A *Borrelia burgdorferi* mini-*vls* system that undergoes antigenic switching in mice: investigation of the role of plasmid topology and the long inverted repeat

Mildred Castellanos[1], Theodore B. Verhey[1], George Chaconas[1,2,*]

[1]Department of Biochemistry and Molecular Biology, Snyder Institute for Chronic Diseases, Cumming School of Medicine, University of Calgary, Alberta, Canada

[2]Department of Microbiology, Immunology and Infectious Diseases, Snyder Institute for Chronic Diseases, Cumming School of Medicine, University of Calgary, Alberta, Canada

* Corresponding author

E-mail: chaconas@ucalgary.ca (GC)

Running Title: A mini-*vls* system for antigenic variation

Key words: antigenic variation; *Borrelia burgdorferi;* Lyme disease*; vlsE*

Summary

Borrelia burgdorferi evades the host immune system by switching the surface antigen VlsE, in a process known as antigenic variation. The DNA mechanisms and genetic elements present on the *vls* locus that participate in the switching process remain to be elucidated. Manipulating the *vls* locus has been difficult due to its instability on *E. coli* plasmids. In this study, we generated for the first time a mini-*vls* system composed of a single silent *vlsE* variable region (silent cassette 2) through the *vlsE* gene by performing some cloning steps directly in a highly transformable *B. burgdorferi* strain. Variants of the mini system were constructed with or without the long inverted repeat (IR) located upstream of *vlsE* and on both circular and linear plasmids to investigate the importance of the IR and plasmid topology on recombinational switching at *vlsE*. Amplicon sequencing using PacBio long read technology and analysis of the data with our recently reported pipeline and VAST software showed that the system undergoes switching in mice in both linear and circular versions and that the presence of the hairpin does not seem to be crucial in the linear version, however it is required when the topology is circular.

Introduction

Borrelia burgdorferi is one of the causative agents of Lyme borreliosis (Steere *et al.*, 2016), the most commonly reported tick-borne disease in North America (Stanek *et al.*, 2012) and temperate Eurasia. *B. burgdorferi* can evade the acquired immune system of the mammalian host through continuous variation of VlsE, a surface-bound lipoprotein, in a process known as antigenic variation (Zhang *et al.*, 1997, Norris, 2014). In B31, the prototype *B. burgdorferi* strain, *vlsE* is found on the right end of the linear plasmid lp28-1 where there is also a collection of 15 contiguous silent cassettes (or pseudogenes) in the opposite orientation that are highly similar to the variable region of *vlsE* (**Fig. 1**, top,

see (Norris, 2014). Antigenic variation in *B. burgdorferi* occurs through the unidirectional transfer of information from the cassettes to the *vlsE* gene in a combinatorial manner that generates a huge number of possible VlsE variants (Zhang & Norris, 1998a, Zhang & Norris, 1998b). *vlsE* switching does not occur *in vitro* or in the tick (Indest *et al.*, 2001), however it has been shown that in mice, sequence variation starts as early as 4 days post- infection (Zhang & Norris, 1998b). Spirochetes lacking the *vlsE* locus are infectious, but they are cleared after two to three weeks and cannot mount a persistent infection in immunocompetent mice; however they can cause long-term infections in SCID mice (Bankhead & Chaconas, 2007, Rogovskyy *et al.*, 2015, Rogovskyy & Bankhead, 2013, Magunda & Bankhead, 2016). Although recombinational switching seems to be a process that depends upon sequence homology (Verhey *et al.*, 2018b), it does not require the key protein for homologous recombination, RecA (Liveris *et al.*, 2008, Dresser *et al.*, 2009) or a wide variety of other DNA replication/recombination/repair proteins (Dresser *et al.*, 2009, Lin *et al.*, 2009) that are involved in antigenic variation at *pilE* in *Neisseria gonorrhoeae* (Vink *et al.*, 2012). The only required protein identified to date (also involved in recombination at *pilE)* is the branch migrase RuvAB (Dresser *et al.*, 2009, Lin *et al.*, 2009) indicating that there is at least the formation of Holliday junctions or other branched structures that can be migrated (Tsaneva *et al.*, 1992, Lloyd & Sharples, 1993).

In addition to the paucity of information on the proteins involved, little is known about the actual recombinational mechanism or the genetic elements required for switching to occur. The *vls* locus in B31 is characterized by certain distinct genetic elements. Among these are a near-perfect 100 bp inverted repeat (IR) upstream of *vlsE*

that contains the -35 box of the *vlsE* promoter (Hudson *et al.*, 2001). Long IRs of similar size have been found in the same position in other *B. burgdorferi* strains: PAbe plasmid lp28-1, Accession CP019923; 297 lp28-1, Accession AB041949; JD1 lp28-1, and BL206 lp28-1, (unpublished results) and even other *Borrelia* species (*B. garinii* lp28-1, unpublished results) suggesting a possible role in the mechanism of antigenic variation. However, IRs have not been reported in other antigenic variation systems (Vink *et al.*, 2012, Li, 2015).

The segmented *B. burgdorferi* genome has a number of linear replicons terminated by covalently closed hairpin ends (Chaconas & Kobryn, 2010) with similar but non-identical sequences (Tourand *et al.*, 2009). Another important feature is that the *vls* locus has always been found at the end of linear plasmids or the linear chromosome (Kingry *et al.*, 2016, Casjens *et al.*, 2012, Casjens *et al.*, 2011b, Casjens *et al.*, 2011a, Schüler *et al.*, 2015), and hence DNA topology may be important for switching at *vlsE*. In contrast other bacterial antigenic variation systems are present on circular molecules (Vink *et al.*, 2012). However, protozoan antigenic variation systems are typically near telomeric regions and proteins involved in telomere binding may participate in the switching process or its regulation (Li, 2015).

The *vls* locus possesses a repetitive nature represented not only by the high similarity among the cassettes and *vlsE* but also by the 17 bp direct repeats flanking the cassettes and the variable region of *vlsE* (Zhang *et al.*, 1997) as well as the long IR. The 17 bp DRs are not conserved in different strains and species and not always preserved within the silent cassettes in a given strain (Wang *et al.*, 2003) . The *vls* loci of all *B. burgdorferi* strains analyzed also contain G-runs on the coding strand, with the

possibility of a role for G-quadruplexes in the switching reaction (Walia & Chaconas, 2013). A G-quartet has been shown to play a role in recombinatonal switching at the *pilE* locus in *Neisseria* species (Obergfell & Seifert, 2015). The presence of some of these elements and perhaps other yet to be identified DNA have limited cloning of the *vls* locus on plasmids in *E. coli. G*enetic manipulation and analysis of the *vls* locus has therefore not been possible (Norris, 2014), with the exception of deletion of the entire locus from lp28-1 (Bankhead & Chaconas, 2007).

In the present work we describe the successful construction and manipulation of a mini-*vls* system using sequential assembly in an *E. coli-B. burgdorferi* shuttle vector, first using *E. coli* followed by direct cloning of otherwise unstable ligation products directly into *B. burgdorferi*. Spirochetes harboring the mini-*vls* system were used to infect mice and recovered at various times after infection for amplicon sequencing using PacBio long read technology coupled with our previously reported pipeline and VAST software (Verhey *et al*., 2018a, Verhey *et al*., 2018b). We report here the construction and manipulation of a functional mini-*vls* system and an investigation into the possible role of the long IR and the plasmid topology of the *vls* locus.

Results

A genetic system for construction of a mini-*vls* plasmid

The ability to construct a *vls* locus that allows modification and/or exchange of the elements present (e.g. the cassettes, the intergenic region, the 100 bp inverted repeat, the 17 bp direct repeats, presence/absence of the telomere, etc. **Fig. 1**) is crucial to understanding their roles in recombinational switching at *vlsE*. However, the

manipulation of some of these elements has been problematic, as noted earlier. To study the DNA elements in the *vls* locus our approach was to simplify the system by generating a mini-*vls* locus that contained *vlsE,* the naturally occurring intergenic region and a single silent cassette (cassette 2). We were able to accomplish this by performing the final cloning steps directly in a highly competent strain of *Borrelia burgdorferi*, high passage HB19 (Coburn *et al.*, 1993) where we were able to clone *vlsE* with the 100 bp inverted repeat, which was unclonable in *E. coli.*

The initial construction steps were performed in *E. coli* DH5-alpha (see **Experimental procedures**) where we first added the MCS, the gentamicin resistance gene and the pUC origin (**Fig. 2**, number 1) from the shuttle vector pBSV2G (Elias *et al.*, 2003) to the replication region of the linear plasmid lp5 (**Fig. 2**, number 2), to generate pBSV5. We chose the lp5 replication region since the *vls* locus in *Borrelia burgdorferi* is always present on a linear plasmid, and because lp5 has the simplest replication region of the linear plasmids, and requires only one plasmid-encoded protein for replication, *bbt04* (Casjens *et al.*, 2000). After generating pBSV5, we added *pncA,* an essential lp25-encoded nicotinamidase (Purser *et al.*, 2003) (**Fig. 2,** number 3). This gene provided two advantages. The first was that it allowed the use of a recipient strain lacking lp25, where transformation rates are greatly enhanced (Lawrenz *et al.*, 2002). The second was that since *pncA* is essential for infectivity in mice (Purser *et al.*, 2003), it provided a selectable marker for maintenance of the mini-*vls* plasmid during mouse infection. The resulting plasmid pMC73 was circular so we added a replicated telomere (**Fig. 2,** number 4) from lp28-1 (Tourand *et al.*, 2009), which results in conversion into

the linear form by telomere resolution once introduced into *B. burgdorferi* (Chaconas *et al.*, 2001, Kobryn & Chaconas, 2002).

The next step was to add silent cassettes to the construct. However, due to the high degree of sequence identity among the cassettes, it was not possible to obtain a clean amplification of cassettes 2 and 3 together. Finally, after several trials of redesign and/or primer combinations, oligonucleotides B2782 and B2797 (**Table S1**) were used to amplify cassette 2 alone, which was inserted into the construct (**Fig. 2**, number 5). The last step to complete the mini-*vls* construct was to clone *vlsE* along with the 100 bp perfect inverted repeat (IR) in its promoter region (**Fig. 2**, number 6). As expected, this was unsuccessful after several attempts in *E. coli*, so we decided to clone this region directly into *B. burgdorferi* high-passage strain HB19. Previous studies in our lab have shown that the transformation efficiencies of this strain are about 5,440 transformants per µg of DNA (unpublished results). We typically recovered about 10 transformants in HB19 from ligations containing approximately 250 ng of DNA.

Generation of various mini-*vls* constructs

Using the methodology outlined above and in **Fig. 2**, we generated four mini-*vls* constructs shown in **Fig. 3** as well as four control plasmids. The experimental plasmids included both a circular and a linear form of the mini-*vls* locus shown in **Fig. 1** containing cassette 2, the intergenic region with the long IR, and *vlsE.* We also constructed the two equivalent mini-*vls* constructs lacking the 100 bp perfect IR in the *vlsE* promoter. In addition, four control constructs (**Fig 3**, top) were assembled: a circular and a linear plasmid that lacked silent cassette 2 and a circular and linear

plasmid that lacked *vlsE.* The specific usage of each control construct is noted with the experiments they were used for.

Once we obtained the final constructs in *B. burgdorferi* HB19, we transformed them into our experimental strain, (**Table 1**) B31 5A17 (Lawrenz *et al.*, 2002), for subsequent mouse infections. This strain lacks lp25 and lp28-1 and transforms reasonably well (see **Experimental procedures**). As an additional characterization step, we estimated the relative copy numbers of the linear and circular mini-*vls* constructs carrying or missing the long IR. We did this because our previous work showed that conversion of a linear plasmid to a circular plasmid or *vice versa* resulted in a dramatic reduction in plasmid copy number due to a decrease in the transcription of the plasmid replication/maintenance proteins upon a change in plasmid topology (Beaurepaire & Chaconas, 2007). **Fig. 4** shows the change in relative copy number in the circular versus linear form of the mini-*vls* system driven by the lp5 replication region, either with or without the inverted repeat in the intergenic region. The circular form had a 10-fold and 25-fold decrease compared to the linear form, with and without the inverted repeat, respectively. This is in agreement with our previous observations of linear to circular conversion. The implications of this in switching experiments are discussed in a later section.

Analysis of switching by mini-*vls* constructs

Three individual 5A17 clones for each of the linear, circular, linear and circular with no cassette 2 and linear and circular with no inverted repeat constructs (**Fig. 3** and

Table 1) were grown to exponential phase and then pooled in equal amounts to infect three mice per pool for a final count of eighteen mice (see **Experimental Procedures**). SCID mice were used for the infections to prevent clearance of spirochetes that had not undergone switching at *vlsE*. To increase our chances of seeing switching with our one cassette system, we recovered tissues at later time points than usual: ear punches at weeks 4 and 8 post-infection, and harvested organs at week 12 (heart, knee joint, bladder, liver, spleen and lung). Recovery of spirochetes from ear at week 12 was unsuccessful because of universal contamination in these cultures. We used stationary-phase cultures to obtain genomic DNAs and used barcoded oligonucleotides (**Table S2**) to amplify the variable region of *vlsE* from the mini-*vls* constructs. We used two different barcodes per amplicon, one in each primer (**Table S3**).

We designed the *vlsE* amplicon to have a size of 808 bp and we used genomic DNA from GCB3500 and GCB3504 carrying pMC80 and pMC81, respectively (constructs with no *vlsE*, **Fig. 3** and **Table 1**) as negative controls in our PCRs. We obtained no bands in those controls (data not shown) confirming that the fragments obtained in the cultures grown after infection, were indeed *vlsE* and not a PCR artifact.

We sequenced a total of 125 pooled, barcoded samples coming from different time points and tissues (**Table S3**) using PacBio long read sequencing. To search for switching we used a pipeline and software (VAST) we previously developed (Verhey *et al.*, 2018a, Verhey *et al.*, 2018b). The analysis by tissue or time point did not show any interesting patterns, however when we grouped the data by construct type we were able to see that the experimental constructs underwent switching to a significantly higher extent than the controls (constructs with no cassette**, Fig. 5**). We used two different

approaches to analyze the data. The first method included every single read with as few as one templated SNP (**Fig. 5A**). As the total switching level was quite low relative to a full-sized wild-type *vls* locus and to help reinforce the conclusions obtained with the first method we also used a second approach that increased the signal-to-noise ratio for switching activity. Rather than counting individual templated SNPs, we looked at the number of recombination tracts with 2 or more templated SNPs, which are less likely to appear by chance than tracts composed of a single templated SNP (multi-SNP switches, **Fig. 5B**). Both approaches produced similar results as shown in the graphs A and B in **Fig. 5** confirming the conclusions above.

Of the four mini-*vls* constructs tested, three displayed significant switching at *vlsE* by both methods of analysis. Only the circular construct lacking the inverted repeat showed switching levels not clearly distinguishable from background.

In our previous studies on SNP usage during switching at *vlsE* in the wild-type *vls* locus we noted that SNPs from all silent cassettes were transferred to *vlsE* in the population of analyzed switch variants at approximately the frequency of their occurrence in the silent cassettes (Verhey *et al.*, 2018a). To further characterize switching in the mini-*vls* system we analyzed SNP usage from cassette 2 (**Fig. 6**). As expected, all 38 SNPs and the single three nucleotide deletion in cassette 2 were found at their corresponding locations in the collection of switch variants analyzed. Interestingly, the most highly utilized SNPs are those directly in the centre, from 200-250 bp from the start of cassette 2.

We also characterized the recombination tract length in the mini-*vls* system. As shown in **Fig. 7**, the average tract length (of all switches, including those composed of a

single SNP) was 11.9 bp, compared to the length of 16.4 that we previously reported for the wild-type *vls* system carrying all 15 silent cassettes.

Discussion

In this work we have developed for the first time a genetic approach that allowed us to generate and manipulate a functional mini-*vls* system. The *vls* locus has been difficult to investigate in terms of genetic manipulation. Its abundant repeats, high percent sequence identity among the cassettes and potential secondary structures as well as other possible elements in *B. burgdorferi* that are incompatible with *E. coli,* have been impediments to cloning the *vls* locus or the *vlsE* gene with its promoter on a plasmid (Norris, 2014). The *vls* locus has only been clonable on λ phage (Zhang *et al.*, 1997). The expression locus *vlsE* has been cloned on a plasmid only in the absence of the 100 bp inverted repeat (Lawrenz *et al.*, 2004) and with a mutation adjacent to the -35 box at position -38 (C instead of T) and another in position +7 after the ATG (G instead of A) changing a Glutamic Acid to a Lysine (unpublished results). Therefore, our current report on the construction of a mini-*vls co*nstruct that we can genetically modify represents a significant advance and opens the door to study switching of the *vlsE* locus at a new level.

To overcome the problems noted above we performed preliminary cloning steps of stable DNA in *E. coli* using a shuttle vector with *E. coli* components derived from pBSV2G (Elias *et al.*, 2003) and the *B. burgdorferi bbt04* gene from the small linear plasmid, lp5 (Casjens *et al.*, 2000) To add *B. burgdorferi* DNA fragments unstable in *E. coli* we used the high passage *B. burgdorferi Bb* strain HB19 (Coburn *et al.*, 1993) as a direct recipient for DNA ligations. High passage HB19 has lost many of its plasmids,

including those containing the restriction-modification systems, and in our hands transforms about 2000 times more efficiently than low passage, infectious B31 5A4 (Lawrenz *et al.*, 2002).

To test for switching of *vlsE in* the mini-systems that we constructed, we performed infections using SCID mice, where unswitched variants would not be cleared at around two weeks post-infection. We recovered spirochetes from the mouse infections, amplified *vlsE,* sequenced the amplicons by PacBio long read sequencing and analyzed the data using our pipeline and VAST software (Verhey *et al.*, 2018a, Verhey *et al.*, 2018b). We found that three of the four mini-*vls* constructs underwent switching at *vlsE* following mouse infection. However, the switching rate that we observed for these constructs was approximately $7x10^{-4}$ switches per amplicon per week compared with our previously reported switching rate of $7x10^{-1}$ switches per amplicon per week for the wild-type *vls* locus on lp28-1 (Verhey *et al.*, 2018b). The reason for this 10^3-fold decrease in activity is likely multi-factorial. Importantly, the mini-systems carry a single silent cassette rather than all 15, thereby limiting the number of available silent cassettes. In addition, silent cassette 2 used for the mini-system is normally one of the least used silent cassettes (Coutte *et al.*, 2009, Verhey *et al.*, 2018b). This may be because of its proximity to *vlsE* and the intrinsic stiffness in the DNA that must be overcome for synapsis of *vlsE* and cassette 2, with only 398 bp between them. It may also result from the truncation of this cassette at the N-terminal coding region, including the absence of one 17 bp direct repeat (Zhang *et al.*, 1997); these direct repeats may influence the efficiency of the switching reaction (Verhey *et al.*,

2018b). Cassette 2 also shares an N-terminal signal peptide sequence with *vlsE,* providing a possible site for unproductive synapsis.

Although the switching frequency that we observed was low, we observed use of the full complement of available SNPs (**Fig. 6**) as expected from our previous switching studies on the wild-type *vls* locus (Verhey *et al.*, 2018b). The reason for the highest level of switching being in the center of *vlsE* remains open to speculation at this time. The average switching tract length of 11.9 bp for the mini-system compared to 16.4 for wild-type *vlsE* is not identical but in good agreement considering the drastic differences between the two systems.

Finally, our ability to manipulate the *vls* locus and to generate a functional mini-*vls* system has allowed us to investigate two longstanding questions about switching at *vlsE.* The first is whether the topology of the *vls* locus is important or not. In *B. burgdorferi,* all *vls* loci sequenced to date are found on linear DNA molecules and near the hairpin telomeres (Kingry *et al.*, 2016, Casjens *et al.*, 2012, Casjens *et al.*, 2011b, Casjens *et al.*, 2011a, Schüler *et al.*, 2015). A direct comparison between a linear and circular mini-*vls* (both carrying the 100 bp inverted repeat) revealed no significant difference in switching between them (**Fig. 5**). Although the relative copy number of the linear plasmid was 10-fold higher, scoring switching per amplicon normalized for copy number differences. The lack of a difference in switching rate, even though there was a marked difference in copy number, suggests that switching occurs in cis: that is, it likely occurs between *vlsE* and silent cassettes on the same plasmid rather than between different plasmids, or higher frequencies would have been observed on the linear molecules with higher copy number.

The second longstanding question is whether the 100 bp inverted repeat just upstream of *vlsE* plays a role in the switching process. Similar long inverted repeats have been observed in *vls* loci from several strains as noted in the **Introduction**. We therefore constructed both circular and linear versions of the mini-*vls* system, both with and without the IR. Removal of the IR from the native linear construct did not result in a decrease in switching (**Fig. 5**) indicating that the IR is not required for switching. Surprisingly, however, removal of the IR from the circular non-native construct did reduce switching to background levels. The reason for this is not clear. The IRs would appear to be an important component of the *vls* system as they are present adjacent to all *vlsE* genes sequenced thus far (see **Introduction**). The question of why the IR is required for switching on a circular but not linear plasmid is vexing, especially since all *vls* loci found to date have been found on linear replicons (Kingry *et al.*, 2016, Casjens *et al.*, 2012, Casjens *et al.*, 2011b, Casjens *et al.*, 2011a, Schüler *et al.*, 2015). Some possible explanations for our results regarding the IR in the mini-*vls* system are:

1) The *vls* locus is normally present on a linear replicon and the IR may not be needed for recombinational switching, but may instead be required for regulated transcription of the *vlsE* gene.

2) Replication of linear replicons in *Borrelia* species involves circular replication intermediates, if telomere resolution does not precede completion of replication at both ends (Chaconas & Kobryn, 2010). The IR would be a requirement if recombinational switching occurs specifically during the circular intermediate stage. One might speculate that the higher level of DNA supercoiling on the circular plasmid that would result from the absence of the IR (which can be extruded as a cruciform and reduce supercoiling)

might be inhibitory to the switching process. In that case the IR would be required only on the circle to reduce the level of DNA supercoiling.

3) The requirement of the IR for switching on a circular but not linear plasmid could result from a failure of the mini-*vls* system to accurately reflect the wild-type situation. We cannot rule out the possibility that generation of the mini-system has resulted in the loss of an important component of the *vls* locus resulting in a reduced switching frequency with altered properties. Further experiments will be required to distinguish between these possibilities.

In conclusion, we report for the first time a genetic approach for the manipulation of the troublesome *vls* locus, the generation of a functional mini-*vls* system and its use to study the role of DNA topology and the long IR in the *vls* locus. Our results show that the topology of the DNA molecule carrying the *vls* system is not important and that the presence of the long IR is not important when *vls* is on a linear DNA molecule but is required when *vls* is on a circle. Further improvements of the mini system to increase the switching frequency will make this a powerful system for the analysis of *cis*-acting elements in recombinational switching at the *vls* antigenic variation locus in *B. burgdorferi.*

Experimental Procedures

Plasmid construction in *Escherichia coli*

All constructs described in this section were done in *E. coli* DH5-α (Taylor *et al.*, 1993). PCR reactions were performed with Phusion HF polymerase (NEB). Restriction

enzymes and T4 ligase were purchased from NEB as well. Genomic DNA from B31 5A4 (Purser & Norris, 2000) clones was used as the template for all *B. burgdorferi* PCRs.

A pBSV2G (Elias *et al*., 2003) fragment harbouring the *E. coli* origin of replication (pUC*ori*), the multi-cloning site (MCS) and the gentamicin resistance cassette was amplified using the oligos B2776 and B2777 (**Table S1**). This fragment was cloned into a 1.8 Kb PCR product containing the lp5 (Casjens *et al*., 2000) replication region (**Fig. 2**) using the restriction sites NcoI and XhoI, sites for which had been added to both primer sets. PCR conditions for these two fragments were as follows: 95℃ initial denaturation for 30 s followed by 30 cycles of 95℃ denaturation for 10 s, 59℃ annealing for 10 s, and 72℃ extension for 1 minute. PCR was ended after a final extension of 5 min at 72℃. The resultant plasmid was named pBSV5 which was then digested with BamHI and SacI to insert a 2 Kb fragment containing *pncA* (pMC73, **Fig. 2**). *pncA* was amplified using a 60ºC annealing temperature and one min extension time. The new vector pMC73 was used to introduce the replicated telomere (r-tel) of the plasmid lp28-1. To do this we added the telomere to both B2788 and B2789 oligos (**Table S1**) and performed a reverse PCR of the entire pMC73. The linear PCR was circularized by blunt end ligation (pMC82). Previous phosphorylation of the primers was performed by incubating them for 30 min at 37ºC with T4 DNA Ligase Buffer as the ATP source and T4 polynucleotide kinase (NEB). The annealing temperature for the reverse PCR was 58ºC and the extension time was 2 min. pMC82 and pMC73 were digested with BamHI and SalI to be able to introduce cassette 2, generating pMC80 and pMC81 respectively. Cassette *2* was amplified with the primers B2782 and B2797 (**Table S1**) at an annealing temperature of 60ºC for 10 s and extension time of 40 s at 72ºC. Since the

cassette 2 PCR gave rise to multiple bands, we purified the correct fragment from an agarose gel and used this in the ligations.

Plasmid construction in *Borrelia burgdorferi*

To be able to clone the *vlsE* gene including the 100 bp inverted repeat or without the inverted repeat into the different vectors we transformed *Borrelia burgdorferi* high passage HB19 (**Table 1**, (Kurtti *et al.*, 1987), a highly transformable strain that has lost most of its plasmids including those harbouring the restriction-modification systems. *vlsE* fragments were amplified using genomic DNA from B31 5A4 with Phusion HF polymerase (NEB) and since they contained non-specific bands, we purified the correct size fragments from agarose gels (Quiagen QIAquick Gel Extraction kit). After successfully obtaining the constructs in HB19, we transformed 20 to 50 µg of *B. burgdorferi* plasmid preps into our experimental strain B31 5A17 (**Table 1**, (Lawrenz *et al.*, 2002). This strain lacks lp25 which increases its transformation efficiency and lp28-1 which we needed to be absent in order to avoid rearrangements with the mini-systems and/or problems with amplification of the correct *vlsE*. HB19 DNA preparations were done with Qiagen QIAfilter Plasmid Midi kit. T4 Ligase and restriction enzymes were from NEB. DNA digestions and PCRs were purified before their use in the ligations with the Qiagen QIAquick PCR Purification kit.

vlsE including the inverted repeat was amplified with the oligos B2690 and B2796 (**Table S1**) producing a 1.3 kb fragment flanked by the restriction sites SphI and XbaI. The conditions used for this PCR were 95°C initial denaturation for 30 s followed by 30 cycles of 95°C denaturation for 15 s, 60°C annealing for 30 s, and 72°C extension for 1 minute. PCR was ended after a final extension of 5 min at 72°C. This insert was then

cloned into the plasmids pMC80, pMC81, pMC82 and pMC73, which were previously digested with SphII and AvrII (compatible cohesive end with XbaI) to generate the plasmids pMC108, pMC109, pMC110 and pMC111, respectively (**Table 1, Fig. 3**). For the ligations we used 250-500 ng of the vector in a ratio 1:5 (vector:insert) in a total volume of 10 or 20 µL. Reactions were incubated for 2 hr at room temperature and 5 µL were transformed into HB19 competent cells without previous purification of the mixture. Clones were selected in BSK-II plus 100 µg/mL gentamicin and took about 8 days to grow. To confirm the genetic constructs we performed PCR amplification for *aacC1, pncA, cass2* and *vlsE*. As noted below, *cass2* and *vlsE* were sequenced in each construct. We also digested the *B. burgdorferi* plasmid preps (Zymo Research Plasmid Miniprep kit) with the unique NotI enzyme (that does not *cut* the *B.* burgdorferi DNA in the preps) to be able to linearize the circular constructs or to generate two bands for the linear plasmids with hairpin telomeres, to confirm plasmid topology and size by visualization on ethidium bromide stained agarose gels.

For those constructs with no inverted repeat we amplified a 1.2 kb fragment containing *vlsE* and its promoter at an annealing temperature of 60ºC for 15 s and extension time of 50 s at 72ºC. Oligos B2842 and B2843 (**Table S1**) have been previously reported (Lawrenz *et al.*, 2004), to which we added AvrII and PvuI. Plasmids pMC80 and pMC81 were digested with the same restriction enzymes for the insert and we followed the same ligation procedures as mentioned above. In this way we generated plasmids pMC112 and pMC113 (**Table 1 and Fig. 3**).

Upon completion of strain construction, the *vlsE* variable region in the mini-system of the final 5A17 strains was sequenced following recovery by PCR

amplification. The reads showed 100% identity to the B31 5A4 *vlsE* reference sequence for all constructs. This was important to verify that all starting *vlsE* sequences were identical. Similarly, we sequenced cassette 2 in the 5A17 strains since this cassette would be the only template for switching to occur. The analysis revealed that cassette 2 was 100% identical to the B31 5A4 sequence in all strains.

Ethics Statement

All animal experimentation was carried out in accordance with the principles outlined in the most recent policies and *Guide to the Care and Use of Experimental Animals* by the Canadian Council on Animal Care. The animal protocol (AC16-0068) was approved by The Animal Care Committee of the University of Calgary.

Mouse infections and isolation *B. burgdorferi* genomic DNA

Fox Chase SCID mice (C.B-17 SCID) were purchased from Charles River Laboratories. Three individual clones for each of six constructs (see **Fig. 3**) were grown individually and pooled in equal amounts to infect three mice for pools of each construct, for a total of 18 mice. Mice were infected by subcutaneous needle inoculation into the back at 33–39 days of age with 10^5 spirochetes in the exponential growth phase. After 1 week, 50 µL of blood was obtained from the tail vein and ear punch samples were taken at 4 and 8 weeks post-infection. At 12 weeks post-infection, the mice were euthanized and the bladder, ear, heart, knee joint, liver, spleen and lung were harvested. Each of these samples was cultured in BSK-II culture medium supplemented with 6% Rabbit Serum and antibiotics (0.02 mg/ml phosphomycin, 0.05 mg/ml rifampicin, and 2.5 µg/ml

amphotericin B) to which *B. burgdorferi* spirochetes are naturally resistant plus 100 µg/mL gentamicin (resistance encoded by the mini-system plasmids). Genomic DNA was extracted from 5 mL cultures by phenyl-chloroform extraction and concentrated by isopropanol precipitation.

vlsE PCR amplification and sequencing

PCR primers were designed for optimal amplification of the *vlsE* variable region, and amplify the 302 – 1078 bp region of the updated B31 gene sequence which includes the telomere sequences, flanked by two of the 16 bp barcode sequences (**Table S2**). The barcodes we used were designed by Pacific Biosciences, and we selected barcodes such that they had an edit distance of at least 6 bp between all possible pairs. Therefore, even with two sequence errors, the correct barcode could still be unambiguously identified. We used a paired barcode scheme, where the combination of the forward and the reverse primers uniquely identified each sample (**Table S3**). All *vlsE* amplicons were generated by PCR with Phusion HF polymerase (NEB) with a 98℃ initial denaturation for 30 s followed by 30 cycles of 98℃ denaturation for 10 s, 60℃ annealing for 30 s, and 72℃ extension for 1 minute. PCR was ended after a final extension of 5 mins at 72℃. GCB3500 and GCB3504 DNAs were used as negative controls. PCR products were quantified by agarose gel densitometry with standard curves, mixed together, purified and concentrated with a Qiagen QIAquick PCR purification kit, run on a gel in small batches, and gel-purified without staining. The final mixture was quantified by gel densitometry. Amplicon mixtures were sequenced using a

Pacific Biosciences RSII instrument and P6C4 chemistry by the Johns Hopkins Deep Sequencing and Microarray Core Facility. A total of 6 SMRTcells were sequenced.

Bioinformatics

We used a previously reported software pipeline to process the sequence data (Verhey *et al.*, 2018a) and analyzed switching activity using the Variable Antigen Sequence Tracer (Verhey *et al.*, 2018b).

Accession numbers

PacBio sequencing data is available in the NCBI sequence read archive (SRA) with accession number SRP135500.

Acknowledgements

We would like to thank Steven Norris and Jenifer Coburn for providing *B. burgdorferi* strains and Genevieve Chaconas for technical assistance. We also thank the JHMI Deep Sequencing and Microarray core for the PacBio sequencing. This work was supported by the Canadian Institutes of Health Research [MOP-53086 and PJT-153336].

Author contributions

Designed experiments, MC, TBV, GC; Executed experiments, MC; Wrote software and performed computer analysis, TBV; Analyzed data, MC, TBV, GC. Wrote manuscript, MC, TBV, GC.

Figure Legends

Figure 1. Schematic of *vls* mini-system. Wild type lp28-1 is shown at the top of the image depicting the arrangement of the silent cassettes and *vlsE* in the *vls* locus. The bottom shows an expanded view of the *vls* mini-system carrying cassette 2 and the rest of the elements normally present in that part of the *vls* locus. CR, constant regions, DR, 17 bp direct repeats. Shown in the intergenic region is the 100 bp perfect inverted repeat and the *vlsE* promoter, P.

Figure 2. Cloning strategy for the construction of mini-*vls* variants used in this study. The *B. burgdorferi-E. coli* shuttle vector pBSV2G (Elias *et al.*, 2003) was used as the template to amplify the region containing the pUC*ori*, the *gent* cassette and the MCS. This PCR product **(1)** was cloned into a 1.8 kb fragment **(2)** carrying the replication region from the 5 kb linear *B. burgdorferi* plasmid lp5, resulting in the new vector pBSV5. A 2 kb fragment harboring *pncA* from lp25 **(3)** was then cloned into pBSV2 resulting in pMC73. The next step was to either clone a replicated telomere **(4)** and cassette 2 **(5)** resulting in pMC80 or cassette 2 without the replicated telomere (pMC81). The final step was to clone the *vlsE* gene **(6)** with or without the inverted repeat (IR) into pMC80 and pMC81. Cloning of *vlsE (*or IR-*vlsE)* was done directly in high passage HB19, a highly transformable *B. burgdorferi* strain (Kurtti *et al.*, 1987) because of the instability of the inverted repeat in *E. coli*. Constructs were then moved into the low passage *B. burgdorferi* strain B31 5A17 (Lawrenz *et al.*, 2002) for

subsequent mouse infections and analysis of switching at *vlsE.* The eight experimental and control plasmids generated are shown in **Fig. 3** and the strains described in **Table 1**.

Figure 3. Schematic of the 8 different constructs used in this study. Controls are missing either *vlsE* (PCR negative controls) or the cassette (infection controls) in both linear and circular versions. The experimental plasmids all carry cassette 2 and contain or are missing the inverted repeat, in both circular and linear versions.

Figure 4. Relative copy numbers of mini-*vls* plasmids in *B. burgdorferi*. The concentration of the different mini-system constructs was calculated relative to the endogenous plasmid lp17 by measuring fluorescence intensities on ethidium bromide stained agarose gels of *B. burgdorferi* plasmid preps. Averages correspond to 6 independent measurements and *p-values* were obtained using the Mann-Whitney test for non-parametric data.

Figure 5. Recombinational switching of SNPs from silent cassette 2 into the mini-*vls* expression locus. A) The frequency of templated SNPs is shown for all post-infection *vlsE* sequences isolated from mice. Sequence data are grouped by the plasmid construct contained in the *B. burgdorferi* strains used for infection of SCID mice. Plasmid constructs with significantly different switching activity ($p < 0.05$ using Dunn-Bonferroni multiple comparison analysis) are shown. Blue bars show significant comparisons between the circular mini-*vls* control system that lacks cassette 2 and constructs carrying both *vlsE* and cassette 2. Red bars show significant comparisons between the linear mini-*vls* control system that lacks cassette 2 and constructs carrying both *vlsE* and cassette 2. Green bars indicate significantly different relationships

between the circular mini-*vls* system that lacked the inverted repeat and other constructs. **B)** Switching was also analyzed using inferred switch events longer than a single SNP; the number of these recombination events is shown for each plasmid construct. As in the first part of the figure, all significant relationships are represented above the bars with their corresponding level of significance. Single asterisks represent $p \leq 0.05$; double asterisks represent $p \leq 0.01$, and triple asterisks represent $p \leq 0.001$. The complete set of P-values for all pairwise comparisons is shown in **Fig. S1**.

Figure 6. Frequency of SNP transfer from silent cassette 2. For each position in the *vlsE* amplicon, the positions of SNPs in silent cassette 2 are shown in red, while the frequency of those SNPs as observed in the sequence data is shown above in blue. A grey bar represents the extent of cassette 2.

Figure 7. Histogram of inferred recombinational switch event length. Switch events with more than one templated SNP were plotted by the inferred recombination tract length as a percentage of all switch events. Lengths are the minimal tract length – representing the smallest possible size of the switch event. In blue, the histogram represents switch events from the 4 mini-*vls* constructs carrying cassette 2 that are reported in this paper. For comparison, the size distribution of inferred recombination events in the wild-type B31 system from our previous work are shown in red (Verhey *et al.*, 2018b). In green, the mean minimal switch length for each population is shown. The means were calculated using all switch events (including those with only 1 SNP) as has been done in previous studies of *vls* recombination tract length.

1

2

3

Table 1. Strains and plasmids used

Strain #	Strain	Plasmid	Drug	Description	Reference
E. coli					
GCE719	DH5-α			*E. coli* strain used for DNA manipulation	(Taylor *et al.*, 1993)
GCE355		pBSV2G	Gent	*Borrelia burgdorferi* shuttle vector harboring cp9 replication region	(Elias *et al.*, 2003)
GCE3739	DH5-α	pBSV5	Gent	*Borrelia burgdorferi* shuttle vector harboring lp5 replication region	This study
GCE3741	DH5-α	pMC73	Gent	pBSV5::*pncA*	This study
GCE3759	DH5-α	pMC82	Gent	pBSV5::*pncA*::*replicated telomere (rtel)*	This study
GCE3757	DH5-α	pMC80	Gent	pBSV5::*pncA*::*rtel*::*cassette2*	This study
GCE3755	DH5-α	pMC81	Gent	pBSV5::*pncA*::*cassette2*	This study
B. burgdorferi					
GCB988	HB19			High-passage wild type strain. Highly transformable.	(Coburn *et al.*, 1993)
GCB2988	HB19	pMC108	Gent	pMC80::InvertedRepeat (IR) *vlsE* (Linear)	This study
GCB2990	HB19	pMC109	Gent	pMC81::IR*vlsE* (Circular)	This study
GCB3477	HB19	pMC110	Gent	pMC82::IR*vlsE* (Linear No cassette)	This study
GCB3480	HB19	pMC111	Gent	pMC73::IR*vlsE* (Circular No cassette)	This study
GCB3527	HB19	pMC112	Gent	pMC80::*vlsE* (Linear No IR)	This study
GCB3529	HB19	pMC113	Gent	pMC81::*vlsE* (Circular No IR)	This study
GCB919	5A17			B31 derivative missing lp28-1 and lp25	(Lawrenz *et al.*, 2002)
GCB3500	5A17	pMC80	Gent	Clone 6 (Linear No *vlsE*)	This study
GCB3504	5A17	pMC81	Gent	Clone 3 (Circular No *vlsE*)	This study
GCB3494	5A17	pMC108	Gent	Clone 1 (Linear)	This study
GCB3495	5A17	pMC108	Gent	Clone 3 (Linear)	This study
GCB3496	5A17	pMC108	Gent	Clone 4 (Linear)	This study
GCB3509	5A17	pMC110	Gent	Clone 1 (Linear No cassette)	This study
GCB3510	5A17	pMC110	Gent	Clone 2 (Linear No cassette)	This study
GCB3513	5A17	pMC110	Gent	Clone 5 (Linear No cassette)	This study
GCB3514	5A17	pMC109	Gent	Clone 1 (Circular)	This study
GCB3515	5A17	pMC109	Gent	Clone 2 (Circular)	This study
GCB3518	5A17	pMC109	Gent	Clone 5 (Circular)	This study
GCB3519	5A17	pMC111	Gent	Clone 1 (Circular No cassette)	This study
GCB3520	5A17	pMC111	Gent	Clone 2 (Circular No cassette)	This study
GCB3521	5A17	pMC111	Gent	Clone 3 (Circular No cassette)	This study
GCB3539	5A17	pMC112	Gent	Clone 5 (Linear No IR)	This study
GCB3540	5A17	pMC112	Gent	Clone 6 (Linear No IR)	This study
GCB3541	5A17	pMC112	Gent	Clone 7 (Linear No IR)	This study
GCB3543	5A17	pMC113	Gent	Clone 1 (Circular No IR)	This study
GCB3544	5A17	pMC113	Gent	Clone 2 (Circular No IR)	This study
GCB3545	5A17	pMC113	Gent	Clone 3 (Circular No IR)	This study

References

Bankhead, T. & G. Chaconas, (2007) The role of VlsE antigenic variation in the Lyme disease spirochete: persistence through a mechanism that differs from other pathogens. *Mol Microbiol* **65**: 1547-1558.

Beaurepaire, C. & G. Chaconas, (2007) Topology-dependent transcription in linear and circular plasmids of the segmented genome of *Borrelia burgdorferi*. *Mol Microbiol* **63**: 443-453.

Casjens, S., N. Palmer, R. van Vugt, W.H. Huang, B. Stevenson, P. Rosa, R. Lathigra, G. Sutton, J. Peterson, R.J. Dodson, D. Haft, E. Hickey, M. Gwinn, O. White & C.M. Fraser, (2000) A bacterial genome in flux: the twelve linear and nine circular extrachromosomal DNAs in an infectious isolate of the Lyme disease spirochete *Borrelia burgdorferi*. *Mol Microbiol* **35**: 490-516.

Casjens, S.R., C.M. Fraser-Liggett, E.F. Mongodin, W.G. Qiu, J.J. Dunn, B.J. Luft & S.E. Schutzer, (2011a) Whole genome sequence of an unusual *Borrelia burgdorferi* sensu lato isolate. *Journal of bacteriology* **193**: 1489-1490.

Casjens, S.R., E.F. Mongodin, W.G. Qiu, J.J. Dunn, B.J. Luft, C.M. Fraser-Liggett & S.E. Schutzer, (2011b) Whole-genome sequences of two *Borrelia afzelii* and two *Borrelia garinii* Lyme disease agent isolates. *Journal of bacteriology* **193**: 6995-6996.

Casjens, S.R., E.F. Mongodin, W.G. Qiu, B.J. Luft, S.E. Schutzer, E.B. Gilcrease, W.M. Huang, M. Vujadinovic, J.K. Aron, L.C. Vargas, S. Freeman, D. Radune, J.F. Weidman, G.I. Dimitrov, H.M. Khouri, J.E. Sosa, R.A. Halpin, J.J. Dunn & C.M. Fraser, (2012) Genome stability of Lyme disease spirochetes: comparative genomics of *Borrelia burgdorferi* plasmids. *PLoS One* **7**: e33280.

Chaconas, G. & K. Kobryn, (2010) Structure, function, and evolution of linear replicons in Borrelia. *Annual review of microbiology* **64**: 185-202.

Chaconas, G., P.E. Stewart, K. Tilly, J.L. Bono & P. Rosa, (2001) Telomere resolution in the Lyme disease spirochete. *EMBO J* **20**: 3229-3237.

Coburn, J., J.M. Leong & J.K. Erban, (1993) Integrin $\alpha_{IIb}\beta_3$ mediates binding of the Lyme disease agent *Borrelia burgdorferi* to human platelets. *Proceedings of the National Academy of Sciences of the United States of America* **90**: 7059-7063.

Coutte, L., D.J. Botkin, L. Gao & S.J. Norris, (2009) Detailed analysis of sequence changes occurring during *vlsE* antigenic variation in the mouse model of *Borrelia burgdorferi* infection. *PLoS pathogens* **5**: e1000293.

Dresser, A.R., P.-O. Hardy & G. Chaconas, (2009) Investigation of the role of DNA replication, recombination and repair genes in antigenic switching at the vlsE locus in *Borrelia burgdorferi*: an essential role for the RuvAB branch migrase. *PLoS pathogens* **5**: e1000680.

Elias, A.F., J.L. Bono, J.J. Kupko, 3rd, P.E. Stewart, J.G. Krum & P.A. Rosa, (2003) New antibiotic resistance cassettes suitable for genetic studies in *Borrelia burgdorferi*. *J Mol Microbiol Biotechnol* **6**: 29-40.

Hudson, C.R., J.G. Frye, F.D. Quinn & F.C. Gherardini, (2001) Increased expression of *Borrelia burgdorferi vlsE* in response to human endothelial cell membranes. *Mol Microbiol* **41**: 229-239.

Indest, K.J., J.K. Howell, M.B. Jacobs, D. Scholl-Meeker, S.J. Norris & M.T. Philipp, (2001) Analysis of *Borrelia burgdorferi vlsE* gene expression and recombination in the tick vector. *Infection and immunity* **69**: 7083-7090.

Kingry, L.C., D. Batra, A. Replogle, L.A. Rowe, B.S. Pritt & J.M. Petersen, (2016) Whole genome sequence and comparative genomics of the novel Lyme borreliosis causing pathogen, *Borrelia mayonii*. *PloS one* **11**: e0168994.

Kobryn, K. & G. Chaconas, (2002) ResT, a telomere resolvase encoded by the Lyme disease spirochete. *Mol Cell* **9**: 195-201.
Kurtti, T.J., U.G. Munderloh, R.C. Johnson & G.G. Ahlstrand, (1987) Colony formation and morphology in *Borrelia burgdorferi*. *Journal of clinical microbiology* **25**: 2054-2058.
Lawrenz, M.B., H. Kawabata, J.E. Purser & S.J. Norris, (2002) Decreased electroporation efficiency in *Borrelia burgdorferi* containing linear plasmids lp25 and lp56: impact on transformation of infectious *B. burgdorferi*. *Infection and immunity* **70**: 4798-4804.
Lawrenz, M.B., R.M. Wooten & S.J. Norris, (2004) Effects of *vlsE* complementation on the infectivity of *Borrelia burgdorferi* lacking the linear plasmid lp28-1. *Infection and immunity* **72**: 6577-6585.
Li, B., (2015) DNA double-strand breaks and telomeres play important roles in trypanosoma brucei antigenic variation. *Eukaryot Cell* **14**: 196-205.
Lin, T., L. Gao, D.G. Edmondson, M.B. Jacobs, M.T. Philipp & S.J. Norris, (2009) Central role of the Holliday junction helicase RuvAB in *vlsE* recombination and infectivity of *Borrelia burgdorferi*. *PLoS pathogens* **12**: e1000679.
Liveris, D., V. Mulay, S. Sandigursky & I. Schwartz, (2008) *Borrelia burgdorferi vlsE* antigenic variation is not mediated by RecA. *Infection and immunity* **76**: 4009-4018.
Lloyd, R.G. & G.J. Sharples, (1993) Processing of recombination intermediates by the RecG and RuvAB proteins of *Escherichia coli*. *Nucleic acids research* **21**: 1719-1725.
Magunda, P.R.H. & T. Bankhead, (2016) Investigating the potential role of non-vls genes on linear plasmid 28–1 in virulence and persistence by *Borrelia burgdorferi*. *BMC microbiology* **16**: 180.
Norris, S.J., (2014) *vls* Antigenic Variation Systems of Lyme Disease *Borrelia*: Eluding Host Immunity through both Random, Segmental Gene Conversion and Framework Heterogeneity. *Microbiology spectrum* **2**.
Obergfell, K.P. & H.S. Seifert, (2015) Mobile DNA in the Pathogenic Neisseria. *Microbiology spectrum* **3**: MDNA3-0015-2014.
Purser, J.E., M.B. Lawrenz, M.J. Caimano, J.K. Howell, J.D. Radolf & S.J. Norris, (2003) A plasmid-encoded nicotinamidase (PncA) is essential for infectivity of *Borrelia burgdorferi* in a mammalian host. *Mol Microbiol* **48**: 753-764.
Purser, J.E. & S.J. Norris, (2000) Correlation between plasmid content and infectivity in *Borrelia burgdorferi*. *Proceedings of the National Academy of Sciences of the United States of America* **97**: 13865-13870.
Rogovskyy, A.S. & T. Bankhead, (2013) Variable VlsE Is Critical for Host Reinfection by the Lyme Disease Spirochete. *PLOS ONE* **8**: e61226.
Rogovskyy, A.S., T. Casselli, Y. Tourand, C.R. Jones, J.P. Owen, K.L. Mason, G.A. Scoles & T. Bankhead, (2015) Evaluation of the Importance of VlsE Antigenic Variation for the Enzootic Cycle of *Borrelia burgdorferi*. *PLOS ONE* **10**: e0124268.
Schüler, W., I. Bunikis, J. Weber-Lehman, P. Comstedt, S. Kutschan-Bunikis, G. Stanek, J. Huber, A. Meinke, S. Bergström & U. Lundberg, (2015) Complete genome sequence of *Borrelia afzelii* K78 and comparative genome analysis. *PloS one* **10**: e0120548.
Stanek, G., G.P. Wormser, J. Gray & F. Strle, (2012) Lyme borreliosis. *Lancet* **379**: 461-473.
Steere, A.C., F. Strle, G.P. Wormser, L.T. Hu, J.A. Branda, J.W. Hovius, X. Li & P.S. Mead, (2016) Lyme borreliosis. *Nat Rev Dis Primers* **2**: 16090.
Taylor, R.G., D.C. Walker & R.R. McInnes, (1993) E. coli host strains significantly affect the quality of small scale plasmid DNA preparations used for sequencing. *Nucleic Acids Res* **21**: 1677-1678.
Tourand, Y., J. Deneke, T.J. Moriarty & G. Chaconas, (2009) Characterization and *in vitro* reaction properties of 19 unique hairpin telomeres from the linear plasmids of the Lyme disease spirochete. *The Journal of biological chemistry* **284**: 7264-7272.

Tsaneva, I.R., B. Müller & S.C. West, (1992) ATP-dependent branch migration of Holliday junctions promoted by the RuvA and RuvB proteins of E. coli. *Cell* **69**: 1171-1180.

Verhey, T.B., M. Castellanos & G. Chaconas, (2018a) Analysis of recombinational switching at the antigenic variation locus of the Lyme spirochete using a novel PacBio sequencing pipeline. *Molecular microbiology* **107**: 104-115.

Verhey, T.B., M. Castellanos & G. Chaconas, (2018b) Antigenic variation in the Lyme disease spirochete: new insights into the mechanism of recombinational switching with a suggested role for error-prone repair. *Cell Reports* **23**: 2595–2605.

Vink, C., G. Rudenko & H.S. Seifert, (2012) Microbial antigenic variation mediated by homologous DNA recombination. *FEMS microbiology reviews* **36**: 917-948.

Walia, R. & G. Chaconas, (2013) Suggested role for G4 DNA in recombinational switching at the antigenic variation locus of the Lyme disease spirochete. *PLoS One* **8**: e57792.

Wang, D., D.J. Botkin & S.J. Norris, (2003) Characterization of the *vls* antigenic variation loci of the Lyme disease spirochaetes *Borrelia garinii* Ip90 and *Borrelia afzelii* ACAI. *Mol Microbiol* **47**: 1407-1417.

Zhang, J.R., J.M. Hardham, A.G. Barbour & S.J. Norris, (1997) Antigenic variation in Lyme disease borreliae by promiscuous recombination of VMP-like sequence cassettes. *Cell* **89**: 275-285.

Zhang, J.R. & S.J. Norris, (1998a) Genetic variation of the *Borrelia burgdorferi* gene *vlsE* involves cassette-specific, segmental gene conversion. *Infection and immunity* **66**: 3698-3704.

Zhang, J.R. & S.J. Norris, (1998b) Kinetics and *in vivo* induction of genetic variation of *vlsE* in *Borrelia burgdorferi*. *Infection and immunity* **66**: 3689-3697.

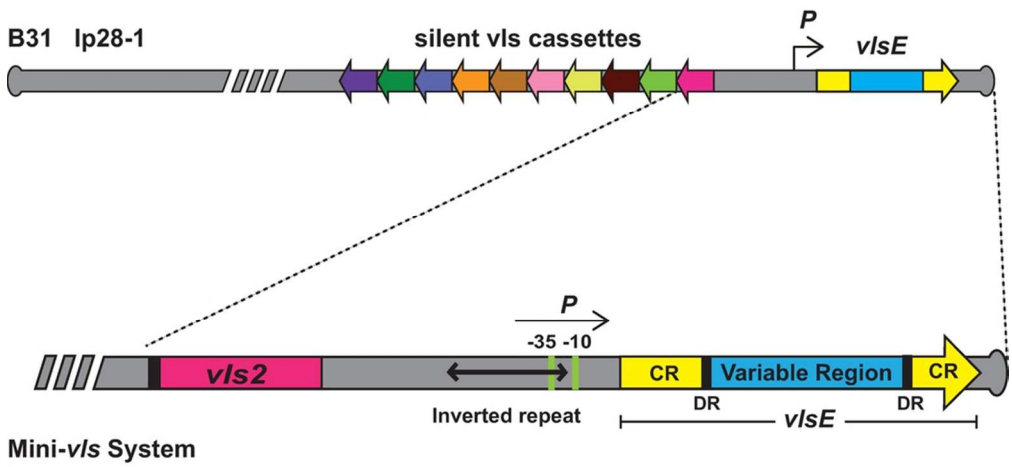

Fig. 1

Figure 1. Schematic of vls mini-system.

88x45mm (300 x 300 DPI)

Lorem ipsum

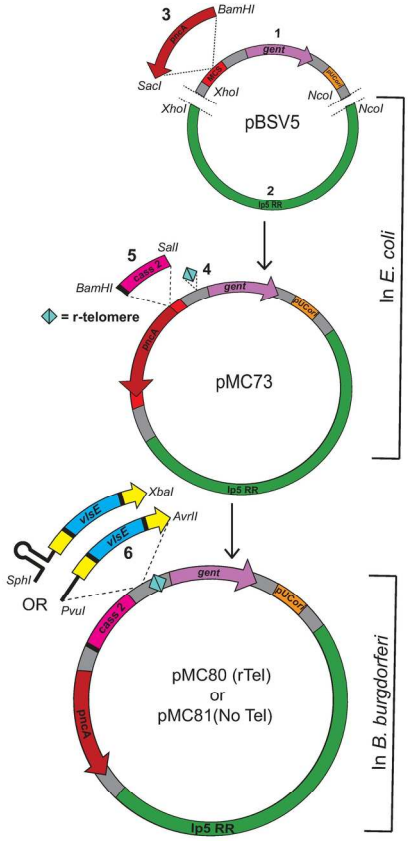

Fig. 2

Figure 2. Cloning strategy for the construction of mini-vls variants used in this study.

182x161mm (300 x 300 DPI)

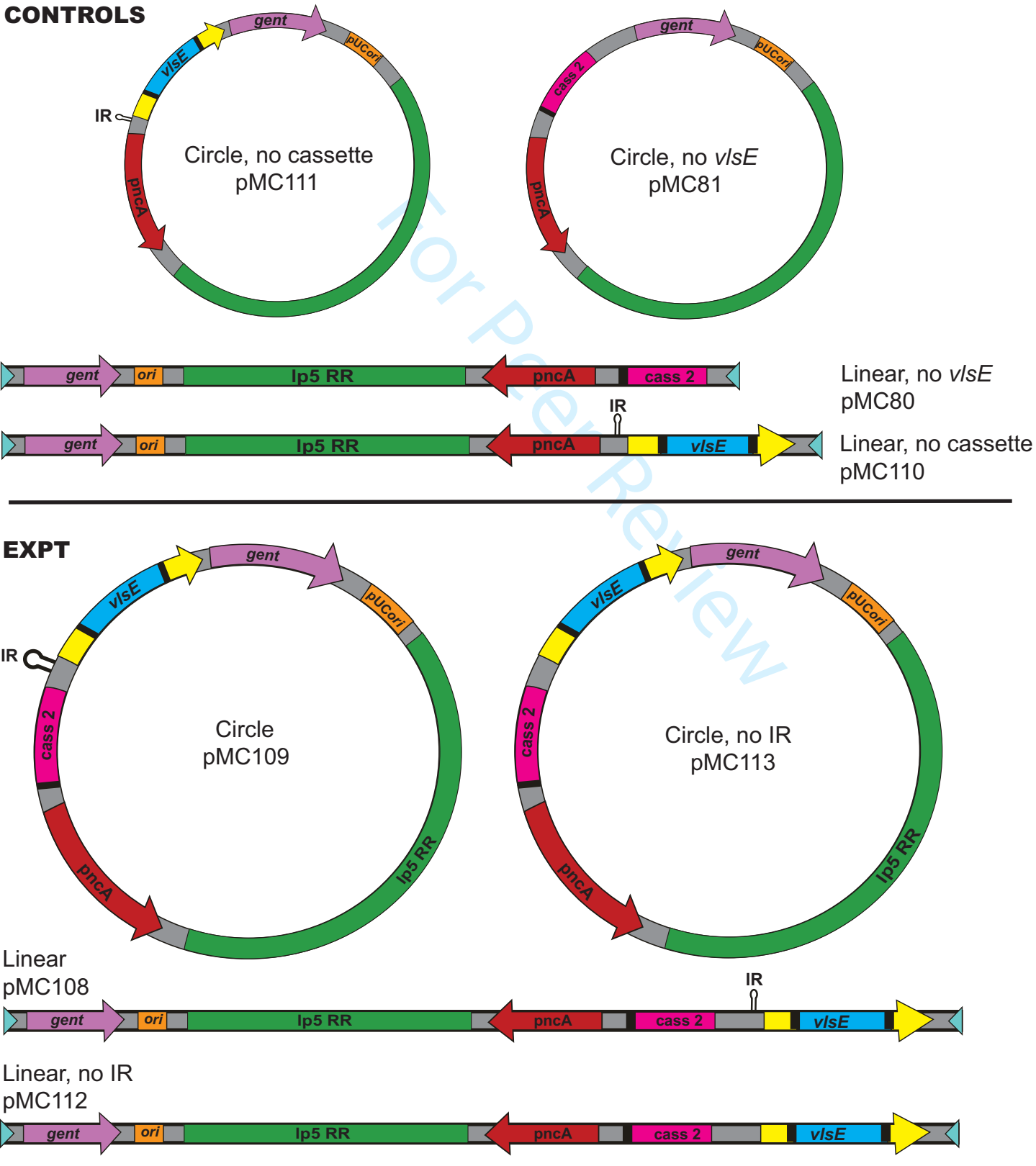

Fig. 3

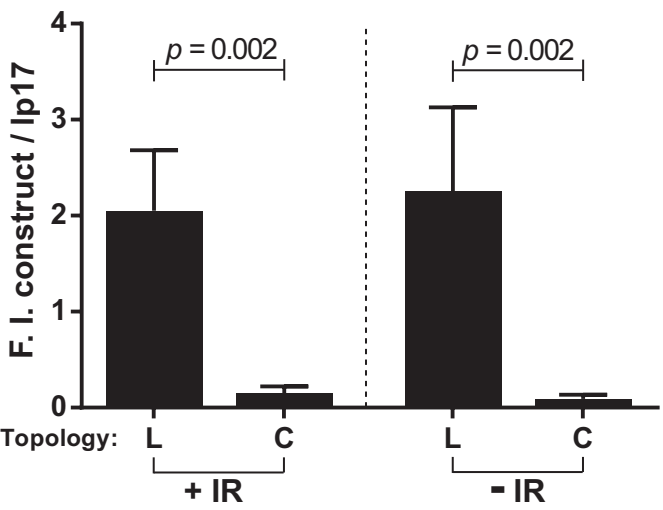

Fig. 4

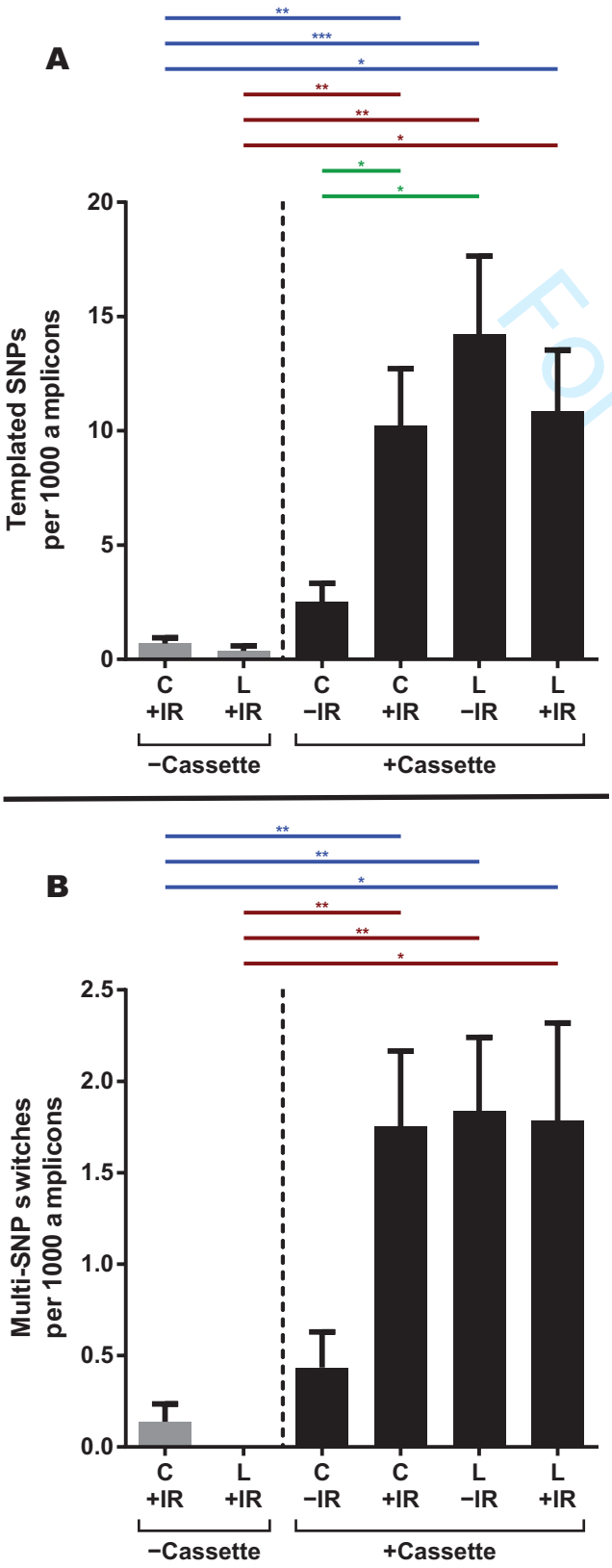

Fig. 5

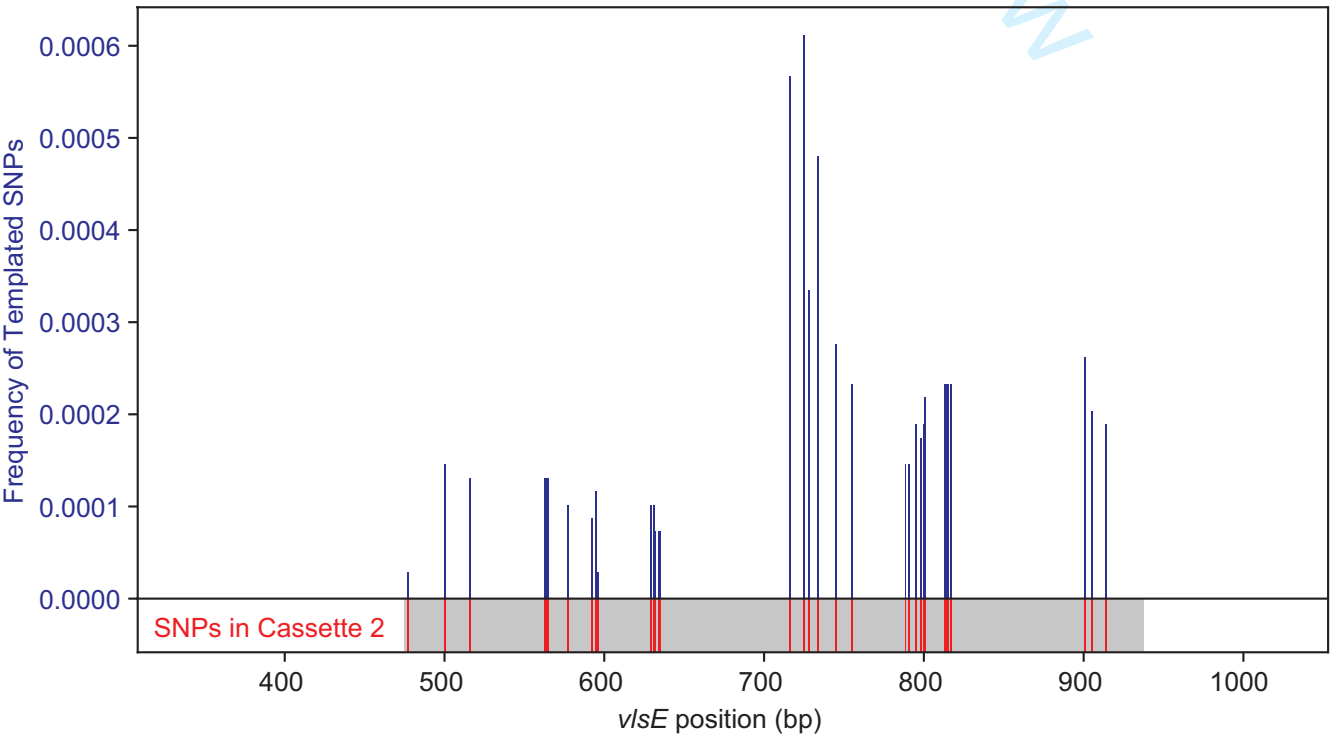

Fig. 6

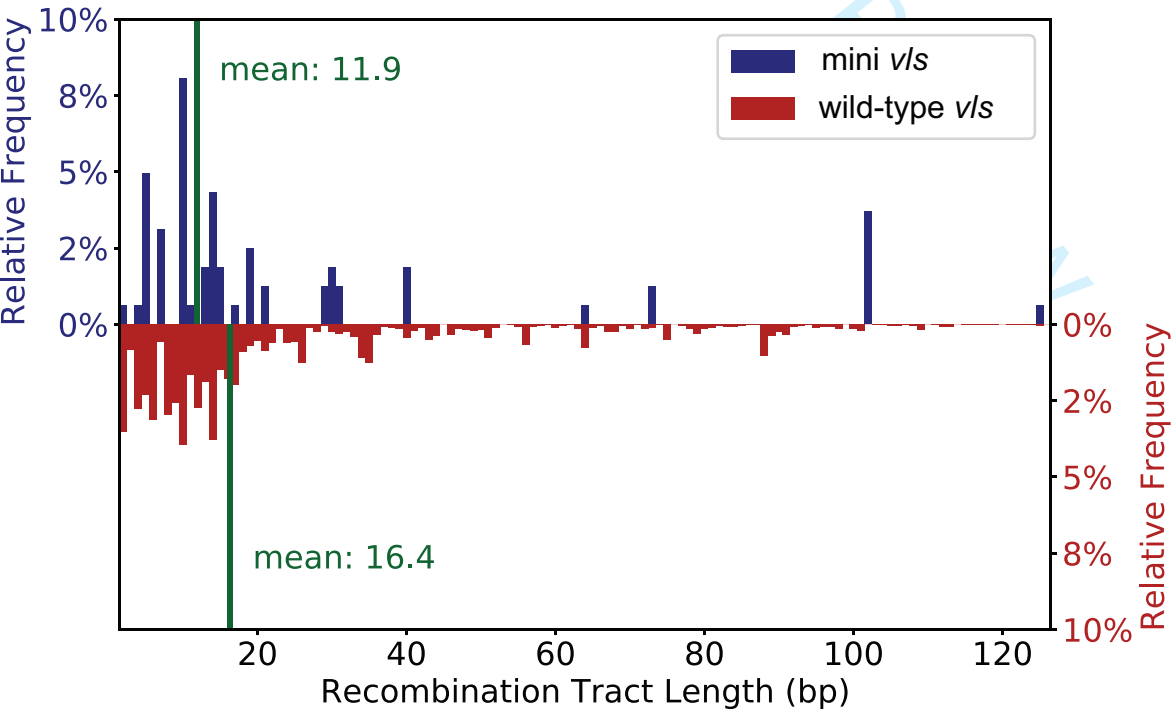
Relative Frequency
10%
8%
5%
2%
0%
mean: 11.9
mini vls
wild-type vls
mean: 16.4
0%
2%
5%
8%
10%
Relative Frequency
20
40
60
80
100
120
Recombination Tract Length (bp)

Fig. 7

We have generated a mini-*vls* system by direct cloning in *Borrelia burgdorferi*. The mini-*vls* undergoes switching in mice. The role of the long inverted repeat (IR) and the plasmid topology in switching were investigated by generating linear and circular constructs that either carried or lacked the IR. The IR is essential for switching to occur only when the topology is circular.