Characterization and genomic analysis of the Lyme disease spirochete bacteriophage ϕBB-1

Dominick R. Faith[1]*, Margie Kinnersley[1]*, Diane M. Brooks[1], Dan Drecktrah[1], Laura S. Hall[1], Eric Luo[2], Andrew Santiago-Frangos[3], Jenny Wachter[2], D. Scott Samuels[1], and Patrick R. Secor[1#]

[1] Division of Biological Sciences, University of Montana, Missoula, MT, USA

[2] Vaccine and Infectious Disease Organization, Saskatoon, SK, Canada

[3] Department of Biology, University of Pennsylvania, Philadelphia, PA, USA

*These authors contributed equally to this work

[#] Correspondence: Patrick.secor@mso.umt.edu

Abstract

Lyme disease is a tick-borne infection caused by the spirochete *Borrelia* (*Borreliella*) *burgdorferi*. *Borrelia* species have highly fragmented genomes composed of a linear chromosome and a constellation of linear and circular plasmids that encode diverse outer membrane lipoproteins, which facilitate movement of the spirochete between its tick vector and a vertebrate host in an enzootic cycle. The *B. burgdorferi* genome shows evidence of horizontal transfer between strains, but the mechanisms remain poorly defined. Almost all Lyme disease spirochetes are infected by 32-kp circular plasmid (cp32) prophages that undergo lytic replication and produce infectious virions called ϕBB-1. In the laboratory, ϕBB-1 transduces cp32s and shuttle vectors between spirochetes. However, the extent that ϕBB-1 participates in horizontal gene transfer between Lyme disease spirochetes is not known. Here, we use proteomics and long-read sequencing to characterize ϕBB-1 virions and the genetic material they package. Our studies reveal that ϕBB-1 packages linear cp32s via a headful mechanism. We identify the cp32 *pac* region and show that plasmids containing the cp32 *pac* region are preferentially packaged into ϕBB-1 virions. Additionally, we find ϕBB-1 packages fragments of the linear chromosome and other plasmids including lp54, cp26, and others. Furthermore, sequencing of ϕBB-1 packaged DNA allowed us to resolve the covalently closed hairpin telomeres for the linear *B. burgdorferi* chromosome and most linear plasmids in strain CA-11.2A. Collectively, our results shed light on the biology of the ubiquitous ϕBB-1 phage and further implicates ϕBB-1 in the generalized transduction of diverse genes between spirochetes and the maintenance of genetic diversity in Lyme disease spirochetes.

Introduction

The bacterium *Borrelia* (*Borreliella*) *burgdorferi* is the causative agent of Lyme disease, the most common tick-borne disease in the Northern Hemisphere [1-3]. Lyme disease spirochetes have complex and highly fragmented genomes composed of a ~900 kb linear chromosome and a constellation of linear and circular plasmids that are similar but not identical across the genospecies [4-6].

A key factor in the ability of *B. burgdorferi* to transmit from its tick vector to a vertebrate host is the differential expression of several outer membrane lipoproteins that assist *B. burgdorferi* in evading both vector and host immune responses [7]. As such, a large fraction of the *B. burgdorferi* genome encodes outer membrane lipoproteins, mostly carried on the plasmids [6, 8, 9].

In natural populations, the recombination rate between genomes of coexisting *B. burgdorferi* strains is exceeds the mutation rate and genetic variation in outer membrane lipoprotein alleles associated with species-level adaptations [6, 8-10]. Variation in outer membrane lipoprotein alleles across the genospecies is driven primarily by horizontal gene transfer [5, 11-17]. However, the mechanism(s) by which heterologous *B. burgdorferi* strains exchange genetic material are not well defined.

Viruses that infect bacteria (phages) are key drivers of horizontal gene transfer between bacteria [18]. The genomes of nearly all Lyme disease spirochetes include the 32-kb circular plasmid (cp32) prophages (**Fig 1A and B**) [4]. The cp32s carry several outer membrane lipoprotein gene families including *bdr, mlp*, and *ospE/ospF/elp* (*erps*), which are all involved in immune evasion [19-23] and exhibit sequence variation that is consistent with historical recombination amongst cp32 isoforms [24-26]. Recent work indicates that cp32 prophages are induced in the tick midgut during a bloodmeal [9, 27, 28]. When induced, cp32 prophages undergo lytic replication where they are packaged into infectious virions designated ϕBB-1 (**Fig 1C**) [29-31].

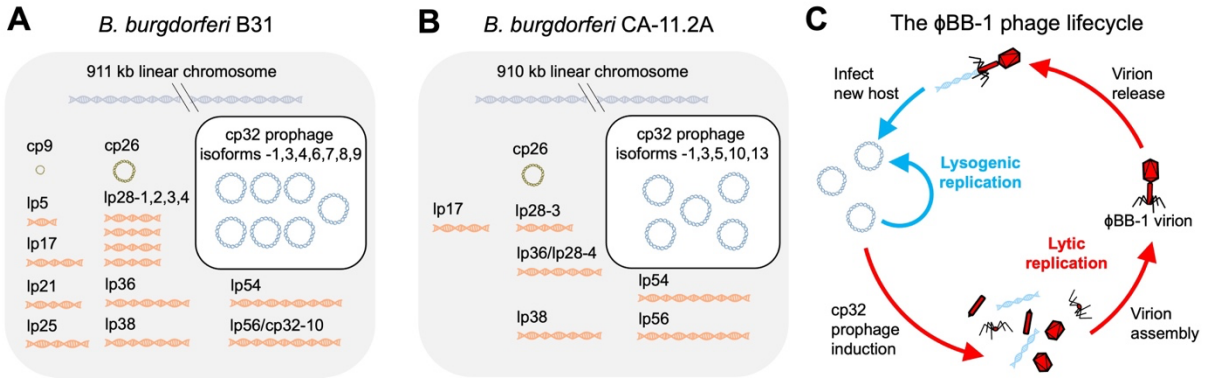


Figure 1. The *B. burgdorferi* genome is highly fragmented and is composed of a linear chromosome, linear and circular plasmids, and cp32 prophages. The genomes of *B. burgdorferi* strains **(A)** B31 and **(B)** CA-11.2A are shown. **(C)** The temperate ϕBB-1 phage lifecycle is depicted.

In addition to horizontally transferring phage genomes between bacterial hosts (transduction), phages frequently package and horizontally transfer pieces of the bacterial chromosome or other non-phage DNA (generalized transduction) [32]. Generalized

transduction was first observed in the *Salmonella* phage P22 in the 1950s [33]. Since then, generalized transduction has been observed in numerous other phage species [32, 34-37]. ϕBB-1 is a generalized transducing phage that can horizontally transfer shuttle vectors carrying antibiotic resistance cassettes between *B. burgdorferi* strains [29]. However, to our knowledge, generalized transduction of anything other than engineered plasmids by ϕBB-1 has not been observed.

Here, we define the genetic material packaged by ϕBB-1 virions isolated from *B. burgdorferi* strain CA-11.2A. Our proteomics studies confirm that ϕBB-1 virions are composed primarily of capsid and other phage structural proteins encoded by the cp32s; however, putative phage structural proteins encoded by lp54 were also detected. Long-read sequencing reveals that ϕBB-1 virions package a variety of genetic material including cp32 isoforms that are linearized at a region immediately upstream of the *erp* locus (o*spE/ospF/elp*) and packaged into ϕBB-1 capsids via a headful genome packaging mechanism at a packaging site (*pac*). When introduced to a shuttle vector, the *pac* region promotes the packaging of shuttle vectors into ϕBB-1 virions, demonstrating the utility of ϕBB-1 as a tool to genetically manipulate Lyme disease spirochetes. Additionally, full-length contigs of cp26, lp17, lp38, lp54, and lp56 are recovered from packaged reads as are fragments of the linear chromosome. Finally, long-read sequencing of packaged DNA allowed us to fully resolve most of the covalently closed hairpin telomeres in the *B. burgdorferi* CA-11.2A genome.

Overall, this study implicates ϕBB-1 in mobilizing large portions of the *B. burgdorferi* genome, which may explain certain aspects of genome stability and diversity observed in Lyme disease spirochetes.

Results

ϕBB-1 phage purification, virion morphology, and proteomic analysis

In the laboratory, lytic ϕBB-1 replication (**Fig 1C**) can be induced by fermentation products such as ethanol [38, 39]. We first measured ϕBB-1 titers in early stationary-phase cultures (~1 × 10^8 cells/mL) of *B. burgdorferi* B31 or CA-11.2A induced with 5% ethanol, as described by Eggers *et al.* [38]. Seventy-two hours after induction, bacteria were removed by centrifugation and filtering and virions were purified from supernatants by chloroform extraction and precipitation with ammonium sulfate. Purified virions were treated overnight with DNase to destroy DNA not protected within a capsid. Phage DNA was then extracted using a proteinase K/SDS/phenol-chloroform DNA extraction protocol [30] and qPCR was used to measure packaged cp32 copy numbers.

B. burgdorferi strain CA-11.2A consistently produced ~10 times more phage than B31 (**Fig 2A**) and was selected for further study. Imaging of purified virions collected from CA-11.2A by transmission electron microscopy reveals virions with an elongated capsid and contractile tail (**Fig 2B**), which is similar to the Myoviridae morphology of ϕBB-1 virions produced by strain B31 *in vitro* [9, 40, 41] and by a human *B. burgdorferi* isolate following ciprofloxacin treatment [42].

Mass spectrometry analysis of purified virions identified ten capsid and other

structural proteins encoded by the cp32s (**Fig 2C, Table S1**). We also detected highly conserved predicted phage capsid proteins encoded by lp54 (**Fig 2D**). While the virions we visualized all appear to have the same elongated capsid morphology, virions with a notably smaller capsid morphology have been isolated and imaged from *B. burgdorferi* CA-11.2A [29]. These observations raise the possibility that there are multiple intact phages inhabiting the CA-11.2A genome.

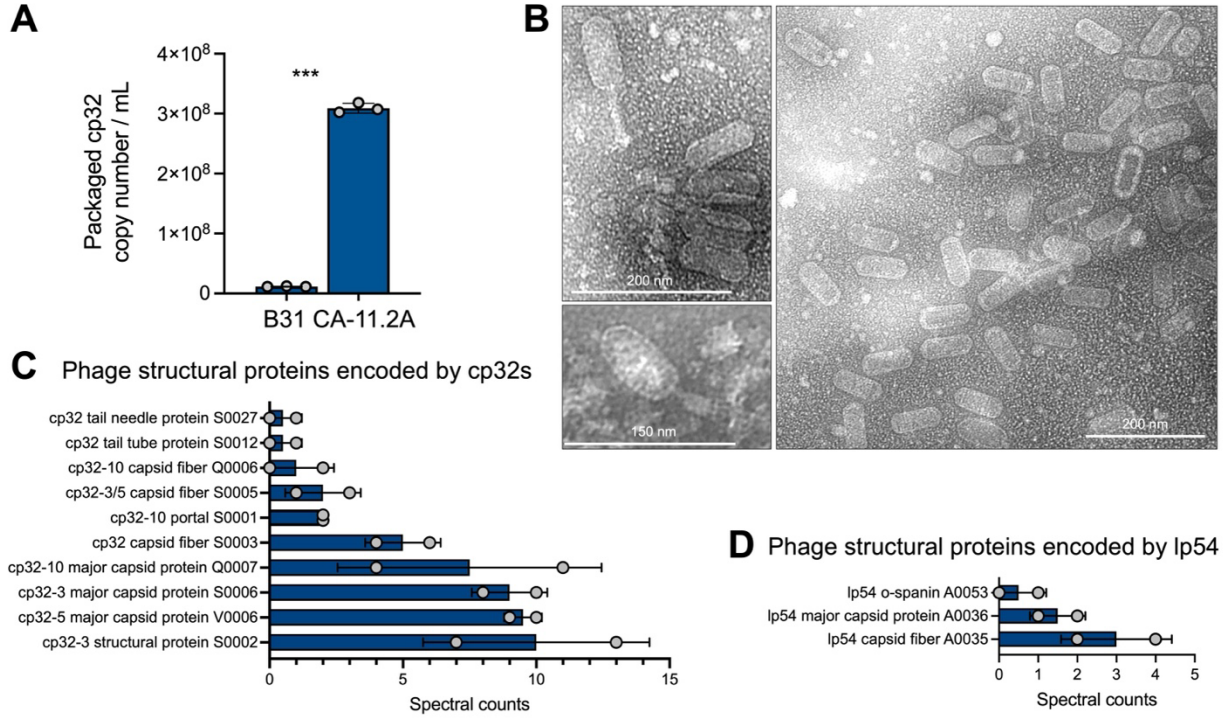


Figure 2. ϕBB-1 phage titer, virion morphology, and proteomic analysis. (A) Packaged, DNase-protected cp32 copy numbers in bacterial supernatants were measured by qPCR. Data are the SE of the mean of three experiments, ***P<0.001. **(B)** Virions were purified from 4-L cultures of *B. burgdorferi* CA-11.2A and imaged by transmission electron microscopy. Representative images from two independent preparations are shown. **(C and D)** HPLC-MS/MS-based proteomics was used to identify proteins in the two purified virion preparations. The SE of the mean of spectral counts for peptides associated with the indicated phage structural proteins are shown for each replicate. See also **Table S1** for the complete proteomics dataset.

ϕBB-1 virions package portions of the *B. burgdorferi* genome

We performed long-read sequencing on DNA packaged in purified ϕBB-1 virions, as outlined in **Figure 3**. There is concern that contaminating unpackaged *B. burgdorferi* chromosomal or plasmid DNA co-purifies with phage virions. To control for this, we spiked purified ϕBB-1 virions with high molecular weight (> 20 kb) salmon sperm DNA (**Fig 4A**) at 1.7 µg/mL, a concentration that approximates the amount of DNA released by 3×10^8 lysed bacterial cells into one milliliter of media [43]. Samples were then treated with DNase overnight followed by phage DNA extraction using a proteinase K/SDS/phenol-chloroform DNA extraction protocol [30]. Recovered DNA was directly sequenced using the Nanopore MinION (long read) platform.

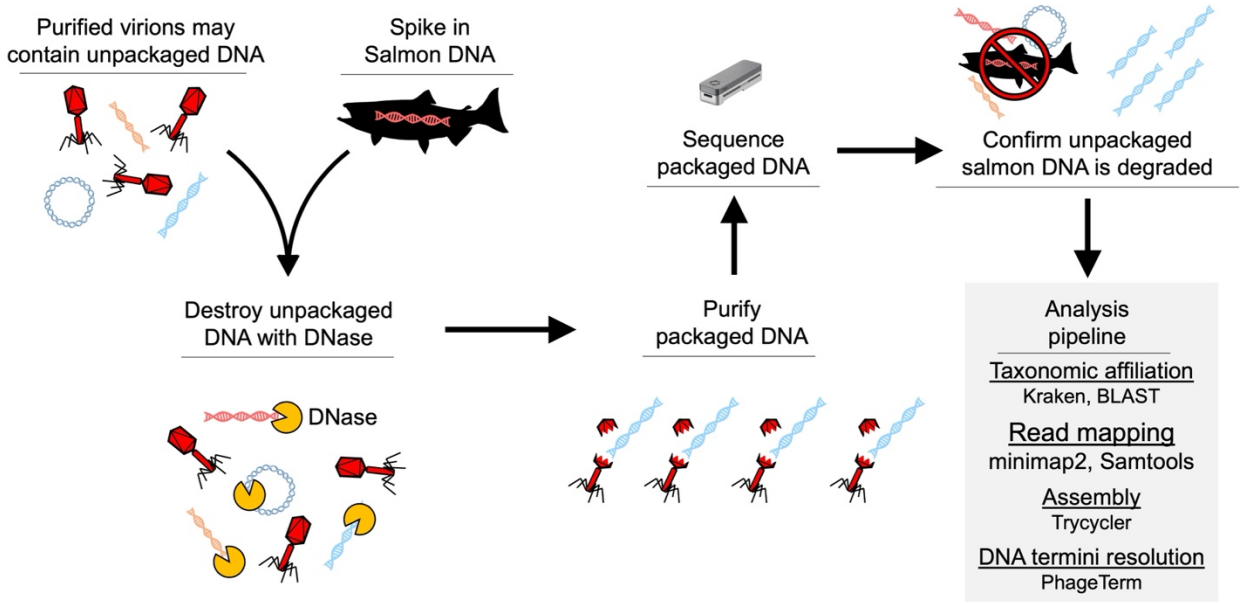


Figure 3. Workflow for sequencing packaged ϕBB-1 DNA.

Kraken [44] was used to classify prokaryotic and eukaryote sequences followed by BLAST to classify reads missed by Kraken (e-value <0.01, query coverage >10%). Reads less than 700 bp were discarded. Of 110,986 reads >700 bp, 0.14% were classified as matching salmon sequences (**Fig 4B**) with an average length of 1.2 kb (**Fig 4B**), indicating the DNase treatment step successfully degraded unpackaged DNA. To exclude reads that may be derived from unpackaged DNA, we removed all reads less than 5 kb from the dataset, leaving a total of 58,399 reads longer than 5 kb (**Fig 4C**) with a median length of 12.3 kb (**Fig 4D**). Note that cp32 prophages are approximately 32 kb in size and we detect a high number of ~32 kb reads in each replicate (**Fig 4D**, dashed line).

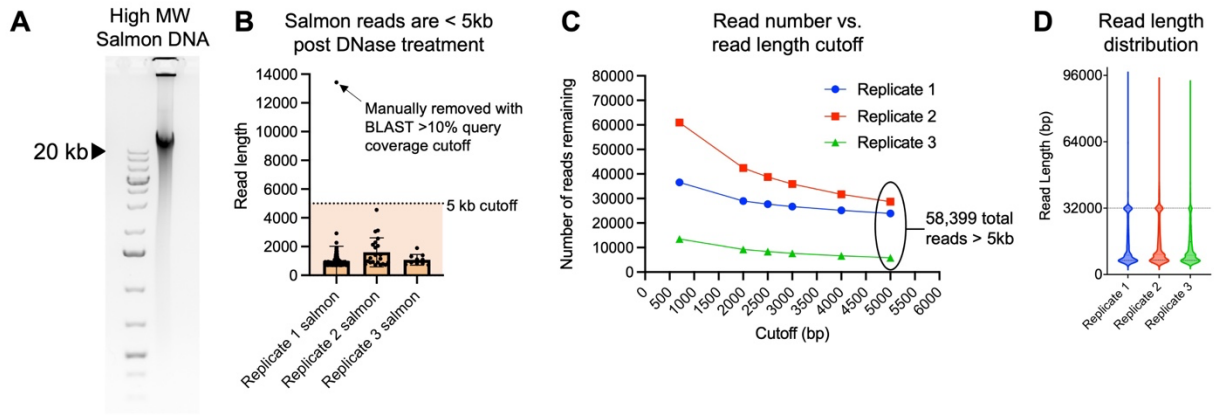


Figure 4. Establishing a 5kb read length cutoff to exclude unpackaged reads. (A) The salmon sperm DNA used to spike purified phages prior to DNase treatment was run on an agarose gel to estimate its size. Note that the majority of salmon DNA is larger than the 20-kb high molecular weight marker in the left lane. **(B)** 0.14% of 110,986 reads > 700 bp, 0.14% were classified as matching salmon sequences. Reads classified as salmon were plotted as a function of their length for each replicate. Error bars represent the SE of the mean of three replicate experiments. All reads except one (arrow) were below 5 kb in length (dashed line) with an average length of 1.2 kb. **(C)** Read length cutoff was plotted as a function of the number of reads remaining in each replicate dataset. In total, 58,399 reads remain after establishing a 5-kb cutoff. **(D)** Read length for all reads >5 kb in each replicate was plotted.

155 Overall, ~99.6% of packaged reads >5 kb were classified as *B. burgdorferi* (**Fig**
156 **5A**), the majority of which (~79%) were cp32 isoforms (**Fig 5B**). Cp32-10 and cp32-3
157 were preferentially packaged (~32% and ~25%, respectively) followed by cp32-13 and
158 cp32-5 (each at ~10%) (**Fig 5B**). Reads mapping to cp32-3, cp32-5, cp32-10, and cp32-
159 13 had a mean coverage of over 1,000× (**Fig 5C**). Cp32-1 reads accounted for only about
160 one percent of all packaged reads (**Fig 5B**) and had lower mean coverage of
161 approximately 36× (**Fig 5C**), suggesting that cp32-1 was not undergoing lytic replication.
162 Read length distributions for the cp32s indicate that full-length ~32 kb molecules were
163 frequently recovered for each cp32 isoform, except for cp32-10 (**Fig 5D**).

164 Additionally, 11.6% of reads > 5 kb mapped to the linear chromosome and ~6.3%
165 of reads >5 kb mapped to lp54 (**Fig 5B**). The remaining reads mapped to *B. burgdorferi*
166 plasmids cp26, lp17, lp36/lp28-4, lp38, lp56, and lp28-3 at 1–2% each (**Fig 5B**). *De novo*
167 assembly of packaged reads produced full-length contigs of all cp32s, lp17, cp26, lp36,
168 lp38, lp54, and lp56 (**Fig S1**), suggesting that full-length versions of these plasmids are
169 packaged by ϕBB-1.

170 Of note, the CA-11.2A genome was reported to contain a unique plasmid,
171 lp36/lp28-4, that is thought to have arisen from the fusion of lp36 with lp28-4 [45]. *De*
172 *novo* assembly of packaged reads resolved lp36/lp28-4 into individual lp36 and lp28-4
173 contigs (**Fig S1E and F**). Additionally, whole genome sequencing of our CA-11.2A strain
174 confirmed that lp36 and lp28-4 are separate as no reads that span the lp36-lp28-4 junction
175 were observed and coverage depth was notably different between lp36 and lp28-4 (~200×
176 vs. 25×, respectively, **Fig S2A**). Furthermore, PCR confirmed the sequencing results (**Fig**
177 **S2B-D**). These data indicate that the lp36/lp28-4 plasmid is two distinct episomes in our
178 CA-11.2A strain.

179 Collectively, these results indicate that in addition to cp32 molecules, ϕBB-1 is
180 capable of packaging non-cp32 portions of the *B. burgdorferi* genome. We discuss the
181 major packaged DNA species in the following sections.

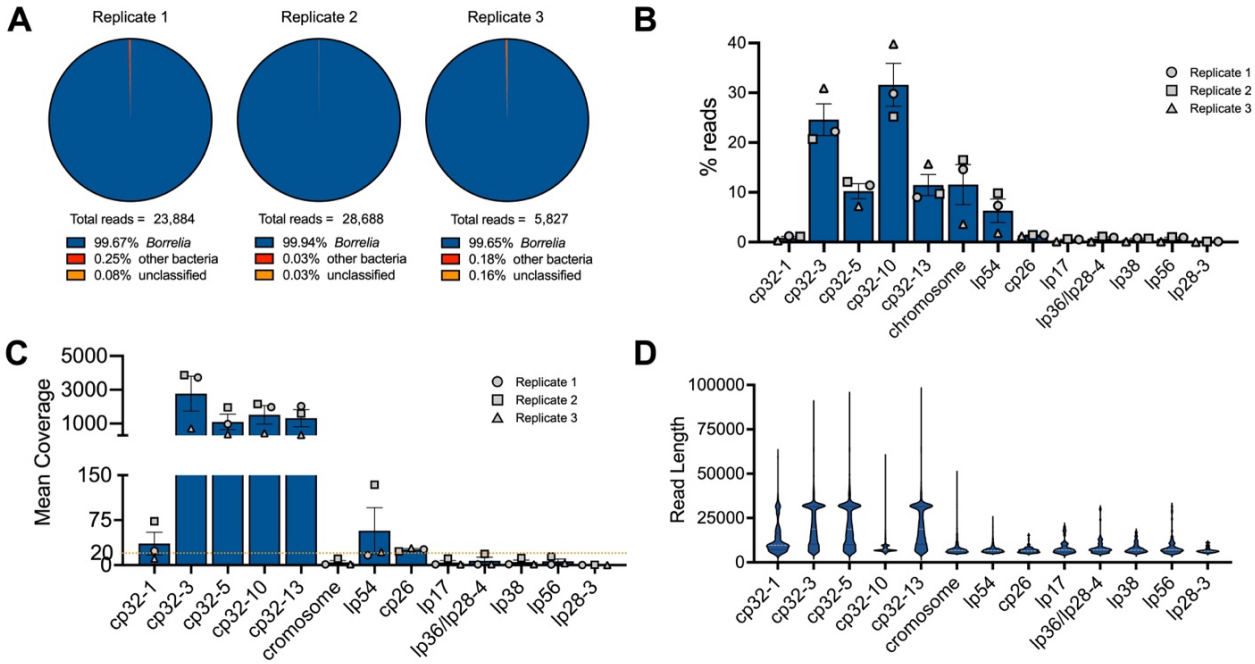


Figure 5. ϕBB-1 virions package cp32 isoforms, chromosome fragments, lp54, and other plasmids. **(A)** Kraken and BLAST were used to determine the taxonomic affiliation of reads >5kb. Note that no eukaryotic reads were identified. **(B and C)** The (B) percent and (C) mean coverage for reads affiliated with the indicated *B. burgdorferi* plasmid or linear chromosome are shown for each replicate. Error bars represent the SE of the mean. **(D)** Read length distributions for the indicated plasmids or chromosome are shown.

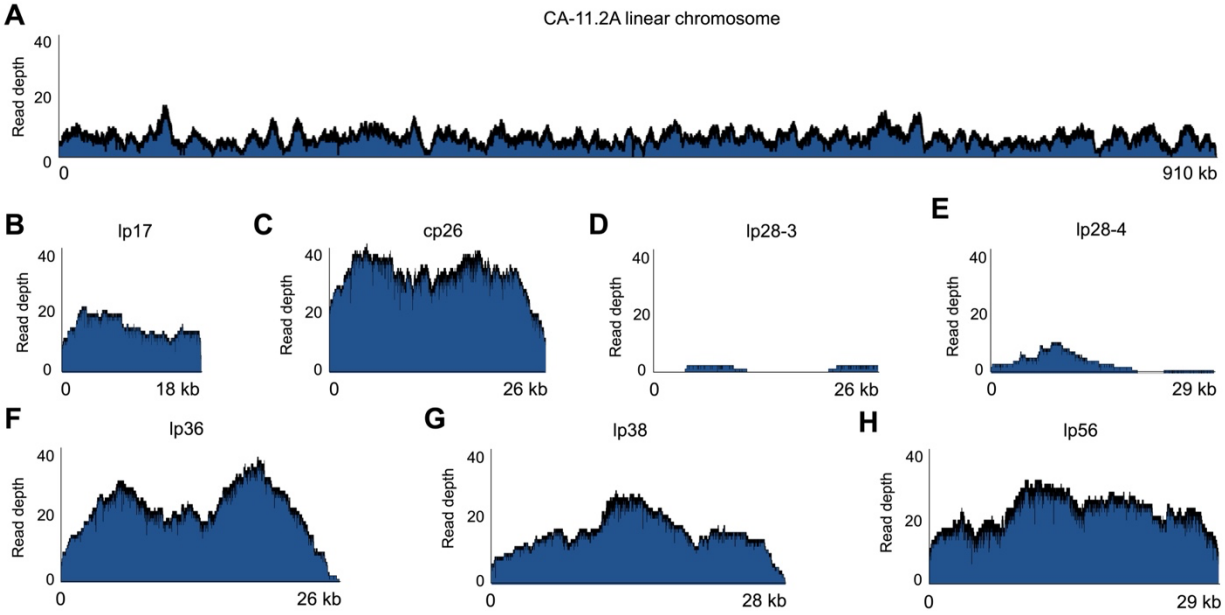


Figure S1. Packaged read depth across the *de novo* CA-11.2A genome assembly. *De novo* assembly of packaged reads >5kb produced the indicated contigs. Read coverage was then mapped to each contig. **(A–H)** Read coverage across the CA-11.2A chromosome or indicated plasmids are shown. Coverage maps for the cp32s and lp54 are shown in Figures 6 and 9, respectively.

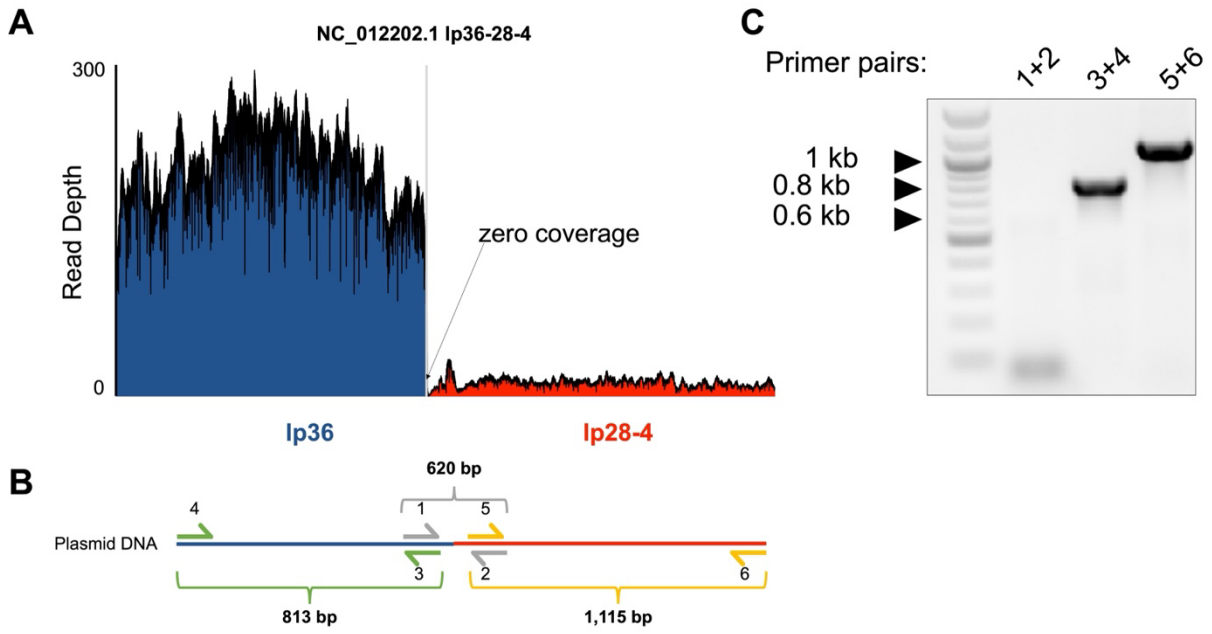
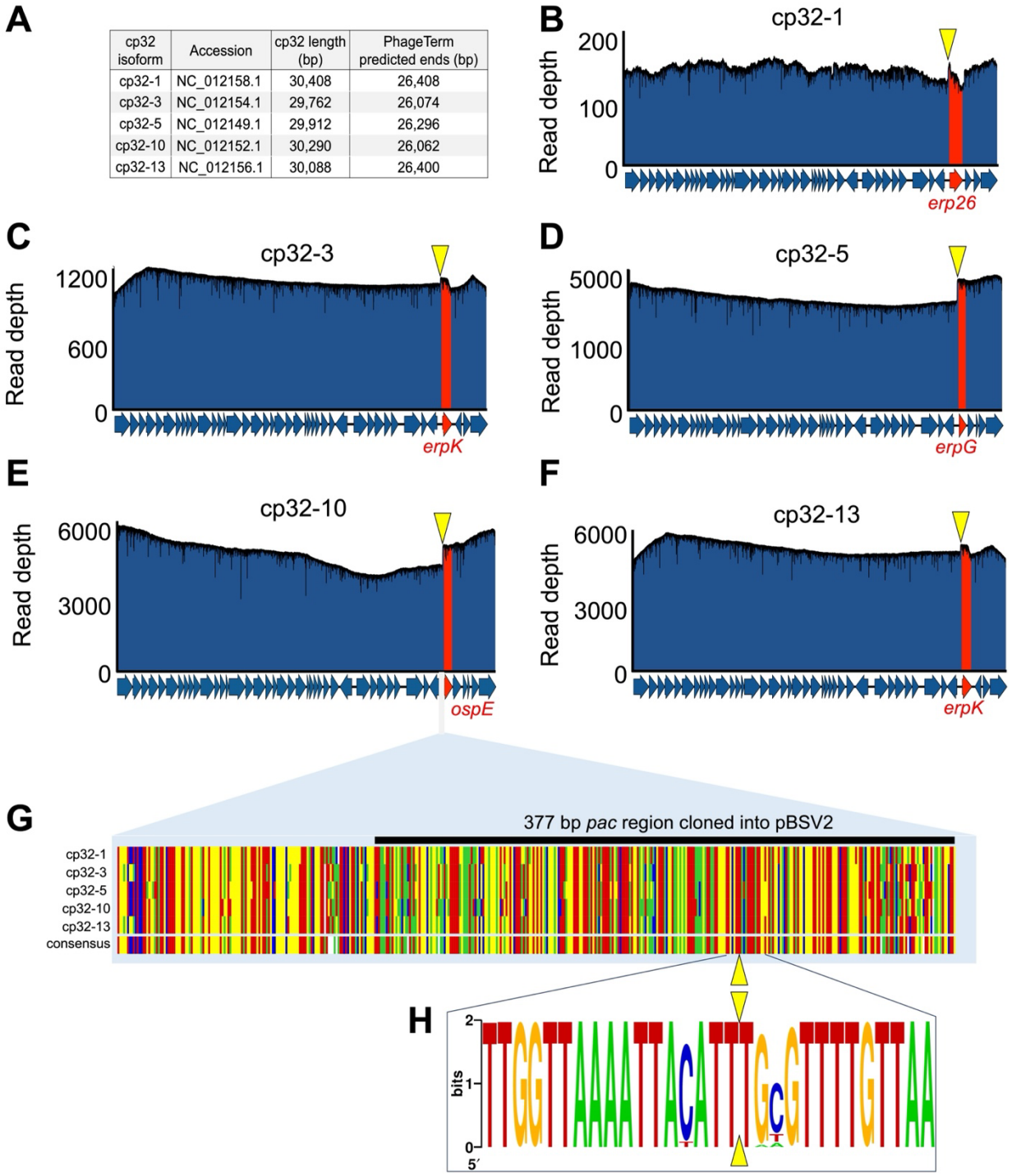


Figure S2. Whole genome sequencing of the CA-11.2A genome reveals that plasmid lp36/lp28-4 resolves into two separate episomes. (A) The CA-11.2A genome was sequenced using long-read technology. Reads were aligned to the lp36/lp28-4 reference sequence (NC_012202.1) and read depth plotted. **(B)** Schematic of PCR design. Primers 1 and 2 flank the lp36/lp28-4 junction, with primer 1 annealing to lp36 and primer 2 annealing to lp28-4, creating a 620 bp product if joined. Primers 3 and 4 anneal to lp36 DNA, creating an 813 bp product if present. Primers 5 and 6 anneal to lp28-4 DNA, creating a 1,115 bp product if present. **(C)** The presence or absence of lp36, lp28-4, or lp36/lp28-4 was confirmed by PCR.

<u>cp32 molecules are linearized near the *erp* locus and packaged via a headful mechanism</u>

Our sequencing data provide insight into how ϕBB-1 packages cp32 molecules. Many phage species package linear double-stranded DNA genomes that circularize after being injected into a host [46] and packaged cp32s are thought to be linearized [30]. We used PhageTerm [47] to predict the linear ends of packaged DNA. Native DNA termini are present once per linear DNA molecule, but non-native DNA ends produced during sequencing are distributed randomly along DNA molecules. Thus, reads that start at native DNA terminal positions occur more frequently than anywhere else in the genome. PhageTerm takes advantage of this to resolve DNA termini and predict phage packaging mechanisms [47]. PhageTerm identified the termini of packaged cp32 molecules at approximately 26 kb in a region lying immediately upstream of the *erp* (*ospE/ospF/elp*) loci (**Fig 6A**). In agreement with the PhageTerm results, when packaged reads were used to map the physical ends of packaged cp32 molecules, a sharp boundary in coverage depth is observed upstream of the *ospE/ospF/elp* loci in all cp32s (**Fig 6B–F**). Notably, the intergenic region upstream of the *erp* loci is conserved across the cp32 isoforms found in diverse strains of Lyme disease spirochetes (**Fig 6G**) [15] and the linear cp32 ends identified by long-read sequencing converge on the same terminal sequence motif (**Fig 6H**).



cp32 isoform	Accession	cp32 length (bp)	PhageTerm predicted ends (bp)
cp32-1	NC_012158.1	30,408	26,408
cp32-3	NC_012154.1	29,762	26,074
cp32-5	NC_012149.1	29,912	26,296
cp32-10	NC_012152.1	30,290	26,062
cp32-13	NC_012156.1	30,088	26,400

Fig 6. cp32s are linearized upstream of the *erp* loci. (A) PhageTerm was used to predict the linear ends of packaged cp32 molecules. **(B–F)** Nanopore reads were mapped to the indicated cp32s. Note the sharp boundary just upstream of the *erp* loci (highlighted in red). The yellow triangles indicate the PhageTerm predicted linear ends. **(G)** Alignments of the intergenic region upstream of the *erp* loci is shown for each cp32. The black line indicates the *pac* region that was cloned into a shuttle vector, as described in Figure 7. **(H)** A nucleic acid logo was constructed from 207 cp32 sequence alignments. Yellow triangles indicate the linear end of cp32 isoforms as predicted by PhageTerm and confirmed by long-read sequencing.

PhageTerm predicts that cp32s are packaged by a headful mechanism. This observation supports the previously proposed headful genome packaging mechanism for cp32s [39]. Phages that use the headful packaging mechanism generate a concatemer

containing several head-to-tail copies of their genome (**Fig 7A**). During headful packaging, a cut is made at a defined packaging site (*pac* site) and a headful (a little more than a full genome) of linear phage DNA is packaged. Once a headful is achieved, the phage genome is cut at non-defined sites, resulting in variable cut positions and size variation in packaged DNA, which we observe in packaged cp32 reads downstream of the initial cut site (**Fig 6B–F**).

Our results suggest that the cp32 *pac* site is upstream of the *erp* (*ospE/ospF/elp*) loci. If the cp32 *pac* site is in this region, then DNA molecules containing the *pac* sequence are expected to be packaged into ϕBB-1 virions. To test this, we cloned the putative cp32-3 *pac* site (**Fig 6G**, black bar) into the pBSV2 shuttle vector [48], transformed *B. burgdorferi* strain CA-11.2A, and induced lytic ϕBB-1 replication with 5% ethanol. Supernatants containing virions were collected, filtered, and DNase treated, and packaged DNA isolated. pBSV2 shuttle vector copy numbers were measured by qPCR using primers that target the pBSV2 kanamycin resistance (*kan*) cassette. To control for possible chromosomal DNA contamination, qPCR was also performed using primers targeting the *flaB* gene. Final packaged pBSV2 copy numbers were calculated by subtracting *flaB* copy numbers from pBSV2 (*kan* cassette) copy numbers.

Copy numbers of packaged pBSV2 encoding the cp32-3 *pac* site were significantly ($P<0.001$) higher compared to virions collected from cells either not carrying the pBSV2 vector or cells carrying an empty pBSV2 vector (**Fig 7B**), indicating that DNA molecules that contain the *pac* site are preferentially packaged by ϕBB-1 virions.

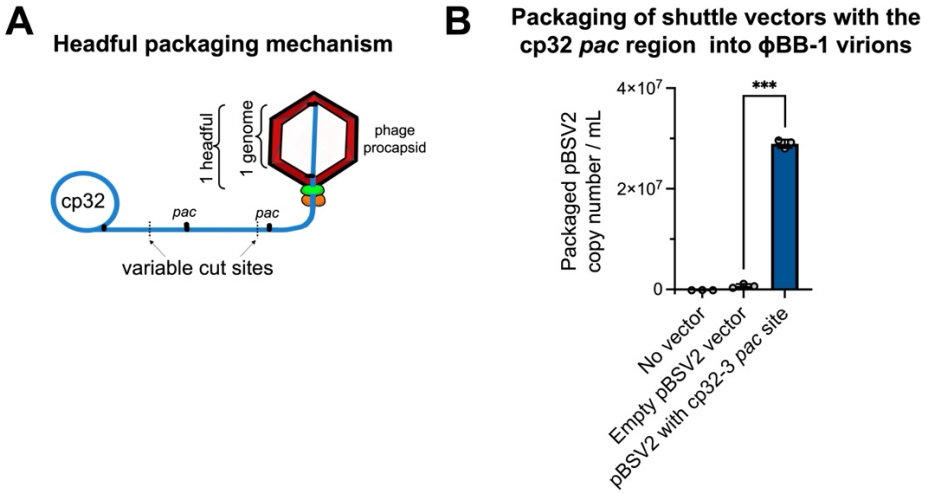


Figure 7. Shuttle vectors containing the cp32 *pac* region are preferentially packaged into ϕBB-1 virions. (A) Schematic depicting the headful genome packaging mechanism. **(B)** After ethanol induction, ϕBB-1 virions were collected from CA-11.2A cells not carrying plasmid pBSV2 (No vector), cells transformed with empty pBSV2, or cells transformed with pBSV2 with the cp32-3 *pac* site (see Fig 6G for the cloned *pac* region). Copy numbers of pBSV2 packaged into ϕBB-1 virions were measured by qPCR. Data are the SE of the mean of three experiments, ***$P<0.001$.

<u>The cp32 prophages have conserved motifs that occur in a specific arrangement not found in other DNA sequences packaged by ϕBB-1 virions</u>

To identify motif(s) that may be shared between the cp32s and other genomic

elements that are packaged into ϕBB-1 virions (*e.g.,* lp54), we first used an iterative BLAST search to identify distantly homologous DNA sequences (**Fig 8**). A non-redundant list of these diverse DNA sequences were then used as an input dataset for sequence motif discovery via MEME [49]. All five cp32 isoforms found in *B. burgdorferi* CA-11.2A have the same specific arrangement of conserved sequence motifs around the *pac* region (**Fig 8A and B**) and these are conserved in cp32 isoforms across *B. burgdorferi* (**Fig 8C**). However, significant matches to these motifs were not identified in other CA-11.2A genetic elements packaged by ϕBB-1 (**Supplementary Data file 1**), suggesting that packaging of non-cp32 DNA may occur spontaneously or through different mechanisms.

The complete or partial arrangement of motifs found around the *pac* site of *B. burgdorferi* cp32 isoforms is conserved in cp32 plasmids and some linear plasmids originating from several different *Borrelia* species (**Fig. 8C**). The iterative BLAST search also revealed that a diverse set of circular and linear plasmids in a broader set of *Borrelia* species share some of the motifs found in *B. burgdorferi* cp32 isoforms. In total, linear or circular plasmid sequences from 21 different *Borrelia* species (both Lyme disease and relapsing fever spirochetes) had homology to the *B. burgdorferi* cp32 *pac*-containing DNA sequences (**Fig 8C**). The motifs that are found most broadly, *e.g.,* blue triangle, and green square, may represent binding sites for conserved host factors that are present in all these *Borrelia* species whereas the other motifs may represent protein-binding sites or regulatory sequences that are specific to given prophage or plasmids.

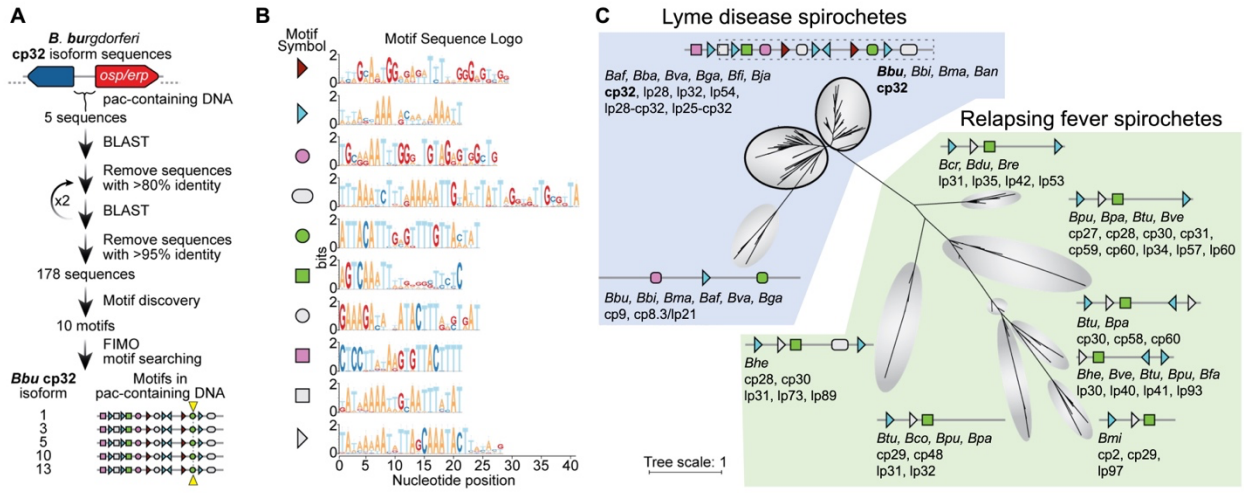


Figure 8. Cp32 prophages have conserved motifs that occur in a specific arrangement around the *pac* site. (A) Outline of bioinformatic strategy to identify motifs enriched in the *pac*-containing DNA sequence of cp32 isoforms. All *B. burgdorferi* cp32 isoforms have the same motifs in the *pac* region. The cp32 cut site is indicated by the yellow triangle. **(B)** Sequence logos of the motifs identified in panel A and schematized in panel C. Nine of the top ten motifs occur at least once in the *pac*-containing region of cp32 DNA sequences. Motifs represented with right or left facing triangles often occur as direct and/or indirect repeats. **(C)** Phylogenetic tree of non-redundant DNA sequences with homology to *B. burgdorferi* cp32 *pac*-region identified in panel A. For each clade, the bacterial species and type of plasmid are listed. For clarity in the figure, bacterial species names have been truncated to a three letter abbreviation consisting of the first letter of the genus and the first two letters of the species (*Borrelia afzelii, Baf; Borrelia andersonii, Ban; Borrelia bavariensis, Bba; Borrelia bissettiae, Bbi; Borrelia burgdorferi, Bbu; Borrelia coriaceae, Bco; Borrelia crocidurae, Bcr; Borrelia duttonii, Bdu; Borrelia fainii, Bfa; Borrelia finlandensis, Bfi; Borrelia garinii,*

Bga; Borrelia hemsii, Bhe; Borrelia japonica, Bja; Borrelia mayonii, Bma; Borrelia miyamotoi, Bmi; Borrelia parkeri, Bpa; Borrelia puertoricensis, Bpu; Borrelia recurrentis, Bre; Borrelia turicatae, Btu; Borrelia valaisiana, Bva; Borrelia venezuelensis, Bve). There is variability in the motif architecture between sequences within a single clade; however, for clarity, a representative motif architecture discovered by MEME is shown [49]. The top two clades of sequences (outlined in black) are dominated by cp32 isoforms and the cp32 motif architecture, therefore a single motif scheme is shown for these two clades. The region of DNA and motifs cloned into the pBSV2 shuttle vector is outlined in dashes.

Deciphering the structure of linear plasmids packaged by ϕBB-1

After the cp32s, lp54 is a major DNA species packaged by ϕBB-1 (**Fig 5C**). Lp54 is a linear plasmid with covalently closed telomeric ends [50]. *De novo* assembly of packaged lp54 reads produces a 67.4 kb contig consisting of full-length lp54 (54,021 bp, NC_012194.1) flanked by sequences containing tail-to-tail (7,310 bp) and head-to-head (6,074 bp) junctions (**Fig 9A**). Read depth for lp54 was >100 for most of the contig; however, read depth drops precipitously at both tail-to-tail and head-to-head junctions (**Fig 9A**), suggesting that the telomeres of lp54 interfere with sequencing.

B. burgdorferi telomeres contain inverted repeat sequences [51] and we identified the CA-11.2A lp54 inverted repeat sequence as 5′–TTTATTAGTATACTAATAAA (**Fig 9B and C**, boxed sequences). The 5′ inverted repeat of the lp54 reference sequence is missing seven nucleotides and our data extends and completes this sequence (**Fig 9B,** underlined). Further, compared to the lp54 reference sequence, the packaged 5′ and 3′ junction-spanning sequences each encode an additional 18 bp of sequence (**Fig 9B and C**). These sequences, although unique at each end (**Fig 9D**), form perfect hairpin structures (**Fig 9E and F**). Overall, these data suggest that lp54 molecules with complete telomere sequences are packaged into virions. However, whether linear lp54 with covalently closed telomeres or lp54 replication intermediates that contain head-to-head and tail-to-tail junctions are packaged is unclear.

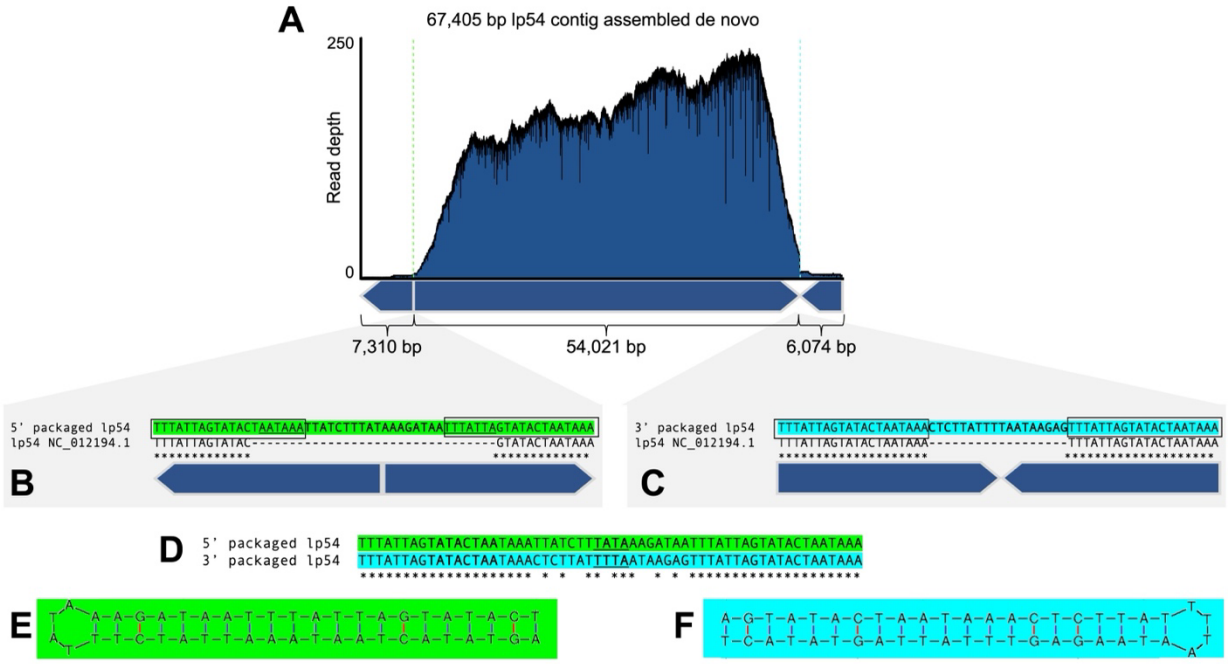


Figure 9. Full-length lp54 with fully resolved telomeres are recovered from ϕBB-1-packaged DNA. (A) *De novo* assembly of packaged reads produced a 67,405 bp contig with tail-to-tail and head-to-head

junctions. **(B and C)** Sequences at the packaged 5′ junction (green) or the 3′ junction (cyan) are compared to the lp54 reference sequence NC_012194.1. The conserved inverted repeat sequence 5′–TTTATTAGTATACTAATAAA is outlined. **(D)** Alignments of the tail-to-tail and head-to-head junctions reveals a variable 18-bp sequence in between the conserved inverted repeats. **(E and F)** Predicted hairpin structures are shown for each end of lp54. The loop sequence for each hairpin is underlined in panel D.

The *de novo* assembly approach applied to lp54 was also successful in resolving the telomeric ends of other linear elements of the CA-11.2A genome, including the linear chromosome and plasmids lp17, lp56, and lp38 (**Fig 10**). Additionally, we were able to resolve left and right telomeres for lp36 (**Fig 10**), providing yet further evidence that lp36 is not fused to lp28-4.

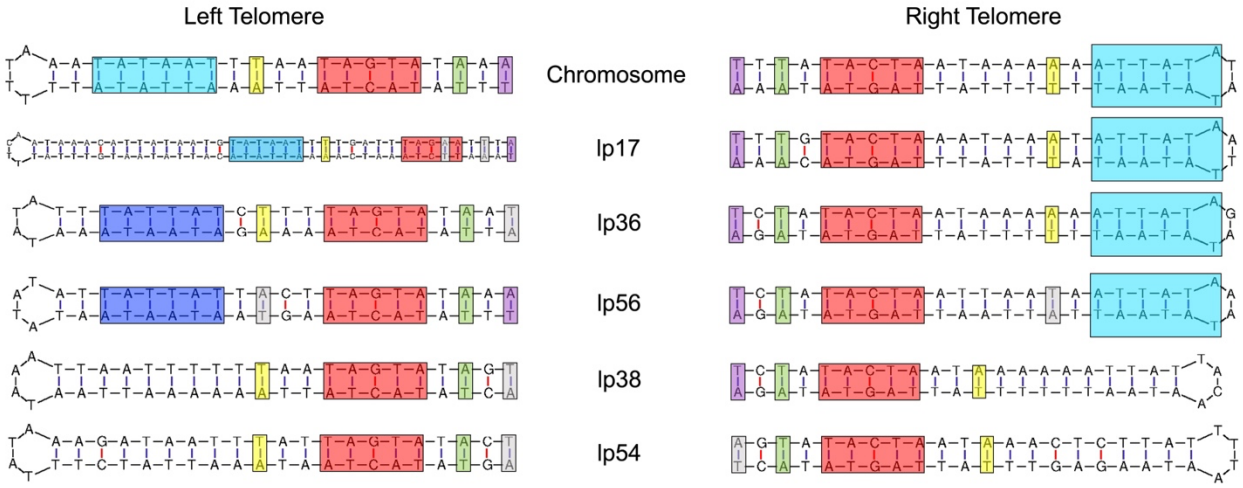


Figure 10. Packaged reads resolve the telomeric ends of the linear chromosome and most linear plasmids in the CA-11.2A genome. Reads spanning tail-to-tail or head-to-head junctions of the linear chromosome or the indicated linear plasmids form perfect hairpin structures. Conserved regulatory elements for each telomere are highlighted [52-57].

342 **Discussion**

343 In nature, Lyme disease spirochetes exist as diverse populations of closely related
344 bacteria that possess sufficient antigenic variability to allow them to co-infect and reinfect
345 non-naïve vertebrate hosts [58-69]. Moreover, horizontal gene transfer between Lyme
346 disease spirochetes has been extensively documented [70-75]. Nevertheless, the
347 mechanism underlying horizontal genetic exchange among Lyme disease spirochetes
348 remains undefined.

349 Horizontal gene transfer between heterologous spirochetes likely occurs in the tick
350 midgut during and immediately after a blood meal when spirochete replication rates and
351 densities are at their highest. ϕBB-1 replication is also induced in the tick midgut during a
352 bloodmeal [9, 27, 28] and homologous recombination between cp32 isoforms [15-17] and
353 the horizontal transfer of cp32s between *Borrelia* strains has been documented [25].
354 These observations implicate ϕBB-1 in mediating horizontal gene transfer between Lyme
355 disease spirochetes.

356 Our study further implicates ϕBB-1 in mediating horizontal gene transfer between
357 Lyme disease spirochetes. Our sequencing data indicate that ϕBB-1 virions package
358 portions of the entire *B. burgdorferi* genome, giving ϕBB-1 the potential to mobilize
359 numerous beneficial alleles during the enzootic cycle via generalized transduction. For
360 example, the circular cp32 prophages are highly conserved across the *Borrelia* genus
361 [22]; however, cp32 isoforms contain variable regions that encode outer membrane
362 lipoproteins such as Bdr, Mlp, and OspE/OspF/Elp, which are known to facilitate the *B.*
363 *burgdorferi* lifecycle [20, 22, 23, 76]. The linear plasmid lp54 encodes the outer membrane
364 lipoproteins OspA and OspB, which are required for *B. burgdorferi* to colonize the tick
365 midgut [77-79]. The outer membrane lipoprotein OspC, which is required for *B.*
366 *burgdorferi* to infect a vertebrate host, is encoded by the circular plasmid cp26 [58, 74,
367 80]. All these alleles (and many others) are packaged by ϕBB-1, which is consistent with
368 a role for phage-mediated transduction of genes encoding essential membrane
369 lipoproteins between heterologous spirochetes.

370 In *B. burgdorferi*, the linear chromosome is highly conserved as are the circular
371 plasmids cp32 and cp26 and the linear plasmids lp17, lp38, lp54, and lp56 are
372 evolutionarily stable [4-6, 16, 81]. However, other plasmids distributed across the
373 genospecies show considerably more variation and encode mostly (87%) pseudogenes
374 and are thought to be in a state of evolutionary decay [6]. The plasmids we identified as
375 packaged by ϕBB-1 virions include the cp32s, cp26, lp17, lp38, lp54, and lp56 —the same
376 plasmids that are evolutionarily stable across the genospecies [4-6, 16, 81]. These
377 observations suggest that genes encoded on ϕBB-1-packaged plasmids are under
378 positive selection, possibly as a result of transduction between Lyme disease spirochetes
379 during the enzootic cycle.

380 In addition to providing evidence that ϕBB-1 virions package large portions of the
381 *B. burgdorferi* genome, our study provides insight into ϕBB-1 virion structure and identifies
382 virion proteins present in ϕBB-1. Using mass spectrometry-based proteomics, we confirm
383 that putative capsid and other predicted structural genes encoded by cp32s such as the

384 major capsid protein P06 are indeed translated and assembled into mature ϕBB-1 virions.

385 Our long-read sequencing studies indicate that ϕBB-1 packages full-length linear
386 cp32 molecules via a headful mechanism using *pac* sites. The headful packaging
387 mechanism is used by numerous phages and was first described for *E. coli* phage T4 in
388 1967 [82]. After injecting linear DNA into a new host, the phage genome re-circularizes
389 before continuing its replication cycle. Genes encoded near the ends of linear phage
390 genomes are subject to copy number variation and recombination as the phage genome
391 re-circularizes [83]. Our data suggest that the conversion of linear cp32 molecules into
392 circular cp32 molecules occurs in the vicinity of the *erp* locus, which could facilitate
393 recombination with polymorphic *erp* alleles encoded by other cp32 isoforms in diverse *B.*
394 *burgdorferi* hosts.

395 In this study, the packaging of specific cp32 isoforms was biased: cp32-3, cp32-5,
396 cp32-10, and cp32-13 were predominantly packaged while cp32-1 was rarely packaged.
397 This result is consistent with observations by Wachter *et al.* where cp32 isoform copy
398 number and transcriptional activity were not uniform across all cp32 isoforms in *B.*
399 *burgdorferi* strain B31: cp32-1, cp32-3, and cp32-6 were predominantly induced (highest
400 copy numbers) and had the highest transcriptional activity while cp32-9 was not induced
401 and was transcriptionally inactive [9]. Variability in the *pac* region or other regulatory
402 elements involved in cp32 induction may explain why different cp32 isoforms replicate
403 and/or are packaged at different rates.

404 In the intergenic region upstream of the *erp* loci, we identified a 377-bp region that
405 contains the cp32 *pac* signal. Introducing the cp32 *pac* region to a shuttle vector facilitated
406 the packaging of the shuttle vector into ϕBB-1 virions. Identification of the cp32 *pac* site
407 will be useful for the engineering of recombinant DNA that can be packaged into virions
408 that infect spirochetes, giving ϕBB-1 the potential for use as a genetic tool.

409 After the cp32s, lp54 was the most frequently packaged plasmid. This may be
410 related to the evolutionary origins of lp54: lp54 contains large blocks of homology with the
411 cp32s and is thought to have emerged from an ancient recombination event between a
412 cp32 and a linear plasmid [6]. In addition, lp54 encodes putative phage proteins including
413 a porin (BBA74) [84] and phage capsid proteins that are highly conserved across the
414 genospecies [85], which we detected in purified virions by mass spectrometry. While we
415 observed virions with a distinct elongated capsid morphology, virions with a notably
416 smaller capsid morphology have been observed after induction *in vitro* [9, 29, 30]. These
417 observations raise the possibility that lp54 may be a prophage, although whether lp54
418 produces its own capsids or relies on cp32-encoded capsids, or if lp54 capsid proteins
419 assemble with capsid proteins encoded by cp32 to produce chimeric capsids, is not clear.

420 Our long-read dataset contained reads that spanned head-to-head and tail-to-tail
421 junctions in lp54. These reads allowed us to define the lp54 telomere sequences;
422 however, whether full-length lp54 molecules are packaged or at which stage of the
423 replication cycle lp54 is packaged is unknown. In *B. burgdorferi*, both the linear
424 chromosome and linear plasmids have covalently closed hairpin telomeres and replicate
425 via a telomere resolution mechanism [53, 55, 86, 87]. Examination of a naturally occurring

426 lp54 dimer in *B. valaisiana* isolate VS116 suggests that a circular head-to-head dimer is
427 produced during lp54 replication prior to telomere resolution and replication completion
428 [88]. Linear, covalently closed lp54 molecules may be packaged or lp54 replication
429 intermediates may be packaged.

430 As obligate vector-borne bacteria, Lyme disease spirochetes live relatively
431 restrictive lifestyles that might be expected to i) limit their exposure to novel gene pools,
432 ii) enhance reductive evolution, and iii) favor the loss of mobile DNA elements. A role for
433 ϕBB-1 in mediating the transduction of beneficial alleles between heterologous
434 spirochetes in local vector and reservoir host populations may explain why cp32
435 prophages are ubiquitous not only among Lyme disease spirochetes, but also relapsing
436 fever spirochetes.

437 **Acknowledgments**

438 We are grateful to Patti Rosa for helpful discussions and to the IDeA National
439 Resource for Quantitative Proteomics Center at the University of Arkansas for their
440 assistance with proteomic analyses of phage virions. PRS is supported by NIH grants
441 R21AI151597 and P30GM140963 (PRS). DRF is supported by NSF GRFP grant 366502.
442 A.S-F. is a M. Jane Williams and Valerie Vargo Presidential Assistant Professor of Biology
443 and is supported by the National Institutes of Health (K99GM147842, R00GM147842),
444 and by the Postdoctoral Enrichment Program Award from the Burroughs Wellcome Fund
445 (G-1021106.01). The funders had no role in study design, data collection and analysis,
446 decision to publish, or preparation of the manuscript. The authors declare no conflicts of
447 interest.

448 **Methods**

449 ϕBB-1 induction. *Borrelia burgdorferi* B31 or CA-11.2A was grown in BSK-II growth
450 medium to 7 × 10^7/mL, transferred to 50 mL tubes and centrifuged at 6,000 × *g*, 10 min,
451 35°C. Supernatants were discarded and the pellet resuspended in fresh media to a
452 density of 2 × 10^8 cells/mL. 100% EtOH was added to a final concentration of 5% (do not
453 use 95% EtOH) and incubated at 35°C for 2 hours. The culture was centrifuged at 6,000
454 x g for 10 min., 35°C. and the pellet resuspended in fresh media to a density of 5 × 10^7
455 cells/mL. The sample was incubated at 35°C for 72 hours followed by filtration through
456 0.2 µm filter. Samples were stored at 4°C or phage were concentrated using a 10 kDa
457 MW cutoff filter, replacing culture medium with SM buffer.

458 cp32 qPCR. Packaged cp32 molecules in DNase-treated supernatants were measured
459 by qPCR using primers that amplify a conserved cp32 intergenic region between
460 *bbp08* and *bbp09* (5′-CTTTACACATATCAAGACCTTAAC, 5′-
461 CAAACCACCCAATTTCCAATTCC). Primers that amplify the *Bb flaB* gene were used to
462 measure *B. burgdorferi* chromosomal DNA contamination (5′-
463 TCTTTTCTCTGGTGAGGGAGCT, 5′- TCCTTCCTGTTGAACACCCTCT) [89]. Absolute
464 cp32/ϕBB-1 copy numbers were determined using a standard curve generated with a
465 cloned copy of the target sequence. Standard curves were performed in duplicate and
466 ranged from 10^{10} copies/well down to 1 copy/well. The lower limit of detection of the assay
467 is 175 copies/mL with a 1 µL input. qPCR was performed in 10 µl reaction volumes: 5 µL
468 SYBR green supermix (BioRad), 1 µL primer 1 + primer 2 at 10 µM, 3 µL H_2O, 1 ul sample.
469 Thermocycling parameters were as follows: Step 1: 95° C 1 min, Step 2: 95° C 15 sec.,
470 Step 3: 55° C 30 sec, Step 4: go to step 2, 45 times, Step 6: Melt curve from 65°C to
471 95°C in 0.5°C increments for 5 seconds.

472 ϕBB-1 virion purification. Phages were precipitated using an ammonium sulfate-based
473 protocol [90] with some modifications. Briefly, cell-free supernatants were filtered and pH
474 adjusted to 7.5 followed by the addition of 0.5M NaCl. Phages were precipitated by adding
475 ammonium sulfate (35% w/vol, mix until completely dissolved) and incubated at 4°C for
476 18 hours. Samples were then centrifuged at 3,000 × *g* for 10 min at 4ºC. The pellicle on
477 the surface was collected and resuspended in 1 mL PBS. Where appropriate, samples
478 were dialyzed using 10-kDa MWCO cassettes against PBS at 4ºC with at least four buffer
479 exchanges.

480 ϕBB-1 virion proteomics. Purified virions (200 µg total protein) were reduced, alkylated,
481 and purified by chloroform/methanol extraction prior to digestion with sequencing grade
482 modified porcine trypsin (Promega). Peptides were separated on an Acquity BEH C18
483 column (100 x 1.0 mm, Waters) using an UltiMate 3000 UHPLC system (Thermo).
484 Peptides were eluted by a 50 min gradient from 99:1 to 60:40 buffer A:B ratio (Buffer A =
485 0.1% formic acid, 0.5% acetonitrile. Buffer B = 0.1% formic acid, 99.9% acetonitrile).
486 Eluted peptides were ionized by electrospray (2.4 kV) followed by mass spectrometric
487 analysis on an Orbitrap Eclipse Tribrid mass spectrometer (Thermo) using multi-notch
488 MS3 parameters. MS data were acquired using the FTMS analyzer over a range of 375
489 to 1500 m/z. Up to 10 MS/MS precursors were selected for HCD activation with
490 normalized collision energy of 65 kV, followed by acquisition of MS3 reporter ion data
491 using the FTMS analyzer over a range of 100-500 m/z. Proteins were identified and

492 quantified using MaxQuant (Max Planck Institute) TMT MS3 reporter ion quantification
493 with a parent ion tolerance of 2.5 ppm and a fragment ion tolerance of 0.5 Da.

494 Packaged ϕBB-1 DNA purification. Virions collected from 4-L cultures per replicate were
495 purified and DNase treated as described above. 10X DNase I reaction buffer (100 mM
496 Tris-HCl, pH 7.6, 25 mM $MgCl_2$, 5 mM $CaCl_2$) was added to each sample to a final
497 concentration of 1×. DNase I was added to a final concentration of 0.1 U/1 µL sample,
498 mixed gently, and incubated at 37°C for 18 h. EDTA was added to a final concentration
499 of 100 mM followed by SDS to a final concentration of 0.3%. Proteinase K was added to
500 a final concentration of 100 µg/mL and incubated at 65°C for 20 min. Samples were
501 extracted twice with an equal volume of phenol-chloroform-isoamyl alcohol (25:24:1)
502 followed by a single extraction with an equal volume of chloroform-isoamyl alcohol (24:1).
503 NaCl was added to 200 mM followed by 2.5 volumes of 100% EtOH. After mixing by
504 inverting the tubes, samples were incubated at -20°C for >30 min. DNA was pelleted by
505 centrifugation (14,000 × *g* for 20 min at 4°C), pellets washed with 70% EtOH, and re-spun
506 for 20 min, at 14,000 × *g* 4°C. The ethanol supernatant was discarded, and the pellet
507 gently air-dried followed by resuspension in in Tris-HCl, pH 8. Samples immediately
508 proceeded to library preparation and sequencing.

509 Long-read sequencing library preparation. Full-length DNA molecules were ligated with
510 multiplexing adapters and sequenced directly using the Nanopore MinION platform (FLO-
511 MIN112 flowcell, sequencing kit SQK-LSK112 and barcoding kit SQK-NBD112.24).
512 Quality control and adapter trimming was performed with bcl2fastq and porechop. Reads
513 were deposited in the NCBI BioProject database accession PRJNA1059007 and in
514 Supplementary Data File 2.

515 Sequence analysis pipeline. Adapter-trimmed long-reads with quality scores ≥7 were
516 used to isolate ≥ 5kb reads using Filtlong (v0.2.1). ≥5kb reads were mapped to the
517 reference *B. burgdorferi* CA-11.2A genome (RefSeq assembly: GCF_000172315.2) with
518 minimap2 (v2.26-r1175) [91]. Primary mapping reads with MAPQ >20 were isolated by
519 contig, filtered, and converted to final file formats using Samtools (v1.17) [92] and SeqKit
520 (v2.5.1) [93]. Read statistics for each replicate were graphed and viewed using GraphPad
521 Prism (v10.1.1). For each contig, *de novo* assemblies were created using Trycycler
522 (v0.5.4) [94], which relied on input assemblies from Flye (v2.9.2-b1786) [95], Raven
523 (v1.8.3) [96], and Minimap2/Miniasm/Minipolish (v2.26-r1175/v0.3-r179/v0.1.2) [91, 97].
524 The long-read *de novo* assemblies were then polished with short reads using Minipolish
525 (v0.1.2) [97]. The telomeres of the linear chromosome and linear plasmids were manually
526 identified in SnapGene (v5.3.3), and the hairpin structures were predicted by the Mfold
527 webserver (http://www.unafold.org/mfold/applications/dna-folding-form.php) [98]. The
528 terminal ends of the cp32 prophage genomes were predicted using PhageTerm through
529 the Galaxy webserver (https://galaxy.pasteur.fr/) [47], via input of the ≥5kb long-read
530 sequences. Coverage maps of the primary mapping or primary and supplementary
531 mapping reads were created by mapping ≥5kb long-reads to the *de novo* assembled CA-
532 11.2A genome or the reference *B. burgdorferi* CA-11.2A genome with Minimap2,
533 converted to final file formats using Samtools, and viewed using R (v4.3.2) and ggplot2

534 (v3.4.4).

535 *Pac* site cloning and qPCR. The shuttle vector pBSV2 [48] was digested with NotI to
536 remove the promoter region and re-ligated. The *pac* region from CA-11.2A genomic DNA
537 was amplified using primers 5′-TGGGTTGTAGAGTGGCTGTG and
538 5′-TCACCACTTGCGTAATTCTTGC and cloned into the NotI site of digested pBSV2. All
539 vectors were sequence verified by long-read sequencing (Plasmidsaurus) and
540 transformed into CA-11.2A. Each clone was grown in triplicate to mid-log in BSK, phage
541 replication was induced with ethanol, virions purified, and DNase-treated as described
542 above, followed by qPCR in triplicate using primers that target the Kan resistance gene
543 on pBSV2 (5′-CACCGGATTCAGTCGTCACT, 5′-GATCCTGGTATCGGTCTGCG, 120
544 bp product).

545 Identification of conserved motifs in *B. burgdorferi* cp32 isoforms. The roughly 430
546 nucleotides upstream of the *erp26*, *erpK*, *erpG*, *ospE* and *erpK* genes of the *B. burgdorferi*
547 CA-11.2A cp32 isoforms cp32-1, cp32-3, cp32-5, cp32-10 and cp32-13 respectively were
548 used as queries for a discontinuous MegaBLAST against the NCBI Nucleotide collection
549 database. The results from these first five BLASTs were combined and sequence hits
550 with more than 80% identity were removed with CD-HIT [99]. The resulting representative
551 sequences were used as queries for discontinuous MegaBLAST against the NCBI
552 Nucleotide collection (nt) database, and sequence hits with more than 80% identity were
553 removed with CD-HIT [99], and this process was iterated twice more for a total of three
554 MegaBLAST searches with a representative list of 80% identity query sequences. The
555 sequence hits from the final MegaBLASTs were combined and sequences with more than
556 95% identity were removed with CD-HIT [99], generating a list of 178 sequences. These
557 178 sequences were used as an input dataset for the MEME webserver [49], with custom
558 parameters of “Maximum Number of Motifs” set to “10”, and “Motif Site Distribution” set
559 to “Any number of sites per sequence”. MEME identified motifs in 160 of the input
560 sequences. The Position Weight Matrices (PWMs) of the 10 motifs identified by MEME
561 were used as inputs for FIMO [100] to search for significant sequence matches (q-value
562 < 0.001) in the *B. burgdorferi* chromosome and the *B. burgdorferi* cp32-1, cp32-3, cp32-
563 5, cp32-10, cp32-13, cp26, lp17, lp54 plasmid DNA sequences. The cp32 isoforms had
564 nine highly conserved sequence motifs, some motifs present in multiple copies and
565 arranged in a conserved architecture. The cp26, lp17, lp54 and chromosome sequences
566 did not contain this conserved architecture of nine motifs (see Supplementary Data file
567 1). The sequence logo of each motif was generated by taking the sequence fragments
568 that MEME used to make each PWM, and submitting these sequence fragments to the
569 WebLogo 3.0 webserver [101]. The iterative discontinuous MegaBLAST searches had
570 introduced eukaryotic sequence fragments into the list of 178 non-redundant sequences,
571 suggesting that the search likely reached an endpoint and found most of the related
572 sequences in the NCBI database. To generate a phylogenetic tree, eukaryotic sequence
573 fragments were first removed, and the remaining 149 non-redundant sequences were
574 aligned using the MAFFT webserver [102], with custom parameters of “Direction of
575 nucleotide sequences” set to “Adjust direction according to the first sequence”, and
576 “Strategy” set to E-INS-2. The resulting aligned sequence fasta file was used as input for

577 the IQ-TREE webserver [103, 104], with the following command-line: path_to_iqtree -s
578 149_aligned_Borrelia__plasmid_sequences.fasta.fasta -st DNA -m TEST -bb 1000 -alrt
579 1000. TreeViewer was used to display the phylogenetic tree [105].

580 **References**

581 1. Mead PS. Epidemiology of Lyme disease. Infect Dis Clin North Am.
582 2015;29(2):187-210. Epub 2015/05/23. doi: 10.1016/j.idc.2015.02.010. PubMed PMID:
583 25999219.
584 2. Adrion ER, Aucott J, Lemke KW, Weiner JP. Health care costs, utilization and
585 patterns of care following Lyme disease. PLoS One. 2015;10(2):e0116767. Epub
586 2015/02/05. doi: 10.1371/journal.pone.0116767. PubMed PMID: 25650808; PubMed
587 Central PMCID: PMC4317177.
588 3. Radolf JD, Strle K, Lemieux JE, Strle F. Lyme Disease in Humans. Curr Issues
589 Mol Biol. 2021;42:333-84. Epub 20201211. doi: 10.21775/cimb.042.333. PubMed PMID:
590 33303701; PubMed Central PMCID: PMCPMC7946767.
591 4. Schwartz I, Margos G, Casjens SR, Qiu WG, Eggers CH. Multipartite Genome of
592 Lyme Disease Borrelia: Structure, Variation and Prophages. Curr Issues Mol Biol.
593 2021;42:409-54. Epub 20201217. doi: 10.21775/cimb.042.409. PubMed PMID:
594 33328355.
595 5. Casjens SR, Mongodin EF, Qiu WG, Luft BJ, Schutzer SE, Gilcrease EB, et al.
596 Genome stability of Lyme disease spirochetes: comparative genomics of Borrelia
597 burgdorferi plasmids. PLoS One. 2012;7(3):e33280. Epub 20120314. doi:
598 10.1371/journal.pone.0033280. PubMed PMID: 22432010; PubMed Central PMCID:
599 PMCPMC3303823.
600 6. Casjens S, Palmer N, van Vugt R, Huang WM, Stevenson B, Rosa P, et al. A
601 bacterial genome in flux: the twelve linear and nine circular extrachromosomal DNAs in
602 an infectious isolate of the Lyme disease spirochete Borrelia burgdorferi. Mol Microbiol.
603 2000;35(3):490-516. Epub 2000/02/15. PubMed PMID: 10672174.
604 7. Samuels DS, Lybecker MC, Yang XF, Ouyang Z, Bourret TJ, Boyle WK, et al.
605 Gene Regulation and Transcriptomics. Curr Issues Mol Biol. 2021;42:223-66. Epub
606 20201210. doi: 10.21775/cimb.042.223. PubMed PMID: 33300497; PubMed Central
607 PMCID: PMCPMC7946783.
608 8. Fraser CM, Casjens S, Huang WM, Sutton GG, Clayton R, Lathigra R, et al.
609 Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi. Nature.
610 1997;390(6660):580-6. doi: 10.1038/37551. PubMed PMID: 9403685.
611 9. Wachter J, Cheff B, Hillman C, Carracoi V, Dorward DW, Martens C, et al. Coupled
612 induction of prophage and virulence factors during tick transmission of the Lyme disease
613 spirochete. Nat Commun. 2023;14(1):198. Epub 20230113. doi: 10.1038/s41467-023-
614 35897-3. PubMed PMID: 36639656; PubMed Central PMCID: PMCPMC9839762.
615 10. Qiu WG, Martin CL. Evolutionary genomics of Borrelia burgdorferi sensu lato:
616 findings, hypotheses, and the rise of hybrids. Infect Genet Evol. 2014;27:576-93. Epub
617 20140403. doi: 10.1016/j.meegid.2014.03.025. PubMed PMID: 24704760; PubMed
618 Central PMCID: PMCPMC4299872.
619 11. Brisson D, Drecktrah D, Eggers CH, Samuels DS. Genetics of Borrelia burgdorferi.
620 Annu Rev Genet. 2012;46:515-36. Epub 2012/09/15. doi: 10.1146/annurev-genet-
621 011112-112140. PubMed PMID: 22974303; PubMed Central PMCID:
622 PMCPMC3856702.
623 12. Qiu WG, Bosler EM, Campbell JR, Ugine GD, Wang IN, Luft BJ, et al. A population
624 genetic study of Borrelia burgdorferi sensu stricto from eastern Long Island, New York,

625 suggested frequency-dependent selection, gene flow and host adaptation. Hereditas.
626 1997;127(3):203-16. doi: 10.1111/j.1601-5223.1997.00203.x. PubMed PMID: 9474903.
627 13. Haven J, Vargas LC, Mongodin EF, Xue V, Hernandez Y, Pagan P, et al. Pervasive
628 recombination and sympatric genome diversification driven by frequency-dependent
629 selection in Borrelia burgdorferi, the Lyme disease bacterium. Genetics.
630 2011;189(3):951-66. Epub 20110902. doi: 10.1534/genetics.111.130773. PubMed PMID:
631 21890743; PubMed Central PMCID: PMCPMC3213364.
632 14. Combs MA, Tufts DM, Adams B, Lin YP, Kolokotronis SO, Diuk-Wasser MA. Host
633 adaptation drives genetic diversity in a vector-borne disease system. PNAS Nexus.
634 2023;2(8):pgad234. Epub 20230808. doi: 10.1093/pnasnexus/pgad234. PubMed PMID:
635 37559749; PubMed Central PMCID: PMCPMC10408703.
636 15. Brisson D, Zhou W, Jutras BL, Casjens S, Stevenson B. Distribution of cp32
637 prophages among Lyme disease-causing spirochetes and natural diversity of their
638 lipoprotein-encoding erp loci. Appl Environ Microbiol. 2013;79(13):4115-28. Epub
639 20130426. doi: 10.1128/AEM.00817-13. PubMed PMID: 23624478; PubMed Central
640 PMCID: PMCPMC3697573.
641 16. Casjens SR, Gilcrease EB, Vujadinovic M, Mongodin EF, Luft BJ, Schutzer SE, et
642 al. Plasmid diversity and phylogenetic consistency in the Lyme disease agent Borrelia
643 burgdorferi. BMC Genomics. 2017;18(1):165. Epub 20170215. doi: 10.1186/s12864-017-
644 3553-5. PubMed PMID: 28201991; PubMed Central PMCID: PMCPMC5310021.
645 17. Margos G, Hepner S, Mang C, Marosevic D, Reynolds SE, Krebs S, et al. Lost in
646 plasmids: next generation sequencing and the complex genome of the tick-borne
647 pathogen Borrelia burgdorferi. BMC Genomics. 2017;18(1):422. Epub 2017/06/01. doi:
648 10.1186/s12864-017-3804-5. PubMed PMID: 28558786; PubMed Central PMCID:
649 PMCPMC5450258.
650 18. Touchon M, Moura de Sousa JA, Rocha EP. Embracing the enemy: the
651 diversification of microbial gene repertoires by phage-mediated horizontal gene transfer.
652 Curr Opin Microbiol. 2017;38:66-73. Epub 20170517. doi: 10.1016/j.mib.2017.04.010.
653 PubMed PMID: 28527384.
654 19. Brissette CA, Haupt K, Barthel D, Cooley AE, Bowman A, Skerka C, et al. Borrelia
655 burgdorferi infection-associated surface proteins ErpP, ErpA, and ErpC bind human
656 plasminogen. Infect Immun. 2009;77(1):300-6. Epub 2008/11/13. doi: 10.1128/IAI.01133-
657 08. PubMed PMID: 19001079; PubMed Central PMCID: PMCPMC2612283.
658 20. Brissette CA, Cooley AE, Burns LH, Riley SP, Verma A, Woodman ME, et al. Lyme
659 borreliosis spirochete Erp proteins, their known host ligands, and potential roles in
660 mammalian infection. Int J Med Microbiol. 2008;298 Suppl 1:257-67. Epub 20080111.
661 doi: 10.1016/j.ijmm.2007.09.004. PubMed PMID: 18248770; PubMed Central PMCID:
662 PMCPMC2596196.
663 21. Stevenson B, Bono JL, Schwan TG, Rosa P. Borrelia burgdorferi erp proteins are
664 immunogenic in mammals infected by tick bite, and their synthesis is inducible in cultured
665 bacteria. Infect Immun. 1998;66(6):2648-54. Epub 1998/05/29. doi:
666 10.1128/IAI.66.6.2648-2654.1998. PubMed PMID: 9596729; PubMed Central PMCID:
667 PMCPMC108251.
668 22. Stevenson B, Zuckert WR, Akins DR. Repetition, conservation, and variation: the
669 multiple cp32 plasmids of Borrelia species. J Mol Microbiol Biotechnol. 2000;2(4):411-22.
670 Epub 2000/11/15. PubMed PMID: 11075913.
671 23. Kenedy MR, Lenhart TR, Akins DR. The role of Borrelia burgdorferi outer surface
672 proteins. FEMS Immunol Med Microbiol. 2012;66(1):1-19. Epub 20120521. doi:
673 10.1111/j.1574-695X.2012.00980.x. PubMed PMID: 22540535; PubMed Central PMCID:
674 PMCPMC3424381.

24. Stevenson B, Casjens S, Rosa P. Evidence of past recombination events among the genes encoding the Erp antigens of Borrelia burgdorferi. Microbiology (Reading). 1998;144 (Pt 7):1869-79. doi: 10.1099/00221287-144-7-1869. PubMed PMID: 9695920.
25. Stevenson B, Miller JC. Intra- and interbacterial genetic exchange of Lyme disease spirochete erp genes generates sequence identity amidst diversity. J Mol Evol. 2003;57(3):309-24. Epub 2003/11/25. doi: 10.1007/s00239-003-2482-x. PubMed PMID: 14629041.
26. Qiu WG, Schutzer SE, Bruno JF, Attie O, Xu Y, Dunn JJ, et al. Genetic exchange and plasmid transfers in Borrelia burgdorferi sensu stricto revealed by three-way genome comparisons and multilocus sequence typing. Proc Natl Acad Sci U S A. 2004;101(39):14150-5. Epub 20040916. doi: 10.1073/pnas.0402745101. PubMed PMID: 15375210; PubMed Central PMCID: PMCPMC521097.
27. Sapiro AL, Hayes BM, Volk RF, Zhang JY, Brooks DM, Martyn C, et al. Longitudinal map of transcriptome changes in the Lyme pathogen Borrelia burgdorferi during tick-borne transmission. Elife. 2023;12. Epub 20230714. doi: 10.7554/eLife.86636. PubMed PMID: 37449477.
28. Tokarz R, Anderton JM, Katona LI, Benach JL. Combined effects of blood and temperature shift on Borrelia burgdorferi gene expression as determined by whole genome DNA array. Infect Immun. 2004;72(9):5419-32. doi: 10.1128/IAI.72.9.5419-5432.2004. PubMed PMID: 15322040; PubMed Central PMCID: PMCPMC517457.
29. Eggers CH, Kimmel BJ, Bono JL, Elias AF, Rosa P, Samuels DS. Transduction by phiBB-1, a bacteriophage of Borrelia burgdorferi. J Bacteriol. 2001;183(16):4771-8. doi: 10.1128/JB.183.16.4771-4778.2001. PubMed PMID: 11466280; PubMed Central PMCID: PMCPMC99531.
30. Eggers CH, Samuels DS. Molecular evidence for a new bacteriophage of Borrelia burgdorferi. J Bacteriol. 1999;181(23):7308-13. Epub 1999/11/26. PubMed PMID: 10572135; PubMed Central PMCID: PMCPMC103694.
31. Zhang H, Marconi RT. Demonstration of cotranscription and 1-methyl-3-nitroso-nitroguanidine induction of a 30-gene operon of Borrelia burgdorferi: evidence that the 32-kilobase circular plasmids are prophages. J Bacteriol. 2005;187(23):7985-95. Epub 2005/11/18. doi: 10.1128/JB.187.23.7985-7995.2005. PubMed PMID: 16291672; PubMed Central PMCID: PMCPMC1291276.
32. Fillol-Salom A, Alsaadi A, Sousa JAM, Zhong L, Foster KR, Rocha EPC, et al. Bacteriophages benefit from generalized transduction. PLoS Pathog. 2019;15(7):e1007888. Epub 2019/07/06. doi: 10.1371/journal.ppat.1007888. PubMed PMID: 31276485; PubMed Central PMCID: PMCPMC6636781.
33. Zinder ND, Lederberg J. Genetic exchange in Salmonella. J Bacteriol. 1952;64(5):679-99. doi: 10.1128/jb.64.5.679-699.1952. PubMed PMID: 12999698; PubMed Central PMCID: PMCPMC169409.
34. Budzik JM, Rosche WA, Rietsch A, O'Toole GA. Isolation and characterization of a generalized transducing phage for Pseudomonas aeruginosa strains PAO1 and PA14. J Bacteriol. 2004;186(10):3270-3. Epub 2004/05/06. doi: 10.1128/JB.186.10.3270-3273.2004. PubMed PMID: 15126493; PubMed Central PMCID: PMCPMC400619.
35. Mandecki W, Krajewska-Grynkiewicz K, Klopotowski T. A quantitative model for nonrandom generalized transduction, applied to the phage P22-Salmonella typhimurium system. Genetics. 1986;114(2):633-57. Epub 1986/10/01. PubMed PMID: 3021574; PubMed Central PMCID: PMCPMC1202961.
36. Ebel-Tsipis J, Botstein D, Fox MS. Generalized transduction by phage P22 in Salmonella typhimurium. I. Molecular origin of transducing DNA. J Mol Biol.

1972;71(2):433-48. Epub 1972/11/14. doi: 10.1016/0022-2836(72)90361-0. PubMed PMID: 4564486.
37. Chiang YN, Penades JR, Chen J. Genetic transduction by phages and chromosomal islands: The new and noncanonical. PLoS Pathog. 2019;15(8):e1007878. Epub 20190808. doi: 10.1371/journal.ppat.1007878. PubMed PMID: 31393945; PubMed Central PMCID: PMCPMC6687093.
38. Eggers CH, Gray CM, Preisig AM, Glenn DM, Pereira J, Ayers RW, et al. Phage-mediated horizontal gene transfer of both prophage and heterologous DNA by varphiBB-1, a bacteriophage of Borrelia burgdorferi. Pathog Dis. 2016;74(9). Epub 2016/11/05. doi: 10.1093/femspd/ftw107. PubMed PMID: 27811049.
39. Eggers CH. Identification and characterization of a bacteriophage of Borrelia burgdorferi Graduate Student Theses, Dissertations, & Professional Papers 10590. 2000.
40. Rumnieks J, Fuzik T, Tars K. Structure of the Borrelia bacteriophage phiBB1 procapsid. J Mol Biol. 2023:168323. Epub 20231020. doi: 10.1016/j.jmb.2023.168323. PubMed PMID: 37866476.
41. Eggers CH, Casjens S, Hayes SF, Garon CF, Damman CJ, Oliver DB, et al. Bacteriophages of spirochetes. J Mol Microbiol Biotechnol. 2000;2(4):365-73. Epub 2000/11/15. PubMed PMID: 11075907.
42. Neubert U, Schaller M, Januschke E, Stolz W, Schmieger H. Bacteriophages induced by ciprofloxacin in a Borrelia burgdorferi skin isolate. Zentralbl Bakteriol. 1993;279(3):307-15. doi: 10.1016/s0934-8840(11)80363-4. PubMed PMID: 8219501.
43. Gillis M, De Ley J, De Cleene M. The determination of molecular weight of bacterial genome DNA from renaturation rates. Eur J Biochem. 1970;12(1):143-53. doi: 10.1111/j.1432-1033.1970.tb00831.x. PubMed PMID: 4984994.
44. Wood DE, Salzberg SL. Kraken: ultrafast metagenomic sequence classification using exact alignments. Genome Biol. 2014;15(3):R46. Epub 20140303. doi: 10.1186/gb-2014-15-3-r46. PubMed PMID: 24580807; PubMed Central PMCID: PMCPMC4053813.
45. Schutzer SE, Fraser-Liggett CM, Casjens SR, Qiu WG, Dunn JJ, Mongodin EF, et al. Whole-genome sequences of thirteen isolates of Borrelia burgdorferi. J Bacteriol. 2011;193(4):1018-20. Epub 20101008. doi: 10.1128/JB.01158-10. PubMed PMID: 20935092; PubMed Central PMCID: PMCPMC3028687.
46. Casjens SR, Gilcrease EB. Determining DNA packaging strategy by analysis of the termini of the chromosomes in tailed-bacteriophage virions. Methods Mol Biol. 2009;502:91-111. doi: 10.1007/978-1-60327-565-1_7. PubMed PMID: 19082553; PubMed Central PMCID: PMCPMC3082370.
47. Garneau JR, Depardieu F, Fortier LC, Bikard D, Monot M. PhageTerm: a tool for fast and accurate determination of phage termini and packaging mechanism using next-generation sequencing data. Sci Rep. 2017;7(1):8292. Epub 20170815. doi: 10.1038/s41598-017-07910-5. PubMed PMID: 28811656; PubMed Central PMCID: PMCPMC5557969.
48. Stewart PE, Thalken R, Bono JL, Rosa P. Isolation of a circular plasmid region sufficient for autonomous replication and transformation of infectious Borrelia burgdorferi. Mol Microbiol. 2001;39(3):714-21. doi: 10.1046/j.1365-2958.2001.02256.x. PubMed PMID: 11169111.
49. Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, et al. MEME SUITE: tools for motif discovery and searching. Nucleic Acids Res. 2009;37(Web Server issue):W202-8. Epub 20090520. doi: 10.1093/nar/gkp335. PubMed PMID: 19458158; PubMed Central PMCID: PMCPMC2703892.

50. Casjens S. Evolution of the linear DNA replicons of the Borrelia spirochetes. Curr Opin Microbiol. 1999;2(5):529-34. doi: 10.1016/s1369-5274(99)00012-0. PubMed PMID: 10508719.
51. Tourand Y, Bankhead T, Wilson SL, Putteet-Driver AD, Barbour AG, Byram R, et al. Differential telomere processing by Borrelia telomere resolvases in vitro but not in vivo. J Bacteriol. 2006;188(21):7378-86. Epub 20060825. doi: 10.1128/JB.00760-06. PubMed PMID: 16936037; PubMed Central PMCID: PMCPMC1636258.
52. Tourand Y, Kobryn K, Chaconas G. Sequence-specific recognition but position-dependent cleavage of two distinct telomeres by the Borrelia burgdorferi telomere resolvase, ResT. Mol Microbiol. 2003;48(4):901-11. doi: 10.1046/j.1365-2958.2003.03485.x. PubMed PMID: 12753185.
53. Kobryn K, Chaconas G. Hairpin Telomere Resolvases. Microbiol Spectr. 2014;2(6). doi: 10.1128/microbiolspec.MDNA3-0023-2014. PubMed PMID: 26104454.
54. Huang WM, Robertson M, Aron J, Casjens S. Telomere exchange between linear replicons of Borrelia burgdorferi. J Bacteriol. 2004;186(13):4134-41. doi: 10.1128/JB.186.13.4134-4141.2004. PubMed PMID: 15205414; PubMed Central PMCID: PMCPMC421586.
55. Chaconas G, Stewart PE, Tilly K, Bono JL, Rosa P. Telomere resolution in the Lyme disease spirochete. Embo J. 2001;20(12):3229-37. doi: 10.1093/emboj/20.12.3229. PubMed PMID: 11406599; PubMed Central PMCID: PMCPMC150187.
56. Casjens S, Murphy M, DeLange M, Sampson L, van Vugt R, Huang WM. Telomeres of the linear chromosomes of Lyme disease spirochaetes: nucleotide sequence and possible exchange with linear plasmid telomeres. Mol Microbiol. 1997;26(3):581-96. doi: 10.1046/j.1365-2958.1997.6051963.x. PubMed PMID: 9402027.
57. Tourand Y, Deneke J, Moriarty TJ, Chaconas G. Characterization and in vitro reaction properties of 19 unique hairpin telomeres from the linear plasmids of the lyme disease spirochete. J Biol Chem. 2009;284(11):7264-72. Epub 20090102. doi: 10.1074/jbc.M808918200. PubMed PMID: 19122193; PubMed Central PMCID: PMCPMC2652300.
58. Wilske B, Preac-Mursic V, Jauris S, Hofmann A, Pradel I, Soutschek E, et al. Immunological and molecular polymorphisms of OspC, an immunodominant major outer surface protein of Borrelia burgdorferi. Infect Immun. 1993;61(5):2182-91. doi: 10.1128/iai.61.5.2182-2191.1993. PubMed PMID: 8478108; PubMed Central PMCID: PMCPMC280819.
59. Probert WS, Crawford M, Cadiz RB, LeFebvre RB. Immunization with outer surface protein (Osp) A, but not OspC, provides cross-protection of mice challenged with North American isolates of Borrelia burgdorferi. J Infect Dis. 1997;175(2):400-5. doi: 10.1093/infdis/175.2.400. PubMed PMID: 9203661.
60. Barthold SW. Specificity of infection-induced immunity among Borrelia burgdorferi sensu lato species. Infect Immun. 1999;67(1):36-42. doi: 10.1128/IAI.67.1.36-42.1999. PubMed PMID: 9864193; PubMed Central PMCID: PMCPMC96274.
61. Brisson D, Vandermause MF, Meece JK, Reed KD, Dykhuizen DE. Evolution of northeastern and midwestern Borrelia burgdorferi, United States. Emerg Infect Dis. 2010;16(6):911-7. doi: 10.3201/eid1606.090329. PubMed PMID: 20507740; PubMed Central PMCID: PMCPMC3086229.
62. Nadelman RB, Hanincova K, Mukherjee P, Liveris D, Nowakowski J, McKenna D, et al. Differentiation of reinfection from relapse in recurrent Lyme disease. N Engl J Med. 2012;367(20):1883-90. doi: 10.1056/NEJMoa1114362. PubMed PMID: 23150958; PubMed Central PMCID: PMCPMC3526003.

822 63. Bhatia B, Hillman C, Carracoi V, Cheff BN, Tilly K, Rosa PA. Infection history of
823 the blood-meal host dictates pathogenic potential of the Lyme disease spirochete within
824 the feeding tick vector. PLoS Pathog. 2018;14(4):e1006959. Epub 20180405. doi:
825 10.1371/journal.ppat.1006959. PubMed PMID: 29621350; PubMed Central PMCID:
826 PMCPMC5886588.
827 64. Kurtti TJ, Munderloh UG, Hughes CA, Engstrom SM, Johnson RC. Resistance to
828 tick-borne spirochete challenge induced by Borrelia burgdorferi strains that differ in
829 expression of outer surface proteins. Infect Immun. 1996;64(10):4148-53. doi:
830 10.1128/iai.64.10.4148-4153.1996. PubMed PMID: 8926082; PubMed Central PMCID:
831 PMCPMC174350.
832 65. Piesman J, Dolan MC, Happ CM, Luft BJ, Rooney SE, Mather TN, et al. Duration
833 of immunity to reinfection with tick-transmitted Borrelia burgdorferi in naturally infected
834 mice. Infect Immun. 1997;65(10):4043-7. doi: 10.1128/iai.65.10.4043-4047.1997.
835 PubMed PMID: 9317005; PubMed Central PMCID: PMCPMC175581.
836 66. Hofmeister EK, Glass GE, Childs JE, Persing DH. Population dynamics of a
837 naturally occurring heterogeneous mixture of Borrelia burgdorferi clones. Infect Immun.
838 1999;67(11):5709-16. doi: 10.1128/IAI.67.11.5709-5716.1999. PubMed PMID:
839 10531219; PubMed Central PMCID: PMCPMC96945.
840 67. Shang ES, Wu XY, Lovett MA, Miller JN, Blanco DR. Homologous and
841 heterologous Borrelia burgdorferi challenge of infection-derived immune rabbits using
842 host-adapted organisms. Infect Immun. 2001;69(1):593-8. doi: 10.1128/IAI.69.1.593-
843 598.2001. PubMed PMID: 11119560; PubMed Central PMCID: PMCPMC97926.
844 68. Derdakova M, Dudioak V, Brei B, Brownstein JS, Schwartz I, Fish D. Interaction
845 and transmission of two Borrelia burgdorferi sensu stricto strains in a tick-rodent
846 maintenance system. Appl Environ Microbiol. 2004;70(11):6783-8. doi:
847 10.1128/AEM.70.11.6783-6788.2004. PubMed PMID: 15528545; PubMed Central
848 PMCID: PMCPMC525125.
849 69. Khatchikian CE, Nadelman RB, Nowakowski J, Schwartz I, Wormser GP, Brisson
850 D. Evidence for strain-specific immunity in patients treated for early lyme disease. Infect
851 Immun. 2014;82(4):1408-13. Epub 20140113. doi: 10.1128/IAI.01451-13. PubMed PMID:
852 24421042; PubMed Central PMCID: PMCPMC3993397.
853 70. Dykhuizen DE, Polin DS, Dunn JJ, Wilske B, Preac-Mursic V, Dattwyler RJ, et al.
854 Borrelia burgdorferi is clonal: implications for taxonomy and vaccine development. Proc
855 Natl Acad Sci U S A. 1993;90(21):10163-7. doi: 10.1073/pnas.90.21.10163. PubMed
856 PMID: 8234271; PubMed Central PMCID: PMCPMC47734.
857 71. Marconi RT, Samuels DS, Landry RK, Garon CF. Analysis of the distribution and
858 molecular heterogeneity of the ospD gene among the Lyme disease spirochetes:
859 evidence for lateral gene exchange. J Bacteriol. 1994;176(15):4572-82. doi:
860 10.1128/jb.176.15.4572-4582.1994. PubMed PMID: 7913928; PubMed Central PMCID:
861 PMCPMC196277.
862 72. Stevenson B, Barthold SW. Expression and sequence of outer surface protein C
863 among North American isolates of Borrelia burgdorferi. FEMS Microbiol Lett.
864 1994;124(3):367-72. doi: 10.1111/j.1574-6968.1994.tb07310.x. PubMed PMID:
865 7851744.
866 73. Jauris-Heipke S, Liegl G, Preac-Mursic V, Rossler D, Schwab E, Soutschek E, et
867 al. Molecular analysis of genes encoding outer surface protein C (OspC) of Borrelia
868 burgdorferi sensu lato: relationship to ospA genotype and evidence of lateral gene
869 exchange of ospC. J Clin Microbiol. 1995;33(7):1860-6. doi: 10.1128/jcm.33.7.1860-
870 1866.1995. PubMed PMID: 7665660; PubMed Central PMCID: PMCPMC228286.

871 74. Livey I, Gibbs CP, Schuster R, Dorner F. Evidence for lateral transfer and
872 recombination in OspC variation in Lyme disease Borrelia. Mol Microbiol. 1995;18(2):257-
873 69. doi: 10.1111/j.1365-2958.1995.mmi_18020257.x. PubMed PMID: 8709845.
874 75. Will G, Jauris-Heipke S, Schwab E, Busch U, Rossler D, Soutschek E, et al.
875 Sequence analysis of ospA genes shows homogeneity within Borrelia burgdorferi sensu
876 stricto and Borrelia afzelii strains but reveals major subgroups within the Borrelia garinii
877 species. Med Microbiol Immunol. 1995;184(2):73-80. doi: 10.1007/BF00221390. PubMed
878 PMID: 7500914.
879 76. Hefty PS, Brooks CS, Jett AM, White GL, Wikel SK, Kennedy RC, et al. OspE-
880 related, OspF-related, and Elp lipoproteins are immunogenic in baboons experimentally
881 infected with Borrelia burgdorferi and in human lyme disease patients. J Clin Microbiol.
882 2002;40(11):4256-65. Epub 2002/11/01. PubMed PMID: 12409407; PubMed Central
883 PMCID: PMCPMC139709.
884 77. Tilly K, Bestor A, Rosa PA. Functional Equivalence of OspA and OspB, but Not
885 OspC, in Tick Colonization by Borrelia burgdorferi. Infect Immun. 2016;84(5):1565-73.
886 Epub 20160422. doi: 10.1128/IAI.00063-16. PubMed PMID: 26953324; PubMed Central
887 PMCID: PMCPMC4862709.
888 78. Battisti JM, Bono JL, Rosa PA, Schrumpf ME, Schwan TG, Policastro PF. Outer
889 surface protein A protects Lyme disease spirochetes from acquired host immunity in the
890 tick vector. Infect Immun. 2008;76(11):5228-37. Epub 20080908. doi: 10.1128/IAI.00410-
891 08. PubMed PMID: 18779341; PubMed Central PMCID: PMCPMC2573341.
892 79. Yang XF, Pal U, Alani SM, Fikrig E, Norgard MV. Essential role for OspA/B in the
893 life cycle of the Lyme disease spirochete. J Exp Med. 2004;199(5):641-8. Epub
894 20040223. doi: 10.1084/jem.20031960. PubMed PMID: 14981112; PubMed Central
895 PMCID: PMCPMC2213294.
896 80. Wang IN, Dykhuizen DE, Qiu W, Dunn JJ, Bosler EM, Luft BJ. Genetic diversity of
897 ospC in a local population of Borrelia burgdorferi sensu stricto. Genetics. 1999;151(1):15-
898 30. doi: 10.1093/genetics/151.1.15. PubMed PMID: 9872945; PubMed Central PMCID:
899 PMCPMC1460459.
900 81. Lemieux JE, Huang W, Hill N, Cerar T, Freimark L, Hernandez S, et al. Whole
901 genome sequencing of human Borrelia burgdorferi isolates reveals linked blocks of
902 accessory genome elements located on plasmids and associated with human
903 dissemination. PLoS Pathog. 2023;19(8):e1011243. Epub 20230831. doi:
904 10.1371/journal.ppat.1011243. PubMed PMID: 37651316; PubMed Central PMCID:
905 PMCPMC10470944.
906 82. Streisinger G, Emrich J, Stahl MM. Chromosome structure in phage t4, iii. Terminal
907 redundancy and length determination. Proc Natl Acad Sci U S A. 1967;57(2):292-5. doi:
908 10.1073/pnas.57.2.292. PubMed PMID: 16591467; PubMed Central PMCID:
909 PMCPMC335503.
910 83. Beaulaurier J, Luo E, Eppley JM, Uyl PD, Dai X, Burger A, et al. Assembly-free
911 single-molecule sequencing recovers complete virus genomes from natural microbial
912 communities. Genome Res. 2020;30(3):437-46. Epub 2020/02/23. doi:
913 10.1101/gr.251686.119. PubMed PMID: 32075851; PubMed Central PMCID:
914 PMCPMC7111524.
915 84. Skare JT, Champion CI, Mirzabekov TA, Shang ES, Blanco DR, Erdjument-
916 Bromage H, et al. Porin activity of the native and recombinant outer membrane protein
917 Oms28 of Borrelia burgdorferi. J Bacteriol. 1996;178(16):4909-18. doi:
918 10.1128/jb.178.16.4909-4918.1996. PubMed PMID: 8759855; PubMed Central PMCID:
919 PMCPMC178274.

85. Kasumba IN, Tilly K, Lin T, Norris SJ, Rosa PA. Strict Conservation yet Non-Essential Nature of Plasmid Gene bba40 in the Lyme Disease Spirochete Borrelia burgdorferi. Microbiol Spectr. 2023;11(3):e0047723. Epub 20230403. doi: 10.1128/spectrum.00477-23. PubMed PMID: 37010416; PubMed Central PMCID: PMCPMC10269632.
86. Barbour AG, Garon CF. Linear plasmids of the bacterium Borrelia burgdorferi have covalently closed ends. Science. 1987;237(4813):409-11. doi: 10.1126/science.3603026. PubMed PMID: 3603026.
87. Hinnebusch J, Tilly K. Linear plasmids and chromosomes in bacteria. Mol Microbiol. 1993;10(5):917-22. doi: 10.1111/j.1365-2958.1993.tb00963.x. PubMed PMID: 7934868.
88. Marconi RT, Casjens S, Munderloh UG, Samuels DS. Analysis of linear plasmid dimers in Borrelia burgdorferi sensu lato isolates: implications concerning the potential mechanism of linear plasmid replication. J Bacteriol. 1996;178(11):3357-61. doi: 10.1128/jb.178.11.3357-3361.1996. PubMed PMID: 8655522; PubMed Central PMCID: PMCPMC178094.
89. Pahl A, Kuhlbrandt U, Brune K, Rollinghoff M, Gessner A. Quantitative detection of Borrelia burgdorferi by real-time PCR. J Clin Microbiol. 1999;37(6):1958-63. doi: 10.1128/JCM.37.6.1958-1963.1999. PubMed PMID: 10325354; PubMed Central PMCID: PMCPMC84995.
90. Armon R, Arella M, Payment P. A highly efficient second-step concentration technique for bacteriophages and enteric viruses using ammonium sulfate and Tween 80. Can J Microbiol. 1988;34(5):651-5. doi: 10.1139/m88-107. PubMed PMID: 2850100.
91. Li H. Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics. 2018;34(18):3094-100. doi: 10.1093/bioinformatics/bty191.
92. Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, et al. Twelve years of SAMtools and BCFtools. GigaScience. 2021;10(2). doi: 10.1093/gigascience/giab008.
93. Shen W, Le S, Li Y, Hu F. SeqKit: A Cross-Platform and Ultrafast Toolkit for FASTA/Q File Manipulation. PLOS ONE. 2016;11(10):e0163962. doi: 10.1371/journal.pone.0163962.
94. Wick RR, Judd LM, Cerdeira LT, Hawkey J, Méric G, Vezina B, et al. Trycycler: consensus long-read assemblies for bacterial genomes. Genome Biology. 2021;22(1). doi: 10.1186/s13059-021-02483-z.
95. Kolmogorov M, Yuan J, Lin Y, Pevzner PA. Assembly of long, error-prone reads using repeat graphs. Nature Biotechnology. 2019;37(5):540-6. doi: 10.1038/s41587-019-0072-8.
96. Vaser R, Šikić M. Time- and memory-efficient genome assembly with Raven. Nature Computational Science. 2021;1(5):332-6. doi: 10.1038/s43588-021-00073-4.
97. Wick RR, Holt KE. Benchmarking of long-read assemblers for prokaryote whole genome sequencing. F1000Research. 2021;8:2138. doi: 10.12688/f1000research.21782.4.
98. Zuker M. Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Research. 2003;31(13):3406-15. doi: 10.1093/nar/gkg595.
99. Fu L, Niu B, Zhu Z, Wu S, Li W. CD-HIT: accelerated for clustering the next-generation sequencing data. Bioinformatics. 2012;28(23):3150-2. Epub 20121011. doi: 10.1093/bioinformatics/bts565. PubMed PMID: 23060610; PubMed Central PMCID: PMCPMC3516142.

100. Grant CE, Bailey TL, Noble WS. FIMO: scanning for occurrences of a given motif. Bioinformatics. 2011;27(7):1017-8. Epub 20110216. doi: 10.1093/bioinformatics/btr064. PubMed PMID: 21330290; PubMed Central PMCID: PMCPMC3065696.
101. Crooks GE, Hon G, Chandonia JM, Brenner SE. WebLogo: a sequence logo generator. Genome Res. 2004;14(6):1188-90. Epub 2004/06/03. doi: 10.1101/gr.849004. PubMed PMID: 15173120; PubMed Central PMCID: PMCPMC419797.
102. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol. 2013;30(4):772-80. Epub 20130116. doi: 10.1093/molbev/mst010. PubMed PMID: 23329690; PubMed Central PMCID: PMCPMC3603318.
103. Nguyen LT, Schmidt HA, von Haeseler A, Minh BQ. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol. 2015;32(1):268-74. Epub 20141103. doi: 10.1093/molbev/msu300. PubMed PMID: 25371430; PubMed Central PMCID: PMCPMC4271533.
104. Minh BQ, Nguyen MA, von Haeseler A. Ultrafast approximation for phylogenetic bootstrap. Mol Biol Evol. 2013;30(5):1188-95. Epub 20130215. doi: 10.1093/molbev/mst024. PubMed PMID: 23418397; PubMed Central PMCID: PMCPMC3670741.
105. Bianchini G, Sánchez-Baracaldo P. TreeViewer - Cross-platform software to draw phylogenetic trees. 2023. doi: https://doi.org/10.5281/zenodo.7768343.