Supporting Information of Di *et al*.

S1. Design of universal primers for full-length amplification of *ospC*

The design of PCR primers for amplifying the full-length *ospC* sequences is aided by a previous study of intergenic sequences in *Borreliella* genomes (1).

The following is an alignment of the upstream region of *ospC* in 9 *Borreliella* species and of a *vsp* locus in *Borrelia miyamotoii* (genome accession CP017137, locus tag AXH25_04790) (2). All sequences end with the start codon "ATG". The forward primer region (in red) is 100% conserved among the *Borreliella* species, while differing at 6 positions from the *Borrelia miyamotoii* locus. This region includes the conserved ribosomal binding site (RBS, "GGAGG", underlined).

```
CP017137:13840-13920        CAAAGTTTTAACTATTTTGTCG--TTATTAATGTA-AAGGAACAAGGAGGCATATAATATG
BVAVS116_B0017-B0018        GAAAA----AACAAAATTGTTGAACTAATAATTCA---ATAAAAAGGAGGCACAAATTATG
BB_B18-B19                  TGAAA----AACAAAATTGTTGGACTAATAATTCATAAATAAAAAGGAGGCACAAATTATG
BSV1_B18-B19                TGAAA----AACAAAATTGTTGAACTAATAATTCATAAATAAAAAGGAGGCACAAATTATG
BGB17-BGB18                 TGAAA----AGAAAAATTGTTGAACTAATAATTCAT--ATAAAAAGGAGGCACAAATTATG
BSPA14S_B0019-B0021         TGAAA----AGAAAAATTGTTGAACTAATAATTCAT--ATAAAAAGGAGGCACAAATTATG
BGAFAR04_B0017-B0018        TGAAA----AGAAAAATTGTTGAACTAATAATTTAT--ATAAAAAGGAGGCACAAATTATG
BafACA1_B18-B19             TGAAA----AGAAAAATTGTTGGACTATTAATTC----ATAAAAAGGAGGCACAAATTATG
BbiDN127_B0018-B0019        TGAAA----AGTAAAATTGTTGGACTAATAATTCATAAATAAAAAGGAGGCACAAATTATG
KK9_2022-2023               TGAAA----AGAAAAATTGTTGGACTAATAATTCAT--ATAAAAAGGAGGCACAAATTATG
                              **     *   *  **** *   ** ****        ** ********* * * ****
```

The following is the alignment of a downstream region of *ospC* in *Borreliella* species and of a *vsp* locus in *B. miyamotoii* (genome accession CP017137, locus tag AXH25_04790) (2). All sequences start with the stop codon “TAA”. The reverse primer region (in red) is nearly 100% conserved among the *Borreliella* species as well as between the *Borreliella* species and *Borrelia miyamotoii*.

```
CP017137:14547-14816      TAATGGTTAAT-----CTTAATAAGGTAAGGGAAAAAGTTAATTTTAGAAGTTATAAGAT
BVAVS116_B0018-B0020      TAA---TTTAG-----ATT-ATA----------------------------TTATAAGAT
BbiDN127_B0019-B0020      TAA---TTAAG-----ATCAATA----------------------------TTATAAGAT
BSPA14S_B0021-B0022       TAA---TCAAG-----ATCAATA----------------------------TTATAAGAT
BafPKo_B0019-B0020        TAA---TTAG------ATCAATA----------------------------TTATAAGAT
BB_B19-B22                TAA---TTAAG-----ATCAATA----------------------------TTATAAGAT
BSV1_B19-B20              TAA---TTAAG-----ATCAATA----------------------------TTATAAGAT
KK9_2023-2024             TAA---TTAAGATCAAATCAATA----------------------------TTATAAGAT
BGB18-BGB19               TAA---TTAAGATCAAATTAATA----------------------------TTATAAGAT
BGAFAR04_B0018-B0020      TAA---TTAAG-----GTC-ATA----------------------------ATATAAGAT
                          ***   *          *  ***                              ********

CP017137:14547-14816      TAGTTTTTTAATTAAAAGTAAGTAACTGG-AAAAATAAAGTCAATAA--GAAGGAAGCTA
BVAVS116_B0018-B0020      TAATTTGT---TTTAAAA-AAGTAACTGGAAAAAATAAAGTCAATAT--AAAGCCAAGAA
BbiDN127_B0019-B0020      TAATTTGT---TTTAAAA-AAGTAACTGG-AAAAATAAAGTCAATAT---AAGTCAAGAA
BSPA14S_B0021-B0022       T-----------TTAAAA-AAGTAACTGGAAAAAATAAAGTCAATAT--AGAGTCAAGAA
BafPKo_B0019-B0020        TAATTTGT---TTTAAAA-AAGTAACTGG-AAAAATAAAGTCAATAT--AGAGTCAAGAA
BB_B19-B22                TAATTTGT---TTTAAAA-AAGTAACTGG-AAAAATAAAGTCAATAT---AAGTCAAGAA
BSV1_B19-B20              TAATTTGT---TTTAAAA-AAGTAACTGG-AAAAATAAAGTCAATAT---AAGTCAAGAA
KK9_2023-2024             -----------TTTAAAA-AAGTAACTGG-AAAAATAAAGTCAATAT--AGAGTCAAGAA
BGB18-BGB19               -----------TTTAAAA-AAGTAACTGG-AAAAATAAAGTCAATATAGAGAGTCAAGAA
BGAFAR04_B0018-B0020      TAATTTGT---TTTAAAA-AAGTAACTGG-AAAAATAAAGTCAATAT--AGAGTCAAGAA
                                    * ***  ********** ****************     **  *   *
```

S2. Reference *ospC* sequences used for allele identification

>B3_F006
ATGAAAAAGAATACATTAAGTGCAATATTAATGACTTTATTTTTATTTATATCTTGTAATAATTCAGGAAAAGATGG
GAATGCATCTGTAAATTCTGCTGATGAGTCTGTTAAAGGGCCTAATCTTACAGAAATAAGTAAAAAAATTACAGATT
CTAACGCAGTTGTTCTGGCTGTGAAAGAAATTGAAACTTTGCTTTTATCTATAGATGAACTTGCTAAAAAATCCCTT
GGTCAAAGAATAGATCAGGGCAATAAATTGGTTGCCGATGCGAATCACAATGGAGCATTATTGGCAGGAGTTTATAC
AATATCAACTCTAATAACGGAAAAATTAAATGGACTAACAATTTCAGAAGGTTTAAAAGAAAAGGTTGAAAAGGCTA
AGAAATACTCTTTAGACTTTACTAAAAAACTATCAGATAGTCACGCAGAACTTGGTCCAGCGGGTGGAGCTGTTACT
GATGAGAATGCAAAAAAAGCTATTTTAATAACAGATGCAACTAAAGATAAGGGCGCTGCAGAGCTTGAAAAGTTATT
TGAATCAGTAAAAAACTTGTCAAAAGCAGCTCAAGAAACACTAAATAATTCAGTTAAAGAACTTACAAGTCCTGTTG
TGGCAAAAACTCCAAAAAAACCTTAA
>N_F004
ATGAAAAAGAATACATTAAGTGCAATATTAATGACTTTATTTTTATTTATATCTTGTAATAATTCAGGGAAAGATGG
GAATGCATCTACAAATTCTGCCGATGAGTCTGTTAAAGGGCCTAATCTTACAGAAATAAGTAAAAAAATTACAGAAT
CTAATGCAGTTGTACTGGCTGTAAAAGAAGTTGCGGCGTTGCTTTCATCTATAGATGAGCTTGCTAAAGCTATTGGT
AAAAAAATAAATAATAATGGTTTAGATGATGTGCAAAACTTCAACGCATCATTATTGGCAGGAGCTCATACAATATC
AAAATTAGTAACAGAAAAATTAAGCAAATTGAAAAATTCAGAAGGATTAAAAGAAAAAATTGAGGACGCCAAAAAAT
GTTCTGATGATTTTACTAAAAAACTACAATCTAGCCATGCACAGCTTGGTGTTGCTGGTGGTGCTACTACTGATGAA
GAGGCTAAAAAAGCTATTTTAAGAACAAACGCAATTAAAGATAAGGGCGCAGATGAACTTGAAAAGTTATTTAAATC
AGTAGAAAGCTTAGCAAAAGCAGCTCAAGACGCACTAGCCAATTCAGTTAACGAGCTTACAGGTCCTGTTGTGGCAG
AAACTCCAAAAAAACCTTAA
>T_F128
ATGAAAAAGAATACATTAAGTGCAATATTGATGACTTTATTTTTATTTATATCTTGTAATAATTCAGGAAAAGATGG
GAATGCATCTGTAAATTCTGCTGATGAGTCTGTTAAAGGGCCTAATCTTACAGAAATAAGTAAAAAAATTACAGAAT
CTAACGCAGTTGTTCTCGCCGTGAAAGAAGTTGAAACTCTGCTTGCATCTATAGATGAACTTGCTAATAAAGCTATT
GGTCAAAAAATAGATCAAAATAATGGTTTGAGTGTCGATGCAGGCCATAATGGACCATTGTTAGCAGGAGCCTACGC
AATATCAGCTCTAATAACAGAAAAATTAAATGGATTAACAATCTCAGAAGAATTAAAGGAAAAGATTGAAAAGGCTA
AAAAATGTTCTACAGGCTTTACCAATAAACTAAAAAGTGGTCATGCAGAGCTTGGTCCAGTGGGTGGGAATGCTACT
GATGAAAATGCAAAACAAGCTATTTTAAAAACTCATGGAAATGTTACTAAGGGCGCTAAAGAGCTTAAAGATTTATC
CGAATCAGTGGAAGCCTTAGCAAAAGCAGCTCAAGCAATGCTAACTAATTCAGTTAAAGAGCTTACAAGTCCTGTTG
TGGCAGAAACTCCAAAAAAACCTTAA
>vsp_N030
AGAAAAACGTTAAGTGCGATAATAATGACTTTATTTTTAATAATAAATATTGTAATGATATCATGTGGAAGTGGGGG
ACCGGCACCTAAGGAAGGGCAGGCAGCAAAGGCTGATGGGACAGTGATTGATTTAAAGGCGGTAAGTAAAAAGATAA
AAGATGCGGTTGAGTTTGCAGCAAGTGTAAAAGAAGTGCATACTTTAGTTAGGTCAATAGATGTGCTTGCTAAAGGG
ATTGGAAAGAAAATCAAAAATGCAGATGAACTTGATACTGTGTCTGACAAGAATGGGACATTAGTTACAGCAGTATT
TAGTTTAATGCTTGATATAAAAACCAAATTGACAAAATTGGAGACAGAATCTGATAAATTTGATGGAATGAAGACAA
AGGTTGCTGCTGCTAAAAGTGAATGTGAAAAATTCATAGCAACAGTGAAAAGCAAAAATACTGATCTTGGTAAAGAT
GGTGTTACTGATACACATGCACAAGAAGTTATGGACATAACATCTAAGCCTTCTGGAGATAAAGGAGCTGAGGAGCT
TGTTAAATTAAACACAGAAATTGGAAAATTATTAACAGCTGCTGATGAGTTAGCAGAAGAGACAATCAAATATCTTA
CAACTCCTTCTACTAAGGTAGCGACCTAA
>C14_N150
ATGAAAAAGAATACATTAAGTGCAATATTGATGACTTTATTTTTATTTATATCTTGTAATAATTCAGGGAAAGATGG
CAATTCTGCATCTAATAATTCTGCTGATGAGTCCGCTAAAGGGCCTAATCTTATAGAAATAAGTAAAAAAATTACAG
ACTCTAATGCAGTTGTACTCGCCGTTAAAGAAATTGAAACTTTGATTTCATCTATAGATGAACTTGCCAATAAAGCC
ATTGGTAAAAGAATACAAGCAAATGGCTTAGAGAACATGCCAAATGAGAACGGATCATTATTAGCAGGAGCTTATGC
AATATCAACTTTAATAACACAAAAATTAGATGGATTGAAAAATGAAGAATTAAAAGAAAAGATTGCCGCAGCTAAGA
AGTGCTCCGAAGAATTTGGTACTAAGCTAAAAAGTTCTAACGTAAATCTTGGGCCAGTGAATGGAGACGCTACTGAC
GAACATGCAAAACAAGCTATTTTAAAAACAAATGGAACTAAAGATAAGGGTGTTGACGAACTTTTAAAGTTATCTGA
AGCAGTAGAAGGCTTGGCAAAAGCATCTCAAGCCGTACTAATTAATGCAGTTAAAGAGCTTACAAAGCCTGTTGTGA
CAGAAGCTCCAAAAAAACCTTAA
>A_B31
ATGAAAAAGAATACATTAAGTGCAATATTAATGACTTTATTTTTATTTATATCTTGTAATAATTCAGGGAAAGATGG
GAATACATCTGCAAATTCTGCTGATGAGTCTGTTAAAGGGCCTAATCTTACAGAAATAAGTAAAAAAATTACGGATT

CTAATGCGGTTTTACTTGCTGTGAAAGAGGTTGAAGCGTTGCTGTCATCTATAGATGAAATTGCTGCTAAAGCTATT
GGTAAAAAAATACACCAAAATAATGGTTTGGATACCGAAAATAATCACAATGGATCATTGTTAGCGGGAGCTTATGC
AATATCAACCCTAATAAAACAAAAATTAGATGGATTGAAAAATGAAGGATTAAAGGAAAAAATTGATGCGGCTAAGA
AATGTTCTGAAACATTTACTAATAAATTAAAAGAAAAACACACAGATCTTGGTAAAGAAGGTGTTACTGATGCTGAT
GCAAAAGAAGCCATTTTAAAAACAAATGGTACTAAAACTAAAGGTGCTGAAGAACTTGGAAAATTATTTGAATCAGT
AGAGGTCTTGTCAAAAGCAGCTAAAGAGATGCTTGCTAATTCAGTTAAAGAGCTTACAAGCCCTGTTGTGGCAGAAA
GTCCAAAAAAACCTTAA
>E_N40
ATGAAAAAGAATACATTAAGTGCAATATTAATGACTTTATTTTTATTTATATCTTGTAATAATTCAGGAAAAGATGG
GAATGCATCTGCAAATTCTGCTGATGAGTCTGTTAAAGGGCCTAATCTTACAGAAATAAGTAAAAAAATTACAGAAT
CTAACGCAGTTGTTCTGGCCGTGAAAGAAGTTGAGACCTTACTTGCATCTATAGATGAACTTGCTACCAAAGCTATT
GGTAAAAAAATAGGCAATAATGGTTTAGAGGCCAATCAGAGTAAAAACACATCATTGTTATCAGGAGCTTATGCAAT
ATCTGACCTAATAGCAGAAAAATTAAATGTATTGAAAAATGAAGAATTAAAGGAAAAGATTGATACAGCTAAGCAAT
GTTCTACAGAATTTACTAATAAACTAAAAAGTGAACATGCAGTGCTTGGTCTGGACAATCTTACTGATGATAATGCA
CAAAGAGCTATTTTAAAAAAACATGCAAATAAAGATAAGGGTGCTGCAGAACTTGAAAAGTTATTTAAAGCGGTAGA
AAACTTATCAAAAGCAGCTCAAGACACATTAAAAAATGCTGTTAAAGAGCTTACAAGTCCTATTGTGGCAGAAAGTC
CAAAAAAACCTTAA
>C_JD1
ATGAAAAAGAATACATTAAGTGCAATATTAATGACTTTATTTTTATTTATATCTTGTAATAATTCAGGAAAAGATGG
GAATGCATCTGCAAATTCTGCTGATGAGTCTGTTAAAGGGCCTAATCTTACAGAAATAAGTAAAAAAATTACAGAAT
CTAACGCAGTTGTTCTCGCCGTGAAAGAAGTTGAGACCTTACTTGCATCTATAGATGAACTTGCTAAAGCTATTGGT
AAAAAAATAAAAAACGATGTTAGTTTAGATAATGAGGCAGATAACAACGGATCATTAATAGCAGGAGCTTATACAAT
ATCAACCTTAATAACACAAAAATTAAGTAAATTAAACGGATCAGAAGGTTTAAAGGAAAAGATTGCCGAAGCTAAGA
AATGCTCTGAAGAATTTACTAAAAAATTAAAAGAAAAACACACAGATCTTGGTAAAAAAGATGCTACTGATGTTCAT
GCAAAAGAAGCTATTTTAAAAACAAATGGTACTAAAGATAAGGGCGCTGCAGAGCTTGAAAAATTATTTGAATCAGT
AGAAAACTTGGCAAAAGCAGCTAAAGAGATGCTATCCAATTCAGTTAAAGAGCTTACAAGCCCTGTTGTGGCAGAAA
GTCCAAAAAACCCTTAA
>K_297
ATGAAAAAGAATACATTAAGTGCAATATTAATGACTTTATTTTTATTTATATCTTGTAATAATTCAGGAAAAGATGG
GAATACATCTGCAAATTCTGCTGATGAGTCTGTTAAAGGGCCTAATCTTACAGAAATAAGTAAAAAAATTACAGAAT
CTAACGCAGTTGTTCTGGCTGTGAAAGAAATTGAAACTTTGCTTGCATCTATAGATGAACTTGCTACTAAAGCTATT
GGTAAAAAAATACAACAAAATGGTGGTTTAGCTGTCGAAGCGGGGCATAATGGAACATTGTTAGCAGGTGCTTATAC
AATATCAAAACTAATAACACAAAAATTAGATGGATTGAAAAATTCAGAAAAATTAAAGGAAAAAATTGAAAATGCTA
AGAAATGTTCTGAAGATTTTACTAAAAAACTAGAAGGAGAACATGCGCAACTTGGAATTGAAAATGTTACTGATGAG
AATGCAAAAAAAGCTATTTTAATAACAGATGCAGCTAAAGATAAGGGCGCTGCAGAGCTTGAAAAGCTATTTAAAGC
AGTAGAAAACTTGGCAAAAGCAGCTAAAGAGATGCTTGCTAATTCAGTTAAAGAGCTTACAAGTCCTATTGTGGCAG
AAAGTCCAAAAAACCCTTAA
>H_156a
ATGAAAAAGAATACATTAAGTGCAATATTAATGACTTTATTTTTATTTATATCTTGTAATAATTCAGGGAAAGATGG
GAATGCATCTGCAAATTCTGCTGATGAGTCTGTTAAAGGGCCTAATCTTACAGAAATAAGTAAAAAAATTACAGAAT
CTAACGCAGTTGTACTGGCTGTAAAAGAAGTTGAAACCTTGCTTGCATCTATAAACCAACTTGCTAAAGCTATTGGT
AAAAAAATAGATCAAAATGGCACTTTGGGTGACGATGGGGGCCAAAATGGATCATTGCTAGCAGGAGCTTATGCAAT
CTCAACTGTAATAATAGAAAAATTGAGCACATTAAAAAATGTAGAAGAATTAAAAGAAAAAATTACAAAGGCTAAGG
ATTGTTCCGAAAAGTTTGCTGGCAAATTAAAAAATGAACACGCAAGTCTTGGTAAAAAAGATGCTACTGATGATGAT
GCAAAAAAAGCTATTTTAAAGACACATGGAAATACCGATAAGGGCGCTAAAGAACTTAAAGATCTATCTGATTCAGT
AGAAAGCTTGGTAAAAGCGGCTAAAGAGATGCTAACTAATTCAGTTAAAGAGCTTACAAGTCCTGTTGTGGCAGAAA
GTCCAAAAAAACCTTAA
>G_72a
ATGAAAAAGAATACATTAAGTGCAATATTAATGACTTTATTTTTATTTATATCTTGTAATAATTCAGGAAAAGATGG
GAATGCATCTACAAATTCTGCCGATGAGTCTGTTAAAGGGCCTAATCTTGCAGAAATAAGTAAAAAAATTACAGAAT
CTAATGCAGTTGTACTGGCTGTAAAAGAAGTTGCGGCGTTGCTTTCATCTATAGATGAACTTGCTAAAGCTATTGGT
AAAAAAATAGAGCAAAATGGTTTGGGTGCTGATGCGAATCACAACACCTCATTGTTAGCAGGAGCTCATGAAATATC
AACCCTAATAAAACAAAAATTAGATGGATTGAAAAATGAAGGATTAAATAAAGAAATTGAAGCAGCTAAGAAATGTT
CTGCAGCATTTACTAAAAAACTAGCAGATAGTAATGCAGATCTTGGAGTGGCTGCTGGTAATGCTACTGATGATAAT
GCAAAAAGAGCTATTTTAAAAACACATGGGCATGAGGACAAGGGTGGTAAAGAACTCAAAGAGTTATCTGAAGCAGT
AAAAAGCTTGCTTAAAGCAGCGCAAGCAGCATTAGCTAATTCAGTTCAAGAGCTTACAAGTCCTGTTGTGGCAGAAA
CTCCAAAAAAACCTTAA

>J_118a
ATGAAAAAGAATACATTAAGTGCAATATTAATGACTTTATTTTTATTTATATCTTGTAATAATTCAGGAAAAGATGG
GAATACATCTGCAAATTCTGCTGATGAGTCTGTTAAAGGGCCTAATCTTACAGAAATAAGTAAAAAAATTACAGAAT
CTAACGCAGTTGTTCTGGCTGTGAAAGAAATTGAAACTTTGCTTGCATCTATAGATGAACTTGCTACTAAAGCTATT
GGTAAAAAAATAGATAACAATGCTGGTTTGGGTGCTGAAGTGGGTCAAAACGGATCATTGCTAGCAGGAGCTTATGC
AATCTCAACTGTAATAATAGAAAAATTGAGCACATTAAAAAATGTAGAAGAATTAAAAGAAAAAATTACAAAGGCTA
AGGATTGTTCTGAAAAATTCACTAAAAAATTAAAAGATAGTCACGCAGAGCTTGGTAAAAAAGATGCCAGTGATGAT
GATGCAAAAAAAGCTATTTTAAAAACAAATCAAGCTAACGATAAGGGTGCTAAAGAACTTAAAGAGTTATTTGAAGC
AGTAGAAAGCTTGTCAAAAGCGGCTAAAGAGATGCTAAACAAGTCAGTTAAAGAGCTTACAAGTCCTATTGTGGCAG
AAAGTCCAAAAAACCCTTAA
>U_94a
ATGAAAAAGAATACATTAAGTGCAATATTAATGACTTTATTTTTATTTATATCTTGTAATAATTCAGGAAAAGATGG
GAATGCATCTGCAAATTCTGCTGATGAGTCTGTTAAAGGGCCTAATCTTGCAGAAATAAGTAAAAAAATTACAGAAT
CTAACGCAGTTGTTCTGGCTGTAAAAGAAGTTGAGGCTTTGCTTGCATCTATAGATGAAATTGGTAGTAAAGCTATT
GGTAAAAGAATACAAGCGAATGGTTTACAGGACTTGCAAGGGCAGAACGGATCATTGTTAGCAGGAGCCTATGCAAT
TTCAAATTTAATAACACAAAAAATAAATGTATTAAATGGACTGAAAAATTCAGAAGAATTGAAAGAGAAAATTAACG
AGGCTAAGGGCTGCTCTGAAAAATTTACTAAAAAACTATCAGAAAGTCATGCAGATATTGGTATACAAGCCGCTACT
GATGCTAATGCAAAAGATGCTATTTTAAAAACAAATCCTACTAAAACTAAGGGTGCTGAAGAACTTGATAAGTTATT
TAAAGCGGTGGAAAACTTGTCAAAAGCGGCTAAAGAGATGCTTGCTAATTCAGTTAAAGAACTTACAAGTCCTGTTG
TGGCAGAAAGTCCAAAAAAACCTTAA
>M_29805
ATGAAAAAGAATACATTAAGTGCAATATTAATGACTTTATTTTTATTTATATCTTGTAATAATTCAGGGAAAGATGG
GAATACATCTGCAAATTCTGCTGATGAGTCTGTTAAAGGGCCTAATCTTACAGAAATAAGTAAAAAAATTACAGAAT
CTAACGCAGTTGTTCTCGCCGTGAAAGAAGTTGAAACTCTGCTTGCATCTATAGATGAAGTTGCTAAGAAAGCTATT
GGGAATTTGATAGCCCAAAATGGTTTAAATGCCGGCGCAAATCAAAACGGATCATTGTTAGCGGGAGCCTACGTAAT
ATCAACCCTAATAGCAGAAAAATTAGATGGATTGAAAAATTCAGAAGAATTAAAGGAAAAAATTGAAGATGCTAAAA
AATGTAACAAAGCATTTACTGATAAACTAAAAAGTAGTCATGCGGAACTCGGTATAGCGAATGGAGCTGCTACTGAT
GCTAATGCAAAAGCGGCTATTTTAAAAACAAATGGTACTAAAGATAAGGGTGCTCAAGAGCTTGAAAAGTTATTTGA
ATCAGTAAAAAACTTGTCAAAAGCAGCTCAAGAAACACTAAATAATTCAGTTAAAGAACTTACAAGTCCTGTTGTGG
CAGAAAATCCAAAAAAACCTTAA
>D_CA-11-2A
ATGAAAAAGAATACATTAAGTGCAATATTAATGACTTTATTTTTATTTATATCTTGTAATAATTCAGGAAAAGATGG
GAATACATCTGCAAATTCTGCTGATGAGTCTGTTAAAGGGCCTAATCTTACAGAAATAAGTAAAAAAATTACGGATT
CTAATGCGGTTTTACTTGCTGTGAAAGAGGTTGAAGTGTTGCTGTCATCTATAGATGAACTTGCTAAGAAAGCTATT
GGTAAAAAAATAGATCAAAACAATGCTTTAGGCACTCTAGATAATCATAACGGATCATTGTTGGCGGGAGCTTATGC
TATATCAGCTCTAATAACAGAAAAATTAAGTTCAATAAAAGATTCAGGAGAATTGAAGGCAGAAATTGAAAAGGCTA
AGAAATGTTCTGAAAGCTTTACTAAAAAACTATCAGATAATCAAGCAGAGCTTGGTATAGAGAATGCTACTGATGAT
AATGCAAAAAAGGCTATTTTAAAAACACATAATGCTAAAGACAAGGGTGCTGAAGAACTTGTAAAGTTATCTGAATC
AGTAGCAGGCTTATTAAAAGCAGCTCAAGCCATACTGGCTAATTCAGTTAAAGAGCTTACAAGTCCTGTTGTGGCAG
AAAGTCCAAAAAAACCTTAA
>B_64b
ATGAAAAAGAATACATTAAGTGCAATATTAATGACTTTATTTTTATTTATATCTTGTAATAATTCAGGGAAAGATGG
GAATACATCTGCAAATTCTGCTGATGAGTCTGTTAAAGGGCCTAATCTTACAGAAATAAGTAAAAAAATTACGGATT
CTAATGCGGTTTTACTTGCTGTGAAAGAGGTTGAAGCGTTGCTGTCATCTATAGATGAGCTTGCTAAAGCTATTGGT
AAAAAAATAAAAAACGATGGTAGTTTAGATAATGAAGCAAATCGCAACGAGTCATTGTTAGCAGGAGCTTATACAAT
ATCAACCTTAATAACACAAAAATTAAGTAAATTAAACGGATCAGAAGGTTTAAAGGAAAAGATTGCCGCAGCTAAGA
AATGCTCTGAAGAGTTTAGTACTAAACTAAAAGATAATCATGCACAGCTTGGTATACAGGGCGTTACTGATGAAAAT
GCAAAAAAAGCTATTTTAAAAGCAAATGCAGCGGGTAAAGATAAGGGCGTTGAAGAACTTGAAAAGTTGTCCGGATC
ATTAGAAAGCTTATCAAAAGCAGCTAAAGAGATGCTTGCTAATTCAGTTAAAGAGCTTACAAGCCCTGTTGTGGTAG
AAAGTCCAAAAAAACCTTAA
>I_WI91-23
ATGAAAAAGAATACATTAAGTGCAATATTAATGACTTTATTTTTATTTATATCTTGTAATAATTCAGGGAAAGATGG
GAATACATCTGCAAATTCTGCTGATGAGTCTGTTAAAGGGCCTAATCTTACAGAAATAAGTAAAAAAATTACAGAAT
CTAACGCAGTTGTTCTCGCCGTGAAAGAAGTTGAAACTTTGCTTACATCTATAGATGAGCTTGCTAAAGCTATTGGT
AAAAAAATAAAAAACGATGTTAGTTTAGATAATGAGGCAGATCACAACGGATCATTAATATCAGGAGCATATTTAAT
TTCAACATTAATAACAAAAAAAATAAGTGCAATAAAAGATTCAGGAGAATTGAAGGCAGAAATTGAAAAGGCTAAGA
AATGTTCTGAAGAATTTACTGCTAAATTAAAAGGTGAACACACAGATCTTGGTAAAGAAGGCGTTACTGATGATAAT

GCAAAAAAAGCCATTTTAAAAACAAATAATGATAAAACTAAGGGCGCTGATGAACTTGAAAAGTTATTTGAATCAGT
AAAAAACTTGTCAAAAGCAGCTAAAGAGATGCTTACTAATTCAGTTAAAGAGCTTACAAGCCCTGTTGTGGCAGAAA
GTCCAAAAAAACCTTAA
>F_F084
ATGAAAAAGAATACATTAAGTGCAATATTAATGACTTTATTTTTATTTATATCTTGTAATAATTCAGGAAAAGATGG
GAATACATCTGCAAATTCTGCTGATGAGTCTGTTAAAGGGCCTAATCTTACAGAAATAAGTAAAAAAATTACAGAAT
CTAACGCAGTTGTTCTGGCTGTAAAAGAAATTGAAACTTTGCTTTCATCTATAGATGAACTTGCTACTAAAGCTATT
GGTCAAAAAATAGATGCAAATGGTTTAGGCGTTCAGGCGAATCAAAACGGATCATTGTTAGCAGGTGCCTATGCAAT
ATCAACCCTAATAACACAAAAATTAAGTGCATTAAATTCAGAAGACTTAAAAGAAAAAGTTGCAAAGGTTAAGAAAT
GTTCCGAAGACTTCACTAACAAGTTAAAAAATGGTAATGCACAGCTTGGCTTAGCTGCTGCTACTGATGATAATGCA
AAAGCGGCGATTTTAAAAACAAATGGGACTAACGATAAGGGCGCTAAAGAACTTAAAGATCTATCTGATTCAGTAGA
AAGCTTGGTAAAAGCAGCTCAAGTAATGCTAACTAATTCAGTTAAAGAGCTTACAAGTCCTGTTGTGGCAGAAAGTC
CAAAAAAACCTTAA
>L_T255
ATGAAAAAGAATACATTAAGTGCGATATTGATGACTTTATTTTTATTTATATCTTGTAATAATTCAGGAAAAGATGG
GAATGCATCTGTAAATTCTGCTGATGAGTCTGTTAAAGGGCCTAATCTTGTAGAAATAAGTAAAAAAATTACCGATT
CTAATGCGGTTGTTATTGCAGTGAAAGAAGTTGAAACTTTGCTTGTATCTATAGATGAGCTTGCTAAAGCTATTGGT
AAAAAAATAGAAGCAGGTGGTACTTTAGGTAGCGATGGGGCACACAACGGATCATTACTAGCAGGGGCCTATAAAAT
AGCAACCGAAATAACAGCAAATTTAAGCAAATTAAAAGCTTCAGAAGACTTAAAAGAAAAAATTACAAAGGCTAAGG
AATGTTCTGAGAAATTTACTGATAAACTAAAAAGTGAAAATGTAGCGCTTGGCAAACAGGATGCTAGTGATGATGAT
GCAAAAAAAAGCTATTTTAAAAACACATAATGATATAACTAAGGGTGCTAAAGAACTTAAAGAGTTATCAGAATCAGT
GGAGACCTTGTTAAAAGCAGCTAAAGAGATGCTTGCTAATTCAGTTAAAGAACTTACAAGTCCTGTTGTGGCAGAAA
GTCCAAAAAAACCTTAA
>O_N045
ATGAAAAAGAATACATTAAGTGCAATATTAATGACTTTATTTTTATTTATATCTTGTAATAATTCAGGGAAAGATGG
GAATACATCTGCAAATTCTGCTGATGAGTCTGTTAAAGGGCCTAATCTTACAGAAATAAGTAAAAAAATTACAGAAT
CTAACGCAGTTGTTCTGGCTGTAAAAGAAGTTGAGGCTTTGCTTTCATCTATAGATGAGCTTGCTAAAGCTATTGGT
AAAGAAATAGGGGCAAATGGCTTAGTTAACCAAGCAAATCATAACGTATCATTGTTAGCAGGAGCCTATGAAATATC
AACTCTAATAACAGAAAAATTAAGTAAATTAAACGGTTCAGAAGGATTAAAGGAAAAGATTGCCGCAGCTAAGAAAT
GCTCTGAAGAGTTTAGTACAAAACTAAAAAGTAGTAATGCACAGCTTAATCAGGCCAATGCTAATGATGCTAATGCA
AAAGCGGCTATTTTAAAAACACATAATACTAAAGACAAGGGTGCTGAAGAACTTGTAAAGTTAGCTGAATCAGTAGC
AGGCTTGTTAAAAGTAGCGCAAGAAATGCTAAACAATTCAGTTAAAGAACTTACAAGTCCTGTTGTGGCAGAAAATC
CAAAAAAACCTTAA

S3. Bioinformatics protocols for allele identification, read quantification, and read simulation

1. **Allele identification.** The following commands use software packages bwa (3), samtools (4), and bedtools (5). The commands run in a Linux BASH environment to align paired-end short reads ("r1.fq" and "r2.fq", in FASTQ format) to a set of reference *ospC* sequences ("ref.nuc" in FASTA format). A separate file ("ref.bed"), containing on each line tab-separated columns of the identifier and the beginning and ending nucleotide positions of a reference allele, is needed.

```
1) bwa index ref.nuc # generate index files
2) bwa mem ref.nuc r1.fq r2.fq > sample.sam # align reads
3) samtools view –b sample.sam > sample.bam # convert to binary file
4) samtools sorts sample.bam sample.sorted # sort reads
5) samtools index sample.sorted.bam # index sorted reads
6) bedtools coverage –abam sample.sorted.bam –b refs.bed –d >
   sample.cov # obtain coverage at each nucleotide site of each
   reference sequence
```

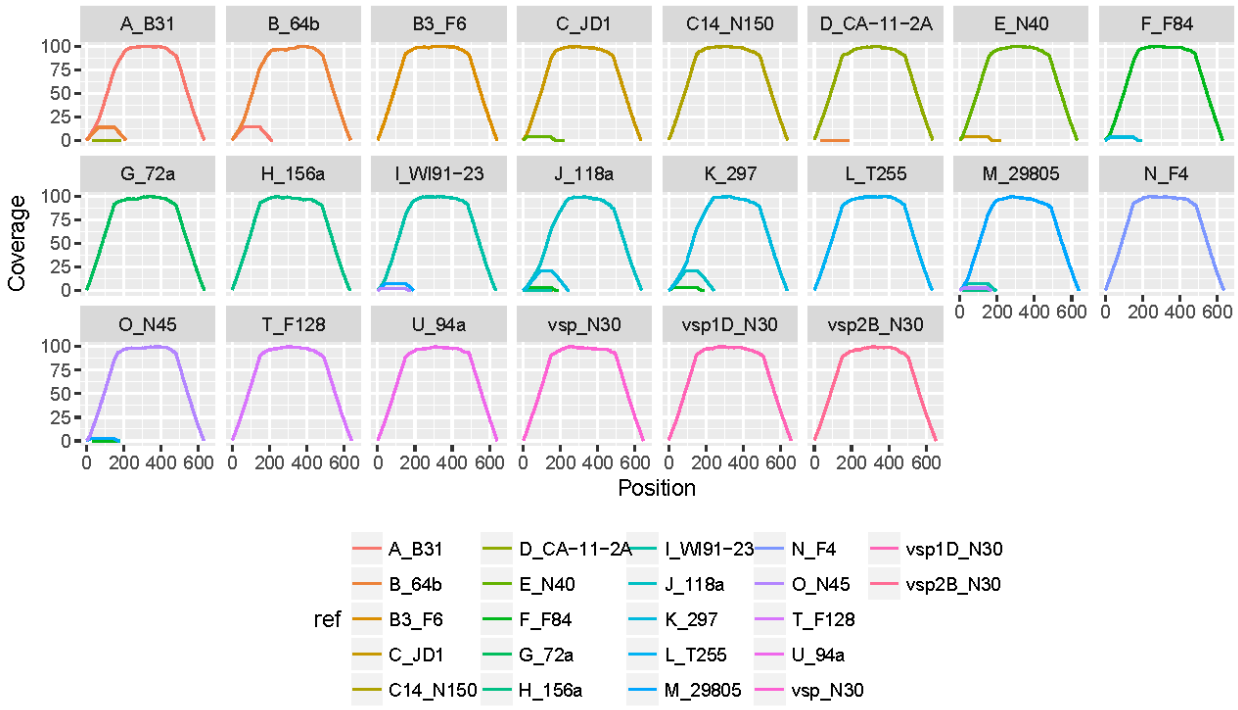
2. **Generation of simulated reads.** The following commands use software wgsim (6) to generate simulated short reads of a hypothetical sample containing a 10:1 mixture of two given *ospC* alleles. The input files are two allele sequences in FASTA format ("A.fas" & "B.fas").

```
1) wgsim -h -N 10000 -1 150 -2 150 -d 150 -s 60 A.fas a-r1.fq a-
   r2.fq # generate 10,000 simulated 150-base paired-end read pairs
   with a standard deviation of 60 bases for the distance between
   the pairs and in haploid mode
2) wgsim -h -N 1000 -1 150 -2 150 -d 150 -s 60 B.fas b-r1.fq b-r2.fq
   # generate 1,000 simulated read pairs for the B allele
3) cat a-r1.fq b-r1.fq > sim.1.fq; cat a-r2.fq b-r2.fq > sim.2.fq; #
   concatenate each file
```

3. ***De novo* assembly of new alleles.** We used the software metaSPAdes (7) to assemble reads from samples in which the majority of reads do not map to provided *ospC* reference sequences, indicating presence of novel alleles. The validity of assembled *ospC* sequences is tested by using the "translated query – protein subject" blast (blastx) from the NCBI BLAST+ package (8).

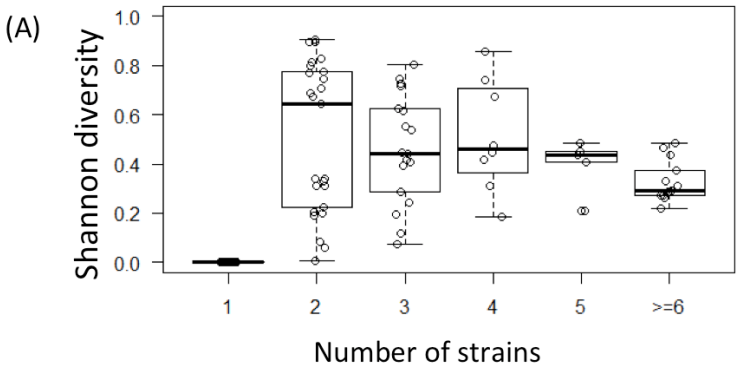
```
1) spades.py -k 21,31,41,51,61,71,81,91,101 -t4 --meta --phred-offset
   33 --pe1-1 R1.fastq.gz --pe1-2 R2.fastq.gz -o output-folder #
   assemble the reads with different k-mer lengths, 4 threads, for a
   sample with mixed amplicons, PHRED quality offset of 33, two paired-
   end input files, and send all outputs to a folder "test"
2) blastx -query output-folder/contigs.fasta -db ref.pep -outfmt 6
```

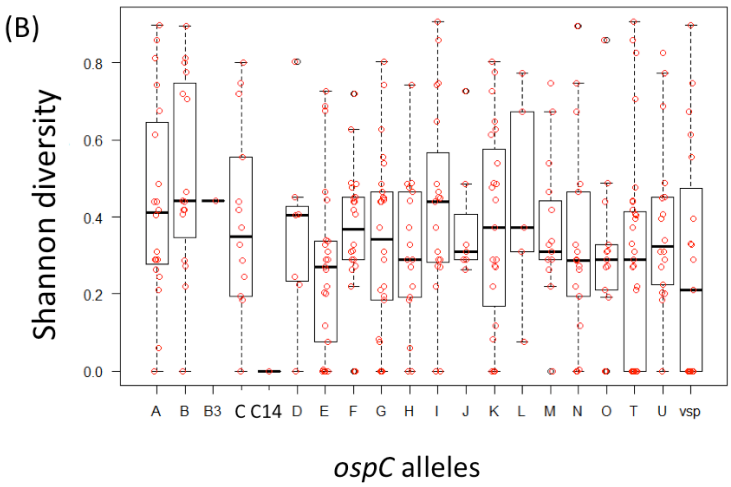
S4. Tests of specificity of allele identification



A sample of 10,000 simulated paired-end reads is generated for each reference sequence using wgsim (6) [see commands in S3(2)]. Each sample is aligned to each reference sequence (*ospC* positions on *x*-axis) to identify presence of alleles as well as to quantify the number of reads (*y*-axis, normalized by the most frequent allele at 100% coverage). Except for the non-specifically aligned reads at the 5' conserved regions for some reference sequences, the bioinformatics protocol identifies each allele without any ambiguity.

S5. Strain distributions within single ticks





(A) Strain diversity within each infected tick (a total of N=55 adult ticks from Sample #9, Figure 1). Shannon index ranges from zero (when a tick is infected by a single strain) to one (when a tick is infected by an equal amount of strains). The median levels of diversity are approximately 0.5, suggesting co-infecting strains are not evenly distributed and consist of dominant strains. (B) Each point represents a tick infected by an *ospC* allele. The median levels of strain diversity are similar for all alleles, suggesting no single strain is consistently dominant.

Reference cited

1. Martin CL, Martin CI, Sukarna TY, Akther S, Ramrattan G, Pagan P, et al. Phylogenomic identification of regulatory sequences in bacteria: an analysis of statistical power and an application to Borrelia burgdorferi sensu lato. mBio. 2015;6(2).

2. Kingry LC, Batra D, Replogle A, Rowe LA, Pritt BS, Petersen JM. Whole Genome Sequence and Comparative Genomics of the Novel Lyme Borreliosis Causing Pathogen, Borrelia mayonii. PloS One. 2016;11(12):e0168994.

3. Li H, Durbin R. Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinforma Oxf Engl. 2010 Mar 1;26(5):589–95.

4. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. Bioinforma Oxf Engl. 2009 Aug 15;25(16):2078–9.

5. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinforma Oxf Engl. 2010 Mar 15;26(6):841–2.

6. Li H. WGSIM [Internet]. 2011. Available from: https://github.com/lh3/wgsim

7. Nurk S, Meleshko D, Korobeynikov A, Pevzner PA. metaSPAdes: a new versatile metagenomic assembler. Genome Res. 2017;27(5):824–34.

8. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, et al. BLAST+: architecture and applications. BMC Bioinformatics. 2009;10:421.