burgdorferi* B31 adapted from work of Gray and colleagues[1]

1.3 GENETIC MANIPULATION

Genetic manipulation of bacteria is based on the thorough knowledge of structure and function of different parts of the genome, growth conditions, and evolutionary relationship with other bacterial species. Different methods of genetic manipulation have been previously applied to *Borrelia* with successful discoveries in the field of tick-borne pathogens[12,13]. About half a dozen methods have been developed becoming standard over the years. These include the use of selectable markers, gene inactivation and complementation, random targeted mutagenesis, use of reporter genes, engineered gene expression and others[12]. Unfortunately there are still strong limitations to using various methods of genetic manipulations which are commonly used for other Gram-positive or Gram-negative bacteria. The source of these limitations is mostly the genetic distance between other bacteria and *Borrelia* and also a unique usage of genetic material by *Borrelia*. During genetic manipulation, transformants tend to lose plasmids compared to the parental strain thus the screening of a particular clone becomes necessary. It is

very important to work on new and improved genetic manipulation methods in order to overcome the limitations in *Borrelia.*

1.4 RESTRICTION MODIFICATION

Methylation patterns are a key feature to many of the processes in both prokaryotic and eukaryotic cells. A methyl group is introduced onto a DNA molecule (usually a palindromic site) by methyltransferase, which moves the group from S-adenosine-L-methionine to the DNA.

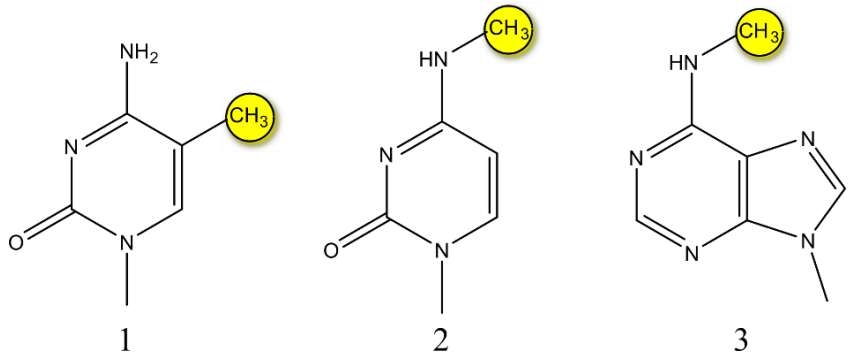

Figure 2. Various types of methylated DNA bases. 1: 5-methylcytosine, 2: N4-methylcytosine, 3: N6- methyladenine

A methyl group can be transferred to adenine or cytosine, the latter nucleotide having two different ways of accepting the methyl group. Thus additional information can be introduced on a DNA molecule, making them so-called additional letters of the genetic alphabet[14]. Importantly, the addition of methyl groups does not interfere with Watson-Crick pairing principles. Due to evolutionary reasons, there exists a number of methyltransferases, which can recognize multiple sequences. Today more than 700 different methyltransferases have been sequenced which recognize up to 300 different sequences. Even though in eukaryotic cells methylation plays a very important role, in prokaryotes it has a bigger variety of applications. Methylation is responsible for self and non-self DNA recognition, direction of post replicative mismatch repair, and control of cell cycle[14]. Also it is important to mention that methylation can play a role in the pathogenicity of certain bacterial species[15,16]. Among known methods to determine methylation patterns, bisulfite treatment is one of the more common methods that allows one to determine precise pattern of CpG methylation. Also one can acquire information about methylation of DNA by performing restriction enzyme digests.

Restriction modification (R-M) system was discovered as one of the main obstacles in the way for transformation in bacteria. This system is present in both Bacteria and Archae[17] and according to recent

research it serves multiple roles such as regulation of genetic uptake between bacteria, adaptation of the bacteria to specific ecological niche[18], evolution[19] and many more[20,21]. However, its main role is to defend the organism against foreign DNA, for instance DNA of bacteriophages[22]. R-M system has a specific enzyme, endonuclease, which will recognize palindromic stretches of DNA, usually 4-8 bp long, and cut it up therefore ensuring successful destruction of the alien DNA[23]. Since the 4-8 bp pieces of DNA are relatively short and it is very likely that the same sequences occur in the bacteria's own genome, there has to be a way to prevent R-M system from splicing its own genetic material. Methyltransferase fixes this problem by methylating specific stretches of DNA with the particular sequence and preventing self-destruction of the organism by its own R-M system. R-M systems are classified into 4 different types according to their build-up, where types I, II and III can cut up unmethylated DNA and type IV deals with foreign methylated DNA[24]. 95% of known bacteria have R-M systems and one third of them have more than one, leading to an even stronger defense system[25]. Concerning R-M in *B. burgdorferi* some work has been done previously. It was determined that upon loss of linear plasmids (lp) 25 and 56, the transformation of shuttle vector is more efficient with limited/no loss of plasmids in the transformants[26], which meant that possibly some gene products on lp25 and lp56 must contribute to the makeup of R-M system in this organism. Data coming from both, transformation experiments and bioinformatic analysis, pointed to *bbeo2* (lp25) and *bbq67* (lp56) genes, probably being responsible for the R-M system in *B. burgdorferi*[26,27]. Some experiments focused on inactivation of *bbeo2* in a clone that was missing lp56, which substantially increased the transformation efficiency with minimum to no loss of other plasmids in the transformants obtained[27]. Other experiments performed showed that shuttle vector DNA that has been CpG methylated *in vitro* overcame the restriction barrier posed by *bbq67* and could thus increase transformation efficiency too[28]. The restriction properties of BBE02 and BBQ67 led to their enzyme designations as *BbuI* and *BbuII* respectively (REBASE communication). As it was determined before in literature[28], a possible sequence methylated by *bbq67* gene is GTA^{met}C, where adenine is methylated at position N6. Another finding was that the TCGA sequence occurs five bases upstream from the GTAC site and might also contribute to the recognition process[10]. In addition, it was shown that *B.burgdorferi* N40 has a RM system homologous to the methyltransferase in *Haemophilus haemolyticus*[28]. This system methylates GC^{met}GC stretches and thus protects it from digestion by *HhaI* and *SfoI*.

1.5 PLASMID ARTIFICIAL MODIFICATION

As it was mentioned before, the R-M system brings considerable difficulty for the introduction of foreign DNA during genetic manipulations. During a transformation in bacteria, restriction enzymes see the introduced DNA as foreign and destroy it immediately. One method to circumvent this problem is called Plasmid Artificial Modification (PAM). The concept was discovered by Nobel Prize winner W. Arber[29], however a big step forward was made by Suzuki and Yasui[30] for bacteria whose genome were already sequenced.

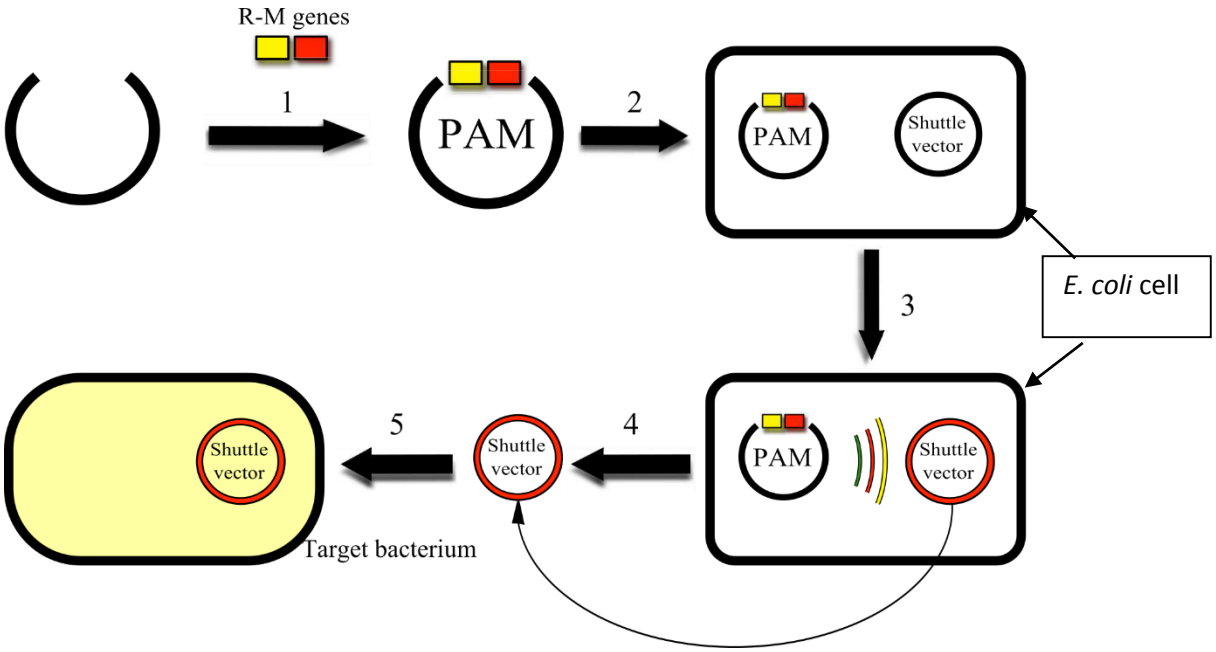

Figure 3. Plasmid Artificial Modification schematic. step 1: Introduction of R-M genes on PAM plasmid; step 2: PAM plasmid is transformed into *E. coli*; step 3: R-M genes activation and methylation of shuttle vector; step 4: Shuttle vector taken out of the cell; 5: Transport of the shuttle vector to the target bacterium

The main point behind PAM process is to methylate the shuttle vector or knock-out construct used for transformation using the methylation enzyme of the target bacteria. If the genome of the bacteria is already sequenced, genes coding for methylation enzymes are identified. Following cloning of these genes into special PAM plasmids, the PAM plasmids can be transformed into an *E. coli* strain that already possesses the vector of interest. By introducing an induction signal present in the medium, the PAM plasmid can be induced to start expressing the methylation genes, thus introducing the methylation pattern on the vector. After that, the methylated vector can be isolated and used for transformation into the target bacteria with a significantly higher efficiency.

1.6 ANTIBIOTIC RESISTANCE IN *B. burgdorferi*

It is essential to study antibiotic resistance in *B. burgdorferi* for various reasons. This topic is not well described, and there are not too many reports about antibiotic resistance in *Borrelia*. The reports that are available often explore different possibilities in which *Borrelia* can develop resistance, underestimating genomic mutations associated with antibiotic resistance[31]. Currently only few antibiotics are available to treat Lyme disease. In addition, research in this field can lead to the discovery of new selectable markers, which will allow for novel genetic experiments[32].

It was discovered that the sequence and secondary structure of 16S rRNA of *B. burgdorferi* resembles that of other species such as *E. coli*, *T. thermophilus* and eukaryotic chloroplast 16S rRNA of *C. reinhardii* and *N. tabacus*[33]. Studies confirmed that mutation at certain sites of 16S rRNA of these species causes resistance to kanamycin, gentamicin, and atreptomycin. Antibiotics that target 16S rRNA either prevent translocation of tRNA during translation or prevent errors in selecting related tRNA. Since 16S rRNA is present only in one copy, mutations in it influence the resistance to antibiotics drastically. In current times when searching for new antibiotics is widespread, it is important to understand how and why the resistance works and the way it develops.

2 MATERIALS AND METHODS

Bacterial Species and Strains

The following strains of bacteria available in the lab were used during this work: *E. coli* ER1821, DH5-α, B. *afzelii* PKo, ACA-I, RU1, RU2, HL5, CB43, *B. garinii* PBr, *B. burgdorferi* N40 Sh, ZS7, and *B. baverensis* PBi.

Number	Name	Sequence
1	EO2 F	5' - ACA TCG TCA AAA CCA ACA ACC CG - 3'
2	EO2 R	5' - TAT GGT AGA AGA TTT TAT TGT TC - 3'
3	Q67 F	5' - CTG ATC CTG AAC TAC ATC TCG - 3'
4	Q67 R	5' - GAT CAG TGA TTT GCC TTT GTT G - 3'
5	T7	5' - TAA TAC GAC TCA CTA TAG GG - 3'
6	SP6	5' - ATT TAG GTG ACA CTA TAG - 3'
7	16S 553F	5’ - TCA AGC CCT GGT AAG GTT CC - 3’
8	16S 659R	5’ - GAG TAT GCT CGC AAG AGT G - 3’
9	16S 1060F	5’ - TCA TCA CTT TGT CAT TTC - 3’
10	16S 993R	5’ - CGT TGT TTC GGG ATT ATT G - 3’
11	rpsE U322F	5’ - AAT AAA GGA CAA AAT AGG G-3’
12	rpsE 441R	5’ - CAA AAC TAA ATC AAA TGC C - 3’
13	16S 1538	5’ - AAA TAA CGA AGA GTT TGA TCC - 3’
14	BB4253	5’ - GGA AGA TGA GAG AAG GGA AG - 3’
15	rpsL 0106F	5' - AAA ATT AAA GTT AGT GAA AAT ATC G - 3’
16	rps 658R	5' - AAT TAT ATC TGG TAT CAA CAA AAA C - 3’
17	4278F	5' - GCA CCC CAG GCT TTA CAC TTT ATG - 3'
18	4278R	5' - CGG GCC TCT TCG CTA TTA CG - 3'

Table 1. Primers used during this work and their names

2.1 PLASMID ARTIFICIAL MODIFICATION

2.1.1 Competent cells preparation

Competent cells of *E. coli* ER1821 strain were prepared using the $CaCl_2$ method. First a tube with 5 mL of LB (lysogeny broth) media was incubated with *E. coli* cells from a glycerol stock and left to grow in shaking incubator overnight. The next day 1 mL of that culture was transferred to 100 mL of fresh LB media and was left to grow to OD600 - 0.2 - 0.4. Then the culture was centrifuged at 5000 rpm for 5 min at 4^{o}C and later re-suspended in 50 mL of 0.1 M ice-cold $MgCl_2$. This mixture was kept on ice for 30

min and later centrifuged at 5000 rpm for 5 min at 4°C. The pellet was re-suspended in 10 mL of ice-cold 0.1 M $CaCl_2$, mixed, left on ice for 30 min and then stored at -80°C for later use.

2.1.2 Gel extraction

Synthesized *bbeo2* and *bbq67* were isolated by gel extraction, following digestion with *HindIII* of their plasmid pUC57 (Synthesized by GenScript –codon optimized for expression in *E. coli*), with Gel Extraction Kit (QIAGEN) and following the protocol provided by manufacturer. The mixture of 2 µL of Buffer, 0.5 µL enzyme, 14.5 µL MilliQ water and 3 µL of plasmid with the insert was kept at 37°C for an hour followed by visualizing the digested product on the gel, cutting of the desired gel size and gel extraction.

2.1.3 Molecular cloning

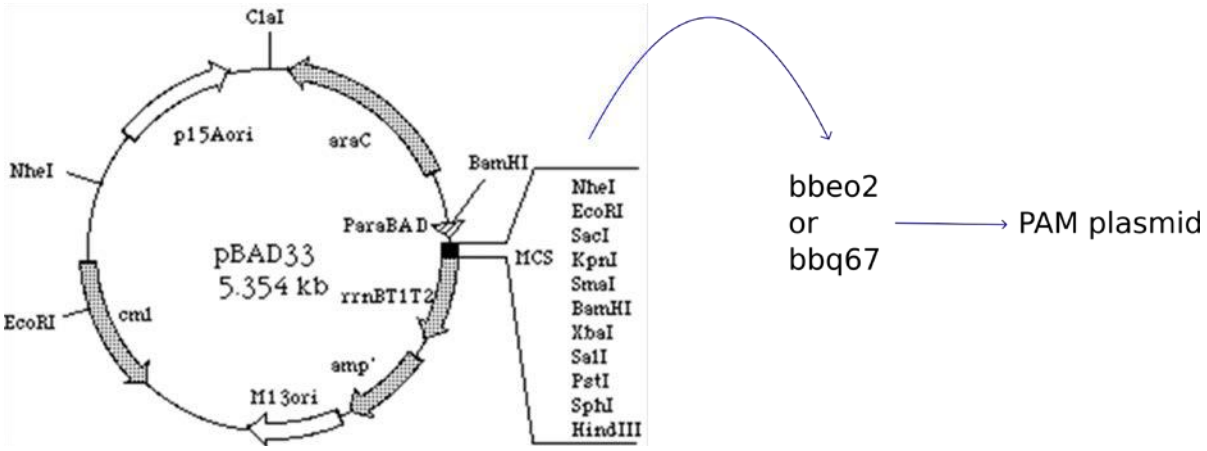

Figure 4. pBAD33 vector map and the basic principle of PAM plasmid construction

The procedure was performed as described below using pBAD33 digested with *HindIII* as a vector and *bbeo2*/*bbq67* as a DNA insert.

Reagents	**Volume**/µL (**Conc.** ng μL^{-1})
2x Rapid Ligation Buffer, T4 DNA Ligase	5
Vector	1 (50)
DNA insert	3(4)
Water	1

Table 2. Ligation reaction set up

For this procedure a slightly modified version of pGEM©-T and pGEM©-T Easy Vector Systems Quick Protocol was used. Vectors were briefly centrifuged and the ligation reaction was set up as described in the table below. The reaction mixture was incubated overnight at 16°C.

50 µL of *E. coli* competent cells were added to the ligation reaction tubes. The mixture was incubated on ice for 30 min. Then it was heat shocked in water bath at 42°C for 40 s and immediately transferred on ice for 2 min. 250 µL of SOC medium was added to the reaction mixture and it was incubated in a shaking incubator for 1 h. Meanwhile plates were pre-warmed to 37°C degrees in an incubator and 40 µL of X-Gal

was added to each plate and spread out. After 15 min 50 µL and 100 µL of *E.coli* cells were plated out. The plates were kept for 5 min turned over and incubated at 37°C overnight. The plates used for this step were containing carbenicillin.

2.1.4 Plasmid DNA isolation

White colonies from the plates incubated previously were inoculated in LB media and were kept overnight in a shaking incubator at 37°C. Plasmid DNA isolation was performed using QIAprep Spin Miniprep Kit by QIAGEN following the protocol provided by the manufacturer.

2.1.5 Screening transformants for desired inserts

Primers	1 µL forward + 1 µL reverse (0.5 µM of each)
Buffer	2 µL (1x)
dNTPs solution	2 µL (200 µM of each)
Template DNA	1 µL (0.8 µg per reaction)
Taq Polymerase	0.1 µL (2.5 units per reaction)
Water	12.9 µL

Table 3. Reaction mixture for PCR reaction. Primers are described in Table 1 and were used according to the insert that was detected; buffer, dNTP solution and Taq Polymerase were purchased from New England Biolabs and were 5× OneTaq Standard Polymerase

Step number	PCR program	Temperature/°C	Time/sec	
1	denaturation	95	30	
2c	denaturation	94	30	30 cycles
3c	annealing	55	45	
4c	elongation	72	60	
5	elongation	72	600	
6	hold	16	∞	

Table 4. PCR program used during a run

Restriction digest and gel electrophoesis. Restriction digest was performed on the isolated plasmid DNA and followed by gel electrophoresis visualisation. Restriction digest was performed according to a modified protocol of New England Biolabs "Optimizing Restriction Endonuclease Reactions". For the restriction digest 3 µL of the isolated plasmid DNA were mixed with 14.5 µL of MilliQ water, 0.5 µL of enzyme (*HindIII*) and 2 µL of 1x NEBuffer 2.1. This mixture was incubated for 1 h at 37°C with no stopping of the reaction afterwards. Agarose gel electrophoresis was performed using 1% agarose gel.

Another method to confirm the presence of the product was PCR screening of the plasmids isolated. The reaction was mixed as described in *Table 3*. Primers used in this experiment are under numbers 1-4 are in *Table 1*. OneTaq® Hot Start DNA Polymerase was used in combination with appropriate buffer provided by manufacturer (5X OneTaq Standard Reaction Buffer) and 10 mM dNTPs. Conditions for PCR reaction are described in *Table 4*.

This procedure was adapted from "Protocol for OneTaq Hot Start DNA Polymerase (M0481)". After the completion of PCR reaction the resulting solutions were run on electrophoresis gel to show the presence of the insert.

2.2 CYTOSINE METHYLATION DETECTION

2.2.1 Dot blot

Two solutions were prepared: blocking buffer (PBS, 5% non-fat dry milk) and Wash buffer (PBS, 0.1% Tween-20). After this the gDNA of different *Borrelia* species/strains (more details in Table 6) was diluted to 200 ng μL^{-1}, 100 ng μL^{-1} and 50 ng μL^{-1}. A drop of 3 µL of each dilution was spotted on a Nitrocellulose BA85 membrane. After this the DNA was fixed by UV cross linking using UV Stratalinker and the membrane was blocked for 1-3 h at room temperature in blocking buffer. Subsequently, the membrane was incubated over night at $4^{o}C$ in purified anti anti 5-methylcytosine antibody (from Zyma research) diluted in blocking buffer (4 µL antibody and 2 mL buffer). On the next day, the membrane was washed 5 times for 15 min with the washing buffer. After this HPR conjugated anti-mouse antibody was diluted in blocking buffer (1:10000) and poured over the membrane, followed by incubation for 1h. Again the washing procedure was repeated 5 times for 15 min. Then a reaction mixture consisting of 2 mL of Reagent A and 2 mL of Reagent B from Pierce ECL Western blotting substrate (an enhanced chemiluminescent substrate for detection of HRP) was added to the membrane and shaken for 5 min. Excess solution was drained and the membrane was exposed using the BioRad Universal Hood III ChemiDoc TM MP Imaging System.

2.2.2 Bisulfite sequencing

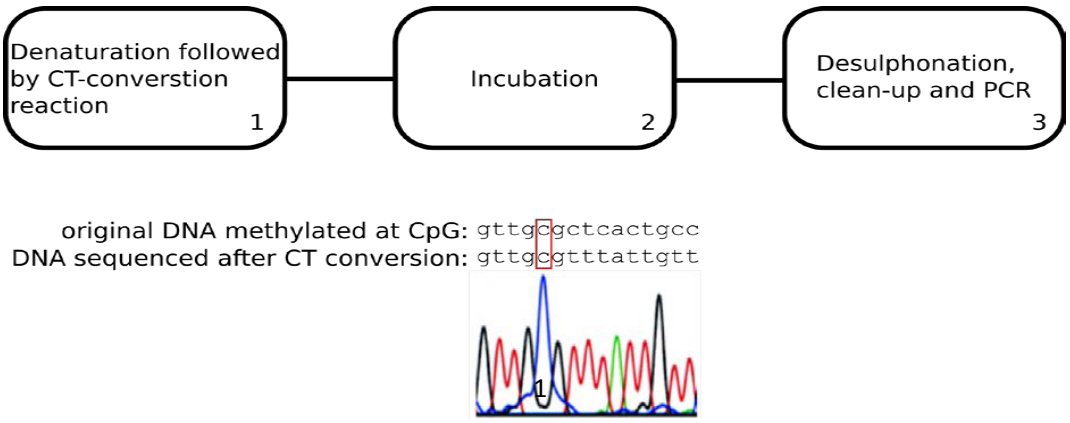

Figure 5. Basic flowchart and principle of bisulfite methylation procedure. Figure adapted from the manufacturer protocol.

In order to determine which cytosines were methylated, the EZ DNA Methylation Gold Kit (Zymo Research) was used. The idea behind bisulfite treatment of DNA is that upon reaction with bisulfite, unmethylated cytosine will turn into uracil whereas methylated cytosine will stay unchanged. Following the treatment and PCR, a sequencing profile is accessed and the cytosines converted into uracils will be detected as thymine. With the help of this kit all the cytosines, which are not protected by methylation will be transformed into uracil. The procedure was performed according to the protocol of the manufacturer. The DNA samples processed in this experiment: N40+pBSV2, pBSV2g (*E.coli*) and pKFSS1(*E.coli*). Sequencing was performed on these DNA samples using 4278 F and R primers (Table 1). The pBSV2g plasmid DNA was in vitro methylated using cytosine methylase from Zymo research following the manufacturer's protocol. This DNA was used as the control in the bisulfite procedure.

2.2.3 *HhaI* digestion

HhaI digestion was performed on a number of plasmid DNA samples (available in the lab) (*Table 5*) in order to determine a presence of specific methylated cytosine that occurs in the sequence $GC^{met}GC$[29]. Reaction mixture was prepared by mixing 2 µL of Buffer, 0.5 µL enzyme, 14.5 µL MilliQ water and 3 µL of DNA sample and keeping the mixture at 37^{o}C for an hour.

Number	Name
1	*B. afzelii* PKo
2	*B. burgdorferi* Sh
3	*B. garinii* PBr
4	*B. afzelii* RU2
5	*B. burgdorferi* N40
6	*B. garinii* PBr ATCC
7	*B. afzeli*i PKo, p94
8	*B. afzelii* HL5
9	*B. afzelii* CB43
10	*B. afzelii* ACA-I
11	*B. baverensis* PBi, P5
12	*B. afzelii* RU1
13	*B. baverensis* PBi, P2
14	*B. burgdorferi* ZS7

Table 5. Strains used in *HhaI* digestion experiment

2.3 CONSTRUCTING PROTEIN 3D STRUCTURES FOR BBEO2 AND BBQ67 AND PHYLOGENETIC ANALYSIS

Nucleotide sequences of *bbeo2* (Unpublished – Rego) and *bbq67* (AE001584.1) were translated to amino acid sequences with the help of Geneious[34]. Then the 3D structure of the amino acids and expected proteins was modeled with the help of Phyre2 web portal for protein modeling[35]. This method uses Hhblits to gather homologous sequences and to construct multiple sequence alignment while PSIPRED is used to predict secondary structure. Following this multiple alignment and the calculated secondary structure is converted to Hidden Markov Models (HMMs) and searched against a database of HMMs. This part is used to generate a rough backbone of the structure, while in the last step the side chains are added in an appropriate manner to avoid steric clashes. Phylogenetic analysis was performed using ATGC South of France Bioinformatics platform PhyML 3.0[35] and visualized with Seaview[36]. The tree was constructed using Maximum Likelihood algorithm, which was used in combination with BLOSUM65. Tree searching operation was performed using best of nearest neighbor interchange (NNI) and sub tree pruning regrafting (SPR). Bootstrapping was performed with 1000 replicas to access significance of the tree branches.

2.4 ANTIBIOTIC RESISTANCE

2.4.1 DNA isolation & sequencing

B. afzelii CB43 was grown in streptomycin environment. The surviving resistant bacteria were used for further studies. In order to prepare DNA for sequencing from resistant clones, it was first isolated from *B. afzelii* CB43, next specific regions of 16s rRNA and *rpsE* (S5) were amplified by PCR using primers 7-16 shown in *Table 1*. More detail on how the primers were grouped and which regions they were targeting are given in *Table 6*. Amplified DNA was separated from the rest by gel extraction, cloned into a pGEM-T Easy vector and transformed into *E. coli.* Subsequently, plasmid DNA was isolated and finally sent for sequencing. Data obtained from sequencing was analyzed using Geneious software[34].

Labeling	Name	Target
a	16S 553F 16S 659R	16s rRNA
b	16S 1060F 16S 993R	16s rRNA
c	*rpsE* U322F *rpsE* 441R	*rpsE* (S5)
d	16S 1538 BB4253	16s rRNA
e	*rpsL* 0106F *rpsL* 658R	*rpsL* (S12)

Table 6. Grouping, names and targets of primers used in the antibiotic resistance experiment. Primer sequences can be found in Table 1

Genomic DNA isolation Procedure was adapted from Promega protocol for Isolating Genomic DNA (Wizard® Genomic DNA Purification Kit). 8 mL of *Borrelia* cultures at mid-log phase (10^7borrelia/ml), grown at 34^{o}C were put into 15 ml falcon tubes. They were then centrifuged at 8000 rpm for 10 min in order to pellet the cells. After this, the supernatant was removed and 600 µL of Nuclei Lysis Solution was added. The pellet was resuspended by pipetting and briefly vortexed. Following this the mixture was incubated at 80^{o}C for 5 min and consequently kept on ice for 5 min. 3 µL of RNase Solution were added to the cell lysate and the tube was mixed 2-5 times.

Then it was incubated for 15 min at 37^{o}C. 200 µL of Protein Precipitation Solution was added to the mixture, which was then mixed lightly by inverting the tube few times. The sample was again incubated for 5 min on ice and centrifuged for 10 min at maximum speed. The supernatant was transferred to a clean 1.5 mL microcentrifuge tube containing 600 µL of room temperature isopropanol. The solution was

gently mixed by inverting until the DNA strands became visible and later centrifuged at maximum speed for 10 min at 10^{o}C. The supernatant was carefully poured away and the tubes were left to drain on an absorbent paper. 600 µL of room temperature 70% ethanol was added and the tubes were gently inverted a few times. They were centrifuged again at maximum speed for 10 min after which they were left open for few minutes to let the ethanol evaporate. 100 µL of DNA Rehydration solution were added to the tube and the DNA was stored at 4^{o}C for further use.

PCR. Parts of the genomic DNA which was assumed to be important based on literature research[31] in the resistance mechanism were accessed by PCR. The reaction mixture is described in *Table 3* and the PCR program was as described in *Table 4*.

Primer sequences (7-16) used in this experiment (*Table 1*) and more details on how they were grouped and what targets they were amplifying for are given in *Table 6*.

In order to isolate amplified DNA, gel extraction was performed, followed by cloning into a plasmid and transforming into competent cells of *E. coli.* Plasmid DNA was isolated and sent for sequencing. Procedure for sequencing was as follows. The template DNA (5 µL) was mixed with primers T7 or SP6 (5 µL) (sequences shown in *Table 1*) and sent for sequencing to Biogen Lightrun sequencing services. Gel extraction, cloning, and plasmid DNA isolation were the same as in Plasmid Artificial Modification section.

3 RESULTS

3.1 PLASMID ARTIFICIAL MODIFICATION

bbeo2 and *bb67* were cut out of their plasmids using *HindIII*, isolated by gel extraction, ligated with pBAD33 which was digested with *HindIII* and cloned into competent cells of *E. coli*. Some colonies with the plasmid containing the insert have been isolated. PCR with primers specific for those genes was performed to check if the plasmids really contained *bbeo2* and *bbq67*. Unfortunately that did not yield satisfactory results, so another measure was implemented, this time sequencing the plasmid DNA. This attempt also proved to yield no results. This procedure was adopted with certain modification to suit the method to our system from previous work using this method[30].

3.2 METHYLATION

During the dot blot experiment specific strains of *B. burgdorferi, B. afzelii, B. garinii, and B. finlandensis,* were tested for presence of cytosine methylation. Concentrations of 200 ng μL^{-1}, 100 ng μL^{-1} and 50 ng μL^{-1} were added onto the membrane. When looking on the dot blot membrane one can see different strains showing different outcomes in the experiment.

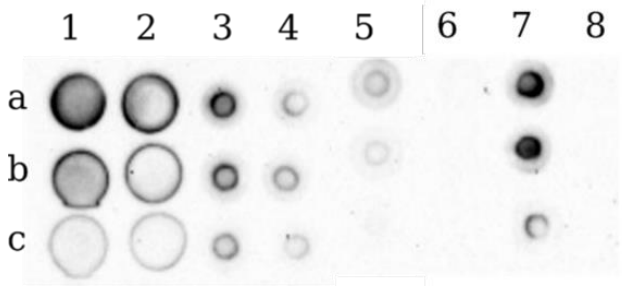

Figure 6. Dot blot assay, where numbers correspond to species mentioned in Table 7 and letter corresponds to various concentrations, where "a" is 200 ng μL^{-1}, "b" - 100 ng μL^{-1} and "c" – 50 ng μL^{-1}. Positive reaction is a presence of a dot and it means DNA has cytosine methylation. Absence of a dot is a negative result.

Positive reaction, or presence of a dot means cytosine methylation in the DNA. Absence of the dot is a negative result. *E. coli* k12 dam- dcm- (sample 8) was used as a negative control and shuttle vector *E. coli* DH5-α pKfss1 (sample 7) as a positive control. As a result of this dot blot experiment it is possible to observe that strains *B. burgdorferi* ZS7, *B. afzelii* PKO have strong cytosine methylation pattern, *B. garinii* PBr, *B. finlandensis* SV1 and *B. afzelii* CB43 show moderate signals whereas strain *B. burgdorferi* A3-S9 (*B. burgdorferi*), which lacks the R-M genes, shows no methylation.

No.	Name
1	*B. burgdorferi* ZS7
2	*B. afzelii* PKo
3	*B. garinii* PBr
4	*B. finlandensis* SV1
5	*B. afzelii* CB43
6	*B. burgdorferi A3-S9*
7	Shuttle vector *E. coli* DH5- α pKfss1
8	*E. coli* k12 dam- dcm-

Table 7. Description of strains used to analyze during a dot blot procedure (Figure 6)

During the bisulfite methylation experiment several types of plasmid DNA were analyzed. Only *pKfss1* yielded the results described in *Figure 7*. On the picture one can see some of the cytosines were methylated versus those that were not suggesting that the protocol used for bisulfite conversion was able to convert the unmethylated cytosines.

Figure 7. Alignment of pKFSS1 sequences (E. coli) obtained using the EZ Methylation Kit. Yellow color cytosines indicate those that did not change after the processing and thus it is assumed to be methylated; red color indicates cytosines which were changed during the experiment, thus assumed to be unprotected or unmethylated

HhaI digestion of various *Borrelia* genospecies genomic DNA (*Figure 8*), identified none of the other *Borrelia* genospecies/strains possessing the methylation enzyme that protects from *HhaI* digestion as the

positive control N40. This confirms previous findings in the literature[37,38,39] that R-M systems present in *B. burgdorferi* present a barrier to digestion by methylation of certain bases and suggests yet another possible site of modification on the DNA which is GCGC or some bases that are adjacent to it.

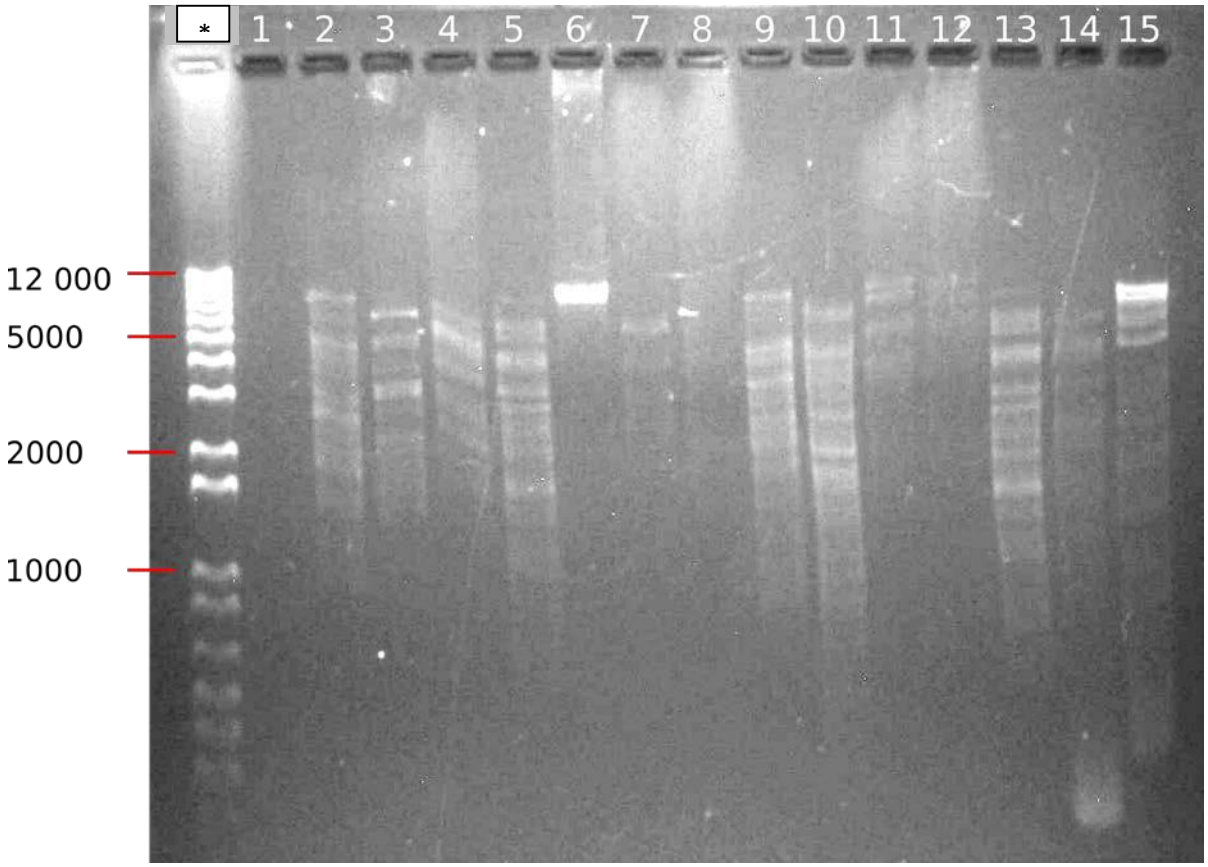

Figure 8. *HhaI* digestion of various *Borrelia* strains: * – 1 kb Plus DNA Ladder by Invitrogen, 1 – blank, 2 - *B. afzelii* PKo, 3 - *B. burgdorferi* Sh, 4 - *B. garinii* PBr, 5 - *B. afzelii* RU2, 6 - *B. burgdorferi* N40, 7 - *B. garinii* PBr ATCC, 8 - *B. afzelii* PKo, p94, 9 – *B. afzelii* HL5, 10 – *B. afzelii* CB43, 11 – *B.afzelii* ACA-I, 12 – *B. bavarensis* PBi, P5, 13 – *B.afzelii* RU1, 14 – *B.baverensis* PBi, P2, 15 – *B. burgdorferi* ZS7

3.3 CONSTRUCTING PROTEIN 3D STRUCTURES FOR *BBEO2* AND BBQ67 AND PHYLOGENETIC ANALYSIS

Nucleotide sequences of *bbeo2* and *bbq67* of *B. burgdorferi* B31 were analyzed with the help of various bioinformatic tools available in order to learn more about the structure and relationship to other representatives of *Borrelia* genus. In *Figure 9*(a) and *Figure 9*(b) one can see the 3D structure of BBEO2 computed by Phyre2 server. The model consists of 30% of the whole sequence which correspond to 384 residues and which was modeled with 100% accuracy.

The structure was modeled using a template with PDB title of "adenine-specific methyltransferase m. taq I/dna complex".

Strain of *Borrelia* for BBEO2	Amino Acid accession number
B. burgdorferi B31	WP_044283683.1
B. burgdorferi 72a	ACM10332
B. burgdorferi N40	ADQ29952
B. burgdorferi 9805	ACO38361
B. bisettii	AEL19468.1
B. bisettii DN127	AEL19469.1
B. garinii Far04	ACL34994
B. afzelii ACA-1	ACJ73320
B. afzelii K78	AJY73058
B. afzelii PKo	AEL70481
B. garinii PBr	ACL34484
B. miyamotoi FR64b	AHH05837
C. jejuni	WP_002860014.1

Table 8. Sequences used for phylogenetic analysis of BBEO2 and their accession number in NCBI database

Strain of *Borrelia* for BBQ67	Amino Acid accession number
B. burgdorferi B31	AAF07736.2
B. valaisiana VS116	ACN52728
B. afzelii	ABH02310.1
B. afzelii PKo	AEL70543
B. afzelii ACA-1	ACJ73483
B. coriaceae Co53	AHH11707
B. hermsii	AAM49818.1
*B. hermsii*_MTW	AHH14727
B. hermsii YOR	AHH04267
B. hermsii YBT	AHH13171.1
B. garinii Far04	ACL35072.1
H. pylori	WP_000910964.1

Table 9. Sequences used for phylogenetic analysis of BBQ67 and their accession number in NCBI database

Concerning the BBQ67 3D model, which can be seen in *Figure 9*(c) and *Figure 9*(d) 572 residues which make up 53% of the sequence were modeled with 100% accuracy. This was modeled based on the template type with PDB title: II G restriction endonuclease *BpuSI*. Both of the calculations confirms previous findings and classifications of these systems in the literature as type II R-M enzymes and their role as N6-adenine methyltrasnferases [10,24] Both BBEO2 and BBQ67 show similar secondary structure (Figure 10). Especially noticeable are the two stretches of β-sheets.

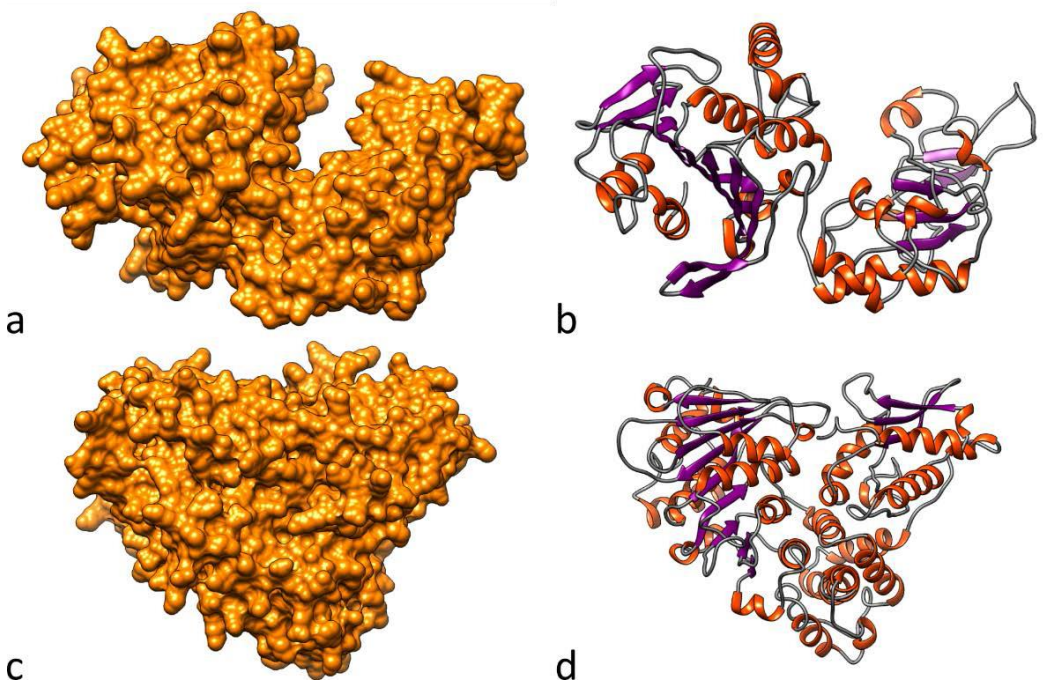

Figure 9. Protein 3D structure structures using Phyre2 server. a - surface structure of BBEO2 as computed by Chimera; b - schematic representation of secondary structure of BBEO2; c - surface structure of BBQ67 computed in Chimera; d - secondary structure schematic of BBQ67

The phylogenetic trees were constructed to show the relatedness of the two genes *bbe02* and *bbq67* within the *Borrelia* genus. *B. burgdorferi* BBQ67 is more closely related to similar restriction modification proteins in *B. garinii* and *B. valaisiana* than it is to *B. afzelii*. *B. hermsii* and *B. coriaceae* are more distantly related to *B. bu*rgdorf*eri (Figure 10).*

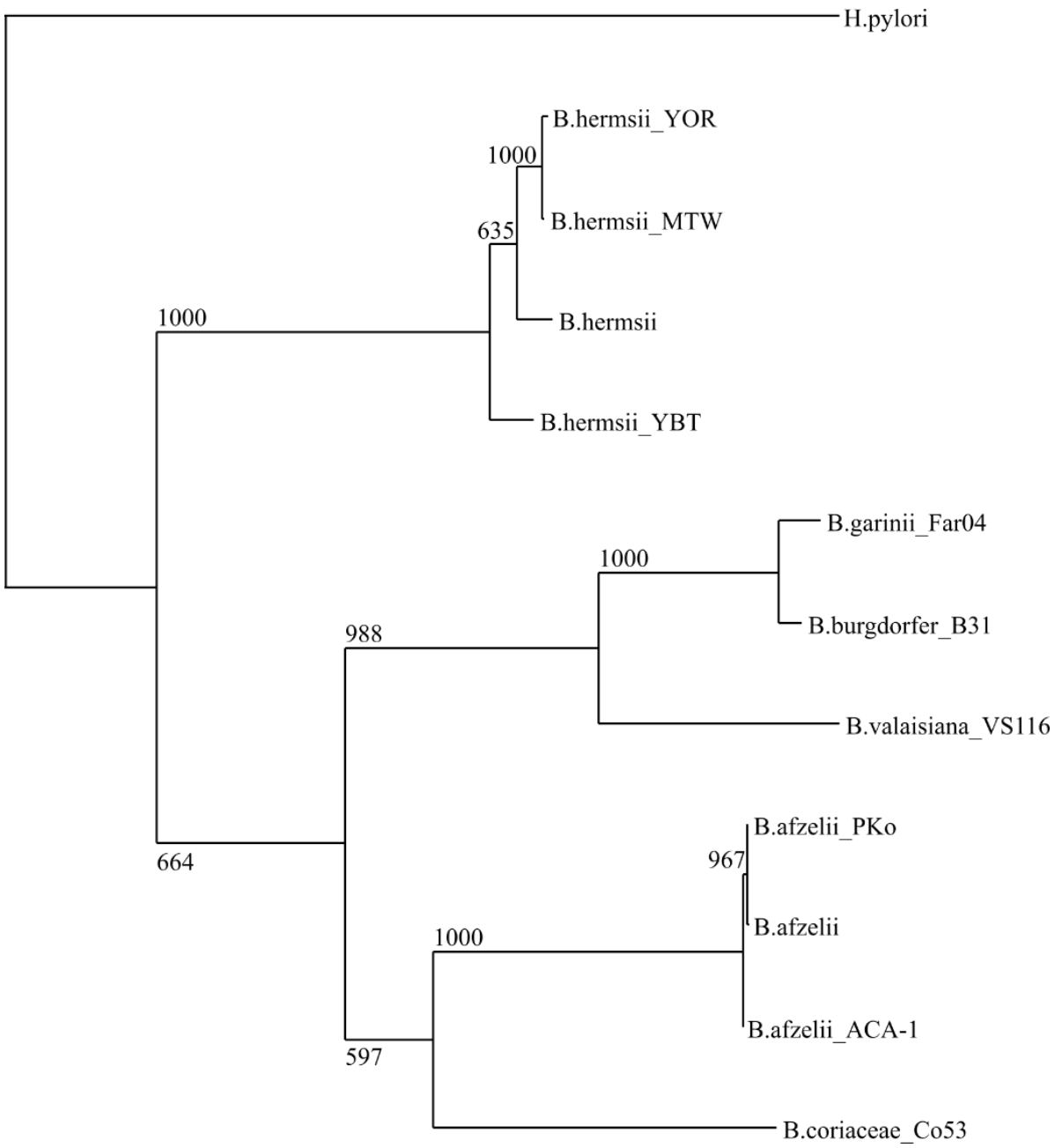

Figure 10. Maximum Likelihood tree for BBQ67

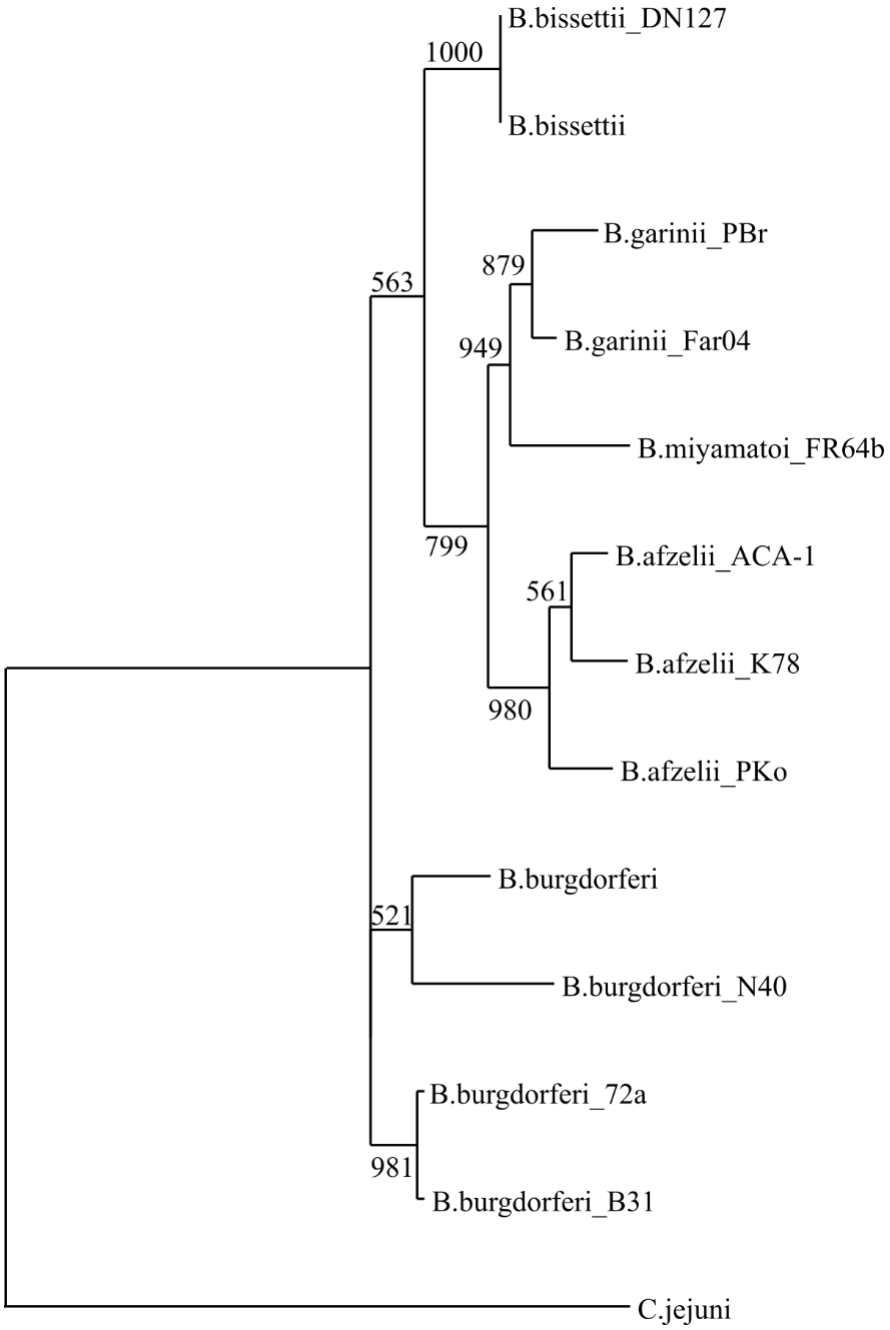

Figure 11. Maximum Likelihood tree for BBEO2

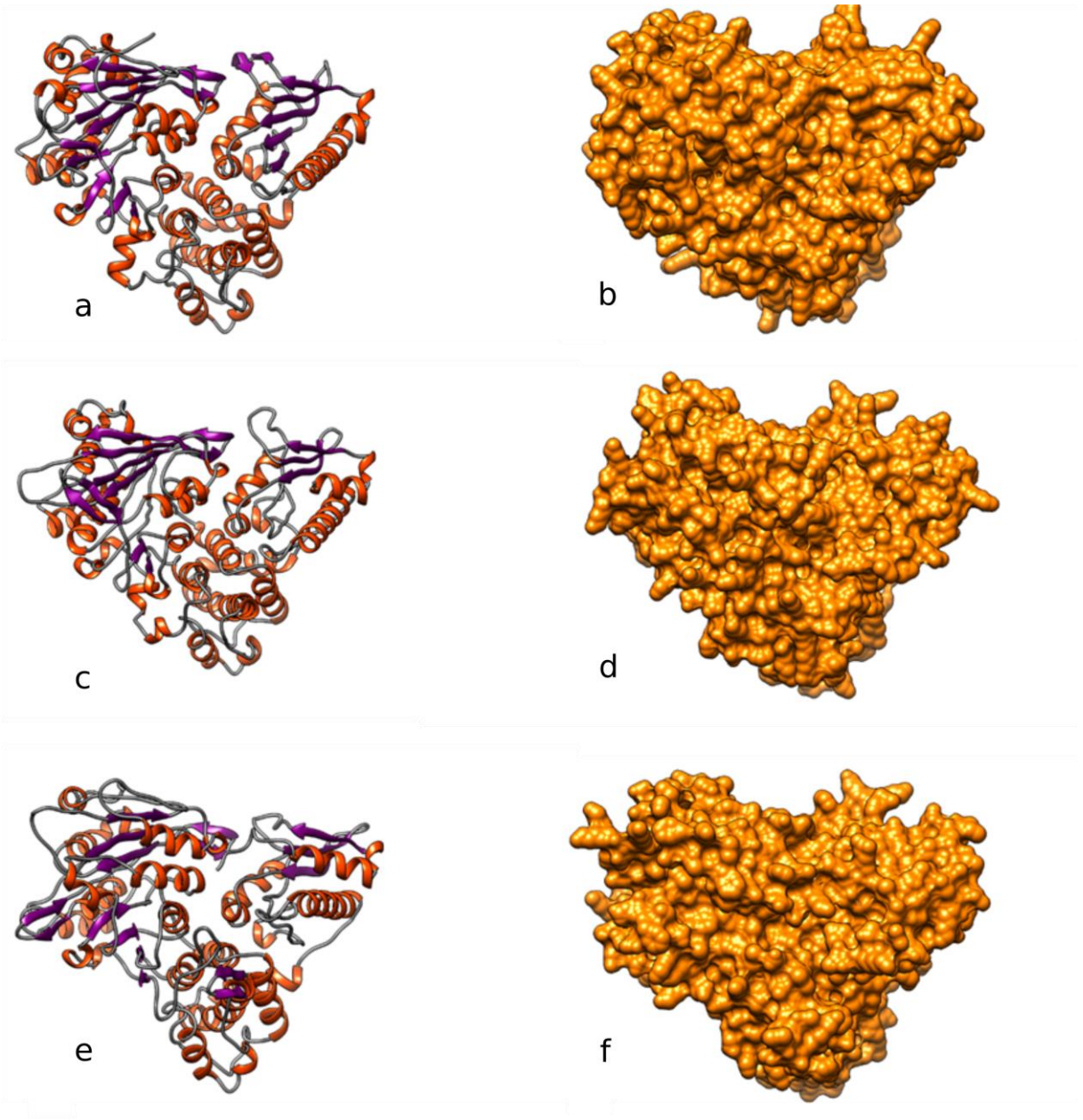

Figure 12. Comparison of the 3D stuctures of BBQ67 homologues representatives; a,b – *B. afzelii*; c,d – *B. hermsii*; e,f – *B. valaisiana*

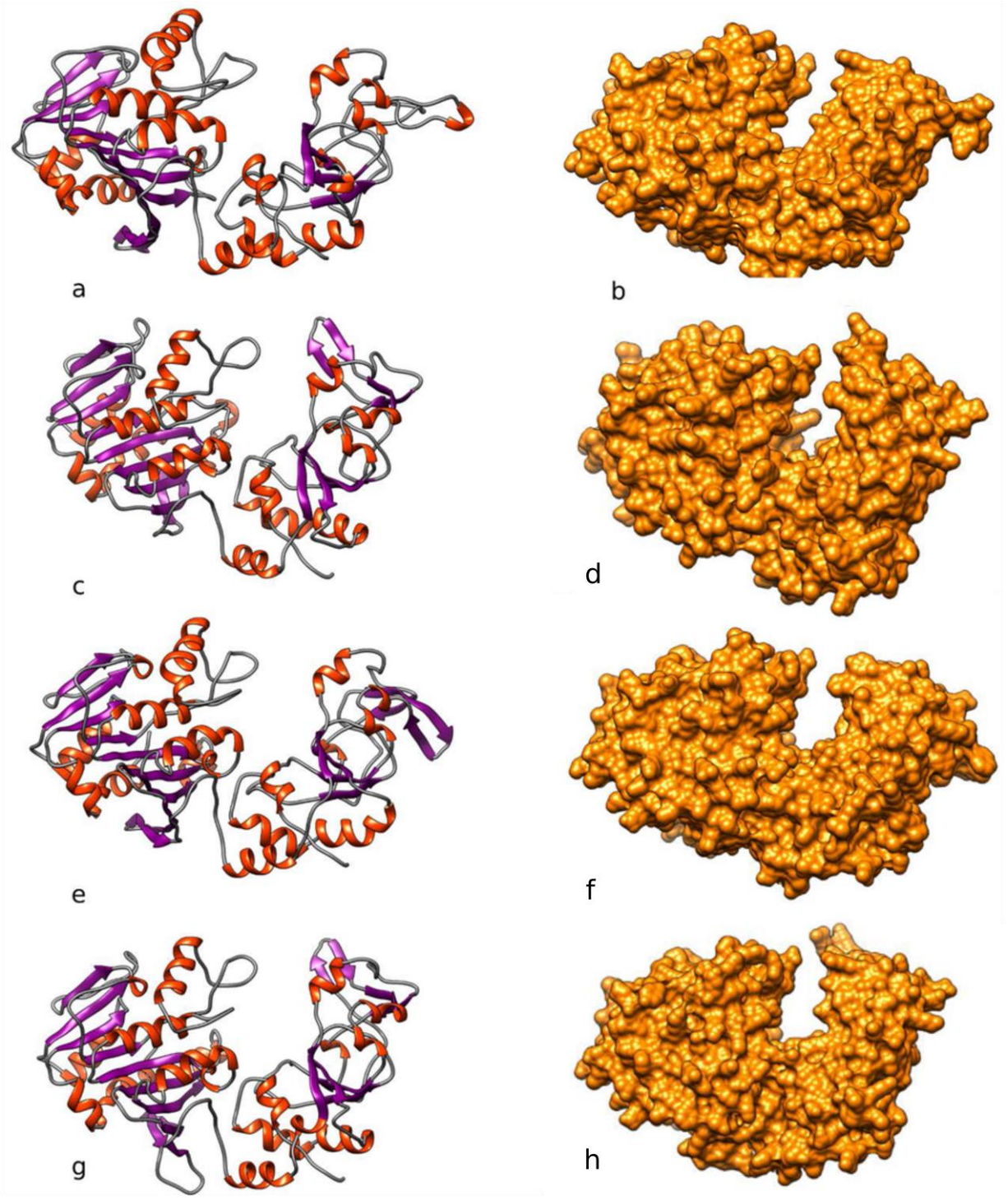

Figure 13. Comparison of the 3D structures of BBEO2 homologue representatives: a,b –*B. bisetiii*; c,d – *B. garinii*; e,f – *B. miyamatoii*; g,h – *B. spielmanii*

3.4 ANTIBIOTIC RESISTANCE

DNA of *B. afzelii* CB43 was isolated and sequenced using primers described in *Table 6*. Out of 5 primer sets satisfactory results were obtained only from primer set c, which was used as a reference sequence when comparing to antibiotic resistant mutants described further in this section.

As a next step *B. afzelii* CB43 was grown in a media which contained streptomycin. Genomic DNA of surviving borrelial clones obtained by plating, was isolated and sequenced in order to compare to the DNA of previously sequenced parental strain. Comparison was made using multiple alignment in Geneious[34]. Aligned sequences were trimmed to have only the homologous parts and the mismatched parts were accessed. Figure 6 shows the results, where resistant strains were compared to the parental strain which is not resistant to the streptomycin. It was discovered that only at 6 positions there are mismatches detected and only 3 of the mismatched sites show a valuable substitution of the nucleotides (position 180, 524, 753).

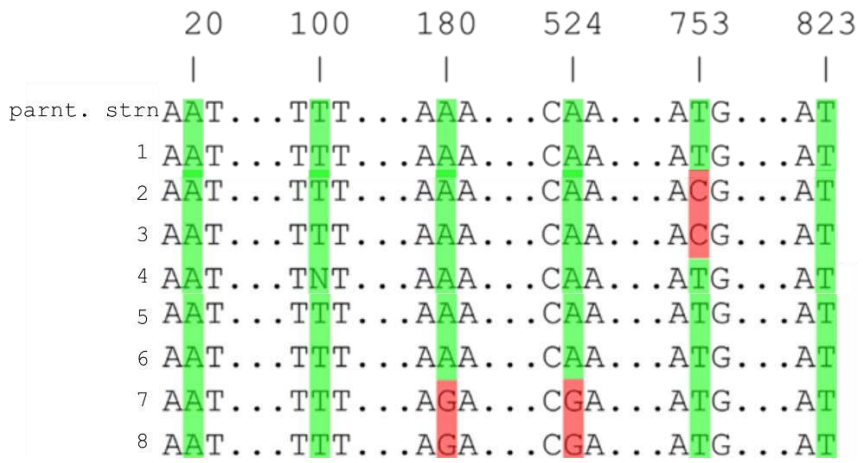

Figure 14. Segments of the multiple alignment of *rpsE* gene, showing the parts where strains resistant to streptomycin showed a mismatch to a nucleotide sequence of parental strain. "Parnt. strn" stands for parental strand and number from 1 to 8 represent samples which are resistant to streptomycin

4 DISCUSSION

During this project various techniques were adapted to understand restriction-modification in *Borrelia* species as well as looking at possible reasons for antibiotic resistance in *B. afzelii*.

Restriction-modification is a very important "defence"system in a bacteria, which among other functions protects bacteria from foreign DNA influences. It works by introducing methyl groups on nucleic bases in certain DNA code stretches employing endonucleases.

Borrelia are recalcitrant to genetic transformations and this has been shown to be due to the presence of R-M system[26]. Deletion or absence of the genes *bbeo2* and *bbq67* in *B. burgdorferi*, which are known to be encoding for proteins that both have endonuclease and methyltransferase properties, makes *B. burgdorferi* easier to transform[27]. Using PAM to overcome the R-M system appears to be an ideal method for success in genetic transformations[30]. Its success in other such recalcitrant bacteria[40] suggests that this method would be useful with *B. burgdorferi*. Although the attempts at cloning the codon optimized *bbeo2* and *bbq67* in the pBAD33 vector were unsuccessful, it clearly is a problem of the cloning method which needs some more adjustments. Once successful, the possibilities of knocking out the genes or complementing could be a lot easier in *Borrelia*.

Chen and colleagues showed that methylating a shuttle vector in vitro with cytosine methyltransferase alleviated the restriction on transformations into infectious strains[28]. There is also evidence of more than one cytosine methylation system within different *Borrelia* species[41]. With most European strains in this project and the variety of *Borrelia* genospecies, there does not appear to be the $GC^{met}GC$ modification as seen with *B. burgdorferi* strain N40. It suggests that the N40 along with other strains may have incorporated this R-M gene into its genome at a more recent evolutionary time point.

The use of the dot blot method to identify the type of methylation is a technique that is rarely used in bacteria[23]. This is the first time that this technique has been used to look at cytosine methylation in *Borrelia* using an anti-cytosine antibody. The result suggests that this method could be used in screening other *Borrelia* genospecies as well as in other bacteria to investigate cytosine methylation.

The bisulfite sequencing method for determination of cytosine methylation was used to look closer at the pKFSS1 shuttle vector derived from various strains of *Borrelia* and *E. coli.* Although this technique is now at the forefront of trying to understand epigenomics in bacteria and provide the required methylation information[42], this method would need to better adapt for work with *Borrelia*.

R-M genes *bbeo2* and *bbq67* were sequenced and their nucleotide sequences were translated into amino acid sequences. These were analyzed and tertiary structures were computed. The analysis shows secondary structure similarities between *bbeo2* and *bbq67* amino acid products, especially the location of the ß-sheets, even though predicted tertiary structure varies. Also the translated amino acid sequences were used to do phylogenetic analysis which showed the relationship of different *Borrelia* strains in terms of their R-M systems. There does appear to be sequence variation within *bbeo2*, showing a diversification in lineage into the *B. burgdorferi* sensu stricto clade and the other Borrelia genospeceis clade which again appears to diverge based on the species. All the *bbq67* homologues of the Lyme borreliosis causing spirochetes appeared to be similar and form a clade separate from the relapsing fever. This suggests an ancestral division of this gene, probably at the time when these two *Borrelia* types were diverging.

Trying to understand the rise of spontaneous antibiotic resistance is of importance when studying all pathogenic bacteria. Surprisingly, it has not been looked at in any detail in *Borrelia*, given the increasing epidemic that is Lyme disease and the use of antibiotics to treat the disease. The work yielded no successful amplification of products for most of the primer sets used when looking at antibiotic resistance in *Borrelia*[32]. However, one of the sequenced regions was practically accessible by amplification using PCR. This region did not show any significant variation from the antibiotic susceptible parental strain, but more work might help provide a better insight into how such antibiotic-resistant mutants arose.

Most of the techniques had to be adapted as they had not been used in studies that concerned *Borrelia*. Future modification in these methods will probably help to improve them. The goal of getting a PAM vector expressing BBEO2 or BBQ67 is essential for easing restriction barrier during genetic manipulations in *Borrelia*. Overall, this work adds another layer of knowledge on the R-M system in various *Borrelia* species.

5 REFERENCES

1. Gray, J., Kahl, O., Lane, R. S. & Stanek, G. *Lyme borreliosis: biology, epidemiology, and control*. (CABI, 2002).

2. Rudenko N., Golovchenko M., Grubhoffer L., O. J. Updates on Borrelia burgdorferi sensi lato complex with respect to public health. **18,** 1199–1216 (2013).

3. Tilly, K., Rosa, P. A. & Stewart, P. E. {B}iology of infection with {B}orrelia burgdorferi. *Infect. Dis. Clin. North Am.* **22,** 217–234 (2008).

4. Schüler, W. *et al.* Complete Genome Sequence of Borrelia afzelii K78 and Comparative Genome Analysis. *PLoS One* **10,** e0120548 (2015).

5. Sykes, R. A. An Estimate of Lyme Borreliosis Incidence in Western Europe. *Res Medica* **22,** 76 (2014).

6. Steere, A. C. Lyme Disease. *N. Engl. J. Med.* **345,** 115–125 (2001).

7. Wormser, G. P. *et al.* Practice guidelines for the treatment of Lyme disease. *Clin. Infect. Dis.* **31,** S1–S14 (2000).

8. Paster, B. J., Stackebrandt, E., Hespell, R. B., Hahn, C. M. & Woese, C. R. The phylogeny of the spirochetes. *Syst. Appl. Microbiol.* **5,** 337–351 (1984).

9. Ferdows, M. S. & Barbour, A. G. Megabase-sized linear DNA in the bacterium Borrelia burgdorferi, the Lyme disease agent. *Proc. Natl. Acad. Sci.* **86,** 5969–5973 (1989).

10. Rego, R. O. M., Bestor, A. & Rosa, P. a. Defining the plasmid-borne restriction-modification systems of the Lyme disease spirochete Borrelia burgdorferi. *J. Bacteriol.* **193,** 1161–71 (2011).

11. Strother, K. O. & de Silva, A. Role of Borrelia burgdorferi linear plasmid 25 in infection of Ixodes scapularis ticks. *J. Bacteriol.* **187,** 5776–5781 (2005).

12. Rosa, P. A., Cabello, F. C. & Samuels, D. S. Genetic manipulation of Borrelia burgdorferi. *Mol. Biol. Host Interact. Pathog.* 189–220 (2010).

13. Tilly, K. *et al.* DNA exchange and insertional inactivation in spirochetes. *J. Mol. Microbiol. Biotechnol.* **2,** 433–442 (2000).

14. Jeltsch, A. Beyond Watson and Crick: DNA methylation and molecular enzymology of DNA methyltransferases. *Chembiochem* **3,** 274–293 (2002).

15. Fish, F., Navon, Y. & Goldman, S. Hydrophobic adherence and phase variation in Bordetella pertussis. *Med. Microbiol. Immunol.* **176,** 37–46 (1987).

16. Hale, W. B., Van Der Woude, M. W., Braaten, B. A. & Low, D. A. Regulation of Uropathogenic Escherichia coli Adhesin Expression by DNA Methylation. *Mol. Genet. Metab.* **65,** 191–196 (1998).

17. Zhang, G. *et al.* A mimicking-of-DNA-methylation-patterns pipeline for overcoming the restriction barrier of bacteria. *PLoS Genet.* **8,** e1002987 (2012).

18. Jeltsch, A. Maintenance of species identity and controlling speciation of bacteria: a new function for restriction/modification systems? *Gene* **317,** 13–16 (2003).

19. Price, C. & Bickle, T. A. A possible role for DNA restriction in bacterial evolution. *Microbiol. Sci.* **3,** 296–299 (1986).

20. Naito, T., Kusano, K. & Kobayashi, I. Selfish behavior of restriction-modification systems. *Science (80-.).* **267,** 897–899 (1995).

21. Low, D. A. & Casadesús, J. Clocks and switches: bacterial gene regulation by DNA adenine methylation. *Curr. Opin. Microbiol.* **11,** 106–112 (2008).

22. Labrie, S. J., Samson, J. E. & Moineau, S. Bacteriophage resistance mechanisms. *Nat. Rev. Microbiol.* **8,** 317–327 (2010).

23. Humbert, O. & Salama, N. R. The Helicobacter pylori HpyAXII restriction--modification system limits exogenous DNA uptake by targeting GTAC sites but shows asymmetric conservation of the DNA methyltransferase and restriction endonuclease components. *Nucleic Acids Res.* **36,** 6893–6906 (2008).

24. Roberts, R. J. *et al.* A nomenclature for restriction enzymes, DNA methyltransferases, homing endonucleases and their genes. *Nucleic Acids Res.* **31,** 1805–1812 (2003).

25. Roberts, R. J., Vincze, T., Posfai, J. & Macelis, D. REBASE—a database for DNA restriction and modification: enzymes, genes and genomes. *Nucleic Acids Res.* gkp874 (2009).

26. Lawrenz, M. B., Kawabata, H., Purser, J. E. & Norris, S. J. Decreased electroporation efficiency in Borrelia burgdorferi containing linear plasmids lp25 and lp56: impact on transformation of infectious B. burgdorferi. *Infect. Immun.* **70,** 4798–4804 (2002).

27. Kawabata, H., Norris, S. J. & Watanabe, H. BBE02 disruption mutants of Borrelia burgdorferi B31 have a highly transformable, infectious phenotype. *Infect. Immun.* **72,** 7147–7154 (2004).

28. Chen, Q. *et al.* In vitro CpG methylation increases the transformation efficiency of Borrelia burgdorferi strains harboring the endogenous linear plasmid lp56. *J. Bacteriol.* **190,** 7885–7891 (2008).

29. Arber, W. & Linn, S. DNA modification and restriction. *Annu. Rev. Biochem.* **38,** 467–500 (1969).

30. Suzuki, T. & Yasui, K. in *Strain Engineering* 309–326 (Springer, 2011).

31. Criswell, D., Tobiason, V. L., Lodmell, J. S. & Samuels, D. S. Mutations conferring aminoglycoside and spectinomycin resistance in Borrelia burgdorferi. *Antimicrob. Agents Chemother.* **50,** 445–452 (2006).

32. Drecktrah, D., Douglas, J. M. & Samuels, D. S. Use of rpsL as a Counterselectable Marker in Borrelia burgdorferi. *Appl. Environ. Microbiol.* **76,** 985–7 (2010).

33. Cannone, J. J. *et al.* The comparative RNA web (CRW) site: an online database of comparative sequence and structure information for ribosomal, intron, and other RNAs. *BMC Bioinformatics* **3,** 2 (2002).

34. Drummond, A. J. *et al.* Geneious, version 5.4. *Geneious, Auckland, New Zeal.* (2011).

35. Guindon, S. *et al.* New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0. *Syst. Biol.* **59,** 307–321 (2010).

36. Gouy, M., Guindon, S. & Gascuel, O. SeaView version 4: a multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Mol. Biol. Evol.* **27,** 221–224 (2010).

37. Rego, R. O. M., Bestor, A. & Rosa, P. A. Defining the plasmid-borne restriction-modification systems of the Lyme disease spirochete Borrelia burgdorferi. *J. Bacteriol.* **193,** 1161–1171 (2011).

38. Casjens, S. *et al.* A bacterial genome in flux: The twelve linear and nine circular extrachromosomal DNAs in an infectious isolate of the Lyme disease spirochete Borrelia burgdorferi. *Mol. Microbiol.* **35,** 490–516 (2000).

39. Fraser, C. M. *et al.* Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi. *Nature* **390,** 580–586 (1997).

40. Jones, M. J., Donegan, N. P., Mikheyeva, I. V & Cheung, A. L. Improving Transformation of Staphylococcus aureus Belonging to the CC1 , CC5 and CC8 Clonal Complexes. 1–14 (2015). doi:10.1371/journal.pone.0119487

41. Casjens, S. R. *et al.* Genome stability of lyme disease spirochetes: Comparative genomics of borrelia burgdorferi plasmids. *PLoS One* **7,** (2012).

42. Huo, W., Adams, H. M., Zhang, M. Q. & Palmer, K. L. Genome modification in Enterococcus faecalis OG1RF assessed by bisulfite sequencing and single-molecule real-time sequencing. *J. Bacteriol.* **197,** 1939–1951 (2015).