

## Table 3. Gene content of genomic islands

REGION 1.1 [7 sequence(s)]:

2603V/R [111524-118753]

|         |        |  |
|---------|--------|--|
| SAG0102 | 154 aa | conserved hypothetical protein                       |
| SAG0103 | 189 aa | conserved hypothetical protein TIGR01440             |
| SAG0104 | 280 aa | conserved hypothetical protein                       |
| SAG0105 | 427 aa | trigger factor                                       |
| SAG0106 | 191 aa | DNA-directed RNA polymerase, delta subunit, putative |
| SAG0107 | 534 aa | CTP synthase   |
| SAG0108 | 308 aa | conserved hypothetical protein                       |

NEM316 [111556-118785]

|         |        |  |
|---------|--------|--|
| gbs0101 | 154 aa | integral membrane protein                |
| gbs0102 | 189 aa | conserved hypothetical protein TIGR01440 |
| gbs0103 | 280 aa | mechanosensitive transport protein       |
| gbs0104 | 427 aa | trigger factor                           |
| gbs0105 | 176 aa | DNA-directed RNA polymerase delta chain  |
| gbs0106 | 534 aa | CTP synthase                             |
| gbs0107 | 308 aa | cell surface hydrolase, membrane-bound   |

A909 [111218-118447]

|               |        |   |
|---------------|--------|---|
| A909.ORF01962 | 154 aa | membrane protein, putative                  |
| A909.ORF01961 | 189 aa | conserved hypothetical protein TIGR01440    |
| A909.ORF01960 | 280 aa | mechanosensitive ion channel family protein |
| A909.ORF01959 | 427 aa | trigger factor                              |
| A909.ORF01958 | 191 aa | DNA-directed RNA polymerase, delta subunit  |
| A909.ORF01957 | 534 aa | CTP synthase                                |
| A909.ORF01956 | 308 aa | conserved hypothetical protein              |

515 [103127-110244]

|              |        |  |
|--------------|--------|--|
| 515.ORF00119 | 154 aa | Unknown                                  |
| 515.ORF00120 | 189 aa | conserved hypothetical protein TIGR01440 |
| 515.ORF00121 | 280 aa | Mechanosensitive ion channel family      |
| 515.ORF00122 | 427 aa | trigger factor                           |
| 515.ORF00123 | 62 aa  | delta subunit RNA polymerase             |
| 515.ORF00124 | 88 aa  | delta subunit RNA polymerase             |
| 515.ORF00125 | 50 aa  | CTP synthase [imported]                  |
| 515.ORF00126 | 443 aa | CTP synthase                             |
| 515.ORF00127 | 44 aa  | CTP synthase                             |
| 515.ORF00128 | 308 aa | Unknown                                  |

CJB111 [85168-92181]

|                 |        |  |
|-----------------|--------|--|
| CJB111.ORF00093 | 154 aa | conserved hypothetical protein           |
| CJB111.ORF00094 | 189 aa | conserved hypothetical protein TIGR01440 |
| CJB111.ORF00095 | 280 aa | Mechanosensitive ion channel family      |

|                 |        |                                |
|-----------------|--------|--------------------------------|
| CJB111.ORF00096 | 427 aa | trigger factor                 |
| CJB111.ORF00097 | 140 aa | delta subunit RNA polymerase   |
| CJB111.ORF00098 | 77 aa  | delta subunit RNA polymerase   |
| CJB111.ORF00099 | 534 aa | CTP synthase                   |
| CJB111.ORF00100 | 300 aa | conserved hypothetical protein |

## COH1 [116302-121476]

|               |        |                              |
|---------------|--------|------------------------------|
| COH1.ORF00130 | 65 aa  | BioY family protein          |
| COH1.ORF00131 | 199 aa | trigger factor               |
| COH1.ORF00132 | 176 aa | delta subunit RNA polymerase |
| COH1.ORF00133 | 534 aa | CTP synthase                 |
| COH1.ORF00134 | 239 aa | Unknown                      |

## H36B [106464-114629]

|               |        |   |
|---------------|--------|---|
| H36B.ORF00110 | 154 aa | Unknown                                     |
| H36B.ORF00111 | 189 aa | conserved hypothetical protein TIGR01440    |
| H36B.ORF00112 | 280 aa | Mechanosensitive ion channel family         |
| H36B.ORF00113 | 91 aa  | heat-inducible transcription repressor HrcA |
| H36B.ORF00114 | 286 aa | trigger factor                              |
| H36B.ORF00115 | 370 aa | trigger factor                              |
| H36B.ORF00116 | 191 aa | delta subunit RNA polymerase                |
| H36B.ORF00117 | 534 aa | CTP synthase                                |
| H36B.ORF00118 | 308 aa | Unknown                                     |

## REGION 1.2 [6 sequence(s)]:

## 2603V/R [231668-244184]

|         |        |   |
|---------|--------|---|
| SAG0216 | 33 aa  | hypothetical protein                              |
| SAG0217 | 384 aa | site-specific recombinase, phage integrase family |
| SAG0218 | 158 aa | transcriptional regulator, Cro/CI family          |
| SAG0219 | 101 aa | hypothetical protein                              |
| SAG0220 | 92 aa  | conserved hypothetical protein                    |
| SAG0221 | 76 aa  | hypothetical protein                              |
| SAG0222 | 108 aa | conserved domain protein                          |
| SAG0223 | 209 aa | conserved hypothetical protein, fusion            |
| SAG0224 | 332 aa | replication initiation protein, putative          |
| SAG0225 | 144 aa | hypothetical protein                              |
| SAG0226 | 418 aa | recombination protein                             |
| SAG0227 | 156 aa | hypothetical protein                              |
| SAG0228 | 111 aa | conserved hypothetical protein                    |
| SAG0229 | 95 aa  | conserved hypothetical protein                    |
| SAG0230 | 96 aa  | conserved hypothetical protein                    |
| SAG0231 | 135 aa | hypothetical protein                              |
| SAG0232 | 186 aa | hypothetical protein                              |
| SAG0233 | 226 aa | hypothetical protein                              |
| SAG0234 | 128 aa | hypothetical protein                              |

SAG0235 93 aa hypothetical protein  
 SAG0236 32 aa hypothetical protein

## NEM316 [232768-245594]

gbs0211 384 aa site-specific recombinase, phage integrase family  
 gbs0212 183 aa Helix-turn-helix domain protein  
 gbs0213 97 aa conserved hypothetical protein  
 gbs0214 93 aa conserved hypothetical protein  
 gbs0215 108 aa Replication initiator protein A (RepA) N-terminus family  
 gbs0216 213 aa conserved hypothetical protein  
 gbs0217 208 aa RepJ  
 gbs0218 144 aa conserved hypothetical protein  
 gbs0219 418 aa recombination protein  
 gbs0220 156 aa conserved hypothetical protein  
 gbs0221 111 aa conserved hypothetical protein  
 gbs0222 95 aa conserved hypothetical protein  
 gbs0223 96 aa a small heat shock protein of Clostridium acetobutylicum

homolog

## lmo0056

gbs0224 135 aa conserved hypothetical protein  
 gbs0225 135 aa RAD2 endonuclease PFB0265c  
 gbs0226 195 aa conserved hypothetical protein  
 gbs0228 128 aa conserved hypothetical protein  
 gbs0229 93 aa conserved hypothetical protein

## A909 [241218-253571]

A909.ORF01840 384 aa site-specific recombinase, phage integrase family  
 A909.ORF01839 183 aa DNA-binding protein  
 A909.ORF01838 97 aa hypothetical protein  
 A909.ORF01837 93 aa conserved hypothetical protein  
 A909.ORF01836 108 aa replication initiation factor, RepA family  
 A909.ORF01835 213 aa conserved domain protein  
 A909.ORF01834 291 aa replication initiation factor family protein  
 A909.ORF01833 144 aa hypothetical protein  
 A909.ORF01832 418 aa plasmid recombination enzyme  
 A909.ORF01831 156 aa hypothetical protein  
 A909.ORF01830 111 aa RelE/ParE family protein  
 A909.ORF01829 95 aa conserved hypothetical protein  
 A909.ORF01828 96 aa conserved hypothetical protein  
 A909.ORF02143 135 aa hypothetical protein  
 A909.ORF01827 186 aa hypothetical protein  
 A909.ORF01826 226 aa conserved hypothetical protein, truncation  
 A909.ORF01825 128 aa hypothetical protein  
 A909.ORF01824 93 aa hypothetical protein  
 A909.ORF01823 32 aa hypothetical protein

## CJB11 [204882-222292]

|                  |        |   |
|------------------|--------|---|
| CJB111.ORF00215  | 384 aa | site-specific recombinase, phage integrase family |
| CJB111.ORF00216  | 196 aa | transcriptional regulator, Cro/CI family          |
| CJB111.ORF00217  | 101 aa | conserved hypothetical protein                    |
| CJB111.ORF00218  | 92 aa  | conserved hypothetical protein                    |
| CJB111.ORF00219  | 76 aa  | conserved hypothetical protein                    |
| CJB111.ORF00220  | 108 aa | conserved hypothetical protein                    |
| CJB111.ORF00221  | 209 aa | conserved hypothetical protein                    |
| CJB111.ORF00222  | 332 aa | replication initiation protein, putative          |
| CJB111.ORF00223  | 144 aa | conserved hypothetical protein                    |
| CJB111.ORFA00217 | 53 aa  | hypothetical protein                              |
| CJB111.ORF00224  | 418 aa | recombination protein                             |
| CJB111.ORF00225  | 156 aa | conserved hypothetical protein                    |
| CJB111.ORF00226  | 111 aa | conserved hypothetical protein                    |
| CJB111.ORF00227  | 95 aa  | conserved hypothetical protein                    |
| CJB111.ORF00228  | 484 aa | adenine specific DNA methyltransferase (mod)      |
| CJB111.ORF00229  | 902 aa | restriction enzyme LlaFI                          |
| CJB111.ORF00230  | 96 aa  | conserved hypothetical protein                    |
| CJB111.ORF00231  | 142 aa | conserved hypothetical protein                    |
| CJB111.ORF00232  | 58 aa  | conserved hypothetical protein                    |
| CJB111.ORF00233  | 121 aa | conserved hypothetical protein                    |
| CJB111.ORF00234  | 226 aa | conserved hypothetical protein                    |
| CJB111.ORF00235  | 128 aa | conserved hypothetical protein                    |
| CJB111.ORF00236  | 93 aa  | conserved hypothetical protein                    |
| CJB111.ORF00237  | 32 aa  | conserved hypothetical protein                    |

## H36B [208398-220750]

|               |        |   |
|---------------|--------|---|
| H36B.ORF00218 | 384 aa | site-specific recombinase, phage integrase family |
| H36B.ORF00219 | 183 aa | Unknown   |
| H36B.ORF00220 | 97 aa  | Unknown   |
| H36B.ORF00221 | 93 aa  | Unknown   |
| H36B.ORF00222 | 108 aa | Unknown   |
| H36B.ORF00223 | 213 aa | Unknown   |
| H36B.ORF00224 | 291 aa | replication initiation protein, putative          |
| H36B.ORF00225 | 144 aa | conserved hypothetical protein                    |
| H36B.ORF00226 | 418 aa | recombination protein                             |
| H36B.ORF00227 | 156 aa | conserved hypothetical protein                    |
| H36B.ORF00228 | 111 aa | conserved hypothetical protein                    |
| H36B.ORF00229 | 95 aa  | conserved hypothetical protein                    |
| H36B.ORF00230 | 96 aa  | conserved hypothetical protein                    |
| H36B.ORF00231 | 142 aa | conserved hypothetical protein                    |
| H36B.ORF00232 | 186 aa | conserved hypothetical protein                    |
| H36B.ORF00233 | 87 aa  | conserved hypothetical protein                    |
| H36B.ORF00234 | 81 aa  | conserved hypothetical protein                    |
| H36B.ORF00235 | 128 aa | conserved hypothetical protein                    |

|                               |        |  |
|-------------------------------|--------|--|
| H36B.ORF00236                 | 93 aa  | conserved hypothetical protein   |
| H36B.ORF00237                 | 32 aa  | conserved hypothetical protein   |
| <b>18RS21 [222344-234698]</b> |        |  |
| 18RS21.ORF00227               | 384 aa | site-specific recombinase, phage integrase family                        |
| 18RS21.ORF00228               | 183 aa | Helix-turn-helix domain protein  |
| 18RS21.ORF00229               | 97 aa  | conserved hypothetical protein   |
| 18RS21.ORF00230               | 93 aa  | conserved hypothetical protein   |
| 18RS21.ORF00231               | 213 aa | conserved hypothetical protein   |
| 18RS21.ORF00232               | 291 aa | RepJ   |
| 18RS21.ORF00233               | 144 aa | conserved hypothetical protein   |
| 18RS21.ORF00234               | 418 aa | recombination protein  |
| 18RS21.ORF00235               | 156 aa | conserved hypothetical protein   |
| 18RS21.ORF00236               | 111 aa | Plasmid stabilization system protein superfamily                         |
| 18RS21.ORF00237               | 95 aa  | conserved hypothetical protein   |
| 18RS21.ORF00238               | 96 aa  | a small heat shock protein of Clostridium acetobutylicum homolog lmo0056 |
| 18RS21.ORF00239               | 135 aa | conserved hypothetical protein   |
| 18RS21.ORF00240               | 186 aa | conserved hypothetical protein   |
| 18RS21.ORF00241               | 226 aa | conserved hypothetical protein   |
| 18RS21.ORF00242               | 128 aa | conserved hypothetical protein   |
| 18RS21.ORF00243               | 93 aa  | conserved hypothetical protein   |
| 18RS21.ORF00244               | 32 aa  | conserved hypothetical protein   |

**REGION 1.3 [4 sequence(s)]:****2603V/R [249947-261676]**

|         |        |   |
|---------|--------|---|
| SAG0245 | 152 aa | protein of unknown function/lipoprotein, putative |
| SAG0246 | 268 aa | hypothetical protein                              |
| SAG0247 | 116 aa | hypothetical protein                              |
| SAG0248 | 90 aa  | hypothetical protein                              |
| SAG0249 | 116 aa | hypothetical protein                              |
| SAG0250 | 193 aa | membrane protein, putative                        |
| SAG0251 | 72 aa  | transcriptional regulator, Cro/CI family          |

**A909 [259334-272624]**

|               |        |                            |
|---------------|--------|----------------------------|
| A909.ORF01815 | 496 aa | ISSag8, transposase        |
| A909.ORF01814 | 152 aa | lipoprotein, putative      |
| A909.ORF01813 | 544 aa | conserved domain protein   |
| A909.ORF01812 | 90 aa  | hypothetical protein       |
| A909.ORF01811 | 116 aa | conserved domain protein   |
| A909.ORF01810 | 193 aa | membrane protein, putative |
| A909.ORF01809 | 72 aa  | DNA-binding protein        |

**CJB111 [333152-344784]**

|                 |        |   |
|-----------------|--------|---|
| CJB111.ORF00351 | 152 aa | protein of unknown function/lipoprotein, putative |
| CJB111.ORF00352 | 544 aa | conserved hypothetical protein                    |

|                        |        |   |
|------------------------|--------|---|
| CJB111.ORF00353        | 90 aa  | conserved hypothetical protein                    |
| CJB111.ORF00354        | 116 aa | conserved hypothetical protein                    |
| CJB111.ORFA00344       | 32 aa  | conserved hypothetical protein                    |
| CJB111.ORF00355        | 193 aa | membrane protein, putative                        |
| CJB111.ORF00356        | 72 aa  | transcriptional regulator, Cro/CI family          |
| 18RS21 [240461-250474] |        |   |
| 18RS21.ORF00252        | 152 aa | protein of unknown function/lipoprotein, putative |
| 18RS21.ORF00253        | 544 aa | conserved hypothetical protein                    |
| 18RS21.ORF00254        | 90 aa  | conserved hypothetical protein                    |
| 18RS21.ORF00255        | 116 aa | conserved hypothetical protein                    |
| 18RS21.ORF00256        | 32 aa  | conserved hypothetical protein                    |
| 18RS21.ORF00257        | 209 aa | membrane protein, putative                        |
| 18RS21.ORF00258        | 48 aa  | DNA-binding protein                               |

## REGION 1.4 [2 sequence(s)]:

2603V/R [555272-600189]

|         |        |   |
|---------|--------|---|
| SAG0540 | 91 aa  | chorismate mutase, putative   |
| SAG0541 | 398 aa | voltage-gated chloride channel family protein                         |
| SAG0542 | 127 aa | IS1381, transposase OrfA  |
| SAG0543 | 129 aa | IS1381, transposase OrfB  |
| SAG0544 | 115 aa | ribosomal protein L19   |
| SAG0545 | 359 aa | prophage LambdaSa1, site-specific recombinase, phage integrase family |
| SAG0546 | 67 aa  | conserved domain protein  |
| SAG0547 | 185 aa | hypothetical protein  |
| SAG0548 | 265 aa | prophage LambdaSa1, repressor protein, putative                       |
| SAG0549 | 47 aa  | hypothetical protein  |
| SAG0550 | 74 aa  | conserved hypothetical protein  |
| SAG0551 | 52 aa  | conserved hypothetical protein  |
| SAG0552 | 62 aa  | hypothetical protein  |
| SAG0553 | 268 aa | hypothetical protein  |
| SAG0554 | 63 aa  | prophage LambdaSa1, transcriptional regulator, Cro/CI family          |
| SAG0555 | 249 aa | prophage LambdaSa1, antirepressor, putative                           |
| SAG0556 | 47 aa  | hypothetical protein  |
| SAG0557 | 76 aa  | hypothetical protein  |
| SAG0558 | 74 aa  | hypothetical protein  |
| SAG0559 | 286 aa | conserved hypothetical protein  |
| SAG0560 | 77 aa  | conserved hypothetical protein  |
| SAG0561 | 46 aa  | hypothetical protein  |
| SAG0562 | 84 aa  | hypothetical protein  |
| SAG0563 | 53 aa  | hypothetical protein  |
| SAG0564 | 160 aa | conserved hypothetical protein  |
| SAG0565 | 224 aa | conserved domain protein  |
| SAG0566 | 138 aa | prophage LambdaSa1, single-strand binding protein                     |

|         |         |   |
|---------|---------|---|
| SAG0567 | 439 aa  | prophage LambdaSa1, reverse transcriptase/maturase family protein |
| SAG0568 | 67 aa   | conserved hypothetical protein                                    |
| SAG0569 | 158 aa  | conserved hypothetical protein                                    |
| SAG0570 | 115 aa  | hypothetical protein  |
| SAG0571 | 43 aa   | hypothetical protein  |
| SAG0572 | 138 aa  | conserved hypothetical protein                                    |
| SAG0573 | 54 aa   | hypothetical protein  |
| SAG0574 | 89 aa   | conserved hypothetical protein                                    |
| SAG0575 | 110 aa  | hypothetical protein  |
| SAG0576 | 43 aa   | hypothetical protein  |
| SAG0577 | 177 aa  | conserved hypothetical protein                                    |
| SAG0578 | 88 aa   | conserved hypothetical protein                                    |
| SAG0579 | 142 aa  | conserved hypothetical protein                                    |
| SAG0580 | 111 aa  | conserved hypothetical protein, truncation                        |
| SAG0581 | 118 aa  | conserved hypothetical protein                                    |
| SAG0582 | 422 aa  | conserved hypothetical protein                                    |
| SAG0583 | 406 aa  | conserved hypothetical protein                                    |
| SAG0584 | 62 aa   | conserved hypothetical protein, truncation                        |
| SAG0585 | 471 aa  | conserved hypothetical protein                                    |
| SAG0586 | 154 aa  | conserved hypothetical protein                                    |
| SAG0587 | 300 aa  | prophage LambdaSa1, structural protein, putative                  |
| SAG0588 | 71 aa   | conserved hypothetical protein                                    |
| SAG0589 | 143 aa  | conserved hypothetical protein                                    |
| SAG0590 | 112 aa  | conserved hypothetical protein                                    |
| SAG0591 | 78 aa   | conserved hypothetical protein                                    |
| SAG0592 | 111 aa  | conserved hypothetical protein                                    |
| SAG0593 | 185 aa  | prophage LambdaSa1, structural protein                            |
| SAG0594 | 81 aa   | conserved hypothetical protein                                    |
| SAG0595 | 123 aa  | conserved hypothetical protein                                    |
| SAG0596 | 670 aa  | prophage LambdaSa1, pblA protein, internal deletion               |
| SAG0597 | 506 aa  | prophage LambdaSa1, minor structural protein, putative            |
| SAG0598 | 1374 aa | prophage LambdaSa1, N-acetylmuramoyl-L-alanine amidase, family 4  |
| SAG0599 | 668 aa  | prophage LambdaSa1, minor structural protein, putative            |
| SAG0600 | 109 aa  | hypothetical protein  |
| SAG0601 | 70 aa   | hypothetical protein  |
| SAG0602 | 100 aa  | conserved hypothetical protein                                    |
| SAG0603 | 111 aa  | conserved hypothetical protein                                    |
| SAG0604 | 239 aa  | prophage LambdaSa1, lysin, putative                               |
| SAG0605 | 323 aa  | conserved hypothetical protein                                    |
| SAG0606 | 66 aa   | conserved hypothetical protein                                    |
| SAG0607 | 56 aa   | conserved hypothetical protein                                    |
| SAG0608 | 59 aa   | hypothetical protein  |

515 [547295-591885]

|               |        |   |
|---------------|--------|---|
| 515.ORF00610  | 191 aa | DNA polymerase III, epsilon subunit/ATP-dependent helicase DinG       |
| 515.ORF00611  | 91 aa  | chorismate mutase   |
| 515.ORF00612  | 142 aa | voltage-gated chloride channel family protein                         |
| 515.ORF00613  | 115 aa | ribosomal protein L19   |
| 515.ORF00614  | 330 aa | prophage LambdaSa1, site-specific recombinase, phage integrase family |
| 515.ORF00615  | 55 aa  | hypothetical protein  |
| 515.ORF00616  | 204 aa | hypothetical protein  |
| 515.ORFA00586 | 31 aa  | hypothetical protein  |
| 515.ORF00617  | 259 aa | prophage LambdaSa1, repressor protein, putative                       |
| 515.ORF00618  | 97 aa  | hypothetical protein  |
| 515.ORF00619  | 56 aa  | conserved protein-related protein                                     |
| 515.ORF00620  | 52 aa  | conserved hypothetical protein  |
| 515.ORF00621  | 80 aa  | hypothetical protein  |
| 515.ORF00622  | 185 aa | hypothetical protein  |
| 515.ORF00623  | 69 aa  | conserved hypothetical protein  |
| 515.ORF00624  | 68 aa  | transcriptional regulator, Cro/CI family                              |
| 515.ORF00625  | 319 aa | conserved hypothetical protein  |
| 515.ORF00626  | 169 aa | hypothetical protein  |
| 515.ORF00627  | 45 aa  | hypothetical protein  |
| 515.ORF00628  | 48 aa  | hypothetical protein  |
| 515.ORF00629  | 102 aa | hypothetical protein  |
| 515.ORF00630  | 76 aa  | conserved hypothetical protein  |
| 515.ORF00632  | 74 aa  | conserved hypothetical protein  |
| 515.ORF00633  | 286 aa | phage replisome organizer, putative, N-terminal region domain protein |
| 515.ORF00634  | 77 aa  | conserved hypothetical protein  |
| 515.ORF00635  | 46 aa  | conserved hypothetical protein  |
| 515.ORF00636  | 53 aa  | conserved hypothetical protein  |
| 515.ORF00637  | 160 aa | Siphovirus Gp157 superfamily  |
| 515.ORF00638  | 222 aa | ERF superfamily family  |
| 515.ORF00639  | 138 aa | prophage LambdaSa1, single-strand binding protein                     |
| 515.ORF00640  | 66 aa  | conserved hypothetical protein  |
| 515.ORF00641  | 162 aa | Endodeoxyribonuclease RusA superfamily                                |
| 515.ORF00642  | 114 aa | conserved hypothetical protein  |
| 515.ORF00643  | 54 aa  | conserved hypothetical protein  |
| 515.ORF00644  | 96 aa  | conserved hypothetical protein  |
| 515.ORF00645  | 47 aa  | hypothetical protein  |
| 515.ORF00646  | 170 aa | conserved hypothetical protein  |
| 515.ORF00647  | 99 aa  | hypothetical protein  |
| 515.ORF00648  | 64 aa  | hypothetical protein  |
| 515.ORF00649  | 88 aa  | conserved hypothetical protein  |
| 515.ORF00650  | 144 aa | transcriptional regulator, ArpU family, putative                      |
| 515.ORF00651  | 144 aa | transcriptional regulator, ArpU family, putative                      |
| 515.ORF00652  | 258 aa | hypothetical phage protein  |

|              |         |  |
|--------------|---------|--|
| 515.ORF00653 | 79 aa   | bacteriophage L54a, HNH endonuclease family protein    |
| 515.ORF00654 | 169 aa  | conserved hypothetical protein                         |
| 515.ORF00655 | 521 aa  | phage terminase, large subunit, putative               |
| 515.ORF00656 | 99 aa   | hypothetical protein                                   |
| 515.ORF00657 | 382 aa  | phage portal protein, HK97 family                      |
| 515.ORF00658 | 243 aa  | ClpP protease family protein                           |
| 515.ORF00659 | 386 aa  | conserved hypothetical protein                         |
| 515.ORF00660 | 97 aa   | conserved hypothetical protein                         |
| 515.ORF00661 | 78 aa   | conserved hypothetical protein                         |
| 515.ORF00662 | 143 aa  | SLT orf 110-like protein                               |
| 515.ORF00663 | 116 aa  | SLT orf 123-like protein, putative                     |
| 515.ORF00664 | 127 aa  | SLT orf 131b-like protein, putative                    |
| 515.ORF00665 | 201 aa  | major tail protein, putative                           |
| 515.ORF00666 | 100 aa  | conserved hypothetical protein                         |
| 515.ORF00667 | 35 aa   | conserved hypothetical protein                         |
| 515.ORF00668 | 1307 aa | reticulocyte binding protein                           |
| 515.ORF00669 | 502 aa  | prophage LambdaSa1, minor structural protein, putative |
| 515.ORF00670 | 1377 aa | reticulocyte binding protein                           |
| 515.ORF00671 | 663 aa  | prophage LambdaSa1, minor structural protein, putative |
| 515.ORF00672 | 99 aa   | conserved hypothetical protein                         |
| 515.ORF00673 | 70 aa   | conserved hypothetical protein                         |
| 515.ORF00674 | 100 aa  | conserved hypothetical protein                         |
| 515.ORF00675 | 111 aa  | phage holin, LL-H family                               |
| 515.ORF00676 | 248 aa  | endolysin domain protein                               |
| 515.ORF00677 | 323 aa  | conserved hypothetical protein                         |
| 515.ORF00678 | 66 aa   | CsbD-like superfamily                                  |
| 515.ORF00679 | 67 aa   | conserved hypothetical protein                         |

## REGION 1.5 [4 sequence(s)]:

2603V/R [624710-631949]

|         |        |  |
|---------|--------|--|
| SAG0634 | 70 aa  | hypothetical protein   |
| SAG0635 | 245 aa | acid phosphatase, class B  |
| SAG0636 | 172 aa | conserved hypothetical protein   |
| SAG0637 | 0 aa   | transcriptional regulator, TetR family, putative, authentic frameshift |
| SAG0638 | 109 aa | cell wall surface anchor family protein, truncation                    |
| SAG0639 | 273 aa | transposase OrfB, IS3 family   |
| SAG0640 | 91 aa  | transposase OrfA, IS3 family   |
| SAG0641 | 0 aa   | Tn5252, Orf 10 protein, degenerate                                     |
| SAG0642 | 59 aa  | hypothetical protein   |
| SAG0643 | 0 aa   | chaperonin, 33 kDa, degenerate   |
| SAG0644 | 402 aa | transcriptional regulator, AraC family                                 |

NEM316 [637887-645133]

|         |       |                                |
|---------|-------|--------------------------------|
| gbs0614 | 70 aa | conserved hypothetical protein |
|---------|-------|--------------------------------|

|         |        |   |
|---------|--------|---|
| gbs0615 | 245 aa | HAD superfamily (subfamily IIIB) phosphatase, TIGR01672   |
| gbs0616 | 150 aa | conserved hypothetical protein                            |
| gbs0617 | 78 aa  | conserved hypothetical protein                            |
| gbs0618 | 126 aa | conserved hypothetical protein                            |
| gbs0619 | 110 aa | surface protein R28                                       |
| gbs0621 | 91 aa  | transposase   |
| gbs0622 | 49 aa  | Tn5252, Orf 10 protein                                    |
| gbs0625 | 115 aa | 33 kDa chaperonin (Heat shock protein 33 homolog) (HSP33) |
| gbs0627 | 262 aa | transcriptional regulator, AraC family                    |

## COH1 [628012-635254]

|                |        |   |
|----------------|--------|---|
| COH1.ORF00686  | 70 aa  | Unknown-related protein                                 |
| COH1.ORF00687  | 245 aa | HAD superfamily (subfamily IIIB) phosphatase, TIGR01672 |
| COH1.ORF00688  | 172 aa | Unknown   |
| COH1.ORF00689  | 38 aa  | hypothetical protein                                    |
| COH1.ORFA00669 | 78 aa  | conserved hypothetical protein                          |
| COH1.ORF00690  | 110 aa | cell wall surface anchor family protein, truncation     |
| COH1.ORF00691  | 273 aa | transposase OrfB, IS3 family, truncation                |
| COH1.ORF00692  | 106 aa | probable transposase TnpA                               |
| COH1.ORF00693  | 49 aa  | Tn5252, Orf 10 protein [imported]                       |
| COH1.ORF00694  | 59 aa  | conserved hypothetical protein                          |
| COH1.ORF00695  | 67 aa  | chaperonin, 33 kDa                                      |
| COH1.ORF00696  | 115 aa | chaperonin, 33 kDa                                      |
| COH1.ORF00697  | 402 aa | transcriptional regulator, AraC family                  |

## 18RS21 [566307-574505]

|                 |        |   |
|-----------------|--------|---|
| 18RS21.ORF00590 | 70 aa  | conserved hypothetical protein                            |
| 18RS21.ORF00591 | 245 aa | HAD superfamily (subfamily IIIB) phosphatase, TIGR01672   |
| 18RS21.ORF00592 | 172 aa | conserved hypothetical protein                            |
| 18RS21.ORF00593 | 78 aa  | conserved hypothetical protein                            |
| 18RS21.ORF00594 | 109 aa | surface protein R28                                       |
| 18RS21.ORF00595 | 273 aa | transposase   |
| 18RS21.ORF00596 | 91 aa  | transposase   |
| 18RS21.ORF00597 | 89 aa  | transposase OrfA, IS3 family                              |
| 18RS21.ORF00598 | 49 aa  | Tn5252, Orf 10 protein                                    |
| 18RS21.ORF00599 | 115 aa | 33 kDa chaperonin (Heat shock protein 33 homolog) (HSP33) |
| 18RS21.ORF00600 | 402 aa | transcriptional regulator, AraC family                    |
| 18RS21.ORF00601 | 32 aa  | hypothetical protein                                      |
| 18RS21.ORF00602 | 213 aa | conserved hypothetical protein                            |

## REGION 1.6 [6 sequence(s)]:

## 2603V/R [659648-672432]

|         |        |                            |
|---------|--------|----------------------------|
| SAG0674 | 113 aa | hypothetical protein       |
| SAG0675 | 171 aa | secreted protein, putative |
| SAG0676 | 885 aa | proteinase, putative       |

|         |         |                             |
|---------|---------|-----------------------------|
| SAG0677 | 1062 aa | hypothetical protein        |
| SAG0678 | 0 aa    | endopeptidase O, degenerate |
| SAG0679 | 343 aa  | protein of unknown function |
| SAG0680 | 339 aa  | protein of unknown function |
| SAG0681 | 353 aa  | conserved domain protein    |

## A909 [739150-751970]

|               |         |                                   |
|---------------|---------|-----------------------------------|
| A909.ORF01345 | 113 aa  | hypothetical protein              |
| A909.ORF01344 | 171 aa  | conserved domain protein          |
| A909.ORF01343 | 885 aa  | peptidase, S8 (subtilisin) family |
| A909.ORF01342 | 1074 aa | hypothetical protein              |
| A909.ORF01341 | 0 aa    | endopeptidase O, degenerate       |
| A909.ORF01338 | 343 aa  | conserved hypothetical protein    |
| A909.ORF01337 | 339 aa  | conserved domain protein          |
| A909.ORF01336 | 353 aa  | conserved hypothetical protein    |

## CJB111 [661635-674490]

|                 |         |                                      |
|-----------------|---------|--------------------------------------|
| CJB111.ORF00692 | 113 aa  | conserved hypothetical protein       |
| CJB111.ORF00693 | 171 aa  | secreted protein, putative           |
| CJB111.ORF00694 | 885 aa  | proteinase, putative                 |
| CJB111.ORF00695 | 1105 aa | conserved hypothetical protein       |
| CJB111.ORF00696 | 361 aa  | endopeptidase O [imported], putative |
| CJB111.ORF00697 | 272 aa  | endopeptidase O [imported]           |
| CJB111.ORF00698 | 36 aa   | endopeptidase O                      |
| CJB111.ORF00699 | 343 aa  | protein of unknown function          |
| CJB111.ORF00700 | 339 aa  | protein of unknown function          |
| CJB111.ORF00701 | 118 aa  | protein of unknown function          |
| CJB111.ORF00702 | 235 aa  | protein of unknown function          |

## COH1 [662957-675793]

|               |        |                                      |
|---------------|--------|--------------------------------------|
| COH1.ORF00727 | 113 aa | conserved hypothetical protein       |
| COH1.ORF00728 | 171 aa | secreted protein, putative           |
| COH1.ORF00729 | 870 aa | proteinase, putative                 |
| COH1.ORF00730 | 147 aa | conserved hypothetical protein       |
| COH1.ORF00731 | 964 aa | conserved hypothetical protein       |
| COH1.ORF00732 | 361 aa | endopeptidase O [imported], putative |
| COH1.ORF00733 | 272 aa | endopeptidase O [imported]           |
| COH1.ORF00734 | 36 aa  | endopeptidase O                      |
| COH1.ORF00735 | 343 aa | protein of unknown function          |
| COH1.ORF00736 | 83 aa  | protein of unknown function          |
| COH1.ORF00737 | 89 aa  | protein of unknown function          |
| COH1.ORF00738 | 189 aa | protein of unknown function          |
| COH1.ORF00739 | 353 aa | protein of unknown function          |

## H36B [652301-665155]

|               |        |                                |
|---------------|--------|--------------------------------|
| H36B.ORF00734 | 113 aa | conserved hypothetical protein |
| H36B.ORF00735 | 171 aa | secreted protein, putative     |

|               |         |                                      |
|---------------|---------|--------------------------------------|
| H36B.ORF00736 | 870 aa  | proteinase, putative                 |
| H36B.ORF00737 | 1105 aa | conserved hypothetical protein       |
| H36B.ORF00738 | 80 aa   | hypothetical protein                 |
| H36B.ORF00739 | 221 aa  | endopeptidase O [imported], putative |
| H36B.ORF00740 | 272 aa  | endopeptidase O [imported]           |
| H36B.ORF00741 | 36 aa   | endopeptidase O                      |
| H36B.ORF00742 | 253 aa  | protein of unknown function          |
| H36B.ORF00743 | 83 aa   | protein of unknown function          |
| H36B.ORF00744 | 339 aa  | protein of unknown function          |
| H36B.ORF00745 | 353 aa  | protein of unknown function          |

## 18RS21 [603205-615971]

|                 |         |                                |
|-----------------|---------|--------------------------------|
| 18RS21.ORF00633 | 113 aa  | conserved hypothetical protein |
| 18RS21.ORF00634 | 171 aa  | Pac protein precursor          |
| 18RS21.ORF00635 | 870 aa  | proteinase, putative           |
| 18RS21.ORF00636 | 1075 aa | conserved hypothetical protein |
| 18RS21.ORF00637 | 361 aa  | endopeptidase O, putative      |
| 18RS21.ORF00638 | 272 aa  | endopeptidase PepO             |
| 18RS21.ORF00639 | 36 aa   | end peptidase                  |
| 18RS21.ORF00640 | 343 aa  | protein of unknown function    |
| 18RS21.ORF00641 | 339 aa  | protein of unknown function    |
| 18RS21.ORF00642 | 353 aa  | conserved hypothetical protein |

## REGION 1.7 [7 sequence(s)]:

## 2603V/R [800810-807331]

|         |        |  |
|---------|--------|--|
| SAG0801 | 239 aa | ribosomal small subunit pseudouridine synthase A, putative |
| SAG0802 | 38 aa  | hypothetical protein                                       |
| SAG0803 | 383 aa | major facilitator family protein                           |
| SAG0804 | 315 aa | competence protein CoiA                                    |
| SAG0805 | 601 aa | oligoendopeptidase B                                       |
| SAG0806 | 208 aa | hydrolase, haloacid dehalogenase-like family               |
| SAG0807 | 235 aa | O-methyltransferase family protein                         |

## NEM316 [845194-851715]

|         |        |  |
|---------|--------|--|
| gbs0821 | 239 aa | Ribosomal small subunit pseudouridine synthase A |
| gbs0822 | 383 aa | major facilitator family protein                 |
| gbs0823 | 315 aa | competence protein CoiA                          |
| gbs0824 | 601 aa | oligoendopeptidase F                             |
| gbs0825 | 208 aa | hydrolase, haloacid dehalogenase-like family     |
| gbs0826 | 223 aa | O-methyltransferase family protein               |

## A909 [877462-883983]

|               |        |   |
|---------------|--------|---|
| A909.ORF01215 | 239 aa | RNA pseudouridine synthase family protein |
| A909.ORF01214 | 383 aa | major facilitator family protein          |
| A909.ORF01213 | 315 aa | competence protein CoiA, putative         |
| A909.ORF01212 | 601 aa | group B oligopeptidase PepB               |

A909.ORF01211 208 aa HAD-superfamily hydrolase, subfamily IA, variant 1 family protein  
 A909.ORF01210 223 aa O-methyltransferase family protein

## 515 [764722-771476]

515.ORF00861 96 aa Unknown  
 515.ORF00862 240 aa hydrolase, haloacid dehalogenase-like family  
 515.ORF00863 220 aa major facilitator family protein  
 515.ORF00864 315 aa competence protein CoIA  
 515.ORF00865 601 aa oligoendopeptidase F  
 515.ORF00866 208 aa hydrolase, haloacid dehalogenase-like family  
 515.ORF00867 223 aa O-methyltransferase family protein

## CJB111 [799466-806092]

CJB111.ORF00827 239 aa ribosomal small subunit pseudouridine synthase A  
 CJB111.ORF00828 383 aa major facilitator family protein  
 CJB111.ORF00829 315 aa competence protein CoIA  
 CJB111.ORF00830 601 aa oligoendopeptidase F  
 CJB111.ORF00831 208 aa hydrolase, haloacid dehalogenase-like family  
 CJB111.ORF00832 223 aa O-methyltransferase family protein

## COH1 [819751-829200]

COH1.ORF00895 239 aa ribosomal small subunit pseudouridine synthase A  
 COH1.ORF00896 217 aa major facilitator family protein  
 COH1.ORF00897 309 aa PTS system, beta-glucosides-specific IIABC components  
 COH1.ORF00898 91 aa phage infection protein, putative  
 COH1.ORF00899 102 aa 6-phospho-beta-glucosidase  
 COH1.ORF00900 139 aa permease, GntP family  
 COH1.ORF00901 147 aa Unknown  
 COH1.ORF00902 135 aa major facilitator family protein  
 COH1.ORF00903 315 aa competence protein CoIA  
 COH1.ORF00904 601 aa oligoendopeptidase F  
 COH1.ORF00905 208 aa hydrolase, haloacid dehalogenase-like family  
 COH1.ORF00906 223 aa O-methyltransferase family protein

## H36B [796428-802871]

H36B.ORF00885 239 aa ribosomal small subunit pseudouridine synthase A  
 H36B.ORF00886 383 aa major facilitator family protein  
 H36B.ORF00887 315 aa competence protein CoIA  
 H36B.ORF00888 601 aa oligoendopeptidase F  
 H36B.ORF00889 208 aa hydrolase, haloacid dehalogenase-like family  
 H36B.ORF00890 188 aa O-methyltransferase family protein

## REGION 1.8 [3 sequence(s)]:

## 2603V/R [922143-943033]

SAG0914 203 aa conserved hypothetical protein  
 SAG0915 405 aa Tn916, transposase  
 SAG0916 67 aa Tn916, excisionase

|         |        |  |
|---------|--------|--|
| SAG0917 | 83 aa  | Tn916, hypothetical protein                                |
| SAG0918 | 76 aa  | Tn916, hypothetical protein                                |
| SAG0919 | 157 aa | Tn916, hypothetical protein                                |
| SAG0920 | 23 aa  | Tn916, hypothetical protein                                |
| SAG0921 | 117 aa | Tn916, transcriptional regulator, putative                 |
| SAG0922 | 61 aa  | Tn916, hypothetical protein                                |
| SAG0923 | 639 aa | Tn916, tetracycline resistance protein                     |
| SAG0924 | 28 aa  | Tn916, tetM leader peptide                                 |
| SAG0925 | 310 aa | Tn916, hypothetical protein                                |
| SAG0926 | 333 aa | Tn916, NLP/P60 family protein                              |
| SAG0927 | 725 aa | membrane protein, putative                                 |
| SAG0928 | 0 aa   | Tn916, hypothetical protein, authentic frameshift          |
| SAG0929 | 168 aa | Tn916, hypothetical protein                                |
| SAG0930 | 165 aa | Tn916, hypothetical protein                                |
| SAG0931 | 73 aa  | Tn916, hypothetical protein                                |
| SAG0932 | 401 aa | Tn916, transcriptional regulator, putative                 |
| SAG0933 | 461 aa | Tn916, FtsK/SpoIIIE family protein                         |
| SAG0934 | 128 aa | Tn916, hypothetical protein                                |
| SAG0935 | 104 aa | Tn916, hypothetical protein                                |
| SAG0936 | 39 aa  | Tn916, hypothetical protein                                |
| SAG0937 | 0 aa   | ABC transporter, ATP-binding protein, authentic frameshift |

## 515 [879366-898394]

|              |        |  |
|--------------|--------|--|
| 515.ORF00975 | 147 aa | D-tyrosyl-tRNA(Tyr) deacylase              |
| 515.ORF00976 | 738 aa | GTP pyrophosphokinase family protein       |
| 515.ORF00977 | 405 aa | transposase                                |
| 515.ORF00978 | 67 aa  | conserved hypothetical protein             |
| 515.ORF00979 | 140 aa | conserved hypothetical protein             |
| 515.ORF00980 | 117 aa | Tn916, transcriptional regulator, putative |
| 515.ORF00981 | 639 aa | tetracycline resistance protein (tetM)     |
| 515.ORF00982 | 304 aa | conserved hypothetical protein             |
| 515.ORF00983 | 333 aa | lipoprotein, NLP/P60 family                |
| 515.ORF00984 | 725 aa | conserved hypothetical protein             |
| 515.ORF00986 | 815 aa | ATP/GTP-binding protein, putative          |
| 515.ORF00987 | 168 aa | conserved hypothetical protein             |
| 515.ORF00988 | 165 aa | conserved hypothetical protein             |
| 515.ORF00989 | 57 aa  | conserved hypothetical protein             |
| 515.ORF00990 | 165 aa | conserved hypothetical protein             |
| 515.ORF00991 | 73 aa  | conserved hypothetical protein             |
| 515.ORF00992 | 267 aa | transcriptional regulator, Cro/CI family   |

## COH1 [979057-999935]

|               |        |  |
|---------------|--------|--|
| COH1.ORF01053 | 104 aa | uncharacterized conserved protein CAC1439 [imported] |
| COH1.ORF01054 | 125 aa | conserved hypothetical protein                       |
| COH1.ORF01055 | 257 aa | conserved hypothetical protein                       |

|               |        |  |
|---------------|--------|--|
| COH1.ORF01056 | 30 aa  | hypothetical protein   |
| COH1.ORF01058 | 467 aa | TN916 ORF21 homolog lmo1112 [imported]   |
| COH1.ORF01057 | 52 aa  | hypothetical protein   |
| COH1.ORF01059 | 394 aa | transcriptional regulator, Cro/CI family   |
| COH1.ORF01060 | 90 aa  | conserved hypothetical protein   |
| COH1.ORF01061 | 73 aa  | conserved hypothetical protein   |
| COH1.ORF01062 | 220 aa | conserved hypothetical protein   |
| COH1.ORF01063 | 166 aa | TN916 ORF18 homolog lmo1108 [imported]   |
| COH1.ORF01065 | 341 aa | hypothetical protein   |
| COH1.ORF01064 | 30 aa  | hypothetical protein   |
| COH1.ORF01066 | 145 aa | TN916 ORF17 homolog lmo1107 [imported]   |
| COH1.ORF01067 | 817 aa | ATP/GTP-binding protein, putative  |
| COH1.ORF01068 | 675 aa | membrane protein, putative   |
| COH1.ORF01069 | 340 aa | TN916 ORF14 and to <i>L. monocytogenes</i> P60 protein homolog<br>lmo1104 [imported] |
| COH1.ORF01070 | 282 aa | TN916 ORF13 homolog lmo1103 [imported]   |
| COH1.ORF01071 | 639 aa | tetracycline resistance protein (tetM)   |
| COH1.ORF01072 | 117 aa | Tn916, transcriptional regulator, putative   |
| COH1.ORF01073 | 188 aa | conserved hypothetical protein   |
| COH1.ORF01074 | 76 aa  | conserved hypothetical protein   |
| COH1.ORF01075 | 81 aa  | conserved hypothetical protein   |
| COH1.ORF01076 | 397 aa | site-specific recombinase, phage integrase family                                    |

## REGION 1.9 [5 sequence(s)]:

2603V/R [1028558-1036565]

|         |        |  |
|---------|--------|--|
| SAG1018 | 40 aa  | lipoprotein, putative                    |
| SAG1019 | 39 aa  | hypothetical protein                     |
| SAG1020 | 227 aa | lipoprotein, putative                    |
| SAG1021 | 107 aa | hypothetical protein                     |
| SAG1022 | 177 aa | hypothetical protein                     |
| SAG1023 | 48 aa  | hypothetical protein                     |
| SAG1024 | 183 aa | lipoprotein, putative                    |
| SAG1025 | 149 aa | hypothetical protein                     |
| SAG1026 | 0 aa   | immunogenic secreted protein, degenerate |
| SAG1027 | 84 aa  | conserved hypothetical protein           |
| SAG1028 | 196 aa | hypothetical protein                     |
| SAG1029 | 101 aa | hypothetical protein                     |

NEM316 [1098347-1106336]

|         |        |                                |
|---------|--------|--------------------------------|
| gbs1053 | 127 aa | lipoprotein, putative          |
| gbs1055 | 107 aa | conserved hypothetical protein |
| gbs1056 | 177 aa | conserved hypothetical protein |
| gbs1057 | 180 aa | lipoprotein, putative          |
| gbs1061 | 142 aa | conserved hypothetical protein |

|                                 |        |   |          |
|---------------------------------|--------|---|----------|
| gbs1062                         | 84 aa  | uncharacterized protein, homolog of <i>Bacillus firmus</i> (gi, | 2654481) |
| gbs1063                         | 196 aa | conserved hypothetical protein                                  |          |
| gbs1064                         | 101 aa | conserved hypothetical protein                                  |          |
| <b>515 [1025263-1033262]</b>    |        |   |          |
| 515.ORF01128                    | 40 aa  | lipoprotein, putative   |          |
| 515.ORF01129                    | 39 aa  | conserved hypothetical protein                                  |          |
| 515.ORF01130                    | 227 aa | lipoprotein, putative   |          |
| 515.ORF01131                    | 107 aa | conserved hypothetical protein                                  |          |
| 515.ORF01132                    | 177 aa | conserved hypothetical protein                                  |          |
| 515.ORF01133                    | 48 aa  | conserved hypothetical protein                                  |          |
| 515.ORF01134                    | 180 aa | lipoprotein, putative   |          |
| 515.ORF01135                    | 149 aa | FtsK/SpoIIIIE family protein                                    |          |
| 515.ORF01136                    | 86 aa  | N-acetylmuramoyl-L-alanine amidase domain protein               |          |
| 515.ORF01137                    | 269 aa | immunogenic secreted protein, putative                          |          |
| 515.ORF01138                    | 142 aa | Unknown   |          |
| 515.ORF01139                    | 84 aa  | conserved hypothetical protein                                  |          |
| 515.ORF01140                    | 196 aa | conserved hypothetical protein                                  |          |
| 515.ORF01141                    | 101 aa | conserved hypothetical protein                                  |          |
| <b>CJB111 [1010521-1021611]</b> |        |   |          |
| CJB111.ORF01027                 | 38 aa  | hypothetical protein  |          |
| CJB111.ORF01028                 | 70 aa  | hypothetical protein  |          |
| CJB111.ORF01029                 | 176 aa | conserved hypothetical protein                                  |          |
| CJB111.ORF01030                 | 238 aa | lipoprotein, putative   |          |
| CJB111.ORF01031                 | 180 aa | FtsK/SpoIIIIE family protein                                    |          |
| CJB111.ORF01032                 | 295 aa | N-acetylmuramoyl-L-alanine amidase domain protein               |          |
| CJB111.ORF01033                 | 138 aa | hypothetical protein  |          |
| CJB111.ORF01034                 | 83 aa  | conserved hypothetical protein                                  |          |
| CJB111.ORF01035                 | 294 aa | hypothetical protein  |          |
| CJB111.ORF01036                 | 247 aa | hypothetical protein  |          |
| CJB111.ORF01037                 | 81 aa  | protein of unknown function                                     |          |
| CJB111.ORF01038                 | 47 aa  | conserved hypothetical protein                                  |          |
| CJB111.ORF01039                 | 97 aa  | conserved hypothetical protein                                  |          |
| CJB111.ORF01040                 | 216 aa | hypothetical protein  |          |
| CJB111.ORF01041                 | 84 aa  | hypothetical protein  |          |
| CJB111.ORF01042                 | 199 aa | hypothetical Cytosolic Protein, putative                        |          |
| CJB111.ORF01043                 | 104 aa | conserved hypothetical protein                                  |          |
| CJB111.ORF01044                 | 181 aa | hypothetical protein  |          |
| <b>18RS21 [921301-930038]</b>   |        |   |          |
| 18RS21.ORF00954                 | 40 aa  | lipoprotein, putative   |          |
| 18RS21.ORF00955                 | 227 aa | lipoprotein, putative   |          |
| 18RS21.ORF00956                 | 107 aa | conserved hypothetical protein                                  |          |
| 18RS21.ORF00957                 | 177 aa | conserved hypothetical protein                                  |          |
| 18RS21.ORF00958                 | 186 aa | lipoprotein, putative   |          |

|                 |        |  |
|-----------------|--------|--|
| 18RS21.ORF00959 | 149 aa | FtsK/SpoIIIE family protein  |
| 18RS21.ORF00960 | 86 aa  | immunogenic secreted protein precursor homolog                           |
| 18RS21.ORF00961 | 269 aa | immunogenic secreted protein precursor homolog                           |
| 18RS21.ORF00962 | 142 aa | conserved hypothetical protein   |
| 18RS21.ORF00963 | 84 aa  | uncharacterized protein, homolog of <i>Bacillus firmus</i> (gi, 2654481) |
| 18RS21.ORF00964 | 196 aa | conserved hypothetical protein   |
| 18RS21.ORF00965 | 101 aa | conserved hypothetical protein   |

## REGION 1.10 [7 sequence(s)]:

2603V/R [1133383-1143006]

|         |        |   |
|---------|--------|---|
| SAG1126 | 228 aa | protein of unknown function   |
| SAG1127 | 446 aa | conserved domain protein  |
| SAG1128 | 65 aa  | transcriptional regulator, Cro/CI family                                      |
| SAG1129 | 36 aa  | hypothetical protein  |
| SAG1130 | 49 aa  | hypothetical protein  |
| SAG1131 | 164 aa | thiol peroxidase  |
| SAG1132 | 219 aa | conserved hypothetical protein  |
| SAG1133 | 254 aa | conserved hypothetical protein  |
| SAG1134 | 213 aa | transcriptional regulator, GntR family/potassioum uptake protein, TrkA family |
| SAG1135 | 183 aa | gls24 protein, putative   |
| SAG1136 | 65 aa  | conserved hypothetical protein  |
| SAG1137 | 180 aa | gls24 protein, putative   |
| SAG1138 | 64 aa  | conserved hypothetical protein  |
| SAG1139 | 193 aa | conserved hypothetical protein  |
| SAG1140 | 82 aa  | conserved hypothetical protein  |
| SAG1141 | 112 aa | conserved hypothetical protein  |

NEM316 [1236464-1246102]

|         |        |   |
|---------|--------|---|
| gbs1194 | 228 aa | conserved hypothetical protein  |
| gbs1195 | 447 aa | conserved hypothetical protein  |
| gbs1197 | 177 aa | conserved hypothetical protein  |
| gbs1198 | 164 aa | thiol peroxidase  |
| gbs1199 | 219 aa | DNA alkylation repair enzyme  |
| gbs1200 | 254 aa | HAD-superfamily hydrolase, subfamily IIB subfamily                            |
| gbs1201 | 213 aa | transcriptional regulator, GntR family/potassioum uptake protein, TrkA family |
| gbs1202 | 183 aa | gls24 protein   |
| gbs1203 | 65 aa  | CsbD-like superfamily   |
| gbs1204 | 180 aa | EF0117  |
| gbs1205 | 64 aa  | conserved hypothetical protein  |
| gbs1206 | 193 aa | conserved hypothetical protein  |
| gbs1207 | 82 aa  | integral membrane protein   |
| gbs1208 | 112 aa | integral membrane protein   |

## A909 [1189137-1198714]

|               |        |  |
|---------------|--------|--|
| A909.ORF00928 | 228 aa | conserved hypothetical protein                             |
| A909.ORF00927 | 0 aa   | conserved hypothetical protein, degenerate                 |
| A909.ORF00925 | 62 aa  | DNA-binding protein  |
| A909.ORF00924 | 177 aa | conserved hypothetical protein                             |
| A909.ORF00923 | 164 aa | thiol peroxidase   |
| A909.ORF00922 | 219 aa | conserved hypothetical protein                             |
| A909.ORF00920 | 0 aa   | conserved hypothetical protein, authentic frameshift       |
| A909.ORF00919 | 213 aa | transcriptional regulator, GntR family/TrkA domain protein |
| A909.ORF00918 | 183 aa | alkaline shock protein 23, putative                        |
| A909.ORF00917 | 65 aa  | CsbD family protein  |
| A909.ORF00916 | 180 aa | alkaline shock protein 23, putative                        |
| A909.ORF00915 | 64 aa  | conserved hypothetical protein                             |
| A909.ORF00914 | 193 aa | conserved hypothetical protein                             |
| A909.ORF00913 | 82 aa  | transglycosylase associated protein                        |
| A909.ORF00912 | 81 aa  | transglycosylase associated protein                        |

## 515 [1132658-1142281]

|              |        |  |
|--------------|--------|--|
| 515.ORF01250 | 228 aa | protein of unknown function  |
| 515.ORF01251 | 356 aa | conserved hypothetical protein   |
| 515.ORF01252 | 35 aa  | conserved hypothetical protein   |
| 515.ORF01253 | 31 aa  | Unknown  |
| 515.ORF01254 | 62 aa  | conserved hypothetical protein   |
| 515.ORF01255 | 177 aa | conserved hypothetical protein   |
| 515.ORF01256 | 164 aa | thiol peroxidase   |
| 515.ORF01257 | 219 aa | Unknown  |
| 515.ORF01259 | 179 aa | conserved hypothetical protein   |
| 515.ORF01260 | 30 aa  | conserved hypothetical protein   |
| 515.ORF01261 | 213 aa | transcriptional regulator, GntR family/potassioum uptake protein, TrkA |

## family

|              |        |  |
|--------------|--------|--|
| 515.ORF01262 | 183 aa | gls24 protein                              |
| 515.ORF01263 | 65 aa  | CsbD-like superfamily                      |
| 515.ORF01264 | 180 aa | gls24 protein                              |
| 515.ORF01265 | 64 aa  | conserved hypothetical protein             |
| 515.ORF01266 | 193 aa | conserved hypothetical protein             |
| 515.ORF01267 | 82 aa  | Transglycosylase associated protein family |
| 515.ORF01268 | 112 aa | Transglycosylase associated protein family |

## CJB111 [1118612-1128433]

|                 |        |                                |
|-----------------|--------|--------------------------------|
| CJB111.ORF01140 | 214 aa | protein of unknown function    |
| CJB111.ORF01141 | 447 aa | Unknown                        |
| CJB111.ORF01142 | 36 aa  | conserved hypothetical protein |
| CJB111.ORF01143 | 177 aa | conserved hypothetical protein |
| CJB111.ORF01144 | 164 aa | thiol peroxidase               |
| CJB111.ORF01145 | 200 aa | Unknown                        |

|                        |        |  |
|------------------------|--------|--|
| CJB111.ORF01146        | 67 aa  | 6-phospho-beta-glucosidase                                       |
| CJB111.ORF01147        | 99 aa  | conserved hypothetical protein                                   |
| CJB111.ORF01148        | 156 aa | conserved hypothetical protein                                   |
| CJB111.ORF01149        | 213 aa | transcriptional regulator, GntR family/potassium uptake protein, |
| TrkA family            |        |  |
| CJB111.ORF01150        | 183 aa | gls24 protein  |
| CJB111.ORF01151        | 65 aa  | CsbD-like superfamily  |
| CJB111.ORF01152        | 180 aa | gls24 protein  |
| CJB111.ORF01153        | 64 aa  | conserved hypothetical protein                                   |
| CJB111.ORF01154        | 193 aa | conserved hypothetical protein                                   |
| CJB111.ORF01155        | 82 aa  | Transglycosylase associated protein family                       |
| CJB111.ORF01156        | 112 aa | Transglycosylase associated protein family                       |
| COH1 [1163446-1170912] |        |  |
| COH1.ORF01255          | 256 aa | spermidine/putrescine ABC transporter, ATP-binding protein       |
| COH1.ORF01256          | 447 aa | Unknown  |
| COH1.ORFA01216         | 62 aa  | transcription regulator, probable -related protein               |
| COH1.ORF01257          | 177 aa | conserved hypothetical protein                                   |
| COH1.ORF01258          | 164 aa | thiol peroxidase   |
| COH1.ORF01259          | 177 aa | Unknown  |
| COH1.ORFA01220         | 38 aa  | hypothetical protein   |
| COH1.ORF01260          | 64 aa  | conserved hypothetical protein                                   |
| COH1.ORF01261          | 193 aa | conserved hypothetical protein                                   |
| COH1.ORF01262          | 82 aa  | Transglycosylase associated protein family                       |
| COH1.ORF01263          | 112 aa | Transglycosylase associated protein family                       |
| H36B [1114116-1123650] |        |  |
| H36B.ORF01213          | 228 aa | protein of unknown function                                      |
| H36B.ORF01214          | 356 aa | conserved hypothetical protein                                   |
| H36B.ORF01215          | 35 aa  | conserved hypothetical protein                                   |
| H36B.ORF01216          | 31 aa  | Unknown  |
| H36B.ORF01217          | 62 aa  | conserved hypothetical protein                                   |
| H36B.ORF01218          | 177 aa | conserved hypothetical protein                                   |
| H36B.ORF01219          | 164 aa | thiol peroxidase   |
| H36B.ORF01220          | 219 aa | Unknown  |
| H36B.ORF01222          | 179 aa | conserved hypothetical protein                                   |
| H36B.ORF01223          | 30 aa  | conserved hypothetical protein                                   |
| H36B.ORF01224          | 179 aa | transcriptional regulator, GntR family/potassium uptake protein, |
| TrkA family            |        |  |
| H36B.ORF01225          | 73 aa  | gls24 protein  |
| H36B.ORF01226          | 38 aa  | conserved hypothetical protein                                   |
| H36B.ORF01227          | 180 aa | gls24 protein  |
| H36B.ORF01228          | 64 aa  | conserved hypothetical protein                                   |
| H36B.ORF01229          | 193 aa | conserved hypothetical protein                                   |
| H36B.ORF01230          | 82 aa  | Transglycosylase associated protein family                       |

H36B.ORF01231 112 aa Transglycosylase associated protein family

**REGION 1.11 [2 sequence(s)]:**

2603V/R [1166651-1174667]

|         |        |   |
|---------|--------|---|
| SAG1163 | 318 aa | polysaccharide biosynthesis protein CpsK(V) |
| SAG1164 | 321 aa | glycosyl transferase CpsJ(V)                |
| SAG1165 | 327 aa | glycosyl transferase CpsO(V)                |
| SAG1166 | 295 aa | glycosyl transferase CpsN(V)                |
| SAG1167 | 241 aa | polysaccharide biosynthesis protein CpsM(V) |
| SAG1168 | 364 aa | polysaccharide biosynthesis protein cpsH(V) |
| SAG1169 | 163 aa | glycosyl transferase CpsG(V)                |
| SAG1170 | 149 aa | polysaccharide biosynthesis protein CpsF    |

CJB111 [1150988-1157652]

|                 |        |  |
|-----------------|--------|--|
| CJB111.ORF01179 | 318 aa | capsular polysaccharide biosynthesis protein cpsL [imported] |
| CJB111.ORF01180 | 321 aa | glycosyl transferase CpsJ                                    |
| CJB111.ORF01181 | 327 aa | glycosyl transferase CpsO                                    |
| CJB111.ORF01182 | 295 aa | glycosyl transferase CpsN                                    |
| CJB111.ORF01183 | 163 aa | glycosyl transferase CpsG                                    |
| CJB111.ORF01184 | 149 aa | conserved hypothetical protein                               |

**REGION 1.12 [6 sequence(s)]:**

2603V/R [1229953-1237290]

SAG1221 0 aa glycerophosphoryl diester phosphodiesterase, putative, authentic point mutation

|         |        |                                |
|---------|--------|--------------------------------|
| SAG1222 | 593 aa | excinuclease ABC, C subunit    |
| SAG1223 | 255 aa | conserved hypothetical protein |
| SAG1224 | 446 aa | MATE efflux family protein     |
| SAG1225 | 136 aa | conserved hypothetical protein |
| SAG1226 | 165 aa | conserved hypothetical protein |

NEM316 [1338964-1344350]

|         |        |  |
|---------|--------|--|
| gbs1294 | 593 aa | excinuclease ABC, C subunit              |
| gbs1295 | 255 aa | ncharacterized conserved protein CAC1577 |
| gbs1296 | 446 aa | MATE efflux family protein               |
| gbs1297 | 136 aa | glyoxalase family protein superfamily    |
| gbs1298 | 165 aa | conserved hypothetical protein           |

A909 [1286183-1291571]

|               |        |                                |
|---------------|--------|--------------------------------|
| A909.ORF00832 | 593 aa | UvrABC system, C protein       |
| A909.ORF00831 | 255 aa | conserved hypothetical protein |
| A909.ORF00830 | 446 aa | MATE efflux family protein     |
| A909.ORF00829 | 136 aa | glyoxalase family protein      |
| A909.ORF00828 | 165 aa | conserved hypothetical protein |

515 [1230984-1237565]

|              |        |                             |
|--------------|--------|-----------------------------|
| 515.ORF01359 | 583 aa | excinuclease ABC, C subunit |
|--------------|--------|-----------------------------|

|                        |        |  |
|------------------------|--------|--|
| 515.ORF01360           | 255 aa | Uncharacterized BCR, COG1636 superfamily             |
| 515.ORF01361           | 310 aa | MATE efflux family protein                           |
| 515.ORF01362           | 128 aa | glucose-inhibited division protein A                 |
| 515.ORF01363           | 124 aa | galactosyl transferase cpsD [imported]               |
| 515.ORF01364           | 186 aa | CpsIaC   |
| 515.ORF01365           | 135 aa | purine nucleoside phosphorylase, family 1 [imported] |
| COH1 [1259168-1265278] |        |  |
| COH1.ORF01354          | 277 aa | IS861, transposase OrfB                              |
| COH1.ORF01355          | 178 aa | IS861, transposase OrfA                              |
| COH1.ORF01356          | 583 aa | excinuclease ABC, C subunit                          |
| COH1.ORF01357          | 255 aa | Uncharacterized BCR, COG1636 superfamily             |
| COH1.ORF01358          | 425 aa | MATE efflux family protein                           |
| COH1.ORFA01314         | 32 aa  | hypothetical protein                                 |
| COH1.ORF01360          | 90 aa  | alpha-acetolactate decarboxylase                     |
| COH1.ORF01361          | 37 aa  | membrane protein, putative                           |
| H36B [1215991-1221561] |        |  |
| H36B.ORF01325          | 593 aa | excinuclease ABC, C subunit                          |
| H36B.ORF01326          | 255 aa | Uncharacterized BCR, COG1636 superfamily             |
| H36B.ORF01327          | 373 aa | MATE efflux family protein                           |
| H36B.ORF01328          | 161 aa | ABC transporter, permease protein                    |
| H36B.ORFA01288         | 31 aa  | hypothetical protein                                 |
| H36B.ORF01329          | 161 aa | ABC transporter, permease protein                    |
| H36B.ORF01330          | 120 aa | conserved hypothetical protein                       |

## REGION 1.13 [3 sequence(s)]:

2603V/R [1239695-1312758]

|         |        |   |
|---------|--------|---|
| SAG1231 | 0 aa   | transposase OrfB, IS3 family, degenerate          |
| SAG1232 | 77 aa  | transposase OrfB, IS3 family, truncation          |
| SAG1233 | 822 aa | streptococcal histidine triad family protein      |
| SAG1234 | 306 aa | laminin-binding surface protein                   |
| SAG1235 | 425 aa | GBSi1, group II intron, maturase                  |
| SAG1236 | 0 aa   | C5a peptidase, authentic frameshift               |
| SAG1237 | 444 aa | hypothetical protein                              |
| SAG1238 | 202 aa | hypothetical protein                              |
| SAG1239 | 76 aa  | conserved hypothetical protein                    |
| SAG1240 | 125 aa | conserved hypothetical protein, truncation        |
| SAG1241 | 91 aa  | transposase OrfA, IS3 family                      |
| SAG1242 | 67 aa  | transposase OrfB, IS3 family, truncation          |
| SAG1243 | 96 aa  | ISSdy1, transposase OrfA                          |
| SAG1244 | 259 aa | ISSdy1, transposase OrfB                          |
| SAG1245 | 38 aa  | hypothetical protein                              |
| SAG1246 | 389 aa | hypothetical protein                              |
| SAG1247 | 399 aa | site-specific recombinase, phage integrase family |

|         |         |  |
|---------|---------|--|
| SAG1248 | 75 aa   | conserved hypothetical protein                     |
| SAG1249 | 74 aa   | transcriptional regulator, Cro/CI family           |
| SAG1250 | 621 aa  | Tn5252, relaxase                                   |
| SAG1251 | 121 aa  | Tn5252, Orf 9 protein                              |
| SAG1252 | 120 aa  | Tn5252, Orf 10 protein                             |
| SAG1253 | 435 aa  | transposase, ISL3 family                           |
| SAG1254 | 546 aa  | mercuric reductase                                 |
| SAG1255 | 130 aa  | mercuric resistance operon regulatory protein MerR |
| SAG1256 | 142 aa  | IS861, transposase OrfB, truncation                |
| SAG1257 | 709 aa  | cation-transporting ATPase, E1-E2 family           |
| SAG1258 | 122 aa  | cadmium efflux system accessory protein            |
| SAG1259 | 99 aa   | conserved hypothetical protein                     |
| SAG1260 | 262 aa  | hypothetical protein                               |
| SAG1261 | 198 aa  | conserved hypothetical protein                     |
| SAG1262 | 695 aa  | cation-transporting ATPase, E1-E2 family           |
| SAG1263 | 0 aa    | conserved domain protein, authentic frameshift     |
| SAG1264 | 148 aa  | transcriptional repressor CopY, putative           |
| SAG1265 | 206 aa  | cadmium resistance transporter, putative           |
| SAG1266 | 152 aa  | hypothetical protein                               |
| SAG1267 | 108 aa  | hypothetical protein                               |
| SAG1268 | 230 aa  | repressor protein, putative                        |
| SAG1269 | 44 aa   | hypothetical protein                               |
| SAG1270 | 471 aa  | ImpB/MucB/SamB family protein                      |
| SAG1271 | 116 aa  | conserved hypothetical protein                     |
| SAG1272 | 102 aa  | conserved hypothetical protein                     |
| SAG1273 | 118 aa  | conserved hypothetical protein                     |
| SAG1274 | 129 aa  | conserved hypothetical protein                     |
| SAG1275 | 75 aa   | hypothetical protein                               |
| SAG1276 | 358 aa  | conserved hypothetical protein                     |
| SAG1277 | 163 aa  | hypothetical protein                               |
| SAG1278 | 96 aa   | hypothetical protein                               |
| SAG1279 | 99 aa   | conserved domain protein                           |
| SAG1280 | 2274 aa | SNF2 family protein                                |
| SAG1281 | 183 aa  | hypothetical protein                               |
| SAG1282 | 63 aa   | calcium-binding protein, putative                  |
| SAG1283 | 1631 aa | agglutinin receptor                                |
| SAG1284 | 196 aa  | abortive infection protein AbiGI                   |
| SAG1285 | 281 aa  | abortive infection protein AbiGII                  |
| SAG1286 | 933 aa  | Tn5252, Orf28                                      |
| SAG1287 | 776 aa  | Tn5252, Orf26                                      |
| SAG1288 | 0 aa    | Tn5252, Orf25, degenerate                          |
| SAG1289 | 284 aa  | Tn5252, Orf23                                      |
| SAG1290 | 80 aa   | hypothetical protein                               |

|         |        |   |
|---------|--------|---|
| SAG1291 | 0 aa   | Tn5252, Orf 21 protein, internal deletion |
| SAG1292 | 162 aa | hypothetical protein                      |
| SAG1293 | 194 aa | protease, putative                        |
| SAG1294 | 77 aa  | conserved hypothetical protein            |
| SAG1295 | 127 aa | conserved hypothetical protein            |
| SAG1296 | 142 aa | conserved hypothetical protein            |
| SAG1297 | 451 aa | C-5 cytosine-specific DNA methylase       |
| SAG1298 | 31 aa  | hypothetical protein                      |
| SAG1299 | 272 aa | conserved hypothetical protein            |
| SAG1300 | 57 aa  | conserved hypothetical protein            |
| SAG1301 | 121 aa | ribosomal protein L7/L12                  |
| SAG1302 | 166 aa | ribosomal protein L10                     |

## H36B [1233895-1300439]

|                |         |  |
|----------------|---------|--|
| H36B.ORF01341  | 80 aa   | ISSdy1, transposase OrfA                           |
| H36B.ORF01342  | 169 aa  | FtsK/SpoIIIE family protein, putative              |
| H36B.ORF01343  | 167 aa  | Unknown  |
| H36B.ORF01344  | 104 aa  | Unknown  |
| H36B.ORF01345  | 401 aa  | conserved hypothetical protein                     |
| H36B.ORF01346  | 163 aa  | Unknown  |
| H36B.ORF01347  | 1203 aa | reticulocyte binding protein                       |
| H36B.ORF01348  | 399 aa  | site-specific recombinase, phage integrase family  |
| H36B.ORF01349  | 75 aa   | conserved hypothetical protein                     |
| H36B.ORF01350  | 70 aa   | transcriptional regulator, Cro/CI family           |
| H36B.ORF01351  | 621 aa  | Tn5252, relaxase                                   |
| H36B.ORF01352  | 121 aa  | Tn5252, Orf 9 protein                              |
| H36B.ORF01353  | 120 aa  | Tn5252, Orf 10 protein                             |
| H36B.ORFA01313 | 41 aa   | hypothetical protein                               |
| H36B.ORF01355  | 440 aa  | transposase, ISL3 family                           |
| H36B.ORF01356  | 546 aa  | carbamoyl-phosphate synthase, large subunit        |
| H36B.ORF01357  | 130 aa  | mercuric resistance operon regulatory protein MerR |
| H36B.ORFA01318 | 160 aa  | hypothetical protein                               |
| H36B.ORFA01317 | 36 aa   | hypothetical protein                               |
| H36B.ORF01359  | 61 aa   | IS861, transposase OrfB, truncation                |
| H36B.ORF01360  | 709 aa  | cation-transporting ATPase, E1-E2 family           |
| H36B.ORF01361  | 122 aa  | cadmium efflux system accessory protein            |
| H36B.ORF01362  | 262 aa  | conserved hypothetical protein                     |
| H36B.ORF01363  | 198 aa  | integral membrane protein TIGR01906                |
| H36B.ORF01364  | 695 aa  | copper-translocating P-type ATPase                 |
| H36B.ORF01365  | 68 aa   | copper-translocating P-type ATPase                 |
| H36B.ORF01366  | 158 aa  | copper-translocating P-type ATPase                 |
| H36B.ORF01367  | 148 aa  | transcriptional repressor CopY                     |
| H36B.ORF01368  | 206 aa  | cadmium resistance protein                         |
| H36B.ORF01369  | 105 aa  | cation-transporting ATPase, E1-E2 family           |

|                          |         |   |
|--------------------------|---------|---|
| H36B.ORF01370            | 63 aa   | conserved hypothetical protein            |
| H36B.ORF01371            | 152 aa  | conserved hypothetical protein            |
| H36B.ORF01372            | 108 aa  | conserved hypothetical protein            |
| H36B.ORF01373            | 230 aa  | repressor protein, putative               |
| H36B.ORF01374            | 44 aa   | conserved hypothetical protein            |
| H36B.ORF01375            | 471 aa  | ImpB/MucB/SamB family protein             |
| H36B.ORF01376            | 116 aa  | conserved hypothetical protein            |
| H36B.ORF01377            | 119 aa  | conserved hypothetical protein            |
| H36B.ORF01378            | 119 aa  | conserved hypothetical protein            |
| H36B.ORF01379            | 129 aa  | conserved hypothetical protein            |
| H36B.ORF01380            | 75 aa   | conserved hypothetical protein            |
| H36B.ORF01381            | 358 aa  | conserved hypothetical protein            |
| H36B.ORF01382            | 212 aa  | conserved hypothetical protein            |
| H36B.ORF01383            | 96 aa   | conserved hypothetical protein            |
| H36B.ORFA01340           | 56 aa   | conserved hypothetical protein            |
| H36B.ORF01384            | 172 aa  | SNF2 family protein                       |
| H36B.ORF01385            | 769 aa  | SNF2 family protein                       |
| H36B.ORF01387            | 1355 aa | reticulocyte binding protein              |
| H36B.ORF01388            | 183 aa  | conserved hypothetical protein            |
| H36B.ORF01389            | 63 aa   | calcium-binding protein, putative         |
| H36B.ORF01390            | 1631 aa | reticulocyte binding protein              |
| H36B.ORF01391            | 196 aa  | abortive infection protein AbiGI          |
| H36B.ORF01392            | 289 aa  | abortive infection protein AbiGII         |
| H36B.ORF01393            | 933 aa  | Tn5252, Orf28                             |
| H36B.ORF01394            | 776 aa  | Tn5252, Orf26                             |
| H36B.ORF01395            | 154 aa  | Tn5252, Orf25, degenerate                 |
| H36B.ORF01396            | 284 aa  | Tn5252, Orf23                             |
| H36B.ORF01397            | 80 aa   | conserved hypothetical protein            |
| H36B.ORF01398            | 605 aa  | Tn5252, Orf 21 protein, internal deletion |
| H36B.ORF01399            | 162 aa  | conserved hypothetical protein            |
| H36B.ORF01400            | 194 aa  | protease, putative                        |
| H36B.ORF01401            | 77 aa   | conserved hypothetical protein            |
| H36B.ORF01402            | 158 aa  | conserved hypothetical protein            |
| H36B.ORF01403            | 451 aa  | C-5 cytosine-specific DNA methylase       |
| H36B.ORF01404            | 31 aa   | hypothetical protein                      |
| H36B.ORF01405            | 92 aa   | conserved hypothetical protein            |
| H36B.ORF01406            | 93 aa   | SNF2 family protein                       |
| H36B.ORF01407            | 57 aa   | conserved hypothetical protein            |
| H36B.ORF01408            | 121 aa  | ribosomal protein L7/L12                  |
| H36B.ORF01409            | 166 aa  | ribosomal protein L10                     |
| 18RS21 [1112876-1169347] |         |   |
| 18RS21.ORF01159          | 202 aa  | SAG1238-like protein                      |
| 18RS21.ORF01160          | 86 aa   | conserved hypothetical protein            |

|                 |         |   |
|-----------------|---------|---|
| 18RS21.ORF01161 | 125 aa  | ParA  |
| 18RS21.ORF01162 | 106 aa  | transposase                                       |
| 18RS21.ORF01163 | 86 aa   | transposase                                       |
| 18RS21.ORF01164 | 96 aa   | transposase                                       |
| 18RS21.ORF01165 | 177 aa  | ISSdy1, transposase OrfB                          |
| 18RS21.ORF01166 | 127 aa  | degenerate transposase                            |
| 18RS21.ORF01167 | 129 aa  | Transposase (IS4 family)                          |
| 18RS21.ORF01168 | 411 aa  | conserved hypothetical protein                    |
| 18RS21.ORF01169 | 399 aa  | site-specific recombinase, phage integrase family |
| 18RS21.ORF01170 | 75 aa   | conserved hypothetical protein                    |
| 18RS21.ORF01171 | 70 aa   | transcriptional regulator, Cro/CI family          |
| 18RS21.ORF01172 | 621 aa  | Tn5252, relaxase                                  |
| 18RS21.ORF01173 | 75 aa   | Tn5252, Orf 9 protein                             |
| 18RS21.ORF01174 | 53 aa   | Tn5252, Orf 10 protein                            |
| 18RS21.ORF01175 | 35 aa   | hypothetical protein                              |
| 18RS21.ORF01176 | 160 aa  | hypothetical protein                              |
| 18RS21.ORF01177 | 355 aa  | cation-transporting ATPase, E1-E2 family          |
| 18RS21.ORF01178 | 385 aa  | cation-transporting ATPase, E1-E2 family          |
| 18RS21.ORF01179 | 122 aa  | cadmium efflux system accessory protein           |
| 18RS21.ORF01180 | 262 aa  | conserved hypothetical protein                    |
| 18RS21.ORF01181 | 198 aa  | integral membrane protein TIGR01906               |
| 18RS21.ORF01182 | 502 aa  | copper-translocating P-type ATPase                |
| 18RS21.ORF01183 | 178 aa  | cation-transporting ATPase, E1-E2 family          |
| 18RS21.ORF01184 | 68 aa   | copper-translocating P-type ATPase                |
| 18RS21.ORF01185 | 127 aa  | copper-translocating P-type ATPase                |
| 18RS21.ORF01186 | 105 aa  | transcriptional repressor CopY                    |
| 18RS21.ORF01187 | 124 aa  | CadD  |
| 18RS21.ORF01188 | 152 aa  | conserved hypothetical protein                    |
| 18RS21.ORF01189 | 230 aa  | phage repressor protein                           |
| 18RS21.ORF01190 | 44 aa   | conserved hypothetical protein                    |
| 18RS21.ORF01191 | 471 aa  | ImpB/MucB/SamB family protein                     |
| 18RS21.ORF01192 | 116 aa  | conserved hypothetical protein                    |
| 18RS21.ORF01193 | 129 aa  | conserved hypothetical protein                    |
| 18RS21.ORF01194 | 75 aa   | conserved hypothetical protein                    |
| 18RS21.ORF01195 | 213 aa  | conserved hypothetical protein                    |
| 18RS21.ORF01196 | 2278 aa | SNF2 family protein                               |
| 18RS21.ORF01197 | 183 aa  | conserved hypothetical protein                    |
| 18RS21.ORF01198 | 63 aa   | calcium-binding protein homolog                   |
| 18RS21.ORF01199 | 1631 aa | agglutinin receptor                               |
| 18RS21.ORF01200 | 196 aa  | abortive infection protein AbiGI                  |
| 18RS21.ORF01201 | 281 aa  | abortive infection protein AbiGII                 |
| 18RS21.ORF01202 | 933 aa  | Tn5252, Orf28                                     |
| 18RS21.ORF01203 | 776 aa  | Tn5252, Orf26                                     |

|                 |        |   |
|-----------------|--------|---|
| 18RS21.ORF01204 | 284 aa | Tn5252, Orf23                           |
| 18RS21.ORF01205 | 80 aa  | conserved hypothetical protein          |
| 18RS21.ORF01206 | 605 aa | TraG family protein                     |
| 18RS21.ORF01207 | 162 aa | conserved hypothetical protein          |
| 18RS21.ORF01208 | 194 aa | protease, putative                      |
| 18RS21.ORF01209 | 77 aa  | conserved hypothetical protein          |
| 18RS21.ORF01210 | 451 aa | C-5 cytosine-specific DNA methylase     |
| 18RS21.ORF01211 | 31 aa  | hypothetical protein                    |
| 18RS21.ORF01212 | 272 aa | IFN-response binding factor 1 (IREBF-1) |
| 18RS21.ORF01213 | 57 aa  | conserved hypothetical protein          |
| 18RS21.ORF01214 | 121 aa | ribosomal protein L7/L12                |
| 18RS21.ORF01215 | 31 aa  | 50S ribosomal protein                   |
| 18RS21.ORF01216 | 60 aa  | 50S ribosomal protein                   |

## REGION 1.14 [1 sequence(s)]:

2603V/R [1334107-1342948]

|         |        |  |
|---------|--------|--|
| SAG1329 | 208 aa | GTP pyrophosphokinase family protein   |
| SAG1330 | 68 aa  | hypothetical protein                   |
| SAG1331 | 979 aa | R5 protein                             |
| SAG1332 | 146 aa | transcriptional regulator, MarR family |
| SAG1333 | 690 aa | 5'-nucleotidase family protein         |

## REGION 1.15 [7 sequence(s)]:

2603V/R [1448128-1457502]

|         |        |  |
|---------|--------|--|
| SAG1437 | 55 aa  | hypothetical protein   |
| SAG1438 | 754 aa | glycogen phosphorylase   |
| SAG1439 | 498 aa | 4-alpha-glucanotransferase   |
| SAG1440 | 342 aa | maltose operon repressor MalR, putative                                    |
| SAG1441 | 415 aa | maltose/maltodextrin ABC transporter, maltose/maltodextrin-binding protein |
| SAG1442 | 456 aa | maltose ABC transporter, permease protein                                  |
| SAG1443 | 278 aa | maltose ABC transporter, permease protein                                  |

NEM316 [1559177-1568551]

|         |        |  |
|---------|--------|--|
| gbs1507 | 754 aa | glycogen phosphorylase                       |
| gbs1508 | 498 aa | 4-alpha-glucanotransferase                   |
| gbs1509 | 342 aa | transcription regulator                      |
| gbs1510 | 415 aa | sugar ABC transporter, sugar-binding protein |
| gbs1511 | 456 aa | sugar ABC transporter, permease protein      |
| gbs1512 | 278 aa | sugar ABC transporter, permease protein      |

A909 [1454279-1463652]

|               |        |   |
|---------------|--------|---|
| A909.ORF00665 | 754 aa | glycogen/starch/alpha-glucan phosphorylase              |
| A909.ORF00664 | 498 aa | 4-alpha-glucanotransferase                              |
| A909.ORF00663 | 342 aa | maltose operon transcriptional repressor MalR, putative |

|                                 |        |  |
|---------------------------------|--------|--|
| A909.ORF00662                   | 415 aa | carbohydrate uptake 1 (CUT1) family, carbohydrate-binding protein          |
| A909.ORF00661                   | 456 aa | cyclodextrin ABC transporter, permease protein                             |
| A909.ORF00660                   | 278 aa | cyclodextrin ABC transporter, permease protein                             |
| <b>515 [1379522-1389251]</b>    |        |  |
| 515.ORF01527                    | 754 aa | glycogen phosphorylase   |
| 515.ORF01528                    | 498 aa | 4-alpha-glucanotransferase   |
| 515.ORF01529                    | 342 aa | sugar-binding transcriptional regulator, LacI family                       |
| 515.ORF01530                    | 415 aa | maltose/maltodextrin ABC transporter, maltose/maltodextrin-binding protein |
| 515.ORF01531                    | 456 aa | maltose ABC transporter, permease protein                                  |
| 515.ORF01532                    | 278 aa | maltose ABC transporter, permease protein                                  |
| <b>CJB111 [1369391-1379127]</b> |        |  |
| CJB111.ORF01403                 | 754 aa | glycogen phosphorylase   |
| CJB111.ORF01404                 | 52 aa  | 4-alpha-glucanotransferase   |
| CJB111.ORF01405                 | 450 aa | 4-alpha-glucanotransferase   |
| CJB111.ORF01406                 | 69 aa  | maltose operon repressor MalR, putative                                    |
| CJB111.ORF01407                 | 195 aa | sugar-binding transcriptional regulator, LacI family                       |
| CJB111.ORF01408                 | 415 aa | maltose/maltodextrin ABC transporter, maltose/maltodextrin-binding protein |
| CJB111.ORF01409                 | 456 aa | maltose ABC transporter, permease protein                                  |
| CJB111.ORF01410                 | 278 aa | maltose ABC transporter, permease protein                                  |
| <b>COH1 [1447605-1461182]</b>   |        |  |
| COH1.ORF01557                   | 773 aa | glycogen phosphorylase   |
| COH1.ORF01558                   | 498 aa | 4-alpha-glucanotransferase   |
| COH1.ORF01559                   | 342 aa | sugar-binding transcriptional regulator, LacI family                       |
| COH1.ORF01560                   | 415 aa | maltose/maltodextrin ABC transporter, maltose/maltodextrin-binding protein |
| COH1.ORF01561                   | 114 aa | maltose ABC transporter, permease protein                                  |
| COH1.ORF01562                   | 149 aa | maltose/maltodextrin ABC transporter, maltose/maltodextrin-binding protein |
| COH1.ORF01563                   | 195 aa | maltose/maltodextrin ABC transporter, maltose/maltodextrin-binding protein |
| COH1.ORF01564                   | 145 aa | maltose ABC transporter, permease protein                                  |
| COH1.ORF01565                   | 245 aa | maltose ABC transporter, permease protein                                  |
| COH1.ORF01566                   | 278 aa | maltose ABC transporter, permease protein                                  |
| COH1.ORF01567                   | 490 aa | proton/peptide symporter family protein                                    |
| COH1.ORF01568                   | 178 aa | IS861, transposase OrfA  |
| COH1.ORF01569                   | 30 aa  | IS861, transposase OrfB  |
| <b>H36B [1433347-1442793]</b>   |        |  |
| H36B.ORF01555                   | 534 aa | glycogen phosphorylase   |
| H36B.ORF01556                   | 209 aa | glycogen phosphorylase   |
| H36B.ORF01557                   | 464 aa | 4-alpha-glucanotransferase   |
| H36B.ORF01558                   | 340 aa | sugar-binding transcriptional regulator, LacI family                       |

|               |        |  |
|---------------|--------|--|
| H36B.ORF01559 | 415 aa | maltose/maltodextrin ABC transporter, maltose/maltodextrin-binding protein |
| H36B.ORF01560 | 456 aa | maltose ABC transporter, permease protein                                  |
| H36B.ORF01561 | 278 aa | maltose ABC transporter, permease protein                                  |

## REGION 1.16 [7 sequence(s)]:

2603V/R [1470298-1483258]

|         |        |  |
|---------|--------|--|
| SAG1454 | 398 aa | glycosyl transferase, putative                                   |
| SAG1455 | 295 aa | glycosyl transferase, group 2 family protein                     |
| SAG1456 | 0 aa   | glycosyl transferase, family 8, degenerate                       |
| SAG1457 | 129 aa | IS1381, transposase OrfB   |
| SAG1458 | 127 aa | IS1381, transposase OrfA   |
| SAG1459 | 413 aa | glycosyl transferase, family 8                                   |
| SAG1460 | 401 aa | glycosyl transferase, family 8                                   |
| SAG1461 | 335 aa | conserved hypothetical protein                                   |
| SAG1462 | 970 aa | cell wall surface anchor family protein                          |
| SAG1463 | 0 aa   | transcriptional regulator, RofA family, authentic point mutation |

NEM316 [1581358-1594478]

|         |         |  |
|---------|---------|--|
| gbs1523 | 398 aa  | glycosyl transferase, family 8 SP1767        |
| gbs1524 | 295 aa  | glycosyl transferase, group 2 family protein |
| gbs1525 | 406 aa  | glycosyl transferase, family 8 SP1767        |
| gbs1526 | 413 aa  | glycosyl transferase, family 8 SP1770        |
| gbs1527 | 401 aa  | glycosyl transferase, family 8 SP1765        |
| gbs1528 | 335 aa  | nucleotide sugar synthetase-like protein     |
| gbs1529 | 1310 aa | cell wall surface anchor family protein      |
| gbs1530 | 498 aa  | RogB protein                                 |

A909 [1476459-1489420]

|               |         |  |
|---------------|---------|--|
| A909.ORF00650 | 398 aa  | conserved hypothetical protein                               |
| A909.ORF00649 | 291 aa  | glycosyl transferase, group 2 family protein                 |
| A909.ORF00648 | 394 aa  | glycosyl transferase, family 8                               |
| A909.ORF00647 | 413 aa  | glycosyl transferase, family 8                               |
| A909.ORF00646 | 401 aa  | glycosyl transferase, family 8                               |
| A909.ORF00645 | 335 aa  | conserved hypothetical protein                               |
| A909.ORF00644 | 1258 aa | cell wall surface anchor family protein                      |
| A909.ORF00643 | 0 aa    | transcriptional regulator, RofA family, authentic frameshift |

515 [1402058-1417717]

|              |        |  |
|--------------|--------|--|
| 515.ORF01543 | 398 aa | glycosyl transferase, family 8 SP1767 [imported] |
| 515.ORF01544 | 295 aa | glycosyl transferase, group 2 family protein     |
| 515.ORF01545 | 394 aa | glycosyl transferase, family 8, degenerate       |
| 515.ORF01546 | 413 aa | glycosyl transferase, family 8                   |
| 515.ORF01547 | 401 aa | glycosyl transferase, family 8                   |
| 515.ORF01548 | 335 aa | conserved hypothetical protein                   |
| 515.ORF01549 | 50 aa  | cell wall surface anchor family protein          |

|               |        |  |
|---------------|--------|--|
| 515.ORF01550  | 536 aa | hypothetical protein                                   |
| 515.ORF01551  | 162 aa | glycosyl transferase, family 8                         |
| 515.ORF01552  | 174 aa | conserved hypothetical protein                         |
| 515.ORF01554  | 939 aa | cell wall surface anchor family protein                |
| 515.ORFA01481 | 256 aa | hypothetical protein                                   |
| 515.ORF01555  | 498 aa | transcription regulator, probable [imported], putative |

## CJB111 [1391932-1404505]

|                 |         |  |
|-----------------|---------|--|
| CJB111.ORF01422 | 398 aa  | glycosyl transferase, family 8 SP1767 [imported]       |
| CJB111.ORF01423 | 295 aa  | glycosyl transferase, group 2 family protein           |
| CJB111.ORF01424 | 394 aa  | glycosyl transferase, family 8, degenerate             |
| CJB111.ORF01425 | 413 aa  | glycosyl transferase, family 8                         |
| CJB111.ORF01426 | 401 aa  | glycosyl transferase, family 8                         |
| CJB111.ORF01427 | 335 aa  | conserved hypothetical protein                         |
| CJB111.ORF01429 | 1130 aa | cell wall surface anchor family protein                |
| CJB111.ORF01430 | 273 aa  | transcription regulator, probable [imported], putative |
| CJB111.ORF01431 | 50 aa   | Unknown  |
| CJB111.ORF01432 | 114 aa  | Unknown  |

## H36B [1456431-1469598]

|               |         |  |
|---------------|---------|--|
| H36B.ORF01573 | 398 aa  | glycosyl transferase, family 8 SP1767 [imported]       |
| H36B.ORF01574 | 295 aa  | glycosyl transferase, group 2 family protein           |
| H36B.ORF01575 | 394 aa  | glycosyl transferase, family 8, degenerate             |
| H36B.ORF01576 | 413 aa  | glycosyl transferase, family 8                         |
| H36B.ORF01577 | 401 aa  | glycosyl transferase, family 8                         |
| H36B.ORF01578 | 335 aa  | conserved hypothetical protein                         |
| H36B.ORF01580 | 1326 aa | cell wall surface anchor family protein                |
| H36B.ORF01581 | 498 aa  | transcription regulator, probable [imported], putative |

## 18RS21 [1302019-1313939]

|                 |        |  |
|-----------------|--------|--|
| 18RS21.ORF01350 | 398 aa | glycosyl transferase, family 8 SP1767        |
| 18RS21.ORF01351 | 295 aa | glycosyl transferase, group 2 family protein |
| 18RS21.ORF01352 | 394 aa | glycosyl transferase, family 8 SP1767        |
| 18RS21.ORF01353 | 413 aa | glycosyl transferase, family 8 SP1770        |
| 18RS21.ORF01354 | 401 aa | glycosyl transferase, family 8 SP1765        |
| 18RS21.ORF01355 | 335 aa | nucleotide sugar synthetase-like protein     |
| 18RS21.ORF01356 | 910 aa | cell wall surface anchor family protein      |
| 18RS21.ORF01357 | 498 aa | RogB protein                                 |

## REGION 1.17 [7 sequence(s)]:

## 2603V/R [1506316-1513839]

|         |        |  |
|---------|--------|--|
| SAG1486 | 548 aa | membrane protein, putative             |
| SAG1487 | 233 aa | ABC transporter, ATP binding protein   |
| SAG1488 | 195 aa | dephospho-CoA kinase                   |
| SAG1489 | 273 aa | formamidopyrimidine-DNA glycosylase    |
| SAG1490 | 282 aa | transcriptional regulator, MutR family |

SAG1491 530 aa hypothetical protein  
 SAG1492 58 aa hypothetical protein  
 SAG1493 66 aa hypothetical protein

## NEM316 [1617517-1624996]

gbs1551 548 aa membrane protein, putative  
 gbs1552 233 aa gliding motility protein  
 gbs1553 197 aa dephospho-CoA kinase  
 gbs1554 273 aa formamidopyrimidine-DNA glycosylase  
 gbs1555 282 aa transcriptional activator, Rgg/GadR/MutR family, C-terminal domain protein  
 gbs1556 530 aa conserved hypothetical protein

## A909 [1512462-1519985]

A909.ORF00620 548 aa membrane protein, putative  
 A909.ORF00619 233 aa ABC transporter, ATP binding protein  
 A909.ORF00618 195 aa dephospho-CoA kinase  
 A909.ORF00617 273 aa formamidopyrimidine-DNA glycosylase  
 A909.ORF00616 282 aa transcriptional activator, Rgg/GadR/MutR family  
 A909.ORF00615 530 aa hypothetical protein

## 515 [1443756-1452197]

515.ORF01584 548 aa membrane protein, putative  
 515.ORF01585 233 aa ABC transporter, ATP binding protein  
 515.ORF01586 197 aa dephospho-CoA kinase  
 515.ORF01588 273 aa formamidopyrimidine-DNA glycosylase  
 515.ORF01589 282 aa transcriptional regulator, MutR family  
 515.ORF01590 349 aa conserved hypothetical protein  
 515.ORF01591 264 aa transcriptional regulator, MutR family  
 515.ORF01592 195 aa conserved hypothetical protein  
 515.ORF01593 58 aa conserved hypothetical protein

## CJB111 [1427543-1435355]

CJB111.ORF01455 366 aa membrane protein, putative  
 CJB111.ORF01456 198 aa membrane protein, putative  
 CJB111.ORF01457 233 aa ABC transporter, ATP binding protein  
 CJB111.ORF01458 197 aa dephospho-CoA kinase  
 CJB111.ORF01460 273 aa formamidopyrimidine-DNA glycosylase  
 CJB111.ORF01461 282 aa transcriptional regulator, MutR family  
 CJB111.ORF01462 530 aa conserved hypothetical protein  
 CJB111.ORF01463 58 aa conserved hypothetical protein  
 CJB111.ORFA01420 66 aa conserved hypothetical protein

## COH1 [1488861-1496367]

COH1.ORF01597 531 aa membrane protein, putative  
 COH1.ORF01598 233 aa ABC transporter, ATP binding protein  
 COH1.ORF01599 197 aa dephospho-CoA kinase  
 COH1.ORF01601 273 aa formamidopyrimidine-DNA glycosylase  
 COH1.ORF01602 282 aa transcriptional regulator, MutR family

|                        |        |  |
|------------------------|--------|--|
| COH1.ORF01603          | 529 aa | conserved hypothetical protein         |
| COH1.ORF01604          | 58 aa  | conserved hypothetical protein         |
| H36B [1492637-1500161] |        |  |
| H36B.ORF01604          | 548 aa | membrane protein, putative             |
| H36B.ORF01605          | 233 aa | ABC transporter, ATP binding protein   |
| H36B.ORF01606          | 197 aa | dephospho-CoA kinase                   |
| H36B.ORF01608          | 273 aa | formamidopyrimidine-DNA glycosylase    |
| H36B.ORF01609          | 190 aa | transcriptional regulator, MutR family |
| H36B.ORF01610          | 100 aa | transcriptional regulator, MutR family |
| H36B.ORF01611          | 529 aa | conserved hypothetical protein         |
| H36B.ORF01612          | 58 aa  | conserved hypothetical protein         |
| H36B.ORFA01554         | 45 aa  | conserved hypothetical protein         |

## REGION 1.18 [7 sequence(s)]:

2603V/R [1633028-1638363]

|         |        |  |
|---------|--------|--|
| SAG1631 | 223 aa | potassium uptake protein, Trk family, putative |
| SAG1632 | 276 aa | cobalt transport family protein                |
| SAG1633 | 558 aa | ABC transporter, ATP-binding protein           |
| SAG1634 | 212 aa | conserved hypothetical protein                 |
| SAG1635 | 402 aa | sodium:dicarboxylate symporter family protein  |

NEM316 [1736140-1741477]

|         |        |   |
|---------|--------|---|
| gbs1678 | 223 aa | potassium uptake protein, Trk family          |
| gbs1679 | 276 aa | cobalt transport family protein               |
| gbs1680 | 558 aa | ABC transporter, ATP-binding protein          |
| gbs1681 | 182 aa | conserved hypothetical protein                |
| gbs1682 | 402 aa | sodium:dicarboxylate symporter family protein |

A909 [1635659-1642474]

|               |        |  |
|---------------|--------|--|
| A909.ORF00489 | 223 aa | potassium uptake protein, K <sup>+</sup> transporter (Trk) family  |
| A909.ORF00488 | 276 aa | nickel/cobalt uptake transporter (NiCoT) family ABC transporter, permease protein                          |
| A909.ORF00487 | 558 aa | ABC transporter, ATP-binding protein   |
| A909.ORF00486 | 182 aa | membrane protein, putative   |
| A909.ORF00485 | 402 aa | dicarboxylate/amino acid,cation (Na <sup>+</sup> or H <sup>+</sup> ) symporter (DAACS) family protein YgjU |

|               |      |  |
|---------------|------|--|
| A909.ORF00484 | 0 aa | branched chain amino acid,cation symporter, authentic frameshift |
|---------------|------|--|

515 [1565663-1570723]

|              |        |   |
|--------------|--------|---|
| 515.ORF01729 | 223 aa | potassium uptake protein, Trk family [imported] |
| 515.ORF01730 | 276 aa | cobalt transport family protein                 |
| 515.ORF01731 | 558 aa | ABC transporter, ATP-binding protein            |
| 515.ORF01732 | 212 aa | conserved hypothetical protein                  |
| 515.ORF01733 | 274 aa | sodium:dicarboxylate symporter family protein   |

CJB111 [1547809-1553054]

|                 |        |   |
|-----------------|--------|---|
| CJB111.ORF01589 | 223 aa | potassium uptake protein, Trk family [imported] |
|-----------------|--------|---|

|                        |        |   |
|------------------------|--------|---|
| CJB111.ORF01590        | 276 aa | cobalt transport family protein                               |
| CJB111.ORF01591        | 558 aa | ABC transporter, ATP-binding protein                          |
| CJB111.ORF01592        | 182 aa | conserved hypothetical protein                                |
| CJB111.ORF01593        | 374 aa | sodium:dicarboxylate symporter family protein                 |
| COH1 [1616635-1624981] |        |   |
| COH1.ORF01745          | 223 aa | potassium uptake protein, Trk family [imported]               |
| COH1.ORF01746          | 276 aa | cobalt transport family protein                               |
| COH1.ORF01747          | 558 aa | ABC transporter, ATP-binding protein                          |
| COH1.ORF01748          | 212 aa | conserved hypothetical protein                                |
| COH1.ORF01749          | 402 aa | sodium:dicarboxylate symporter family protein                 |
| COH1.ORF01750          | 241 aa | amidase family protein  |
| COH1.ORF01751          | 105 aa | conserved hypothetical protein                                |
| COH1.ORFA01686         | 100 aa | conserved hypothetical protein                                |
| COH1.ORF01752          | 143 aa | Unknown   |
| COH1.ORF01753          | 443 aa | branched-chain amino acid transport system II carrier protein |
| H36B [1616792-1622629] |        |   |
| H36B.ORF01750          | 223 aa | potassium uptake protein, Trk family [imported]               |
| H36B.ORF01751          | 276 aa | cobalt transport family protein                               |
| H36B.ORF01752          | 259 aa | ABC transporter, ATP-binding protein                          |
| H36B.ORF01753          | 282 aa | ABC transporter, ATP-binding protein                          |
| H36B.ORF01754          | 79 aa  | conserved hypothetical protein                                |
| H36B.ORF01755          | 0 aa   | conserved hypothetical protein                                |
| H36B.ORF01756          | 324 aa | sodium:dicarboxylate symporter family protein                 |
| H36B.ORF01757          | 286 aa | GTP-binding protein, Era/ThdF family                          |

## REGION 1.19 [7 sequence(s)]:

2603V/R [1675066-1683135]

|         |        |  |
|---------|--------|--|
| SAG1679 | 320 aa | asparaginase family protein                |
| SAG1680 | 292 aa | shikimate 5-dehydrogenase                  |
| SAG1681 | 304 aa | oxidoreductase, aldo/keto reductase family |
| SAG1682 | 671 aa | ATP-dependent DNA helicase RecG            |
| SAG1683 | 512 aa | immunogenic secreted protein, putative     |

NEM316 [1778183-1786252]

|         |        |  |
|---------|--------|--|
| gbs1723 | 297 aa | asparaginase family protein                      |
| gbs1724 | 292 aa | shikimate 5-dehydrogenase                        |
| gbs1725 | 304 aa | oxidoreductase, aldo/keto reductase family       |
| gbs1726 | 671 aa | ATP-dependent DNA helicase RecG                  |
| gbs1727 | 512 aa | immunogenic secreted protein precursor homologue |

A909 [1677696-1685765]

|               |        |  |
|---------------|--------|--|
| A909.ORF00443 | 320 aa | asparaginase family protein                |
| A909.ORF00442 | 292 aa | shikimate 5-dehydrogenase                  |
| A909.ORF00441 | 304 aa | oxidoreductase, aldo/keto reductase family |
| A909.ORF00440 | 671 aa | ATP-dependent DNA helicase RecG            |

|                              |         |   |
|------------------------------|---------|---|
| A909.ORF00439                | 512 aa  | immunogenic secreted protein, putative            |
| 515 [1608018-1616087]        |         |   |
| 515.ORF01779                 | 297 aa  | asparaginase family protein                       |
| 515.ORF01780                 | 292 aa  | shikimate 5-dehydrogenase                         |
| 515.ORF01781                 | 304 aa  | oxidoreductase, aldo/keto reductase family        |
| 515.ORF01782                 | 671 aa  | ATP-dependent DNA helicase RecG                   |
| 515.ORF01783                 | 512 aa  | N-acetylmuramoyl-L-alanine amidase, family 4      |
| CJB111 [1590282-1598278]     |         |   |
| CJB111.ORF01637              | 93 aa   | asparaginase family protein                       |
| CJB111.ORF01638              | 78 aa   | asparaginase family protein                       |
| CJB111.ORF01639              | 137 aa  | asparaginase family protein                       |
| CJB111.ORF01640              | 292 aa  | shikimate 5-dehydrogenase                         |
| CJB111.ORF01641              | 304 aa  | oxidoreductase, aldo/keto reductase family        |
| CJB111.ORF01642              | 671 aa  | ATP-dependent DNA helicase RecG                   |
| CJB111.ORF01643              | 86 aa   | N-acetylmuramoyl-L-alanine amidase domain protein |
| CJB111.ORFA01600             | 30 aa   | conserved hypothetical protein                    |
| CJB111.ORF01644              | 317 aa  | immunogenic secreted protein, putative            |
| COH1 [1664831-1672900]       |         |   |
| COH1.ORF01806                | 297 aa  | asparaginase family protein                       |
| COH1.ORF01807                | 292 aa  | shikimate 5-dehydrogenase                         |
| COH1.ORF01808                | 304 aa  | oxidoreductase, aldo/keto reductase family        |
| COH1.ORF01809                | 671 aa  | ATP-dependent DNA helicase RecG                   |
| COH1.ORF01810                | 512 aa  | N-acetylmuramoyl-L-alanine amidase, family 4      |
| H36B [1659913-1667982]       |         |   |
| H36B.ORF01804                | 297 aa  | asparaginase family protein                       |
| H36B.ORF01805                | 292 aa  | shikimate 5-dehydrogenase                         |
| H36B.ORF01806                | 304 aa  | oxidoreductase, aldo/keto reductase family        |
| H36B.ORF01807                | 671 aa  | ATP-dependent DNA helicase RecG                   |
| H36B.ORF01808                | 512 aa  | N-acetylmuramoyl-L-alanine amidase, family 4      |
| REGION 1.20 [1 sequence(s)]: |         |   |
| 2603V/R [1832899-1867259]    |         |   |
| SAG1835                      | 134 aa  | conserved hypothetical protein                    |
| SAG1836                      | 61 aa   | conserved hypothetical protein                    |
| SAG1837                      | 468 aa  | prophage LambdaSa2, lysin, putative               |
| SAG1838                      | 109 aa  | prophage LambdaSa2, holin, putative               |
| SAG1839                      | 136 aa  | conserved hypothetical protein                    |
| SAG1840                      | 112 aa  | hypothetical protein                              |
| SAG1841                      | 76 aa   | conserved domain protein                          |
| SAG1842                      | 1224 aa | prophage LambdaSa2, PblB, putative                |
| SAG1843                      | 240 aa  | conserved hypothetical protein                    |
| SAG1844                      | 911 aa  | conserved hypothetical protein                    |
| SAG1845                      | 42 aa   | hypothetical protein                              |

|         |        |   |
|---------|--------|---|
| SAG1846 | 158 aa | hypothetical protein  |
| SAG1847 | 227 aa | conserved hypothetical protein  |
| SAG1848 | 114 aa | conserved hypothetical protein  |
| SAG1849 | 115 aa | hypothetical protein  |
| SAG1850 | 101 aa | hypothetical protein  |
| SAG1851 | 111 aa | conserved domain protein  |
| SAG1852 | 420 aa | conserved domain protein  |
| SAG1853 | 180 aa | prophage LambdaSa2, protease, putative  |
| SAG1854 | 380 aa | conserved hypothetical protein  |
| SAG1855 | 570 aa | prophage LambdaSa2, terminase large subunit, putative   |
| SAG1856 | 161 aa | hypothetical protein  |
| SAG1857 | 119 aa | prophage LambdaSa2, HNH endonuclease family protein   |
| SAG1858 | 95 aa  | hypothetical protein  |
| SAG1859 | 180 aa | prophage LambdaSa2, site-specific recombinase, phage integrase family                         |
| SAG1860 | 154 aa | conserved hypothetical protein  |
| SAG1861 | 119 aa | prophage LambdaSa2, transcriptional regulator, Cro/CI family                                  |
| SAG1862 | 86 aa  | hypothetical protein  |
| SAG1863 | 138 aa | prophage LambdaSa2, single-strand binding protein   |
| SAG1864 | 68 aa  | hypothetical protein  |
| SAG1865 | 74 aa  | conserved hypothetical protein  |
| SAG1866 | 109 aa | conserved hypothetical protein  |
| SAG1867 | 163 aa | conserved hypothetical protein  |
| SAG1868 | 134 aa | hypothetical protein  |
| SAG1869 | 437 aa | prophage LambdaSa2, type II DNA modification methyltransferase, putative                      |
| SAG1870 | 273 aa | prophage LambdaSa2, DNA replication protein DnaC, putative                                    |
| SAG1871 | 248 aa | prophage LambdaSa2, bacteriophage replication protein/hypothetical protein, truncation/fusion |
| SAG1872 | 200 aa | hypothetical protein  |
| SAG1873 | 443 aa | prophage LambdaSa2, replicative DNA helicase  |
| SAG1874 | 87 aa  | hypothetical protein  |
| SAG1875 | 94 aa  | conserved hypothetical protein  |
| SAG1876 | 176 aa | prophage LambdaSa2, HNH endonuclease family protein   |
| SAG1877 | 236 aa | prophage LambdaSa2, antirepressor protein, putative   |
| SAG1878 | 102 aa | conserved domain protein  |
| SAG1879 | 156 aa | hypothetical protein  |
| SAG1880 | 54 aa  | hypothetical protein  |
| SAG1881 | 51 aa  | hypothetical protein  |
| SAG1882 | 120 aa | prophage LambdaSa2, repressor protein, putative   |
| SAG1883 | 128 aa | conserved hypothetical protein  |
| SAG1884 | 134 aa | hypothetical protein  |
| SAG1885 | 356 aa | prophage LambdaSa2, site-specific recombinase, phage integrase family                         |
| SAG1886 | 32 aa  | hypothetical protein  |

## REGION 1.21 [6 sequence(s)]:

2603V/R [1879464-1887918]

|         |        |  |
|---------|--------|--|
| SAG1898 | 271 aa | PTS system, IID component  |
| SAG1899 | 288 aa | PTS system, IIC component  |
| SAG1900 | 164 aa | PTS system, IIB component  |
| SAG1901 | 398 aa | glucuronyl hydrolase   |
| SAG1902 | 144 aa | PTS system, IIA component  |
| SAG1903 | 34 aa  | hypothetical protein   |
| SAG1904 | 270 aa | oxidoreductase, short-chain dehydrogenase/reductase family                   |
| SAG1905 | 212 aa | conserved hypothetical protein   |
| SAG1906 | 335 aa | carbohydrate kinase, PfkB family   |
| SAG1907 | 212 aa | 2-dehydro-3-deoxyphosphogluconate aldolase/4-hydroxy-2-oxoglutarate aldolase |

NEM316 [1953707-1961808]

|         |        |                                   |
|---------|--------|-----------------------------------|
| gbs1886 | 271 aa | PTS system, IID component         |
| gbs1887 | 288 aa | PTS system, IIC component         |
| gbs1888 | 164 aa | PTS system, IIB component         |
| gbs1889 | 398 aa | unsaturated glucuronyl hydrolase  |
| gbs1890 | 144 aa | PTS system                        |
| gbs1891 | 270 aa | gluconate 5-dehydrogenase         |
| gbs1892 | 212 aa | sugar-phosphate isomerase         |
| gbs1893 | 335 aa | 2-dehydro-3-deoxygluconokinase    |
| gbs1894 | 212 aa | 4-hydroxy-2-oxoglutarate aldolase |

515 [1777377-1785446]

|              |        |  |
|--------------|--------|--|
| 515.ORF01949 | 271 aa | PTS system, IID component  |
| 515.ORF01950 | 288 aa | PTS system, IIC component  |
| 515.ORF01951 | 164 aa | PTS system, IIB component  |
| 515.ORF01952 | 398 aa | glucuronyl hydrolase   |
| 515.ORF01954 | 54 aa  | PTS system, IIA component  |
| 515.ORF01955 | 240 aa | oxidoreductase, short-chain dehydrogenase/reductase family                   |
| 515.ORF01956 | 212 aa | conserved hypothetical protein   |
| 515.ORF01957 | 335 aa | carbohydrate kinase, PfkB family   |
| 515.ORF01959 | 212 aa | 2-dehydro-3-deoxyphosphogluconate aldolase/4-hydroxy-2-oxoglutarate aldolase |

CJB111 [1777345-1784725]

|                 |        |  |
|-----------------|--------|--|
| CJB111.ORF01831 | 271 aa | PTS system, IID component                                  |
| CJB111.ORF01832 | 288 aa | PTS system, IIC component                                  |
| CJB111.ORF01833 | 164 aa | PTS system, IIB component                                  |
| CJB111.ORF01834 | 398 aa | glucuronyl hydrolase                                       |
| CJB111.ORF01835 | 135 aa | PTS system, IIA component                                  |
| CJB111.ORF01836 | 270 aa | oxidoreductase, short-chain dehydrogenase/reductase family |
| CJB111.ORF01837 | 213 aa | conserved hypothetical protein                             |

|                                 |        |  |
|---------------------------------|--------|--|
| CJB111.ORF01838                 | 335 aa | carbohydrate kinase, PfkB family   |
| CJB111.ORF01839                 | 212 aa | 2-dehydro-3-deoxyphosphogluconate aldolase/4-hydroxy-2-oxoglutarate aldolase |
| <b>COH1 [1840727-1848775]</b>   |        |  |
| COH1.ORF01993                   | 271 aa | PTS system, IID component  |
| COH1.ORF01994                   | 288 aa | PTS system, IIC component  |
| COH1.ORF01995                   | 164 aa | PTS system, IIB component  |
| COH1.ORF01996                   | 398 aa | glucuronyl hydrolase   |
| COH1.ORF01998                   | 106 aa | PTS system, IIA component  |
| COH1.ORF01999                   | 240 aa | oxidoreductase, short-chain dehydrogenase/reductase family                   |
| COH1.ORF02000                   | 212 aa | conserved hypothetical protein   |
| COH1.ORF02001                   | 335 aa | carbohydrate kinase, PfkB family   |
| COH1.ORF02003                   | 212 aa | 2-dehydro-3-deoxyphosphogluconate aldolase/4-hydroxy-2-oxoglutarate aldolase |
| <b>18RS21 [1627906-1636125]</b> |        |  |
| 18RS21.ORF01702                 | 271 aa | PTS system, IID component  |
| 18RS21.ORF01703                 | 288 aa | PTS system, IIC component  |
| 18RS21.ORF01704                 | 164 aa | PTS system, IIB component  |
| 18RS21.ORF01705                 | 398 aa | unsaturated glucuronyl hydrolase   |
| 18RS21.ORF01706                 | 106 aa | PTS system, IIA component  |
| 18RS21.ORF01707                 | 270 aa | gluconate 5-dehydrogenase  |
| 18RS21.ORF01708                 | 212 aa | sugar-phosphate isomerase  |
| 18RS21.ORF01709                 | 335 aa | 2-dehydro-3-deoxygluconokinase   |
| 18RS21.ORF01710                 | 212 aa | 4-hydroxy-2-oxoglutarate aldolase  |

**REGION 1.22 [3 sequence(s)]:****2603V/R [1961881-2005812]**

|         |        |   |
|---------|--------|---|
| SAG1979 | 253 aa | membrane protein, putative                        |
| SAG1980 | 300 aa | ABC transporter, ATP-binding protein              |
| SAG1981 | 68 aa  | hypothetical protein                              |
| SAG1982 | 359 aa | transcriptional regulator, Cro/CI family          |
| SAG1983 | 105 aa | conserved hypothetical protein                    |
| SAG1984 | 188 aa | conserved hypothetical protein TIGR00730          |
| SAG1985 | 51 aa  | hypothetical protein                              |
| SAG1986 | 375 aa | site-specific recombinase, phage integrase family |
| SAG1987 | 61 aa  | conserved hypothetical protein                    |
| SAG1988 | 342 aa | conserved hypothetical protein                    |
| SAG1989 | 139 aa | hypothetical protein                              |
| SAG1990 | 127 aa | hypothetical protein                              |
| SAG1991 | 204 aa | transcriptional regulator, Cro/CI family          |
| SAG1992 | 518 aa | protein of unknown function                       |
| SAG1993 | 373 aa | site-specific recombinase, phage integrase family |
| SAG1994 | 108 aa | conserved hypothetical protein                    |

|                       |        |  |
|-----------------------|--------|--|
| SAG1995               | 210 aa | hypothetical protein                               |
| SAG1996               | 263 aa | cell wall surface anchor family protein, putative  |
| SAG1997               | 182 aa | hypothetical protein                               |
| SAG1998               | 457 aa | hypothetical protein                               |
| SAG1999               | 47 aa  | hypothetical protein                               |
| SAG2000               | 666 aa | membrane protein, putative                         |
| SAG2001               | 756 aa | conjugal transfer protein, interruption-C          |
| SAG2002               | 129 aa | IS1381, transposase OrfB                           |
| SAG2003               | 127 aa | IS1381, transposase OrfA                           |
| SAG2004               | 67 aa  | conjugal transfer protein, interruption-N          |
| SAG2005               | 136 aa | conserved hypothetical protein                     |
| SAG2006               | 88 aa  | conserved hypothetical protein                     |
| SAG2007               | 317 aa | conserved hypothetical protein                     |
| SAG2008               | 84 aa  | conserved hypothetical protein                     |
| SAG2009               | 88 aa  | conserved hypothetical protein                     |
| SAG2010               | 157 aa | hypothetical protein                               |
| SAG2011               | 160 aa | conserved hypothetical protein                     |
| SAG2012               | 90 aa  | hypothetical protein                               |
| SAG2013               | 189 aa | hypothetical protein                               |
| SAG2014               | 449 aa | hypothetical protein                               |
| SAG2015               | 99 aa  | transcriptional regulator, Cro/CI family           |
| SAG2016               | 125 aa | hypothetical protein                               |
| SAG2017               | 429 aa | transcriptional regulator, Cro/CI family           |
| SAG2018               | 553 aa | FtsK/SpoIIIE family protein                        |
| SAG2019               | 153 aa | hypothetical protein                               |
| SAG2020               | 98 aa  | hypothetical protein                               |
| SAG2021               | 826 aa | cell wall surface anchor family protein            |
| SAG2022               | 417 aa | transposase, ISL3 family                           |
| SAG2023               | 546 aa | mercuric reductase                                 |
| SAG2024               | 130 aa | mercuric resistance operon regulatory protein MerR |
| SAG2025               | 522 aa | Mn2+/Fe2+ transporter, NRAMP family                |
| SAG2026               | 240 aa | membrane protein, putative                         |
| SAG2027               | 205 aa | ABC transporter, ATP-binding protein               |
| SAG2028               | 36 aa  | conserved hypothetical protein                     |
| 515 [1860750-1907868] |        |  |
| 515.ORF02036          | 373 aa | site-specific recombinase, phage integrase family  |
| 515.ORF02037          | 108 aa | Domain of unknown function (DUF771) superfamily    |
| 515.ORF02038          | 210 aa | conserved hypothetical protein                     |
| 515.ORF02039          | 180 aa | Unknown  |
| 515.ORF02040          | 53 aa  | cell wall surface anchor family protein, putative  |
| 515.ORF02041          | 438 aa | conserved hypothetical protein                     |
| 515.ORF02042          | 289 aa | DNA adenine methylase, putative                    |
| 515.ORF02043          | 259 aa | cell wall surface anchor signal protein            |

|              |        |  |
|--------------|--------|--|
| 515.ORF02045 | 182 aa | conserved hypothetical protein                               |
| 515.ORF02046 | 457 aa | CHAP domain family   |
| 515.ORF02047 | 666 aa | membrane protein, putative                                   |
| 515.ORF02048 | 833 aa | conjugal transfer protein, interruption-C                    |
| 515.ORF02049 | 124 aa | conserved hypothetical protein                               |
| 515.ORF02050 | 88 aa  | conserved hypothetical protein                               |
| 515.ORF02051 | 317 aa | conserved hypothetical protein                               |
| 515.ORF02052 | 84 aa  | conserved hypothetical protein                               |
| 515.ORF02053 | 88 aa  | Uncharacterized ACR, COG2161 subfamily                       |
| 515.ORF02054 | 157 aa | conserved hypothetical protein                               |
| 515.ORF02055 | 161 aa | conserved hypothetical protein                               |
| 515.ORF02056 | 90 aa  | conserved hypothetical protein                               |
| 515.ORF02057 | 117 aa | transcriptional regulator, Cro/CI family                     |
| 515.ORF02058 | 76 aa  | transcriptional regulator, Cro/CI family                     |
| 515.ORF02059 | 125 aa | conserved hypothetical protein                               |
| 515.ORF02060 | 410 aa | transcriptional regulator, Cro/CI family                     |
| 515.ORF02061 | 553 aa | FtsK/SpoIIIE family protein                                  |
| 515.ORF02062 | 153 aa | conserved hypothetical protein                               |
| 515.ORF02063 | 98 aa  | conserved hypothetical protein                               |
| 515.ORF02064 | 863 aa | cell wall surface anchor family protein                      |
| 515.ORF02065 | 753 aa | Unknown  |
| 515.ORF02066 | 87 aa  | conserved hypothetical protein                               |
| 515.ORF02067 | 153 aa | Unknown  |
| 515.ORF02068 | 103 aa | conserved hypothetical protein                               |
| 515.ORF02069 | 143 aa | transcriptional regulator, Cro/CI family                     |
| 515.ORF02070 | 238 aa | FtsK/SpoIIIE family protein                                  |
| 515.ORF02071 | 297 aa | FtsK/SpoIIIE family protein                                  |
| 515.ORF02072 | 279 aa | FtsK/SpoIIIE family protein                                  |
| 515.ORF02073 | 250 aa | FtsK/SpoIIIE family protein                                  |
| 515.ORF02074 | 72 aa  | transcriptional regulator, Cro/CI family                     |
| 515.ORF02075 | 104 aa | transcriptional regulator, Cro/CI family                     |
| 515.ORF02076 | 50 aa  | conserved hypothetical protein                               |
| 515.ORF02077 | 112 aa | transcriptional regulator, Cro/CI family                     |
| 515.ORF02078 | 140 aa | Unknown  |
| 515.ORF02079 | 122 aa | lipoprotein, putative  |
| 515.ORF02080 | 112 aa | Unknown  |
| 515.ORF02081 | 97 aa  | DNA-damage-inducible protein J, putative                     |
| 515.ORF02082 | 252 aa | CAMP factor  |
| 515.ORF02083 | 435 aa | transposase, ISL3 family                                     |
| 515.ORF02084 | 546 aa | carbamoyl-phosphate synthase, large subunit                  |
| 515.ORF02085 | 130 aa | mercuric resistance operon regulatory protein MerR           |
| 515.ORF02086 | 522 aa | Mn <sup>2+</sup> /Fe <sup>2+</sup> transporter, NRAMP family |
| 515.ORF02087 | 240 aa | membrane protein, putative                                   |

|                          |        |  |
|--------------------------|--------|--|
| 515.ORF02088             | 205 aa | ABC transporter, ATP-binding protein   |
| 515.ORF02089             | 36 aa  | conserved hypothetical protein   |
| 18RS21 [1712379-1746497] |        |  |
| 18RS21.ORF01792          | 373 aa | site-specific recombinase, phage integrase family  |
| 18RS21.ORF01793          | 108 aa | Domain of unknown function (DUF771) superfamily  |
| 18RS21.ORF01794          | 210 aa | conserved hypothetical protein   |
| 18RS21.ORF01795          | 263 aa | cell surface protein A   |
| 18RS21.ORF01796          | 65 aa  | conserved hypothetical protein   |
| 18RS21.ORF01797          | 121 aa | CHAP domain superfamily  |
| 18RS21.ORF01798          | 423 aa | CHAP domain family   |
| 18RS21.ORF01799          | 357 aa | membrane protein, putative   |
| 18RS21.ORF01800          | 473 aa | conjugal transfer protein, interruption-C  |
| 18RS21.ORF01801          | 215 aa | conjugal transfer protein, interruption-C  |
| 18RS21.ORF01802          | 136 aa | conserved hypothetical protein   |
| 18RS21.ORF01803          | 88 aa  | conserved hypothetical protein   |
| 18RS21.ORF01804          | 306 aa | conserved hypothetical protein   |
| 18RS21.ORF01805          | 175 aa | conserved hypothetical protein   |
| 18RS21.ORF01806          | 449 aa | Transporter  |
| 18RS21.ORF01807          | 99 aa  | transcriptional regulator, Cro/CI family   |
| 18RS21.ORF01808          | 125 aa | conserved hypothetical protein   |
| 18RS21.ORF01809          | 410 aa | transcriptional regulator, Cro/CI family   |
| 18RS21.ORF01810          | 379 aa | site-specific recombinase, phage integrase family  |
| 18RS21.ORF01811          | 67 aa  | conserved hypothetical protein   |
| 18RS21.ORF01812          | 330 aa | replication protein  |
| 18RS21.ORF01813          | 60 aa  | hypothetical protein   |
| 18RS21.ORF01814          | 97 aa  | transcription regulator  |
| 18RS21.ORF01815          | 406 aa | mercuric reductase   |
| 18RS21.ORF01816          | 553 aa | FtsK/SpoIIIE family protein  |
| 18RS21.ORF01817          | 85 aa  | Macrophage colony stimulating factor I receptor precursor (CSF-1-R)<br>(EC 2.7.1.112) (Fms proto-oncogene) (c-fms) |
| 18RS21.ORF01818          | 70 aa  | conserved hypothetical protein   |
| 18RS21.ORF01819          | 98 aa  | conserved hypothetical protein   |
| 18RS21.ORF01820          | 835 aa | cell wall surface anchor family protein  |
| 18RS21.ORF01821          | 417 aa | transposase, ISL3 family   |
| 18RS21.ORF01822          | 546 aa | mercuric reductase   |
| 18RS21.ORF01823          | 130 aa | Hg(II)-responsive transcriptional regulator  |
| 18RS21.ORF01824          | 90 aa  | Mn <sup>2+</sup> /Fe <sup>2+</sup> transporter, NRAMP family   |
| 18RS21.ORF01825          | 522 aa | Mn <sup>2+</sup> /Fe <sup>2+</sup> transporter, NRAMP family   |
| 18RS21.ORF01826          | 240 aa | membrane protein, putative   |
| 18RS21.ORF01827          | 205 aa | ABC transporter (ATP-binding protein) homolog yydI   |

REGION 1.23 [7 sequence(s)]:

2603V/R [2062261-2069087]

|         |        |  |
|---------|--------|--|
| SAG2080 | 286 aa | glyoxalase family protein  |
| SAG2081 | 243 aa | conserved hypothetical protein                                     |
| SAG2082 | 205 aa | anaerobic ribonucleoside-triphosphate reductase activating protein |
| SAG2083 | 163 aa | acetyltransferase, GNAT family                                     |
| SAG2084 | 310 aa | virulence factor MviM, putative                                    |
| SAG2085 | 47 aa  | conserved hypothetical protein                                     |
| SAG2086 | 723 aa | anaerobic ribonucleoside-triphosphate reductase                    |

## NEM316 [2108986-2115812]

|         |        |   |
|---------|--------|---|
| gbs2035 | 286 aa | glyoxalase family protein   |
| gbs2036 | 230 aa | Hypothetical UPF0246 protein gbs2036                              |
| gbs2037 | 205 aa | anaerobic ribonucleotide triphosphate reductase activator protein |
| gbs2038 | 163 aa | acetyltransferase, GNAT family                                    |
| gbs2039 | 310 aa | oxidoreductase ysjB   |
| gbs2040 | 47 aa  | conserved hypothetical protein                                    |
| gbs2041 | 723 aa | anaerobic ribonucleoside-triphosphate reductase                   |

## A909 [2013670-2020495]

|               |        |  |
|---------------|--------|--|
| A909.ORF00115 | 286 aa | glyoxalase family protein  |
| A909.ORF00114 | 243 aa | conserved hypothetical protein                                     |
| A909.ORF00113 | 205 aa | anaerobic ribonucleoside-triphosphate reductase activating protein |
| A909.ORF00112 | 163 aa | acetyltransferase, GNAT family                                     |
| A909.ORF00111 | 310 aa | oxidoreductase, Gfo/Idh/MocA family                                |
| A909.ORF00110 | 47 aa  | conserved hypothetical protein                                     |
| A909.ORF00109 | 723 aa | anaerobic ribonucleoside-triphosphate reductase                    |

## 515 [1965559-1972385]

|              |        |  |
|--------------|--------|--|
| 515.ORF02143 | 286 aa | glyoxalase family protein  |
| 515.ORF02144 | 230 aa | Protein of unknown function (DUF328) superfamily                   |
| 515.ORF02145 | 205 aa | anaerobic ribonucleoside-triphosphate reductase activating protein |
| 515.ORF02146 | 163 aa | acetyltransferase, GNAT family                                     |
| 515.ORF02147 | 310 aa | virulence factor MviM, putative                                    |
| 515.ORF02148 | 47 aa  | conserved hypothetical protein                                     |
| 515.ORF02149 | 723 aa | anaerobic ribonucleoside-triphosphate reductase                    |

## CJB111 [1924379-1931204]

|                 |        |  |
|-----------------|--------|--|
| CJB111.ORF01986 | 286 aa | glyoxalase family protein  |
| CJB111.ORF01987 | 230 aa | Unknown  |
| CJB111.ORF01988 | 205 aa | anaerobic ribonucleoside-triphosphate reductase activating protein |
| CJB111.ORF01989 | 163 aa | acetyltransferase, GNAT family                                     |
| CJB111.ORF01990 | 310 aa | virulence factor MviM, putative                                    |
| CJB111.ORF01991 | 47 aa  | conserved hypothetical protein                                     |
| CJB111.ORF01992 | 723 aa | anaerobic ribonucleoside-triphosphate reductase                    |

## COH1 [2045624-2052562]

|               |        |  |
|---------------|--------|--|
| COH1.ORF02229 | 286 aa | glyoxalase family protein  |
| COH1.ORF02230 | 230 aa | Unknown  |
| COH1.ORF02231 | 205 aa | anaerobic ribonucleoside-triphosphate reductase activating protein |

|                        |        |  |
|------------------------|--------|--|
| COH1.ORF02232          | 163 aa | acetyltransferase, GNAT family                                     |
| COH1.ORF02233          | 310 aa | virulence factor MviM, putative                                    |
| COH1.ORF02234          | 47 aa  | conserved hypothetical protein                                     |
| COH1.ORF02235          | 648 aa | anaerobic ribonucleoside-triphosphate reductase                    |
| COH1.ORF02236          | 96 aa  | anaerobic ribonucleoside-triphosphate reductase                    |
| H36B [1979166-1985991] |        |  |
| H36B.ORF02128          | 286 aa | glyoxalase family protein  |
| H36B.ORF02129          | 230 aa | Unknown  |
| H36B.ORF02130          | 205 aa | anaerobic ribonucleoside-triphosphate reductase activating protein |
| H36B.ORF02131          | 163 aa | acetyltransferase, GNAT family                                     |
| H36B.ORF02132          | 310 aa | virulence factor MviM, putative                                    |
| H36B.ORF02133          | 47 aa  | conserved hypothetical protein                                     |
| H36B.ORF02134          | 723 aa | anaerobic ribonucleoside-triphosphate reductase                    |

## REGION 1.24 [1 sequence(s)]:

2603V/R [2090618-2100040]

|         |        |   |
|---------|--------|---|
| SAG2109 | 60 aa  | ribosomal protein L32                             |
| SAG2110 | 49 aa  | ribosomal protein L33                             |
| SAG2111 | 173 aa | conserved hypothetical protein                    |
| SAG2112 | 494 aa | site-specific recombinase, phage integrase family |
| SAG2113 | 82 aa  | conserved hypothetical protein                    |
| SAG2114 | 342 aa | conserved hypothetical protein                    |
| SAG2115 | 143 aa | hypothetical protein                              |
| SAG2116 | 151 aa | conserved hypothetical protein                    |
| SAG2117 | 71 aa  | hypothetical protein                              |
| SAG2118 | 306 aa | transcriptional regulator, Cro/CI family          |
| SAG2119 | 373 aa | conserved domain protein                          |
| SAG2120 | 269 aa | hypothetical protein                              |

## REGION 2.1 [1 sequence(s)]:

NEM316 [251357-264482]

|         |        |   |
|---------|--------|---|
| gbs0236 | 260 aa | conserved hypothetical protein          |
| gbs0237 | 377 aa | integrase homolog ydcL                  |
| gbs0238 | 68 aa  | conserved hypothetical protein          |
| gbs0239 | 262 aa | Rep protein                             |
| gbs0240 | 104 aa | conserved hypothetical protein          |
| gbs0241 | 452 aa | FtsK/SpoIIIE family subfamily, putative |
| gbs0242 | 141 aa | conserved hypothetical protein          |
| gbs0243 | 93 aa  | conserved hypothetical protein          |
| gbs0244 | 152 aa | conserved hypothetical protein          |

## REGION 2.2 [3 sequence(s)]:

NEM316 [385777-432905]

|         |         |  |
|---------|---------|--|
| gbs0361 | 62 aa   | conserved hypothetical protein                       |
| gbs0362 | 164 aa  | conserved hypothetical protein                       |
| gbs0363 | 109 aa  | conserved hypothetical protein                       |
| gbs0364 | 30 aa   | conserved hypothetical protein                       |
| gbs0365 | 39 aa   | conserved hypothetical protein                       |
| gbs0366 | 75 aa   | conserved hypothetical protein                       |
| gbs0367 | 497 aa  | repE protein   |
| gbs0369 | 67 aa   | conserved hypothetical protein                       |
| gbs0370 | 137 aa  | conserved hypothetical protein                       |
| gbs0371 | 148 aa  | conserved hypothetical protein                       |
| gbs0372 | 149 aa  | conserved hypothetical protein                       |
| gbs0373 | 136 aa  | conserved hypothetical protein                       |
| gbs0374 | 56 aa   | conserved hypothetical protein                       |
| gbs0375 | 82 aa   | conserved hypothetical protein                       |
| gbs0376 | 318 aa  | conserved hypothetical protein                       |
| gbs0377 | 334 aa  | conserved hypothetical protein                       |
| gbs0378 | 282 aa  | conserved hypothetical protein                       |
| gbs0379 | 91 aa   | Myosin II heavy chain, non muscle                    |
| gbs0380 | 105 aa  | conserved hypothetical protein                       |
| gbs0381 | 115 aa  | conserved hypothetical protein                       |
| gbs0382 | 266 aa  | prophage Lp1 protein 65                              |
| gbs0383 | 152 aa  | conserved hypothetical protein                       |
| gbs0384 | 788 aa  | conserved hypothetical protein                       |
| gbs0385 | 85 aa   | conserved hypothetical protein                       |
| gbs0386 | 1576 aa | conserved hypothetical protein                       |
| gbs0387 | 582 aa  | DNA topoisomerase III topB                           |
| gbs0388 | 639 aa  | negative regulator of genetic competence ClpC/MecB   |
| gbs0389 | 115 aa  | conserved hypothetical protein                       |
| gbs0390 | 360 aa  | conserved hypothetical protein                       |
| gbs0391 | 753 aa  | Gram positive anchor domain protein                  |
| gbs0392 | 240 aa  | surface protein PrgC                                 |
| gbs0393 | 933 aa  | cell wall surface anchor family protein              |
| gbs0394 | 143 aa  | EF0030   |
| gbs0395 | 85 aa   | conserved hypothetical protein                       |
| gbs0396 | 1049 aa | pXO2-16  |
| gbs0397 | 160 aa  | lipoprotein, putative                                |
| gbs0398 | 700 aa  | membrane protein, putative                           |
| gbs0399 | 93 aa   | DNA ligase (Polydeoxyribonucleotide synthase [NAD+]) |
| gbs0400 | 101 aa  | conserved hypothetical protein                       |
| gbs0401 | 215 aa  | conserved hypothetical protein                       |
| gbs0402 | 649 aa  | conserved hypothetical protein                       |
| gbs0403 | 85 aa   | conserved hypothetical protein                       |
| gbs0404 | 445 aa  | EF0011   |

|                        |         |  |
|------------------------|---------|--|
| gbs0405                | 194 aa  | conserved hypothetical protein                     |
| gbs0406                | 272 aa  | ParA   |
| gbs0407                | 92 aa   | conserved hypothetical protein                     |
| gbs0408                | 127 aa  | replication protein                                |
| gbs0409                | 103 aa  | conserved hypothetical protein                     |
| gbs0410                | 442 aa  | conserved hypothetical protein                     |
| NEM316 [711626-759043] |         |  |
| gbs0692                | 62 aa   | conserved hypothetical protein                     |
| gbs0693                | 164 aa  | conserved hypothetical protein                     |
| gbs0694                | 109 aa  | conserved hypothetical protein                     |
| gbs0695                | 30 aa   | conserved hypothetical protein                     |
| gbs0696                | 39 aa   | conserved hypothetical protein                     |
| gbs0697                | 75 aa   | conserved hypothetical protein                     |
| gbs0698                | 497 aa  | repE protein                                       |
| gbs0699                | 67 aa   | conserved hypothetical protein                     |
| gbs0700                | 137 aa  | conserved hypothetical protein                     |
| gbs0701                | 148 aa  | conserved hypothetical protein                     |
| gbs0702                | 149 aa  | conserved hypothetical protein                     |
| gbs0703                | 136 aa  | conserved hypothetical protein                     |
| gbs0704                | 56 aa   | conserved hypothetical protein                     |
| gbs0705                | 82 aa   | conserved hypothetical protein                     |
| gbs0706                | 318 aa  | conserved hypothetical protein                     |
| gbs0707                | 334 aa  | conserved hypothetical protein                     |
| gbs0708                | 282 aa  | conserved hypothetical protein                     |
| gbs0709                | 91 aa   | Myosin II heavy chain, non muscle                  |
| gbs0710                | 105 aa  | conserved hypothetical protein                     |
| gbs0711                | 115 aa  | conserved hypothetical protein                     |
| gbs0712                | 266 aa  | prophage Lp1 protein 65                            |
| gbs0713                | 152 aa  | conserved hypothetical protein                     |
| gbs0714                | 788 aa  | conserved hypothetical protein                     |
| gbs0715                | 85 aa   | conserved hypothetical protein                     |
| gbs0716                | 1576 aa | conserved hypothetical protein                     |
| gbs0717                | 582 aa  | DNA topoisomerase III topB                         |
| gbs0718                | 639 aa  | negative regulator of genetic competence ClpC/MecB |
| gbs0719                | 115 aa  | conserved hypothetical protein                     |
| gbs0720                | 360 aa  | conserved hypothetical protein                     |
| gbs0721                | 753 aa  | Gram positive anchor domain protein                |
| gbs0722                | 240 aa  | surface protein PrgC                               |
| gbs0723                | 933 aa  | cell wall surface anchor family protein            |
| gbs0724                | 143 aa  | EF0030   |
| gbs0725                | 85 aa   | conserved hypothetical protein                     |
| gbs0726                | 1049 aa | pXO2-16  |
| gbs0727                | 160 aa  | lipoprotein, putative                              |

|         |        |  |
|---------|--------|--|
| gbs0728 | 700 aa | membrane protein, putative                           |
| gbs0729 | 93 aa  | DNA ligase (Polydeoxyribonucleotide synthase [NAD+]) |
| gbs0730 | 101 aa | conserved hypothetical protein                       |
| gbs0731 | 215 aa | conserved hypothetical protein                       |
| gbs0732 | 649 aa | conserved hypothetical protein                       |
| gbs0733 | 85 aa  | conserved hypothetical protein                       |
| gbs0734 | 445 aa | EF0011   |
| gbs0735 | 194 aa | conserved hypothetical protein                       |
| gbs0736 | 272 aa | ParA   |
| gbs0737 | 92 aa  | conserved hypothetical protein                       |
| gbs0738 | 127 aa | replication protein                                  |
| gbs0739 | 103 aa | conserved hypothetical protein                       |
| gbs0740 | 442 aa | conserved hypothetical protein                       |

## NEM316 [1012958-1060101]

|         |         |  |
|---------|---------|--|
| gbs0969 | 442 aa  | conserved hypothetical protein                       |
| gbs0970 | 103 aa  | conserved hypothetical protein                       |
| gbs0971 | 127 aa  | replication protein                                  |
| gbs0972 | 92 aa   | conserved hypothetical protein                       |
| gbs0973 | 272 aa  | ParA   |
| gbs0974 | 194 aa  | conserved hypothetical protein                       |
| gbs0975 | 445 aa  | EF0011   |
| gbs0976 | 85 aa   | conserved hypothetical protein                       |
| gbs0977 | 649 aa  | conserved hypothetical protein                       |
| gbs0978 | 215 aa  | conserved hypothetical protein                       |
| gbs0979 | 101 aa  | conserved hypothetical protein                       |
| gbs0980 | 93 aa   | DNA ligase (Polydeoxyribonucleotide synthase [NAD+]) |
| gbs0981 | 700 aa  | membrane protein, putative                           |
| gbs0982 | 160 aa  | lipoprotein, putative                                |
| gbs0983 | 1049 aa | pXO2-16  |
| gbs0984 | 85 aa   | conserved hypothetical protein                       |
| gbs0985 | 143 aa  | EF0030   |
| gbs0986 | 933 aa  | cell wall surface anchor family protein              |
| gbs0987 | 240 aa  | surface protein PrgC                                 |
| gbs0988 | 753 aa  | Gram positive anchor domain protein                  |
| gbs0989 | 360 aa  | conserved hypothetical protein                       |
| gbs0990 | 115 aa  | conserved hypothetical protein                       |
| gbs0991 | 639 aa  | negative regulator of genetic competence ClpC/MecB   |
| gbs0992 | 582 aa  | DNA topoisomerase III topB                           |
| gbs0993 | 1576 aa | conserved hypothetical protein                       |
| gbs0994 | 85 aa   | conserved hypothetical protein                       |
| gbs0995 | 788 aa  | conserved hypothetical protein                       |
| gbs0996 | 152 aa  | conserved hypothetical protein                       |
| gbs0997 | 266 aa  | prophage Lp1 protein 65                              |

|         |        |                                   |
|---------|--------|-----------------------------------|
| gbs0998 | 115 aa | conserved hypothetical protein    |
| gbs0999 | 105 aa | conserved hypothetical protein    |
| gbs1000 | 91 aa  | Myosin II heavy chain, non muscle |
| gbs1001 | 282 aa | conserved hypothetical protein    |
| gbs1002 | 334 aa | conserved hypothetical protein    |
| gbs1003 | 318 aa | conserved hypothetical protein    |
| gbs1004 | 82 aa  | conserved hypothetical protein    |
| gbs1005 | 56 aa  | conserved hypothetical protein    |
| gbs1006 | 136 aa | conserved hypothetical protein    |
| gbs1007 | 149 aa | conserved hypothetical protein    |
| gbs1008 | 148 aa | conserved hypothetical protein    |
| gbs1009 | 137 aa | conserved hypothetical protein    |
| gbs0368 | 67 aa  | conserved hypothetical protein    |
| gbs1010 | 497 aa | repE protein                      |
| gbs1011 | 75 aa  | conserved hypothetical protein    |
| gbs1012 | 39 aa  | conserved hypothetical protein    |
| gbs1013 | 30 aa  | conserved hypothetical protein    |
| gbs1014 | 109 aa | conserved hypothetical protein    |
| gbs1015 | 164 aa | conserved hypothetical protein    |
| gbs1016 | 62 aa  | conserved hypothetical protein    |

## REGION 2.3 [1 sequence(s)]:

NEM316 [500392-510252]

|         |        |   |
|---------|--------|---|
| gbs0471 | 97 aa  | EF0032  |
| gbs0472 | 112 aa | conserved hypothetical protein                    |
| gbs0473 | 118 aa | lipoprotein, putative                             |
| gbs0474 | 140 aa | Domain of unknown function (DUF955) superfamily   |
| gbs0475 | 150 aa | immunity repressor protein                        |
| gbs0476 | 156 aa | conserved hypothetical protein                    |
| gbs0477 | 126 aa | conserved hypothetical protein                    |
| gbs0478 | 185 aa | conserved hypothetical protein                    |
| gbs0479 | 253 aa | cell surface protein A                            |
| gbs0480 | 180 aa | conserved hypothetical protein                    |
| gbs0481 | 108 aa | Domain of unknown function (DUF771) superfamily   |
| gbs0482 | 144 aa | site-specific recombinase, phage integrase family |
| gbs0484 | 101 aa | conserved protein                                 |
| gbs0485 | 112 aa | conserved protein                                 |
| gbs0486 | 254 aa | methyltransferase                                 |

## REGION 2.4 [1 sequence(s)]:

NEM316 [1163291-1197444]

|         |        |                                |
|---------|--------|--------------------------------|
| gbs1118 | 397 aa | conserved hypothetical protein |
| gbs1119 | 99 aa  | conserved hypothetical protein |

|         |         |   |
|---------|---------|---|
| gbs1120 | 415 aa  | transferase                                 |
| gbs1121 | 543 aa  | relaxase                                    |
| gbs1122 | 124 aa  | conserved hypothetical protein              |
| gbs1123 | 187 aa  | conserved hypothetical protein              |
| gbs1125 | 99 aa   | conserved hypothetical protein              |
| gbs1126 | 1078 aa | conserved hypothetical protein              |
| gbs1127 | 121 aa  | conserved hypothetical protein              |
| gbs1128 | 681 aa  | TraG family protein                         |
| gbs1129 | 163 aa  | conserved hypothetical protein              |
| gbs1130 | 224 aa  | conserved hypothetical protein              |
| gbs1131 | 90 aa   | conserved hypothetical protein              |
| gbs1132 | 201 aa  | conserved hypothetical protein              |
| gbs1133 | 896 aa  | Tn5252, Orf28                               |
| gbs1134 | 186 aa  | conserved hypothetical protein              |
| gbs1135 | 783 aa  | Tn5252, Orf26                               |
| gbs1136 | 119 aa  | conserved hypothetical protein              |
| gbs1137 | 152 aa  | conserved hypothetical protein              |
| gbs1138 | 75 aa   | conserved hypothetical protein              |
| gbs1139 | 165 aa  | conserved hypothetical protein              |
| gbs1140 | 250 aa  | membrane protein, putative                  |
| gbs1141 | 161 aa  | conserved hypothetical protein              |
| gbs1142 | 75 aa   | conserved hypothetical protein              |
| gbs1143 | 932 aa  | cell wall surface anchor family protein     |
| gbs1144 | 236 aa  | surface protein PrgC                        |
| gbs1145 | 743 aa  | LPXTG-motif cell wall anchor domain protein |
| gbs1146 | 99 aa   | conserved hypothetical protein              |
| gbs1147 | 65 aa   | transcriptional regulator, Cro/CI family    |
| gbs1148 | 61 aa   | conserved hypothetical protein              |
| gbs1149 | 341 aa  | IFN-response binding factor 1 (IREBF-1)     |
| gbs1150 | 54 aa   | conserved hypothetical protein              |
| gbs1151 | 147 aa  | conserved hypothetical protein              |
| gbs1152 | 85 aa   | EF0029                                      |
| gbs1153 | 123 aa  | conserved hypothetical protein              |

## REGION 2.5 [1 sequence(s)]:

NEM316 [1255493-1262954]

|         |        |                                |
|---------|--------|--------------------------------|
| gbs1217 | 51 aa  | conserved hypothetical protein |
| gbs1218 | 143 aa | ATP synthase A chain subunit 6 |
| gbs1219 | 48 aa  | conserved hypothetical protein |
| gbs1220 | 144 aa | conserved hypothetical protein |
| gbs1221 | 262 aa | UmuD MucA homolog              |
| gbs1222 | 96 aa  | conserved hypothetical protein |
| gbs1223 | 321 aa | conserved hypothetical protein |

|         |        |                          |
|---------|--------|--------------------------|
| gbs1224 | 358 aa | integrase                |
| gbs1225 | 466 aa | 30S ribosomal protein S1 |

## REGION 2.6 [1 sequence(s)]:

NEM316 [1361643-1424899]

|         |         |   |         |
|---------|---------|---|---------|
| gbs1314 | 393 aa  | site-specific recombinase, phage integrase family   |         |
| gbs1315 | 77 aa   | conserved hypothetical protein  |         |
| gbs1316 | 223 aa  | Plasmid replication protein   |         |
| gbs1320 | 167 aa  | conserved hypothetical protein  |         |
| gbs1321 | 104 aa  | conserved hypothetical protein  |         |
| gbs1322 | 401 aa  | conserved hypothetical protein  |         |
| gbs1323 | 163 aa  | conserved hypothetical protein  |         |
| gbs1324 | 1203 aa | type IIS restriction enzyme R and M protein   |         |
| gbs1325 | 399 aa  | site-specific recombinase, phage integrase family   |         |
| gbs1326 | 82 aa   | conserved hypothetical protein  |         |
| gbs1327 | 73 aa   | transcriptional regulator, Cro/CI family  |         |
| gbs1328 | 297 aa  | LacX protein, plasmid   |         |
| gbs1329 | 468 aa  | 6-phospho-beta-galactosidase  |         |
| gbs1330 | 568 aa  | PTS system, lactose-specific IIBC component (EIIBC-LAC) (Lactose-permease IIBC component) (Phosphotransferase enzyme II, BC component)(EC 2.7.1.69) (EII-LAC) |         |
| gbs1331 | 105 aa  | PTS system, lactose-specific IIA component  |         |
| gbs1332 | 277 aa  | antiterminator protein  |         |
| gbs1333 | 326 aa  | tagatose 1,6-diphosphate aldolase   |         |
| gbs1334 | 303 aa  | tagatose-6-phosphate kinase   |         |
| gbs1335 | 171 aa  | galactose-6-phosphate isomerase, LacB subunit   |         |
| gbs1336 | 141 aa  | galactose-6-phosphate isomerase, LacA subunit   |         |
| gbs1337 | 251 aa  | Lactose phosphotransferase system repressor   |         |
| gbs1338 | 625 aa  | Tn5252, relaxase  |         |
| gbs1339 | 121 aa  | Tn5252, Orf 9 protein   |         |
| gbs1340 | 137 aa  | Tn5252, Orf 9 protein   |         |
| gbs1341 | 672 aa  | conserved hypothetical protein  |         |
| gbs1342 | 746 aa  | conserved hypothetical protein  |         |
| gbs1343 | 256 aa  | Zeta toxin  |         |
| gbs1344 | 158 aa  | Bacillus phage phi-105 immunity repressor protein homolog   | lin1762 |
| gbs1345 | 96 aa   | conserved hypothetical protein  |         |
| gbs1346 | 129 aa  | conserved hypothetical protein  |         |
| gbs1347 | 76 aa   | conserved hypothetical protein  |         |
| gbs1348 | 361 aa  | pXO2-78   |         |
| gbs1349 | 212 aa  | conserved hypothetical protein  |         |
| gbs1350 | 99 aa   | conserved hypothetical protein  |         |
| gbs1351 | 99 aa   | conserved hypothetical protein  |         |
| gbs1352 | 2014 aa | SNF2 family protein   |         |
| gbs1353 | 197 aa  | SNF2 family protein   |         |

|         |         |   |
|---------|---------|---|
| gbs1354 | 183 aa  | conserved hypothetical protein                        |
| gbs1355 | 63 aa   | Calcium-binding acidic-repeat protein precursor (ARP) |
| gbs1356 | 1595 aa | agglutinin receptor                                   |
| gbs1357 | 196 aa  | abortive infection protein AbiGI                      |
| gbs1358 | 282 aa  | abortive infection protein AbiGII                     |
| gbs1359 | 931 aa  | Tn5252, Orf28   |
| gbs1360 | 774 aa  | Tn5252, Orf26   |
| gbs1361 | 117 aa  | conserved hypothetical protein                        |
| gbs1362 | 284 aa  | Tn5252, Orf23   |
| gbs1363 | 80 aa   | conserved hypothetical protein                        |
| gbs1364 | 605 aa  | TraG family protein                                   |
| gbs1365 | 162 aa  | conserved hypothetical protein                        |
| gbs1366 | 193 aa  | CAAX amino terminal protease family                   |
| gbs1367 | 77 aa   | conserved hypothetical protein                        |
| gbs1368 | 128 aa  | conserved hypothetical protein                        |
| gbs1369 | 148 aa  | conserved hypothetical protein                        |
| gbs1370 | 450 aa  | C-5 cytosine-specific DNA methylase                   |
| gbs1372 | 271 aa  | IFN-response binding factor 1 (IREBF-1)               |
| gbs1373 | 57 aa   | conserved hypothetical protein                        |
| gbs1374 | 121 aa  | ribosomal protein L7/L12                              |
| gbs1375 | 166 aa  | 50S ribosomal protein                                 |

## REGION 2.7 [2 sequence(s)]:

NEM316 [1798480-1808451]

|         |        |  |
|---------|--------|--|
| gbs1740 | 568 aa | ABC transporter, ATP-binding/permease protein  |
| gbs1741 | 585 aa | ABC transporter, ATP-binding/permease protein  |
| gbs1742 | 459 aa | ABC transporter, ATP-binding protein           |
| gbs1743 | 237 aa | conserved hypothetical protein                 |
| gbs1744 | 198 aa | conserved hypothetical protein                 |
| gbs1745 | 143 aa | conserved hypothetical protein                 |
| gbs1746 | 51 aa  | conserved hypothetical protein                 |
| gbs1748 | 271 aa | 2,3-dihydroxybenzoate-AMP ligase               |
| gbs1749 | 124 aa | iron-dependent transcription repressor CAC1469 |

515 [1629883-1639853]

|              |        |                                      |
|--------------|--------|--------------------------------------|
| 515.ORF01801 | 43 aa  | conserved hypothetical protein       |
| 515.ORF01802 | 568 aa | Unknown                              |
| 515.ORF01803 | 300 aa | Unknown                              |
| 515.ORF01804 | 293 aa | Unknown                              |
| 515.ORF01805 | 459 aa | ABC transporter, ATP-binding protein |
| 515.ORF01806 | 237 aa | Unknown                              |
| 515.ORF01807 | 198 aa | Unknown                              |
| 515.ORF01808 | 51 aa  | conserved hypothetical protein       |
| 515.ORF01809 | 147 aa | 2,3-dihydroxybenzoate-AMP ligase     |

515.ORF01810 271 aa yersiniabactin synthetase, salycilate ligase component  
 515.ORF01811 124 aa iron-dependent repressor, putative

## REGION 2.8 [3 sequence(s)]:

## NEM316 [2036175-2052498]

|                        |        |  |         |
|------------------------|--------|--|---------|
| gbs1965                | 101 aa | conserved hypothetical protein                           |         |
| gbs1966                | 133 aa | conserved hypothetical protein                           |         |
| gbs1967                | 402 aa | conserved hypothetical protein                           |         |
| gbs1968                | 110 aa | conserved hypothetical protein                           |         |
| gbs1970                | 48 aa  | conserved hypothetical protein                           |         |
| gbs1972                | 164 aa | transcriptional regulator, Cro/CI family                 |         |
| gbs1973                | 450 aa | conserved hypothetical protein                           |         |
| gbs1974                | 363 aa | conserved hypothetical protein                           |         |
| gbs1975                | 314 aa | conserved hypothetical protein                           |         |
| gbs1976                | 109 aa | conserved hypothetical protein                           |         |
| gbs1978                | 42 aa  | conserved hypothetical protein                           |         |
| gbs1979                | 96 aa  | a small heat shock protein of Clostridium acetobutylicum | homolog |
| lmo0056                |        |  |         |
| gbs1980                | 124 aa | Protein of unknown function (DUF1310) subfamily          |         |
| gbs1981                | 242 aa | conserved hypothetical protein                           |         |
| gbs1982                | 84 aa  | EF0032   |         |
| gbs1983                | 101 aa | conserved hypothetical protein                           |         |
| gbs1984                | 132 aa | conserved hypothetical protein                           |         |
| gbs1985                | 240 aa | membrane protein, putative                               |         |
| gbs1986                | 205 aa | ABC transporter (ATP-binding protein) homolog yydI       |         |
| A909 [1938736-1957697] |        |  |         |
| A909.ORF00196          | 369 aa | hypothetical protein                                     |         |
| A909.ORF00195          | 69 aa  | hypothetical protein                                     |         |
| A909.ORF00194          | 46 aa  | hypothetical protein                                     |         |
| A909.ORF00193          | 376 aa | site-specific recombinase, phage integrase family        |         |
| A909.ORF00192          | 61 aa  | conserved hypothetical protein                           |         |
| A909.ORF00191          | 342 aa | replication initiation factor family protein             |         |
| A909.ORF00190          | 139 aa | conserved domain protein                                 |         |
| A909.ORF00189          | 57 aa  | hypothetical protein                                     |         |
| A909.ORF00188          | 164 aa | transcriptional regulator, putative                      |         |
| A909.ORF00187          | 316 aa | conserved hypothetical protein                           |         |
| A909.ORF00186          | 196 aa | DJ-1/PfpI family protein                                 |         |
| A909.ORF00185          | 0 aa   | conserved domain protein, authentic frameshift           |         |
| A909.ORF00183          | 128 aa | conserved hypothetical protein                           |         |
| A909.ORF00182          | 190 aa | conserved domain protein                                 |         |
| A909.ORF00181          | 110 aa | hypothetical protein                                     |         |
| A909.ORF00180          | 42 aa  | hypothetical protein                                     |         |
| A909.ORF00179          | 96 aa  | conserved hypothetical protein                           |         |

|               |        |                                      |
|---------------|--------|--------------------------------------|
| A909.ORF00178 | 124 aa | conserved hypothetical protein       |
| A909.ORF00177 | 41 aa  | hypothetical protein                 |
| A909.ORF00176 | 127 aa | conserved hypothetical protein       |
| A909.ORF00175 | 34 aa  | hypothetical protein                 |
| A909.ORF00174 | 114 aa | hypothetical protein                 |
| A909.ORF00173 | 102 aa | hypothetical protein                 |
| A909.ORF02073 | 67 aa  | hypothetical protein                 |
| A909.ORF00172 | 133 aa | hypothetical protein                 |
| A909.ORF00171 | 240 aa | membrane protein, putative           |
| A909.ORF00170 | 205 aa | ABC transporter, ATP-binding protein |

## CJB111 [1861144-1869248]

|                  |        |  |
|------------------|--------|--|
| CJB111.ORF01919  | 101 aa | Unknown                                  |
| CJB111.ORF01920  | 133 aa | Unknown                                  |
| CJB111.ORF01921  | 330 aa | Unknown                                  |
| CJB111.ORF01922  | 71 aa  | Unknown                                  |
| CJB111.ORF01923  | 110 aa | Unknown                                  |
| CJB111.ORFA01870 | 48 aa  | hypothetical protein                     |
| CJB111.ORF01924  | 42 aa  | Unknown                                  |
| CJB111.ORF01925  | 34 aa  | hypothetical protein                     |
| CJB111.ORF01926  | 69 aa  | Unknown                                  |
| CJB111.ORF01927  | 127 aa | protein of unknown function              |
| CJB111.ORF01928  | 34 aa  | hypothetical protein                     |
| CJB111.ORF01929  | 242 aa | Unknown                                  |
| CJB111.ORF01930  | 39 aa  | hypothetical protein                     |
| CJB111.ORF01931  | 84 aa  | DNA-damage-inducible protein J, putative |
| CJB111.ORF01932  | 127 aa | degenerate transposase (orf1) [imported] |

## REGION 2.9 [6 sequence(s)]:

## NEM316 [2137909-2151330]

|         |        |   |
|---------|--------|---|
| gbs2064 | 204 aa | cadmium resistance protein NMB1955      |
| gbs2065 | 112 aa | CadX                                    |
| gbs2066 | 471 aa | conserved hypothetical protein          |
| gbs2067 | 149 aa | conserved hypothetical protein          |
| gbs2068 | 106 aa | conserved hypothetical protein          |
| gbs2069 | 366 aa | FtsK/SpoIIIE family subfamily, putative |
| gbs2070 | 201 aa | conserved hypothetical protein          |
| gbs2071 | 201 aa | Plasmid replication protein             |
| gbs2072 | 82 aa  | prophage ps1 protein 10                 |
| gbs2073 | 421 aa | prophage ps1 protein 23, integrase      |
| gbs2074 | 331 aa | conserved hypothetical protein          |
| gbs2076 | 353 aa | conserved hypothetical protein          |
| gbs2077 | 127 aa | conserved hypothetical protein          |

## A909 [2042593-2052556]

|                          |        |  |
|--------------------------|--------|--|
| A909.ORF00084            | 402 aa | IS10R, transposase   |
| A909.ORF00083            | 204 aa | cadmium resistance protein CadD                                    |
| A909.ORF00082            | 112 aa | cadmium resistance accessory protein CadX                          |
| A909.ORF00081            | 471 aa | conserved domain protein   |
| A909.ORF00080            | 149 aa | conserved domain protein   |
| A909.ORF00079            | 68 aa  | hypothetical protein   |
| A909.ORF00078            | 366 aa | FtsK/SpoIIIE family protein  |
| A909.ORF00077            | 201 aa | conserved hypothetical protein                                     |
| A909.ORF00076            | 201 aa | plasmid replication protein  |
| A909.ORF00075            | 421 aa | site-specific recombinase, phage integrase family                  |
| 515 [1976774-1986364]    |        |  |
| 515.ORF02157             | 258 aa | serine protease, subtilase family                                  |
| 515.ORF02158             | 151 aa | penicillin-binding protein 2A                                      |
| 515.ORF02159             | 457 aa | Unknown  |
| 515.ORF02160             | 149 aa | conserved hypothetical protein                                     |
| 515.ORF02161             | 106 aa | Unknown  |
| 515.ORF02162             | 366 aa | Unknown  |
| 515.ORF02163             | 201 aa | conserved hypothetical protein                                     |
| 515.ORF02164             | 201 aa | Plasmid replication protein  |
| 515.ORF02165             | 82 aa  | conserved hypothetical protein                                     |
| 515.ORF02166             | 421 aa | integrase/recombinase, phage integrase family [imported], putative |
| CJB111 [1953207-1962543] |        |  |
| CJB111.ORF02019          | 204 aa | cadmium resistance protein   |
| CJB111.ORF02020          | 112 aa | cadmium efflux system accessory protein, putative                  |
| CJB111.ORF02021          | 471 aa | Unknown  |
| CJB111.ORF02022          | 149 aa | conserved hypothetical protein                                     |
| CJB111.ORF02023          | 106 aa | Unknown  |
| CJB111.ORF02024          | 366 aa | Unknown  |
| CJB111.ORF02025          | 228 aa | conserved hypothetical protein                                     |
| CJB111.ORF02026          | 201 aa | Plasmid replication protein  |
| CJB111.ORF02027          | 82 aa  | conserved hypothetical protein                                     |
| CJB111.ORF02028          | 83 aa  | prophage ps1 protein 23, integrase [imported]                      |
| CJB111.ORF02029          | 283 aa | integrase/recombinase, phage integrase family [imported], putative |
| CJB111.ORF02030          | 129 aa | Transposase (IS4 family)   |
| CJB111.ORF02031          | 161 aa | degenerate transposase (orf1) [imported]                           |
| COH1 [2074499-2082980]   |        |  |
| COH1.ORF02262            | 204 aa | cadmium resistance protein   |
| COH1.ORF02263            | 112 aa | cadmium efflux system accessory protein, putative                  |
| COH1.ORF02264            | 471 aa | Unknown  |
| COH1.ORF02265            | 149 aa | conserved hypothetical protein                                     |
| COH1.ORF02266            | 106 aa | Unknown  |
| COH1.ORF02267            | 366 aa | Unknown  |
| COH1.ORF02268            | 201 aa | conserved hypothetical protein                                     |

|                        |        |  |
|------------------------|--------|--|
| COH1.ORF02269          | 201 aa | Plasmid replication protein  |
| COH1.ORF02270          | 82 aa  | conserved hypothetical protein                                     |
| COH1.ORF02271          | 421 aa | integrase/recombinase, phage integrase family [imported], putative |
| H36B [2031170-2039653] |        |  |
| H36B.ORF02190          | 204 aa | cadmium resistance protein   |
| H36B.ORF02191          | 112 aa | cadmium efflux system accessory protein, putative                  |
| H36B.ORF02192          | 471 aa | Unknown  |
| H36B.ORFA02119         | 43 aa  | hypothetical protein   |
| H36B.ORF02193          | 149 aa | conserved hypothetical protein                                     |
| H36B.ORF02194          | 106 aa | Unknown  |
| H36B.ORF02195          | 366 aa | Unknown  |
| H36B.ORF02196          | 201 aa | conserved hypothetical protein                                     |
| H36B.ORF02197          | 201 aa | Plasmid replication protein  |
| H36B.ORF02198          | 82 aa  | conserved hypothetical protein                                     |
| H36B.ORF02199          | 421 aa | integrase/recombinase, phage integrase family [imported], putative |

## REGION 3.1 [1 sequence(s)]:

A909 [143616-152999]

|               |         |  |
|---------------|---------|--|
| A909.ORF01931 | 340 aa  | ISSag3, transposase  |
| A909.ORF01930 | 72 aa   | hypothetical protein   |
| A909.ORF01929 | 1164 aa | IgA-binding beta antigen   |
| A909.ORF01927 | 175 aa  | isoprenylcysteine carboxyl methyltransferase (ICMT) family protein |
| A909.ORF01926 | 446 aa  | sensor histidine kinase  |
| A909.ORF01925 | 217 aa  | DNA-binding response regulator                                     |
| A909.ORF01924 | 129 aa  | IS1381, transposase orfB   |
| A909.ORF01923 | 127 aa  | IS1381, transposase orfA   |

## REGION 3.2 [1 sequence(s)]:

A909 [462318-485250]

|               |        |   |
|---------------|--------|---|
| A909.ORF01629 | 0 aa   | site-specific recombinase, phage integrase family, degenerate |
| A909.ORF01628 | 101 aa | conserved hypothetical protein                                |
| A909.ORF01627 | 112 aa | conserved hypothetical protein                                |
| A909.ORF01626 | 254 aa | conserved hypothetical protein                                |
| A909.ORF01625 | 686 aa | PTS system IIA domain protein                                 |
| A909.ORF01624 | 149 aa | PTS system, galactitol-specific IIA component, putative       |
| A909.ORF01623 | 92 aa  | PTS system, galactitol-specific IIB component, putative       |
| A909.ORF01622 | 445 aa | PTS system, galactitol-specific IIC component, putative       |
| A909.ORF01621 | 276 aa | rhamnulose-1-phosphate aldolase                               |
| A909.ORF01620 | 154 aa | PTS system, galactitol-specific IIA component, putative       |
| A909.ORF01619 | 443 aa | PTS system, galactitol-specific IIC component                 |
| A909.ORF01618 | 94 aa  | PTS system, galactitol-specific IIB component, putative       |
| A909.ORF01617 | 276 aa | transcriptional regulator, AraC family                        |
| A909.ORF01616 | 425 aa | sugar ABC transporter, sugar-binding protein, putative        |

|               |        |   |
|---------------|--------|---|
| A909.ORF01615 | 296 aa | sugar ABC transporter, permease protein, putative |
| A909.ORF01614 | 277 aa | ABC transporter, permease protein                 |
| A909.ORF01613 | 733 aa | alpha-galactosidase, putative                     |
| A909.ORF01612 | 0 aa   | galactokinase, authentic point mutation           |
| A909.ORF01610 | 493 aa | galactose-1-phosphate uridylyltransferase         |
| A909.ORF01609 | 331 aa | UDP-glucose 4-epimerase                           |
| A909.ORF01608 | 289 aa | aldose 1-epimerase, interruption-N                |
| A909.ORF01607 | 127 aa | IS1381, transposase orfA                          |
| A909.ORF01606 | 129 aa | IS1381, transposase orfB                          |
| A909.ORF01605 | 64 aa  | aldose 1-epimerase, interruption-C                |

## REGION 3.3 [1 sequence(s)]:

A909 [548498-586484]

|                      |        |   |
|----------------------|--------|---|
| A909.ORF01541        | 91 aa  | DNA-binding protein HU  |
| A909.ORF01540 family | 380 aa | prophage LambdaSa03, site-specific recombinase, phage integrase |
| A909.ORF01539        | 69 aa  | conserved hypothetical protein                                  |
| A909.ORF01538        | 126 aa | conserved hypothetical protein                                  |
| A909.ORF01537        | 119 aa | prophage LambdaSa03, transcriptional regulator, Cro/CI family   |
| A909.ORF01536        | 72 aa  | prophage LambdaSa03, transcriptional regulator, Cro/CI family   |
| A909.ORF01535        | 110 aa | conserved hypothetical protein                                  |
| A909.ORF01534        | 111 aa | hypothetical protein  |
| A909.ORF01533        | 49 aa  | conserved hypothetical protein                                  |
| A909.ORF01532        | 96 aa  | conserved hypothetical protein                                  |
| A909.ORF01531        | 227 aa | conserved hypothetical protein TIGR01618                        |
| A909.ORF01530        | 454 aa | prophage LambdaSa03, helicase, putative                         |
| A909.ORF01529        | 160 aa | conserved hypothetical protein                                  |
| A909.ORF01527        | 517 aa | conserved hypothetical protein/bacteriophage resistance protein |
| A909.ORF01526        | 289 aa | conserved hypothetical protein                                  |
| A909.ORF01525        | 93 aa  | conserved hypothetical protein                                  |
| A909.ORF02151        | 98 aa  | conserved hypothetical protein                                  |
| A909.ORF01524        | 54 aa  | hypothetical protein  |
| A909.ORF01523        | 47 aa  | hypothetical protein  |
| A909.ORF01522        | 170 aa | conserved hypothetical protein                                  |
| A909.ORF01521        | 99 aa  | hypothetical protein  |
| A909.ORF01520        | 64 aa  | hypothetical protein  |
| A909.ORF01519        | 88 aa  | conserved hypothetical protein                                  |
| A909.ORF02145        | 142 aa | conserved hypothetical protein                                  |
| A909.ORF01518        | 124 aa | prophage LambdaSa03, HNH endonuclease family protein            |
| A909.ORF01517        | 118 aa | conserved hypothetical protein                                  |
| A909.ORF01516        | 422 aa | conserved hypothetical protein                                  |
| A909.ORF01515        | 406 aa | conserved hypothetical protein                                  |
| A909.ORF01514        | 62 aa  | hypothetical protein  |

|               |         |   |
|---------------|---------|---|
| A909.ORF01513 | 471 aa  | prophage LambdaSa03, terminase, large subunit, putative |
| A909.ORF01512 | 154 aa  | conserved hypothetical protein                          |
| A909.ORF01511 | 300 aa  | prophage LambdaSa03, structural protein, putative       |
| A909.ORF01510 | 71 aa   | conserved hypothetical protein                          |
| A909.ORF01509 | 143 aa  | conserved hypothetical protein                          |
| A909.ORF01508 | 112 aa  | conserved hypothetical protein                          |
| A909.ORF01507 | 111 aa  | conserved hypothetical protein                          |
| A909.ORF01506 | 185 aa  | prophage LambdaSa03, structural protein, putative       |
| A909.ORF01505 | 81 aa   | conserved hypothetical protein                          |
| A909.ORF01503 | 123 aa  | conserved hypothetical protein                          |
| A909.ORF01502 | 670 aa  | prophage LambdaSa03, pblA protein, internal deletion    |
| A909.ORF01501 | 510 aa  | prophage LambdaSa03, tail component, putative           |
| A909.ORF01500 | 1373 aa | prophage LambdaSa03, minor structural protein, putative |
| A909.ORF01499 | 667 aa  | prophage LambdaSa03, minor structural protein, putative |
| A909.ORF01498 | 130 aa  | conserved hypothetical protein                          |
| A909.ORF01497 | 42 aa   | conserved hypothetical protein                          |
| A909.ORF01496 | 100 aa  | conserved hypothetical protein                          |
| A909.ORF01495 | 75 aa   | prophage LambdaSa03, holin, phi LC3 family              |
| A909.ORF01494 | 443 aa  | prophage LambdaSa03, peptidoglycan endolysin            |
| A909.ORF01493 | 67 aa   | conserved hypothetical protein                          |
| A909.ORF01492 | 59 aa   | conserved hypothetical protein                          |

## REGION 3.4 [1 sequence(s)]:

A909 [652861-711447]

|               |         |  |
|---------------|---------|--|
| A909.ORF01427 | 172 aa  | conserved hypothetical protein                           |
| A909.ORF01426 | 32 aa   | hypothetical protein                                     |
| A909.ORF01425 | 176 aa  | transcriptional regulator, TetR family                   |
| A909.ORF01424 | 1774 aa | collagen-like surface protein, putative                  |
| A909.ORF01423 | 622 aa  | conserved hypothetical protein                           |
| A909.ORF01422 | 189 aa  | conserved hypothetical protein                           |
| A909.ORF01421 | 651 aa  | prophage LambdaSa04, DNA polymerase                      |
| A909.ORF01420 | 54 aa   | conserved hypothetical protein                           |
| A909.ORF01419 | 187 aa  | conserved hypothetical protein                           |
| A909.ORF01418 | 373 aa  | conserved hypothetical protein                           |
| A909.ORF01417 | 107 aa  | conserved hypothetical protein                           |
| A909.ORF01416 | 761 aa  | prophage LambdaSa04, DNA primase, P4 family              |
| A909.ORF01415 | 34 aa   | hypothetical protein                                     |
| A909.ORF01414 | 93 aa   | conserved hypothetical protein                           |
| A909.ORF01413 | 458 aa  | prophage LambdaSa04, helicase, SNF2 family               |
| A909.ORF01412 | 165 aa  | conserved hypothetical protein                           |
| A909.ORF01411 | 0 aa    | conserved hypothetical protein, authentic point mutation |
| A909.ORF01410 | 345 aa  | prophage LambdaSa04, S-adenosylmethionine synthetase     |
| A909.ORF01409 | 121 aa  | prophage LambdaSa04, HNH endonuclease family protein     |

|               |         |   |
|---------------|---------|---|
| A909.ORF01408 | 412 aa  | prophage LambdaSa04, DNA methylase  |
| A909.ORF01407 | 417 aa  | prophage LambdaSa04, methyltransferase, C-5 cytosine-specific                             |
| family        |         |   |
| A909.ORF01406 | 149 aa  | conserved domain protein  |
| A909.ORF01405 | 137 aa  | prophage LambdaSa04, terminase, small subunit, P27 family                                 |
| A909.ORF01404 | 530 aa  | prophage LambdaSa04, terminase, large subunit   |
| A909.ORF01403 | 100 aa  | prophage LambdaSa04, RelB antitoxin family protein  |
| A909.ORF01402 | 101 aa  | prophage LambdaSa04, RelE/ParE family protein   |
| A909.ORF01401 | 90 aa   | conserved hypothetical protein  |
| A909.ORF01400 | 429 aa  | prophage LambdaSa04, portal protein, HK97 family  |
| A909.ORF01399 | 232 aa  | prophage LambdaSa04, ClpP endopeptidase (S14) family, non-peptidase homologue, putative   |
| A909.ORF01398 | 402 aa  | prophage LambdaSa04, major capsid protein, HK97 family                                    |
| A909.ORF01397 | 85 aa   | conserved hypothetical protein TIGR01560  |
| A909.ORF01396 | 112 aa  | prophage LambdaSa04, head-tail adaptor, putative  |
| A909.ORF01395 | 122 aa  | conserved hypothetical protein  |
| A909.ORF01394 | 108 aa  | conserved hypothetical protein  |
| A909.ORF01393 | 189 aa  | prophage LambdaSa04, major tail protein, phi13 family                                     |
| A909.ORF02161 | 139 aa  | conserved hypothetical protein  |
| A909.ORF01392 | 31 aa   | conserved hypothetical protein  |
| A909.ORF01391 | 1039 aa | prophage LambdaSa04, tail tape measure protein, TP901 family                              |
| A909.ORF01390 | 241 aa  | prophage LambdaSa04, tail protein, putative   |
| A909.ORF01389 | 971 aa  | prophage LambdaSa04, minor structural protein   |
| A909.ORF01388 | 618 aa  | prophage LambdaSa04, minor structural protein, putative                                   |
| A909.ORF01387 | 132 aa  | prophage LambdaSa04, holin  |
| A909.ORF01386 | 336 aa  | prophage LambdaSa04, mannosyl-glycoprotein endo-beta-N-acetylglucosamidase family protein |
| A909.ORF01385 | 146 aa  | prophage LambdaSa04, LysM domain protein  |
| A909.ORF01384 | 413 aa  | prophage LambdaSa04, site-specific recombinase, resolvase family                          |
| A909.ORF01383 | 448 aa  | prophage LambdaSa04, site-specific recombinase, resolvase family                          |
| A909.ORF01382 | 257 aa  | hypothetical protein  |
| A909.ORF01381 | 501 aa  | serine/threonine protein kinase, putative   |
| A909.ORF01380 | 70 aa   | conserved domain protein  |
| A909.ORF01379 | 245 aa  | HAD-superfamily phosphatase, subfamily IIIB   |
| A909.ORF01378 | 172 aa  | conserved hypothetical protein  |
| A909.ORF01377 | 78 aa   | hypothetical protein  |
| A909.ORF01376 | 110 aa  | cell wall surface anchor family protein, truncation                                       |
| A909.ORF01375 | 249 aa  | ISSag5, transposase orfB  |
| A909.ORF01374 | 91 aa   | ISSag5, transposase orfA  |
| A909.ORF01373 | 115 aa  | chaperone protein HslO, putative  |
| A909.ORF01372 | 402 aa  | transcriptional regulator, AraC family  |

REGION 3.5 [1 sequence(s)]:

## A909 [1303948-1319058]

|               |         |  |
|---------------|---------|--|
| A909.ORF00818 | 96 aa   | ISSag4, transposase orfA                                 |
| A909.ORF00817 | 253 aa  | ISSag4, transposase orfB                                 |
| A909.ORF00816 | 604 aa  | exonuclease, DNA polymerase III, epsilon subunit family  |
| A909.ORF00815 | 411 aa  | hypothetical protein                                     |
| A909.ORF00814 | 393 aa  | site-specific recombinase, phage integrase family        |
| A909.ORF00813 | 188 aa  | plasmid replication protein                              |
| A909.ORF02181 | 0 aa    | conserved hypothetical protein, degenerate               |
| A909.ORF00812 | 0 aa    | conserved hypothetical protein, authentic point mutation |
| A909.ORF00810 | 104 aa  | hypothetical protein                                     |
| A909.ORF00809 | 401 aa  | conserved domain protein                                 |
| A909.ORF00808 | 163 aa  | hypothetical protein                                     |
| A909.ORF00807 | 1203 aa | type IIG restriction enzyme and methyltransferase        |

## REGION 3.6 [2 sequence(s)]:

## A909 [1746807-1764491]

|               |        |  |
|---------------|--------|--|
| A909.ORF00380 | 526 aa | type I restriction-modification system, M subunit                      |
| A909.ORF00379 | 590 aa | conserved domain protein   |
| A909.ORF00378 | 661 aa | transketolase  |
| A909.ORF00377 | 214 aa | transaldolase, putative  |
| A909.ORF00376 | 229 aa | D-allulose-6-phosphate 3-epimerase                                     |
| A909.ORF00375 | 367 aa | PTS system, fructose-specific, IIC component                           |
| A909.ORF00374 | 149 aa | PTS system, fructose-specific, IIA component                           |
| A909.ORF00373 | 103 aa | PTS system, fructose-specific, IIB component                           |
| A909.ORF00372 | 629 aa | PRD domain/PTS system IIA domain protein                               |
| A909.ORF00371 | 0 aa   | conserved domain protein, authentic frameshift                         |
| A909.ORF00370 | 0 aa   | oxidoreductase, short chain dehydrogenase/reductase family, degenerate |
| A909.ORF00369 | 0 aa   | oxidoreductase, aldo/keto reductase family, authentic frameshift       |
| A909.ORF00367 | 176 aa | conserved hypothetical protein   |
| A909.ORF00366 | 263 aa | phenazine biosynthesis protein, PhzF family                            |
| A909.ORF00365 | 136 aa | acetyltransferase, GNAT family, FRAMESHIFT                             |
| A909.ORF00364 | 32 aa  | hypothetical protein   |

## CJB111 [1658722-1676270]

|                 |        |   |
|-----------------|--------|---|
| CJB111.ORF01706 | 526 aa | type I restriction-modification system, M subunit |
| CJB111.ORF01707 | 590 aa | conserved hypothetical protein                    |
| CJB111.ORF01709 | 661 aa | transketolase                                     |
| CJB111.ORF01710 | 214 aa | transaldolase, putative                           |
| CJB111.ORF01711 | 229 aa | ribulose-phosphate 3-epimerase                    |
| CJB111.ORF01712 | 367 aa | PTS system, IIC component [imported]              |
| CJB111.ORF01713 | 149 aa | PTS system, IIA component [imported]              |
| CJB111.ORF01714 | 103 aa | PTS system, IIB component [imported]              |
| CJB111.ORF01715 | 629 aa | MW0309, putative                                  |
| CJB111.ORF01716 | 75 aa  | transcription regulator, putative                 |

|                  |        |  |
|------------------|--------|--|
| CJB111.ORF01717  | 149 aa | probable bud operon transcriptional regulator , fragment   |
| CJB111.ORF01718  | 116 aa | oxidoreductase, short chain dehydrogenase/reductase family |
| CJB111.ORF01719  | 245 aa | oxidoreductase, aldo/keto reductase family                 |
| CJB111.ORF01720  | 82 aa  | conserved hypothetical protein                             |
| CJB111.ORF01721  | 176 aa | Glutamate-rich protein grpB                                |
| CJB111.ORF01722  | 230 aa | phenazine biosynthesis protein PhzF family                 |
| CJB111.ORF01723  | 136 aa | acetyltransferase, GNAT family family                      |
| CJB111.ORFA01679 | 32 aa  | hypothetical protein                                       |

## REGION 3.7 [4 sequence(s)]:

A909 [2069793-2086270]

|               |        |  |
|---------------|--------|--|
| A909.ORF00060 | 179 aa | prophage Sa05, transcriptional regulator, putative               |
| A909.ORF00059 | 98 aa  | conserved hypothetical protein                                   |
| A909.ORF00058 | 207 aa | hypothetical protein   |
| A909.ORF00056 | 168 aa | conserved hypothetical protein                                   |
| A909.ORF00055 | 36 aa  | hypothetical protein   |
| A909.ORF00054 | 128 aa | prophage Sa05, transcriptional regulator, ArpU family            |
| A909.ORF00053 | 120 aa | conserved hypothetical protein                                   |
| A909.ORF02193 | 162 aa | conserved hypothetical protein                                   |
| A909.ORF00052 | 181 aa | conserved hypothetical protein                                   |
| A909.ORF00051 | 480 aa | prophage Sa05, DNA primase, P4 family                            |
| A909.ORF00050 | 285 aa | prophage Sa05, DNA replication protein, putative                 |
| A909.ORF00049 | 109 aa | conserved hypothetical protein                                   |
| A909.ORF00048 | 110 aa | conserved hypothetical protein                                   |
| A909.ORF00047 | 189 aa | hypothetical protein   |
| A909.ORF00046 | 57 aa  | prophage Sa05, transcriptional regulator, CopG family            |
| A909.ORF00045 | 212 aa | prophage Sa05, BRO domain protein                                |
| A909.ORF00044 | 200 aa | prophage Sa05, BRO domain protein                                |
| A909.ORF00043 | 204 aa | prophage Sa05, DNA-binding protein                               |
| A909.ORF00042 | 34 aa  | hypothetical protein   |
| A909.ORF00041 | 321 aa | conserved domain protein   |
| A909.ORF00040 | 388 aa | prophage Sa05, site-specific recombinase, phage integrase family |

CJB111 [1979919-1995368]

|                  |        |  |
|------------------|--------|--|
| CJB111.ORF02049  | 179 aa | transcriptional regulator, TetR family, putative     |
| CJB111.ORF02050  | 98 aa  | conserved hypothetical protein                       |
| CJB111.ORF02051  | 115 aa | conserved hypothetical protein                       |
| CJB111.ORF02052  | 95 aa  | Unknown, putative                                    |
| CJB111.ORFA01995 | 123 aa | Phage encoded transcriptional regulator, ArpU family |
| CJB111.ORF02053  | 39 aa  | conserved hypothetical protein                       |
| CJB111.ORF02054  | 162 aa | conserved hypothetical protein                       |
| CJB111.ORF02055  | 185 aa | conserved hypothetical protein                       |
| CJB111.ORF02056  | 464 aa | virulence-associated protein E                       |
| CJB111.ORF02057  | 285 aa | DNA primase domain protein                           |

|                        |        |  |
|------------------------|--------|--|
| CJB111.ORF02058        | 90 aa  | conserved hypothetical protein                           |
| CJB111.ORF02059        | 63 aa  | conserved hypothetical protein                           |
| CJB111.ORF02060        | 73 aa  | conserved hypothetical protein                           |
| CJB111.ORF02061        | 110 aa | conserved hypothetical protein                           |
| CJB111.ORF02062        | 150 aa | hypothetical protein                                     |
| CJB111.ORF02063        | 66 aa  | conserved hypothetical protein                           |
| CJB111.ORF02064        | 212 aa | BRO family, N-terminal domain protein                    |
| CJB111.ORF02065        | 200 aa | BRO family, N-terminal domain protein                    |
| CJB111.ORF02066        | 61 aa  | repressor-related protein                                |
| CJB111.ORF02067        | 195 aa | transcriptional regulator, Cro/CI family, putative       |
| CJB111.ORF02068        | 385 aa | ebhA protein, putative                                   |
| CJB111.ORF02069        | 56 aa  | hypothetical protein                                     |
| CJB111.ORF02070        | 394 aa | integrase/recombinase, phage integrase family [imported] |
| COH1 [2103650-2118703] |        |  |
| COH1.ORF02295          | 168 aa | transcriptional regulator, TetR family, putative         |
| COH1.ORF02296          | 98 aa  | conserved hypothetical protein                           |
| COH1.ORF02297          | 31 aa  | hypothetical protein                                     |
| COH1.ORF02298          | 207 aa | hypothetical protein                                     |
| COH1.ORF02299          | 91 aa  | conserved hypothetical protein                           |
| COH1.ORF02301          | 36 aa  | hypothetical protein                                     |
| COH1.ORFA02210         | 126 aa | Phage encoded transcriptional regulator, ArpU family     |
| COH1.ORF02302          | 120 aa | conserved hypothetical protein                           |
| COH1.ORF02303          | 59 aa  | hypothetical protein                                     |
| COH1.ORF02304          | 162 aa | conserved hypothetical protein                           |
| COH1.ORF02305          | 201 aa | conserved hypothetical protein                           |
| COH1.ORF02306          | 273 aa | DnaD and phage-associated domain protein                 |
| COH1.ORF02307          | 55 aa  | hypothetical protein                                     |
| COH1.ORF02308          | 113 aa | conserved hypothetical protein                           |
| COH1.ORF02309          | 120 aa | conserved hypothetical protein                           |
| COH1.ORF02310          | 64 aa  | conserved hypothetical protein                           |
| COH1.ORF02311          | 110 aa | conserved hypothetical protein                           |
| COH1.ORF02312          | 150 aa | hypothetical protein                                     |
| COH1.ORF02313          | 66 aa  | conserved hypothetical protein                           |
| COH1.ORF02314          | 212 aa | BRO family, N-terminal domain protein                    |
| COH1.ORF02315          | 255 aa | Unknown  |
| COH1.ORF02316          | 50 aa  | conserved hypothetical protein                           |
| COH1.ORF02317          | 57 aa  | hypothetical protein                                     |
| COH1.ORF02318          | 34 aa  | conserved hypothetical protein                           |
| COH1.ORF02319          | 80 aa  | hypothetical protein                                     |
| COH1.ORF02320          | 113 aa | conserved hypothetical protein                           |
| COH1.ORF02321          | 321 aa | conserved hypothetical protein                           |
| COH1.ORF02322          | 390 aa | integrase/recombinase, phage integrase family [imported] |
| H36B [2057031-2071241] |        |  |

|                |        |  |
|----------------|--------|--|
| H36B.ORF02216  | 179 aa | transcriptional regulator, TetR family, putative         |
| H36B.ORF02217  | 83 aa  | conserved hypothetical protein                           |
| H36B.ORF02218  | 230 aa | Gp21 protein   |
| H36B.ORF02219  | 75 aa  | hypothetical protein                                     |
| H36B.ORF02220  | 52 aa  | hypothetical protein                                     |
| H36B.ORF02221  | 61 aa  | hypothetical protein                                     |
| H36B.ORFA02146 | 123 aa | Phage encoded transcriptional regulator, ArpU family     |
| H36B.ORF02222  | 120 aa | conserved hypothetical protein                           |
| H36B.ORF02223  | 185 aa | conserved hypothetical protein                           |
| H36B.ORF02224  | 273 aa | DnaD and phage-associated domain protein                 |
| H36B.ORF02225  | 55 aa  | hypothetical protein                                     |
| H36B.ORF02226  | 113 aa | hyp, putative  |
| H36B.ORF02227  | 120 aa | conserved hypothetical protein                           |
| H36B.ORF02228  | 63 aa  | conserved hypothetical protein                           |
| H36B.ORF02229  | 110 aa | conserved hypothetical protein                           |
| H36B.ORF02230  | 103 aa | hypothetical protein                                     |
| H36B.ORF02231  | 78 aa  | membrane protein, putative                               |
| H36B.ORF02232  | 113 aa | membrane protein, putative                               |
| H36B.ORF02233  | 206 aa | BRO family, N-terminal domain protein                    |
| H36B.ORF02234  | 204 aa | Helix-turn-helix domain protein                          |
| H36B.ORF02235  | 34 aa  | conserved hypothetical protein                           |
| H36B.ORF02236  | 321 aa | conserved hypothetical protein                           |
| H36B.ORF02237  | 390 aa | integrase/recombinase, phage integrase family [imported] |

## REGION 4.1 [1 sequence(s)]:

515 [53005-63274]

|              |        |   |
|--------------|--------|---|
| 515.ORF00052 | 85 aa  | Unknown   |
| 515.ORF00053 | 367 aa | Unknown   |
| 515.ORF00054 | 432 aa | adenylosuccinate lyase                                    |
| 515.ORF00055 | 300 aa | transcriptional regulator, Cro/CI family                  |
| 515.ORF00056 | 318 aa | Holliday junction DNA helicase RuvB                       |
| 515.ORF00057 | 145 aa | phosphotyrosine protein phosphatase, low molecular weight |
| 515.ORF00058 | 277 aa | sugar ABC transporter, sugar-binding protein              |
| 515.ORF00059 | 178 aa | sugar ABC transporter, sugar-binding protein              |
| 515.ORF00060 | 225 aa | sugar ABC transporter, sugar-binding protein              |
| 515.ORF00061 | 50 aa  | sugar ABC transporter, permease protein                   |
| 515.ORF00062 | 228 aa | sugar ABC transporter, permease protein                   |
| 515.ORF00063 | 128 aa | MORN motif family protein                                 |

## REGION 4.2 [1 sequence(s)]:

515 [226384-240579]

|              |        |  |
|--------------|--------|--|
| 515.ORF00256 | 284 aa | protein gp49 [Bacteriophage A118] homolog lmo2317 [imported] |
| 515.ORF00257 | 96 aa  | conserved hypothetical protein                               |

|              |        |   |
|--------------|--------|---|
| 515.ORF00258 | 121 aa | Tn5252, Orf 9 protein [imported]                  |
| 515.ORF00259 | 118 aa | Tn5252, Orf 9 protein [imported]                  |
| 515.ORF00260 | 774 aa | type I restriction-modification system, R subunit |
| 515.ORF00261 | 487 aa | N-6 DNA Methylase family                          |
| 515.ORF00262 | 192 aa | type I restriction-modification system, S subunit |
| 515.ORF00263 | 523 aa | putative helicase                                 |
| 515.ORF00264 | 447 aa | hypothetical protein                              |
| 515.ORF00265 | 96 aa  | Unknown   |
| 515.ORF00266 | 135 aa | Unknown   |
| 515.ORF00267 | 195 aa | Unknown   |
| 515.ORF00268 | 55 aa  | conserved hypothetical protein                    |
| 515.ORF00269 | 81 aa  | conserved hypothetical protein                    |
| 515.ORF00270 | 128 aa | Unknown   |
| 515.ORF00271 | 93 aa  | Unknown   |
| 515.ORF00272 | 32 aa  | conserved hypothetical protein                    |

## REGION 4.3 [2 sequence(s)]:

515 [449812-457176]

|              |        |   |
|--------------|--------|---|
| 515.ORF00509 | 234 aa | Fic protein family family                 |
| 515.ORF00510 | 56 aa  | conserved hypothetical protein            |
| 515.ORF00511 | 698 aa | ATP-dependent RNA helicase SrmB, putative |
| 515.ORF00512 | 216 aa | CRISPR-associated protein, CT1134 family  |
| 515.ORF00513 | 190 aa | oxidoreductase, Gfo/Idh/MocA family       |
| 515.ORF00514 | 300 aa | magnesium transporter, CorA family        |
| 515.ORF00515 | 354 aa | Unknown                                   |

CJB111 [452163-457290]

|                 |        |                                     |
|-----------------|--------|-------------------------------------|
| CJB111.ORF00469 | 309 aa | Unknown                             |
| CJB111.ORF00470 | 319 aa | oxidoreductase, Gfo/Idh/MocA family |
| CJB111.ORF00471 | 300 aa | magnesium transporter, CorA family  |
| CJB111.ORF00472 | 354 aa | Unknown                             |
| CJB111.ORF00473 | 57 aa  | aspartate--ammonia ligase           |

## REGION 4.4 [2 sequence(s)]:

515 [1250101-1257453]

|              |        |   |
|--------------|--------|---|
| 515.ORF01378 | 96 aa  | ISSdy1, transposase OrfA                        |
| 515.ORF01379 | 247 aa | ISSdy1, transposase OrfB                        |
| 515.ORF01380 | 604 aa | Unknown   |
| 515.ORF01381 | 411 aa | Unknown   |
| 515.ORF01382 | 121 aa | ribosomal protein L7/L12                        |
| 515.ORF01383 | 166 aa | ribosomal protein L10                           |
| 515.ORF01385 | 156 aa | ATP-dependent Clp protease, ATP-binding subunit |
| 515.ORF01386 | 80 aa  | conserved hypothetical protein                  |

CJB111 [1231555-1237140]

|                 |        |                          |
|-----------------|--------|--------------------------|
| CJB111.ORF01259 | 96 aa  | ISSdy1, transposase OrfA |
| CJB111.ORF01260 | 237 aa | ISSdy1, transposase OrfB |
| CJB111.ORF01261 | 604 aa | Unknown                  |
| CJB111.ORF01262 | 334 aa | Unknown                  |

## REGION 4.5 [1 sequence(s)]:

515 [2047004-2108009]

|               |        |  |
|---------------|--------|--|
| 515.ORF02225  | 453 aa | chromosomal replication initiator protein DnaA                 |
| 515.ORF02226  | 378 aa | DNA polymerase III, beta subunit                               |
| 515.ORF02227  | 116 aa | diacylglycerol kinase catalytic domain protein                 |
| 515.ORF02228  | 65 aa  | amino acid ABC transporter, permease protein                   |
| 515.ORF02229  | 76 aa  | conserved hypothetical protein                                 |
| 515.ORF02230  | 60 aa  | Unknown  |
| 515.ORF02231  | 112 aa | cadmium efflux system accessory protein, putative              |
| 515.ORF02232  | 204 aa | cadmium resistance protein                                     |
| 515.ORF02233  | 49 aa  | ribosomal protein L33  |
| 515.ORF02234  | 60 aa  | ribosomal protein L32  |
| 515.ORFA02144 | 32 aa  | histidyl-tRNA synthetase                                       |
| 515.ORFA02146 | 54 aa  | DNA primase  |
| 515.ORF02235  | 497 aa | Type I restriction modification DNA specificity domain protein |
| 515.ORF02236  | 58 aa  | type I restriction enzyme [imported]                           |
| 515.ORF02237  | 285 aa | major facilitator family protein                               |
| 515.ORF02238  | 239 aa | ribosomal small subunit pseudouridine synthase A               |
| 515.ORF02239  | 146 aa | transcriptional regulator, MarR family                         |
| 515.ORF02240  | 605 aa | ABC transporter, ATP-binding/permease protein                  |
| 515.ORF02241  | 584 aa | ABC transporter, ATP-binding/permease protein                  |
| 515.ORF02242  | 208 aa | GTP pyrophosphokinase family protein                           |
| 515.ORF02243  | 226 aa | DNA-binding response regulator                                 |
| 515.ORF02244  | 46 aa  | hypothetical phage protein-related protein                     |
| 515.ORF02245  | 125 aa | hypothetical phage protein                                     |
| 515.ORF02246  | 83 aa  | voltage-gated chloride channel family protein                  |
| 515.ORF02247  | 127 aa | type I restriction enzyme S protein (hsdS)                     |
| 515.ORF02248  | 477 aa | ATP-dependent DNA helicase recG, putative                      |
| 515.ORF02249  | 111 aa | transcriptional regulator, PadR family domain protein          |
| 515.ORF02250  | 189 aa | conserved hypothetical protein                                 |
| 515.ORF02251  | 100 aa | conserved hypothetical protein                                 |
| 515.ORF02252  | 79 aa  | conserved hypothetical protein                                 |
| 515.ORF02253  | 392 aa | Type I restriction modification DNA specificity domain protein |
| 515.ORF02254  | 265 aa | integrase/recombinase, phage integrase family [imported]       |
| 515.ORF02255  | 283 aa | Unknown  |
| 515.ORFA02168 | 44 aa  | cppA protein   |
| 515.ORF02256  | 36 aa  | methionyl-tRNA synthetase                                      |
| 515.ORF02257  | 180 aa | c protein alpha-antigen precursor                              |

|               |        |   |
|---------------|--------|---|
| 515.ORF02258  | 379 aa | c protein alpha-antigen precursor   |
| 515.ORF02259  | 282 aa | Type I restriction modification DNA specificity domain protein                      |
| 515.ORFA02174 | 65 aa  | seryl-tRNA synthetase   |
| 515.ORF02260  | 140 aa | Unknown   |
| 515.ORF02261  | 69 aa  | Unknown   |
| 515.ORF02262  | 294 aa | C protein alpha antigen   |
| 515.ORFA02177 | 90 aa  | LPXTG-motif cell wall anchor domain protein   |
| 515.ORF02263  | 117 aa | ABC transporter, permease protein, putative   |
| 515.ORFA02179 | 54 aa  | ribose operon repressor RbsR  |
| 515.ORFA02180 | 48 aa  | Binding-protein-dependent transport systems inner membrane component domain protein |
| 515.ORF02264  | 90 aa  | uracil permease   |
| 515.ORF02265  | 53 aa  | phosphoribosylaminoimidazolecarboxamide formyltransferase/IMP cyclohydrolase        |
| 515.ORF02266  | 33 aa  | c protein alpha-antigen precursor   |
| 515.ORF02267  | 200 aa | c protein alpha-antigen precursor   |
| 515.ORFA02185 | 62 aa  | hypothetical protein  |
| 515.ORF02268  | 78 aa  | HD domain protein   |
| 515.ORF02269  | 37 aa  | PTS system, IIC component, putative   |
| 515.ORFA02189 | 41 aa  | hypothetical protein  |
| 515.ORF02270  | 88 aa  | HD domain protein   |
| 515.ORF02271  | 164 aa | sensor histidine kinase   |
| 515.ORF02272  | 280 aa | ATP-dependent RNA helicase, DEAD/DEAH box family                                    |
| 515.ORF02273  | 269 aa | c protein alpha-antigen precursor   |
| 515.ORF02276  | 234 aa | hypothetical protein  |
| 515.ORF02277  | 48 aa  | xanthine permease   |
| 515.ORF02278  | 264 aa | c protein alpha-antigen precursor   |
| 515.ORF02279  | 66 aa  | hydrolase, haloacid dehalogenase-like family  |
| 515.ORF02280  | 194 aa | c protein alpha-antigen precursor   |
| 515.ORF02281  | 258 aa | amidase family protein  |
| 515.ORF02282  | 417 aa | 5'-nucleotidase family protein  |
| 515.ORF02283  | 33 aa  | hypothetical protein  |
| 515.ORF02284  | 86 aa  | hypothetical protein  |
| 515.ORF02285  | 219 aa | c protein alpha-antigen precursor   |
| 515.ORF02286  | 57 aa  | GTP-binding protein, GTP1/Obg family  |
| 515.ORFA02206 | 38 aa  | stage V sporulation protein B   |
| 515.ORF02287  | 100 aa | large conductance mechanosensitive channel protein                                  |
| 515.ORF02288  | 224 aa | DNA primase   |
| 515.ORF02289  | 192 aa | c protein alpha-antigen precursor   |
| 515.ORF02290  | 266 aa | c protein alpha-antigen precursor   |
| 515.ORF02291  | 194 aa | glycosyl transferase, group 2 family protein  |
| 515.ORF02292  | 104 aa | conserved hypothetical protein  |
| 515.ORF02293  | 195 aa | c protein alpha-antigen precursor   |

|               |        |  |
|---------------|--------|--|
| 515.ORFA02213 | 52 aa  | uncharacterized conserved protein CAC1439                        |
| 515.ORF02294  | 62 aa  | cell division protein DivIVA [imported]                          |
| 515.ORFA02215 | 30 aa  | hypothetical protein   |
| 515.ORF02295  | 126 aa | peptidase, M20/M25/M40 family                                    |
| 515.ORF02296  | 58 aa  | oxidoreductase, aldo/keto reductase family                       |
| 515.ORF02297  | 40 aa  | O-sialoglycoprotein endopeptidase family protein                 |
| 515.ORF02298  | 46 aa  | Unknown  |
| 515.ORF02299  | 44 aa  | Unknown  |
| 515.ORF02300  | 56 aa  | peptide chain release factor 1                                   |
| 515.ORFA02223 | 37 aa  | phosphopentomutase   |
| 515.ORF02301  | 161 aa | RNA methyltransferase, TrmA family                               |
| 515.ORF02302  | 146 aa | ABC transporter, permease protein, putative                      |
| 515.ORFA02225 | 37 aa  | hypothetical protein   |
| 515.ORF02303  | 57 aa  | DNA-3-methyladenine glycosylase I                                |
| 515.ORFA02228 | 31 aa  | conserved hypothetical protein                                   |
| 515.ORF02306  | 49 aa  | FtsK/SpoIIIE family protein                                      |
| 515.ORFA02230 | 57 aa  | C3-degrading proteinase  |
| 515.ORF02307  | 77 aa  | unnamed protein product  |
| 515.ORF02308  | 150 aa | ATP-dependent DNA helicase recG, putative                        |
| 515.ORFA02233 | 30 aa  | hypothetical protein   |
| 515.ORF02309  | 197 aa | polyribonucleotide nucleotidyltransferase                        |
| 515.ORFA02236 | 34 aa  | hypothetical protein   |
| 515.ORF02310  | 268 aa | prophage LambdaSa1, N-acetylmuramoyl-L-alanine amidase, family 4 |
| 515.ORF02311  | 174 aa | essential recombination function protein                         |
| 515.ORF02312  | 48 aa  | conserved hypothetical protein                                   |

## REGION 5.1 [2 sequence(s)]:

CJB111 [228054-235974]

|                 |        |  |
|-----------------|--------|--|
| CJB111.ORF00247 | 450 aa | aspartate kinase family protein              |
| CJB111.ORF00248 | 220 aa | hydrolase, haloacid dehalogenase-like family |

COH1 [341611-352205]

|               |        |  |
|---------------|--------|--|
| COH1.ORF00377 | 450 aa | aspartate kinase family protein              |
| COH1.ORF00378 | 216 aa | hydrolase, haloacid dehalogenase-like family |
| COH1.ORF00379 | 100 aa | enoyl-CoA hydratase/isomerase family protein |
| COH1.ORF00380 | 67 aa  | Stp1   |
| COH1.ORF00381 | 162 aa | serine/threonine protein kinase              |
| COH1.ORF00382 | 68 aa  | conserved hypothetical protein               |
| COH1.ORF00383 | 109 aa | sensor histidine kinase, putative            |
| COH1.ORF00384 | 39 aa  | sensor histidine kinase, putative            |
| COH1.ORF00385 | 236 aa | sensor histidine kinase, putative            |

## REGION 5.2 [1 sequence(s)]:

CJB111 [545119-553178]

|                 |        |   |
|-----------------|--------|---|
| CJB111.ORF00567 | 335 aa | ribosomal protein S1  |
| CJB111.ORF00568 | 263 aa | prophage LambdaSa2, site-specific recombinase, phage integrase family |
| CJB111.ORF00569 | 208 aa | prophage pi3 protein 59 [imported]                                    |
| CJB111.ORF00570 | 52 aa  | repressor protein, putative   |
| CJB111.ORF00571 | 55 aa  | repressor protein, putative   |
| CJB111.ORF00572 | 91 aa  | Unknown   |
| CJB111.ORF00573 | 47 aa  | chorismate mutase, putative   |
| CJB111.ORF00574 | 398 aa | voltage-gated chloride channel family protein                         |
| CJB111.ORF00575 | 115 aa | ribosomal protein L19   |
| CJB111.ORF00577 | 169 aa | prophage LambdaSa2, site-specific recombinase, phage integrase family |

## REGION 5.3 [1 sequence(s)]:

CJB111 [577660-633597]

|                 |         |  |
|-----------------|---------|--|
| CJB111.ORF00604 | 172 aa  | conserved hypothetical protein                   |
| CJB111.ORF00605 | 405 aa  | transposase                                      |
| CJB111.ORF00606 | 67 aa   | conserved hypothetical protein                   |
| CJB111.ORF00607 | 140 aa  | conserved hypothetical protein                   |
| CJB111.ORF00608 | 117 aa  | Tn916, transcriptional regulator, putative       |
| CJB111.ORF00609 | 639 aa  | tetracycline resistance protein (tetM)           |
| CJB111.ORF00610 | 304 aa  | conserved hypothetical protein                   |
| CJB111.ORF00611 | 333 aa  | lipoprotein, NLP/P60 family                      |
| CJB111.ORF00612 | 175 aa  | conserved hypothetical protein                   |
| CJB111.ORF00613 | 725 aa  | membrane protein, putative                       |
| CJB111.ORF00614 | 815 aa  | ATP/GTP-binding protein, putative                |
| CJB111.ORF00615 | 168 aa  | conserved hypothetical protein                   |
| CJB111.ORF00616 | 165 aa  | conserved hypothetical protein                   |
| CJB111.ORF00617 | 73 aa   | conserved hypothetical protein                   |
| CJB111.ORF00618 | 401 aa  | transcriptional regulator, Cro/CI family         |
| CJB111.ORF00619 | 461 aa  | FtsK/SpoIIIE family protein                      |
| CJB111.ORF00620 | 128 aa  | conserved hypothetical protein                   |
| CJB111.ORF00621 | 104 aa  | conserved hypothetical protein                   |
| CJB111.ORF00622 | 32 aa   | transposase, IS256 family                        |
| CJB111.ORF00623 | 198 aa  | transcriptional regulator, TetR family, putative |
| CJB111.ORF00624 | 1583 aa | reticulocyte binding protein                     |
| CJB111.ORF00625 | 45 aa   | 3-phosphoshikimate 1-carboxyvinyltransferase     |
| CJB111.ORF00626 | 238 aa  | phage terminase, large subunit, putative         |
| CJB111.ORF00627 | 85 aa   | hypothetical protein                             |
| CJB111.ORF00628 | 123 aa  | Doc protein                                      |
| CJB111.ORF00629 | 108 aa  | hypothetical protein                             |
| CJB111.ORF00630 | 358 aa  | phage portal protein, HK97 family                |
| CJB111.ORF00631 | 71 aa   | hypothetical protein                             |

|                  |         |   |
|------------------|---------|---|
| CJB111.ORF00632  | 232 aa  | ClpP protease family protein                            |
| CJB111.ORF00633  | 402 aa  | phage major capsid protein, HK97 family                 |
| CJB111.ORF00634  | 85 aa   | hypothetical phage protein, putative                    |
| CJB111.ORF00635  | 122 aa  | prophage pi2 protein 37 [imported]                      |
| CJB111.ORF00636  | 108 aa  | prophage pi2 protein 38 [imported], putative            |
| CJB111.ORF00637  | 189 aa  | prophage pi2 protein 39 [imported]                      |
| CJB111.ORF00638  | 139 aa  | conserved hypothetical protein                          |
| CJB111.ORF00639  | 31 aa   | hypothetical protein                                    |
| CJB111.ORF00640  | 1039 aa | tail tape measure protein                               |
| CJB111.ORF00642  | 241 aa  | conserved hypothetical protein                          |
| CJB111.ORF00643  | 971 aa  | prophage LambdaSa2, PblB, putative                      |
| CJB111.ORF00645  | 618 aa  | prophage LambdaSa1, minor structural protein, putative  |
| CJB111.ORFA00621 | 94 aa   | hypothetical protein                                    |
| CJB111.ORF00646  | 132 aa  | holin, putative   |
| CJB111.ORF00647  | 466 aa  | N-acetylmuramoyl-L-alanine amidase, family 4            |
| CJB111.ORF00648  | 35 aa   | conserved hypothetical protein                          |
| CJB111.ORF00649  | 413 aa  | site-specific recombinase, resolvase family, putative   |
| CJB111.ORFA00625 | 66 aa   | hypothetical protein                                    |
| CJB111.ORF00650  | 448 aa  | site-specific recombinase, resolvase family, putative   |
| CJB111.ORF00651  | 257 aa  | hypothetical protein                                    |
| CJB111.ORF00652  | 501 aa  | Protein kinase domain protein                           |
| CJB111.ORF00653  | 70 aa   | Unknown-related protein                                 |
| CJB111.ORF00654  | 245 aa  | HAD superfamily (subfamily IIIB) phosphatase, TIGR01672 |
| CJB111.ORF00655  | 172 aa  | Unknown   |
| CJB111.ORF00656  | 78 aa   | conserved hypothetical protein                          |
| CJB111.ORF00657  | 110 aa  | cell wall surface anchor family protein, truncation     |
| CJB111.ORF00658  | 273 aa  | transposase OrfB, IS3 family, truncation                |
| CJB111.ORF00659  | 116 aa  | probable transposase TnpA                               |
| CJB111.ORF00660  | 49 aa   | Tn5252, Orf 10 protein [imported]                       |
| CJB111.ORF00661  | 53 aa   | chaperonin, 33 kDa                                      |
| CJB111.ORF00662  | 115 aa  | chaperonin, 33 kDa                                      |
| CJB111.ORF00663  | 402 aa  | transcriptional regulator, AraC family                  |

## REGION 5.4 [1 sequence(s)]:

CJB111 [2037780-2118007]

|                 |        |  |
|-----------------|--------|--|
| CJB111.ORF02115 | 453 aa | chromosomal replication initiator protein DnaA           |
| CJB111.ORF02116 | 378 aa | DNA polymerase III, beta subunit                         |
| CJB111.ORF02117 | 111 aa | diacylglycerol kinase catalytic domain protein, putative |
| CJB111.ORF02118 | 197 aa | diacylglycerol kinase catalytic domain protein, putative |
| CJB111.ORF02119 | 112 aa | conserved hypothetical protein                           |
| CJB111.ORF02120 | 127 aa | Helix-turn-helix domain protein                          |
| CJB111.ORF02121 | 63 aa  | hypothetical protein                                     |
| CJB111.ORF02122 | 74 aa  | hypothetical protein                                     |

|                  |         |   |
|------------------|---------|---|
| CJB111.ORF02123  | 101 aa  | modification methylase LlaFI                                |
| CJB111.ORF02124  | 36 aa   | conserved hypothetical protein                              |
| CJB111.ORF02125  | 30 aa   | conserved hypothetical protein                              |
| CJB111.ORF02126  | 406 aa  | site-specific recombinase, phage integrase family, putative |
| CJB111.ORF02127  | 91 aa   | conserved hypothetical protein                              |
| CJB111.ORF02128  | 248 aa  | Rep protein   |
| CJB111.ORF02129  | 263 aa  | FtsK/SpoIIIE family protein, putative                       |
| CJB111.ORF02130  | 99 aa   | hypothetical protein  |
| CJB111.ORF02131  | 307 aa  | phage terminase, large subunit, putative                    |
| CJB111.ORF02132  | 157 aa  | phage terminase, small subunit, putative                    |
| CJB111.ORF02133  | 149 aa  | hypothetical protein  |
| CJB111.ORF02134  | 394 aa  | C-5 cytosine-specific DNA methylase                         |
| CJB111.ORF02135  | 418 aa  | prophage LambdaW4, DNA methylase                            |
| CJB111.ORFA02078 | 62 aa   | Phage endonuclease  |
| CJB111.ORF02136  | 345 aa  | S-adenosylmethionine synthetase                             |
| CJB111.ORF02137  | 64 aa   | conserved hypothetical protein                              |
| CJB111.ORF02138  | 169 aa  | hypothetical protein  |
| CJB111.ORF02139  | 458 aa  | SNF2 domain protein   |
| CJB111.ORF02140  | 109 aa  | P44   |
| CJB111.ORF02141  | 759 aa  | bacteriophage protein, putative                             |
| CJB111.ORF02142  | 60 aa   | hypothetical protein  |
| CJB111.ORF02143  | 107 aa  | hypothetical protein  |
| CJB111.ORF02144  | 373 aa  | phi APSE P51-like protein                                   |
| CJB111.ORF02145  | 187 aa  | phage protein   |
| CJB111.ORF02146  | 54 aa   | hypothetical protein  |
| CJB111.ORF02147  | 651 aa  | DNA polymerase, putative                                    |
| CJB111.ORF02149  | 189 aa  | hypothetical protein  |
| CJB111.ORF02148  | 45 aa   | hypothetical protein  |
| CJB111.ORF02150  | 44 aa   | hypothetical protein  |
| CJB111.ORF02151  | 622 aa  | conserved hypothetical protein                              |
| CJB111.ORF02152  | 76 aa   | peptidase, M20/M25/M40 family [imported], putative          |
| CJB111.ORF02153  | 333 aa  | peptidase, M20/M25/M40 family [imported]                    |
| CJB111.ORF02154  | 247 aa  | surface protein Rib   |
| CJB111.ORF02155  | 188 aa  | ribosomal protein L10                                       |
| CJB111.ORF02156  | 58 aa   | 50S ribosomal protein L7/l12 , fragment-related protein     |
| CJB111.ORF02157  | 1203 aa | reticulocyte binding protein                                |
| CJB111.ORF02158  | 163 aa  | Unknown   |
| CJB111.ORF02159  | 401 aa  | conserved hypothetical protein                              |
| CJB111.ORF02160  | 104 aa  | Unknown   |
| CJB111.ORF02161  | 167 aa  | Unknown   |
| CJB111.ORF02162  | 169 aa  | FtsK/SpoIIIE family protein, putative                       |
| CJB111.ORF02163  | 101 aa  | conserved hypothetical protein                              |
| CJB111.ORF02165  | 329 aa  | surface protein Rib   |

|                  |        |  |
|------------------|--------|--|
| CJB111.ORF02166  | 55 aa  | tail tape measure protein                                    |
| CJB111.ORF02167  | 72 aa  | phosphoribosylaminoimidazole carboxylase, catalytic subunit  |
| CJB111.ORF02168  | 368 aa | site-specific recombinase, phage integrase family            |
| CJB111.ORF02169  | 77 aa  | Unknown-related protein                                      |
| CJB111.ORF02170  | 207 aa | Plasmid replication protein                                  |
| CJB111.ORFA02111 | 31 aa  | hypothetical protein   |
| CJB111.ORF02171  | 251 aa | surface protein Rib  |
| CJB111.ORFA02113 | 125 aa | hypothetical protein   |
| CJB111.ORF02172  | 216 aa | surface protein Rib  |
| CJB111.ORF02173  | 0 aa   | surface protein Rib  |
| CJB111.ORF02174  | 42 aa  | cation uptake protein, Trk family                            |
| CJB111.ORF02175  | 134 aa | polysaccharide biosynthesis protein CpsM                     |
| CJB111.ORF02176  | 99 aa  | IS1381, transposase OrfB [imported]                          |
| CJB111.ORF02177  | 408 aa | site-specific recombinase, phage integrase family            |
| CJB111.ORF02178  | 272 aa | hypothetical protein   |
| CJB111.ORF02179  | 469 aa | hypothetical protein   |
| CJB111.ORF02180  | 145 aa | conserved hypothetical protein                               |
| CJB111.ORF02181  | 220 aa | hypothetical protein   |
| CJB111.ORF02182  | 40 aa  | Unknown  |
| CJB111.ORF02183  | 77 aa  | Unknown-related protein                                      |
| CJB111.ORF02184  | 117 aa | amino acid ABC transporter, amino acid-binding protein       |
| CJB111.ORF02186  | 184 aa | surface protein Rib  |
| CJB111.ORFA02127 | 198 aa | hypothetical protein   |
| CJB111.ORF02187  | 197 aa | surface protein Rib  |
| CJB111.ORF02188  | 58 aa  | pyridine nucleotide-disulphide oxidoreductase family protein |
| CJB111.ORF02189  | 107 aa | surface protein Rib  |
| CJB111.ORF02190  | 176 aa | translation initiation factor IF-3                           |
| CJB111.ORF02191  | 77 aa  | surface protein Rib  |
| CJB111.ORF02192  | 36 aa  | alpha-like protein 3   |
| CJB111.ORF02193  | 50 aa  | voltage-gated chloride channel family protein                |
| CJB111.ORF02194  | 62 aa  | surface protein Rib  |
| CJB111.ORF02195  | 171 aa | surface protein Rib  |
| CJB111.ORF02196  | 55 aa  | DNA-directed RNA polymerase, beta subunit                    |
| CJB111.ORF02197  | 165 aa | surface protein Rib  |
| CJB111.ORFA02141 | 54 aa  | conserved hypothetical protein                               |
| CJB111.ORF02200  | 193 aa | rib protein  |
| CJB111.ORF02198  | 175 aa | surface protein Rib  |
| CJB111.ORF02204  | 199 aa | hypothetical protein   |
| CJB111.ORFA02145 | 199 aa | hypothetical protein   |
| CJB111.ORFA02146 | 246 aa | hypothetical protein   |
| CJB111.ORF02205  | 262 aa | surface protein Rib  |
| CJB111.ORFA02148 | 31 aa  | hypothetical protein   |
| CJB111.ORF02207  | 230 aa | surface protein Rib  |

|                 |        |  |
|-----------------|--------|--|
| CJB111.ORF02208 | 61 aa  | surface protein Rib                                      |
| CJB111.ORF02209 | 59 aa  | surface protein Rib                                      |
| CJB111.ORF02210 | 209 aa | surface protein Rib                                      |
| CJB111.ORF02211 | 221 aa | Leucine Rich Repeat domain protein                       |
| CJB111.ORF02212 | 56 aa  | hypothetical protein                                     |
| CJB111.ORF02213 | 248 aa | surface protein Rib                                      |
| CJB111.ORF02214 | 185 aa | lemA protein   |
| CJB111.ORF02216 | 124 aa | degenerate transposase (orf1) [imported]-related protein |
| CJB111.ORF02217 | 138 aa | IS1381, transposase OrfA/OrfB, truncation                |
| CJB111.ORF02218 | 255 aa | surface protein Rib                                      |
| CJB111.ORF02219 | 289 aa | rib protein  |
| CJB111.ORF02220 | 78 aa  | rib protein  |
| CJB111.ORF02221 | 125 aa | surface protein Rib                                      |
| CJB111.ORF02222 | 244 aa | surface protein Rib                                      |
| CJB111.ORF02223 | 35 aa  | hypothetical protein                                     |
| CJB111.ORF02224 | 94 aa  | polyribonucleotide nucleotidyltransferase                |
| CJB111.ORF02225 | 123 aa | rib protein  |
| CJB111.ORF02226 | 262 aa | surface protein Rib                                      |

## REGION 6.1 [1 sequence(s)]:

COH1 [0-17139]

|               |         |  |
|---------------|---------|--|
| COH1.ORF00001 | 225 aa  | protein of unknown function                    |
| COH1.ORF00002 | 249 aa  | MesJ/Ycf62 family protein                      |
| COH1.ORF00003 | 60 aa   | hypoxanthine-guanine phosphoribosyltransferase |
| COH1.ORF00004 | 43 aa   | hypothetical protein                           |
| COH1.ORF00005 | 116 aa  | hypoxanthine-guanine phosphoribosyltransferase |
| COH1.ORF00006 | 74 aa   | cell division protein FtsH                     |
| COH1.ORF00007 | 58 aa   | cell division protein FtsH [imported]          |
| COH1.ORF00008 | 183 aa  | cell division protein FtsH                     |
| COH1.ORF00009 | 229 aa  | type III restriction system methylase          |
| COH1.ORF00010 | 1456 aa | reticulocyte binding protein                   |
| COH1.ORF00011 | 30 aa   | type II restriction endonuclease, putative     |
| COH1.ORF00012 | 585 aa  | conserved hypothetical protein                 |
| COH1.ORF00013 | 664 aa  | conserved hypothetical protein                 |
| COH1.ORF00014 | 477 aa  | ATP-dependent DNA helicase recG                |

## REGION 6.2 [1 sequence(s)]:

COH1 [226148-234496]

|                |        |  |
|----------------|--------|--|
| COH1.ORF00245  | 161 aa | glutamyl-aminopeptidase                                    |
| COH1.ORF00246  | 94 aa  | glutamyl-aminopeptidase                                    |
| COH1.ORF00247  | 208 aa | tRNA binding domain protein                                |
| COH1.ORFA00238 | 67 aa  | Predicted S-adenosylmethionine-dependent methyltransferase |
| COH1.ORF00248  | 167 aa | tRNA binding domain protein                                |

|               |        |  |
|---------------|--------|--|
| COH1.ORF00249 | 89 aa  | conserved hypothetical protein                             |
| COH1.ORF00250 | 351 aa | oligopeptide ABC transporter, oligopeptide-binding protein |
| COH1.ORF00251 | 131 aa | oligopeptide ABC transporter, oligopeptide-binding protein |
| COH1.ORF00252 | 36 aa  | oligopeptide ABC transporter, oligopeptide-binding protein |
| COH1.ORF00253 | 190 aa | oligopeptide ABC transporter, permease protein             |
| COH1.ORF00254 | 0 aa   | oligopeptide ABC transporter, permease protein             |
| COH1.ORF00255 | 34 aa  | oligopeptide ABC transporter, permease protein             |
| COH1.ORF00256 | 138 aa | peptide ABC transporter, ATP-binding protein               |
| COH1.ORF00257 | 257 aa | peptide ABC transporter, ATP-binding protein               |

## REGION 6.3 [1 sequence(s)]:

COH1 [300911-306701]

|                |        |  |
|----------------|--------|--|
| COH1.ORFA00320 | 50 aa  | Hypothetical UPF0291 protein                         |
| COH1.ORFA00321 | 30 aa  | hypothetical protein                                 |
| COH1.ORF00333  | 148 aa | glycerol uptake facilitator protein                  |
| COH1.ORF00334  | 115 aa | oxidoreductase, pyridine nucleotide-disulfide family |
| COH1.ORF00335  | 252 aa | Unknown  |
| COH1.ORF00336  | 42 aa  | PTS system, IIBC components                          |
| COH1.ORF00337  | 304 aa | PTS system, IIBC components                          |
| COH1.ORF00338  | 254 aa | penicillin-binding protein 2X                        |
| COH1.ORF00339  | 159 aa | aminopeptidase C                                     |

## REGION 6.4 [1 sequence(s)]:

COH1 [536688-545262]

|                |        |   |
|----------------|--------|---|
| COH1.ORF00584  | 76 aa  | conserved hypothetical protein                  |
| COH1.ORF00585  | 77 aa  | Unknown   |
| COH1.ORFA00564 | 77 aa  | conserved hypothetical protein                  |
| COH1.ORFA00565 | 76 aa  | hypothetical protein                            |
| COH1.ORF00586  | 331 aa | ABC transporter, ATP-binding/permease protein   |
| COH1.ORF00587  | 192 aa | ABC transporter, ATP-binding/permease protein   |
| COH1.ORF00588  | 75 aa  | anthranilate synthase component II              |
| COH1.ORF00589  | 307 aa | UDP-N-acetylmuramoylalanine--D-glutamate ligase |
| COH1.ORF00590  | 218 aa | UDP-N-acetylmuramoylalanine--D-glutamate ligase |
| COH1.ORF00591  | 196 aa | isoleucyl-tRNA synthetase                       |
| COH1.ORF00592  | 250 aa | isoleucyl-tRNA synthetase                       |
| COH1.ORF00593  | 219 aa | isoleucyl-tRNA synthetase                       |
| COH1.ORF00594  | 127 aa | isoleucyl-tRNA synthetase                       |
| COH1.ORFA00575 | 47 aa  | conserved hypothetical protein                  |

## REGION 6.5 [1 sequence(s)]:

COH1 [1280188-1295761]

|               |       |                                |
|---------------|-------|--------------------------------|
| COH1.ORF01374 | 89 aa | conserved hypothetical protein |
| COH1.ORF01375 | 86 aa | conserved hypothetical protein |

|                |        |   |
|----------------|--------|---|
| COH1.ORF01376  | 113 aa | conserved hypothetical protein                    |
| COH1.ORF01377  | 55 aa  | transposase OrfA, IS3 family                      |
| COH1.ORF01378  | 37 aa  | ISSdy1, transposase OrfB                          |
| COH1.ORF01379  | 411 aa | Unknown   |
| COH1.ORF01380  | 400 aa | site-specific recombinase, phage integrase family |
| COH1.ORF01381  | 99 aa  | Unknown-related protein                           |
| COH1.ORF01383  | 289 aa | hypothetical protein                              |
| COH1.ORF01382  | 48 aa  | hypothetical protein                              |
| COH1.ORF01384  | 481 aa | hypothetical protein                              |
| COH1.ORF01385  | 141 aa | conserved hypothetical protein                    |
| COH1.ORF01386  | 248 aa | hypothetical protein                              |
| COH1.ORF01387  | 398 aa | putative serine/threonine protein kinase          |
| COH1.ORF01388  | 269 aa | type II restriction enzyme eco47ii                |
| COH1.ORF01389  | 333 aa | C-5 cytosine-specific DNA methylase               |
| COH1.ORFA01341 | 33 aa  | hypothetical protein                              |
| COH1.ORF01391  | 150 aa | conserved hypothetical protein                    |
| COH1.ORF01390  | 121 aa | ribosomal protein L7/L12                          |
| COH1.ORF01392  | 166 aa | ribosomal protein L10                             |
| COH1.ORF01394  | 306 aa | ATP-dependent Clp protease, ATP-binding subunit   |

## REGION 6.6 [1 sequence(s)]:

COH1 [1932369-1983223]

|                |        |   |
|----------------|--------|---|
| COH1.ORF02099  | 56 aa  | hypothetical protein                                      |
| COH1.ORF02100  | 373 aa | site-specific recombinase, phage integrase family         |
| COH1.ORF02101  | 108 aa | Domain of unknown function (DUF771) superfamily           |
| COH1.ORF02102  | 210 aa | conserved hypothetical protein                            |
| COH1.ORF02103  | 263 aa | surface protein PrgC                                      |
| COH1.ORF02105  | 113 aa | conserved hypothetical protein                            |
| COH1.ORF02106  | 209 aa | phosphate ABC transporter, permease protein               |
| COH1.ORF02107  | 72 aa  | phosphate ABC transporter, permease protein               |
| COH1.ORF02108  | 251 aa | probable hemolysin precursor                              |
| COH1.ORF02109  | 129 aa | transcriptional regulator, MerR family                    |
| COH1.ORFA02026 | 129 aa | transcription regulator, MerR family , putative           |
| COH1.ORF02110  | 86 aa  | acetyltransferase, GNAT family                            |
| COH1.ORF02111  | 182 aa | conserved hypothetical protein                            |
| COH1.ORF02112  | 457 aa | CHAP domain family  |
| COH1.ORF02113  | 666 aa | membrane protein, putative                                |
| COH1.ORF02114  | 756 aa | conjugal transfer protein, interruption-C                 |
| COH1.ORF02115  | 66 aa  | conjugal transfer protein, interruption-N-related protein |
| COH1.ORF02116  | 87 aa  | conserved hypothetical protein                            |
| COH1.ORF02117  | 88 aa  | conserved hypothetical protein                            |
| COH1.ORF02118  | 317 aa | conserved hypothetical protein                            |
| COH1.ORF02119  | 74 aa  | conserved hypothetical protein                            |

|                |        |  |
|----------------|--------|--|
| COH1.ORF02120  | 90 aa  | Uncharacterized ACR, COG2161 subfamily                                   |
| COH1.ORF02121  | 136 aa | conserved hypothetical protein   |
| COH1.ORF02122  | 116 aa | conserved hypothetical protein   |
| COH1.ORF02123  | 90 aa  | conserved hypothetical protein   |
| COH1.ORF02124  | 267 aa | conserved hypothetical protein   |
| COH1.ORFA02042 | 62 aa  | transcriptional regulator, Cro/CI family                                 |
| COH1.ORF02125  | 278 aa | conserved hypothetical protein   |
| COH1.ORF02126  | 99 aa  | transcriptional regulator, Cro/CI family                                 |
| COH1.ORF02127  | 125 aa | conserved hypothetical protein   |
| COH1.ORF02128  | 410 aa | transcriptional regulator, Cro/CI family                                 |
| COH1.ORF02129  | 553 aa | FtsK/SpoIIIE family protein  |
| COH1.ORF02130  | 153 aa | conserved hypothetical protein   |
| COH1.ORF02131  | 98 aa  | conserved hypothetical protein   |
| COH1.ORF02132  | 863 aa | cell wall surface anchor family protein                                  |
| COH1.ORF02133  | 753 aa | Unknown  |
| COH1.ORF02134  | 87 aa  | conserved hypothetical protein   |
| COH1.ORF02135  | 153 aa | Unknown  |
| COH1.ORF02136  | 103 aa | conserved hypothetical protein   |
| COH1.ORF02137  | 143 aa | transcriptional regulator, Cro/CI family                                 |
| COH1.ORF02138  | 104 aa | transcriptional regulator, Cro/CI family                                 |
| COH1.ORF02139  | 50 aa  | hypothetical protein   |
| COH1.ORF02140  | 112 aa | transcriptional regulator, Cro/CI family                                 |
| COH1.ORF02141  | 140 aa | Unknown  |
| COH1.ORF02142  | 122 aa | lipoprotein, putative  |
| COH1.ORF02143  | 112 aa | Unknown  |
| COH1.ORF02144  | 97 aa  | DNA-damage-inducible protein J, putative                                 |
| COH1.ORF02145  | 252 aa | CAMP factor  |
| COH1.ORF02146  | 435 aa | transposase, ISL3 family   |
| COH1.ORF02147  | 546 aa | carbamoyl-phosphate synthase, large subunit                              |
| COH1.ORF02148  | 130 aa | mercuric resistance operon regulatory protein MerR                       |
| COH1.ORF02149  | 522 aa | Mn <sup>2+</sup> /Fe <sup>2+</sup> transporter, NRAMP family             |
| COH1.ORF02150  | 240 aa | membrane protein, putative   |
| COH1.ORF02151  | 205 aa | ABC transporter, ATP-binding protein                                     |
| COH1.ORF02152  | 52 aa  | conserved hypothetical protein   |
| COH1.ORF02153  | 249 aa | Sterol-regulatory element binding protein (SREBP) site 2 protease family |
| COH1.ORF02154  | 315 aa | radical SAM domain protein protein                                       |
| COH1.ORF02155  | 101 aa | Unknown  |
| COH1.ORF02156  | 133 aa | Unknown  |
| COH1.ORF02157  | 405 aa | Unknown  |
| COH1.ORF02158  | 128 aa | protein of unknown function, putative                                    |
| COH1.ORF02159  | 169 aa | Unknown  |
| COH1.ORF02160  | 110 aa | Unknown  |

COH1.ORF02161 117 aa Unknown

**REGION 6.7 [1 sequence(s)]:**

COH1 [2135577-2235585]

|                    |         |  |
|--------------------|---------|--|
| COH1.ORF02341      | 414 aa  | Peptidase M16 inactive domain family                             |
| COH1.ORF02342      | 197 aa  | magnesium transporter, CorA family                               |
| COH1.ORF02343      | 319 aa  | oxidoreductase, Gfo/Idh/MocA family                              |
| COH1.ORF02344      | 473 aa  | ATP-dependent RNA helicase SrmB, putative                        |
| COH1.ORF02345      | 259 aa  | conserved hypothetical protein                                   |
| COH1.ORF02346      | 147 aa  | valyl-tRNA synthetase  |
| COH1.ORF02347      | 207 aa  | NOL1/NOP2/sun family protein                                     |
| COH1.ORF02348      | 300 aa  | rib protein  |
| COH1.ORF02349      | 280 aa  | rib protein  |
| COH1.ORF02350      | 91 aa   | hypothetical protein   |
| COH1.ORF02351      | 1385 aa | reticulocyte binding protein                                     |
| COH1.ORF02352      | 353 aa  | LPXTG-motif cell wall anchor domain protein                      |
| COH1.ORF02353      | 174 aa  | MutT/nudix family protein [imported]                             |
| COH1.ORF02354      | 264 aa  | IS861, transposase OrfB  |
| COH1.ORF02355      | 172 aa  | transcriptional regulator, GntR family/potassium uptake protein, |
| <b>TrkA family</b> |         |  |
| COH1.ORF02356      | 203 aa  | Unknown  |
| COH1.ORF02357      | 66 aa   | Unknown  |
| COH1.ORFA02262     | 50 aa   | L-serine dehydratase, iron-sulfur-dependent, alpha subunit       |
| COH1.ORF02358      | 110 aa  | Tn5252, Orf 9 protein [imported]                                 |
| COH1.ORF02359      | 119 aa  | Tn5252, Orf 9 protein  |
| COH1.ORF02360      | 279 aa  | protein gp49 [Bacteriophage A118] homolog lmo2317 [imported]     |
| COH1.ORF02361      | 93 aa   | conserved hypothetical protein                                   |
| COH1.ORF02362      | 172 aa  | transcriptional regulator, Cro/CI family, putative               |
| COH1.ORF02363      | 266 aa  | site-specific recombinase, phage integrase family                |
| COH1.ORF02364      | 39 aa   | glycosyl transferase, group 1 family protein                     |
| COH1.ORF02365      | 89 aa   | sugar transporter, putative                                      |
| COH1.ORFA02270     | 51 aa   | alkylphosphonate utilization operon protein PhnA                 |
| COH1.ORF02366      | 154 aa  | transposase, IS30 family, putative, truncation                   |
| COH1.ORF02367      | 47 aa   | ribosomal protein S5   |
| COH1.ORF02368      | 109 aa  | amino acid ABC transporter, ATP-binding protein                  |
| COH1.ORF02369      | 46 aa   | pullulanase, putative  |
| COH1.ORF02370      | 74 aa   | conserved hypothetical protein                                   |
| COH1.ORF02371      | 396 aa  | ABC transporter, ATP-binding protein                             |
| COH1.ORF02372      | 32 aa   | hypothetical protein   |
| COH1.ORF02373      | 56 aa   | xylulose-5-phosphate/fructose-6-phosphate phosphoketolase        |
| COH1.ORF02374      | 226 aa  | rib protein  |
| COH1.ORF02375      | 245 aa  | Putative transposable element                                    |
| COH1.ORF02376      | 51 aa   | acetolactate synthase, catabolic                                 |

|                |        |   |
|----------------|--------|---|
| COH1.ORF02377  | 66 aa  | NifR3/Smm1 family protein                                     |
| COH1.ORFA02281 | 30 aa  | hypothetical protein  |
| COH1.ORF02378  | 383 aa | threonine synthase  |
| COH1.ORFA02283 | 39 aa  | threonine synthase (EC 4.2.3.1)                               |
| COH1.ORFA02284 | 49 aa  | hypothetical protein  |
| COH1.ORF02379  | 76 aa  | conserved hypothetical protein                                |
| COH1.ORF02380  | 36 aa  | hypothetical protein  |
| COH1.ORF02381  | 56 aa  | oxidoreductase, short chain dehydrogenase/reductase family    |
| COH1.ORF02382  | 142 aa | oxidoreductase, short chain dehydrogenase/reductase family    |
| COH1.ORF02383  | 110 aa | transposase, IS30 family, putative, truncation                |
| COH1.ORF02384  | 226 aa | transposase, IS30 family, putative, truncation                |
| COH1.ORF02385  | 62 aa  | membrane protein, putative                                    |
| COH1.ORF02386  | 68 aa  | conserved hypothetical protein                                |
| COH1.ORFA02292 | 57 aa  | recA protein  |
| COH1.ORF02387  | 281 aa | rib protein   |
| COH1.ORF02388  | 73 aa  | DegV family protein   |
| COH1.ORF02389  | 69 aa  | cold shock protein, CSD family-related protein                |
| COH1.ORF02390  | 51 aa  | hypothetical protein  |
| COH1.ORF02391  | 180 aa | Putative TNP-like transposable element                        |
| COH1.ORF02392  | 252 aa | branched-chain amino acid transport system II carrier protein |
| COH1.ORF02393  | 168 aa | IS861, transposase OrfB                                       |
| COH1.ORF02394  | 32 aa  | large conductance mechanosensitive channel protein            |
| COH1.ORFA02302 | 41 aa  | peptide methionine sulfoxide reductase                        |
| COH1.ORF02395  | 54 aa  | peptide methionine sulfoxide reductase                        |
| COH1.ORF02396  | 204 aa | rib protein   |
| COH1.ORF02398  | 218 aa | rib protein   |
| COH1.ORF02399  | 125 aa | permease, putative  |
| COH1.ORF02400  | 54 aa  | amidophosphoribosyltransferase                                |
| COH1.ORF02401  | 117 aa | unnamed protein product                                       |
| COH1.ORF02402  | 207 aa | Unknown   |
| COH1.ORF02403  | 70 aa  | cell wall surface anchor family protein                       |
| COH1.ORF02404  | 240 aa | phosphoribosylformylglycinamide synthase, putative            |
| COH1.ORF02405  | 77 aa  | sensor histidine kinase                                       |
| COH1.ORF02406  | 99 aa  | cylB protein  |
| COH1.ORF02407  | 55 aa  | DNA-directed RNA polymerase beta' subunit                     |
| COH1.ORF02408  | 220 aa | Unknown protein   |
| COH1.ORF02409  | 266 aa | phosphoribosylformylglycinamide synthase I                    |
| COH1.ORF02410  | 156 aa | transcriptional regulator, Cro/CI family                      |
| COH1.ORF02411  | 206 aa | BRO family, N-terminal domain protein                         |
| COH1.ORF02412  | 70 aa  | conserved hypothetical protein                                |
| COH1.ORF02413  | 54 aa  | cytochrome d oxidase, subunit I                               |
| COH1.ORF02414  | 75 aa  | glycerol dehydrogenase  |
| COH1.ORF02415  | 53 aa  | oxidoreductase, aldo/keto reductase family                    |

|                |        |  |
|----------------|--------|--|
| COH1.ORFA02324 | 62 aa  | GMP synthase   |
| COH1.ORF02416  | 73 aa  | oxidoreductase, FMN-binding  |
| COH1.ORF02417  | 267 aa | rib protein  |
| COH1.ORFA02327 | 30 aa  | alpha-like protein 3   |
| COH1.ORF02418  | 50 aa  | cylJ protein   |
| COH1.ORF02419  | 52 aa  | cylI protein   |
| COH1.ORFA02329 | 32 aa  | hypothetical protein   |
| COH1.ORF02420  | 50 aa  | peptidase, M20/M25/M40 family  |
| COH1.ORF02421  | 60 aa  | protein of unknown function  |
| COH1.ORF02422  | 74 aa  | sortase family protein   |
| COH1.ORF02423  | 31 aa  | hypothetical protein   |
| COH1.ORFA02334 | 49 aa  | Glycerol dehydrogenase   |
| COH1.ORF02424  | 106 aa | mercuric resistance operon regulatory protein MerR                     |
| COH1.ORF02425  | 66 aa  | conserved hypothetical protein   |
| COH1.ORF02426  | 200 aa | transposase, IS30 family, putative, truncation                         |
| COH1.ORF02427  | 213 aa | transcriptional regulator, Cro/CI family                               |
| COH1.ORFA02338 | 62 aa  | transposase, IS30 family, putative, truncation                         |
| COH1.ORF02428  | 402 aa | phosphoribosylformylglycinamide synthase, putative                     |
| COH1.ORFA02340 | 56 aa  | signal recognition particle protein                                    |
| COH1.ORF02429  | 237 aa | rib protein  |
| COH1.ORF02430  | 269 aa | surface protein Rib  |
| COH1.ORF02431  | 170 aa | rib protein  |
| COH1.ORF02432  | 335 aa | rib protein  |
| COH1.ORF02433  | 212 aa | rib protein  |
| COH1.ORFA02346 | 123 aa | isochorismatase family protein   |
| COH1.ORF02434  | 54 aa  | conserved hypothetical protein   |
| COH1.ORF02435  | 253 aa | Unknown  |
| COH1.ORF02436  | 162 aa | phosphoribosylamine--glycine ligase                                    |
| COH1.ORF02437  | 72 aa  | conserved hypothetical protein   |
| COH1.ORF02438  | 51 aa  | D-mannose oxidoreductase   |
| COH1.ORF02439  | 155 aa | putative fimbria-associated protein                                    |
| COH1.ORF02440  | 209 aa | ornithine carbamoyltransferase   |
| COH1.ORFA02354 | 63 aa  | Actin  |
| COH1.ORF02442  | 84 aa  | capsule biosynthesis protein CapA,                                     |
| COH1.ORF02444  | 66 aa  | amidophosphoribosyltransferase   |
| COH1.ORF02445  | 90 aa  | sugar ABC transporter, sugar-binding protein                           |
| COH1.ORF02446  | 143 aa | N-acetylmannosamine-6-phosphate epimerase, putative                    |
| COH1.ORF02447  | 240 aa | preprotein translocase SecY family protein [imported]                  |
| COH1.ORF02448  | 225 aa | membrane protein, putative   |
| COH1.ORF02449  | 294 aa | rib protein  |
| COH1.ORFA02362 | 47 aa  | Single-strand binding protein 1 (SSB 1) (Helix-destabilizing protein1) |
| COH1.ORF02450  | 79 aa  | ribosomal protein S18  |
| COH1.ORF02451  | 169 aa | pullulanase, putative  |

|                |        |   |
|----------------|--------|---|
| COH1.ORF02452  | 158 aa | sensor histidine kinase                               |
| COH1.ORF02453  | 50 aa  | glycosyl hydrolase, family 3                          |
| COH1.ORF02454  | 71 aa  | competence protein F                                  |
| COH1.ORF02455  | 326 aa | hypothetical protein                                  |
| COH1.ORF02456  | 318 aa | site-specific recombinase, phage integrase family     |
| COH1.ORF02457  | 328 aa | rib protein   |
| COH1.ORF02458  | 263 aa | Unknown   |
| COH1.ORF02459  | 118 aa | restriction endonuclease related protein              |
| COH1.ORF02460  | 77 aa  | restriction endonuclease related protein              |
| COH1.ORF02461  | 30 aa  | type II restriction endonuclease, putative            |
| COH1.ORF02462  | 54 aa  | conserved hypothetical protein                        |
| COH1.ORF02463  | 290 aa | unnamed protein product                               |
| COH1.ORFA02375 | 46 aa  | conserved hypothetical protein                        |
| COH1.ORF02464  | 107 aa | amino acid ABC transporter, ATP-binding protein       |
| COH1.ORF02466  | 338 aa | DNA polymerase III, alpha subunit, Gram-positive type |
| COH1.ORF02467  | 308 aa | rib protein   |
| COH1.ORF02468  | 333 aa | conserved hypothetical protein                        |
| COH1.ORF02469  | 267 aa | phosphoribosylamine--glycine ligase                   |
| COH1.ORF02470  | 101 aa | ribosomal protein S1                                  |
| COH1.ORF02471  | 37 aa  | hypothetical protein                                  |
| COH1.ORF02472  | 60 aa  | transport ATP-binding protein coma                    |
| COH1.ORF02473  | 177 aa | ABC transporter, ATP-binding protein                  |
| COH1.ORF02474  | 133 aa | ABC transporter, substrate-binding protein            |

## REGION 7.1 [1 sequence(s)]:

H36B [0-12658]

|               |        |   |
|---------------|--------|---|
| H36B.ORF00001 | 304 aa | oligopeptide ABC transporter, permease protein    |
| H36B.ORF00002 | 343 aa | oligopeptide ABC transporter, permease protein    |
| H36B.ORF00003 | 348 aa | oligopeptide ABC transporter, ATP-binding protein |
| H36B.ORF00004 | 310 aa | oligopeptide ABC transporter, ATP-binding protein |

## REGION 7.2 [1 sequence(s)]:

H36B [50290-55433]

|                |        |  |
|----------------|--------|--|
| H36B.ORF00041  | 62 aa  | MORN motif family protein                                  |
| H36B.ORFA00041 | 53 aa  | hypothetical protein                                       |
| H36B.ORF00042  | 95 aa  | phosphoribosylaminoimidazole-succinocarboxamide synthase   |
| H36B.ORF00043  | 111 aa | phosphoribosylaminoimidazole-succinocarboxamide synthase   |
| H36B.ORF00044  | 61 aa  | phosphoribosylformylglycinamidine synthase, putative       |
| H36B.ORF00045  | 112 aa | Phosphoribosylglycinamide synthetase, ATP-grasp (A) domain |
| H36B.ORF00046  | 63 aa  | phosphoribosylamine--glycine ligase                        |

## REGION 7.3 [1 sequence(s)]:

H36B [226513-232564]

|               |        |   |
|---------------|--------|---|
| H36B.ORF00246 | 143 aa | alpha amylase family protein                |
| H36B.ORF00247 | 229 aa | polyribonucleotide nucleotidyltransferase   |
| H36B.ORF00248 | 178 aa | RNA methyltransferase, TrmH family, group 3 |

## REGION 7.4 [1 sequence(s)]:

H36B [427585-461812]

|                |        |   |
|----------------|--------|---|
| H36B.ORFA00453 | 31 aa  | C protein alpha-antigen precursor                             |
| H36B.ORF00471  | 218 aa | serine protease, subtilase family                             |
| H36B.ORF00472  | 48 aa  | protease, putative  |
| H36B.ORF00473  | 137 aa | glycosyl transferase, group 2 family protein                  |
| H36B.ORF00474  | 34 aa  | hypothetical protein  |
| H36B.ORF00475  | 223 aa | c protein alpha-antigen precursor                             |
| H36B.ORF00476  | 108 aa | site-specific recombinase, phage integrase family, truncation |
| H36B.ORF00477  | 45 aa  | site-specific recombinase, phage integrase family, truncation |
| H36B.ORF00478  | 101 aa | Unknown   |
| H36B.ORF00479  | 112 aa | Unknown   |
| H36B.ORF00480  | 254 aa | methyltransferase   |
| H36B.ORF00481  | 686 aa | MW0309, putative  |
| H36B.ORF00482  | 149 aa | MW, putative  |
| H36B.ORF00483  | 92 aa  | GatB  |
| H36B.ORF00484  | 386 aa | PTS system, IIC component, putative                           |
| H36B.ORF00485  | 39 aa  | hypothetical protein  |
| H36B.ORF00487  | 276 aa | rhamnulose-1-phosphate aldolase                               |
| H36B.ORF00488  | 154 aa | MW  |
| H36B.ORF00489  | 443 aa | PTS system component  |
| H36B.ORF00490  | 94 aa  | galactitol PTS, EIIB  |
| H36B.ORF00491  | 276 aa | msm operon regulatory protein [imported], putative            |
| H36B.ORF00492  | 425 aa | sugar ABC transporter, sugar-binding protein, putative        |
| H36B.ORF00493  | 296 aa | ABC transporter, permease protein                             |
| H36B.ORF00494  | 277 aa | sugar ABC transporter, permease protein                       |
| H36B.ORF00495  | 733 aa | alpha-galactosidase   |
| H36B.ORF00496  | 390 aa | galactokinase   |
| H36B.ORF00497  | 494 aa | galactose-1-phosphate uridylyltransferase                     |
| H36B.ORF00498  | 331 aa | UDP-glucose 4-epimerase                                       |
| H36B.ORF00499  | 230 aa | aldose 1-epimerase  |
| H36B.ORF00500  | 64 aa  | aldose 1-epimerase  |
| H36B.ORF00501  | 84 aa  | conserved hypothetical protein                                |
| H36B.ORF00502  | 103 aa | Unknown   |
| H36B.ORF00503  | 192 aa | acetyltransferase, GNAT family                                |
| H36B.ORF00504  | 216 aa | hypothetical protein  |
| H36B.ORF00505  | 883 aa | valyl-tRNA synthetase   |
| H36B.ORF00506  | 319 aa | oxidoreductase, Gfo/Idh/MocA family                           |
| H36B.ORF00507  | 55 aa  | magnesium transporter, CorA family                            |

|               |        |   |
|---------------|--------|---|
| H36B.ORF00508 | 218 aa | magnesium transporter, CorA family          |
| H36B.ORF00509 | 317 aa | transposase, Mutator family                 |
| H36B.ORF00510 | 350 aa | Protein of unknown function (DUF419) family |

## REGION 7.5 [1 sequence(s)]:

H36B [515790-558957]

|                |        |   |
|----------------|--------|---|
| H36B.ORF00575  | 91 aa  | DNA-binding protein HU  |
| H36B.ORF00576  | 380 aa | integrase-like protein  |
| H36B.ORF00577  | 183 aa | conserved hypothetical protein                                    |
| H36B.ORF00578  | 128 aa | conserved hypothetical protein                                    |
| H36B.ORF00579  | 139 aa | transcriptional regulator, Cro/CI family                          |
| H36B.ORF00580  | 81 aa  | hypothetical protein  |
| H36B.ORF00581  | 262 aa | hypothetical protein  |
| H36B.ORF00582  | 63 aa  | conserved hypothetical protein                                    |
| H36B.ORF00583  | 103 aa | prophage ps2 protein 07, excisionase [imported]                   |
| H36B.ORF00584  | 49 aa  | conserved hypothetical protein                                    |
| H36B.ORF00585  | 75 aa  | conserved hypothetical protein                                    |
| H36B.ORF00586  | 79 aa  | hypothetical protein  |
| H36B.ORFA00560 | 42 aa  | hypothetical protein  |
| H36B.ORF00588  | 439 aa | chromosome segregation protein, putative                          |
| H36B.ORF00589  | 357 aa | conserved hypothetical protein                                    |
| H36B.ORF00590  | 201 aa | conserved hypothetical protein                                    |
| H36B.ORF00591  | 202 aa | conserved hypothetical protein                                    |
| H36B.ORF00592  | 530 aa | helicase, putative  |
| H36B.ORF00593  | 65 aa  | conserved hypothetical protein                                    |
| H36B.ORF00594  | 83 aa  | hypothetical protein  |
| H36B.ORF00595  | 765 aa | conserved hypothetical protein                                    |
| H36B.ORF00596  | 93 aa  | RNA polymerase sigma factor                                       |
| H36B.ORF00597  | 135 aa | Endodeoxyribonuclease RusA superfamily                            |
| H36B.ORF00598  | 71 aa  | conserved hypothetical protein                                    |
| H36B.ORF00599  | 104 aa | hypothetical protein  |
| H36B.ORF00600  | 110 aa | conserved hypothetical protein                                    |
| H36B.ORF00601  | 137 aa | putative transcriptional activator - phage associated             |
| H36B.ORF00602  | 143 aa | terminase small subunit   |
| H36B.ORF00603  | 426 aa | phage terminase, large subunit, PBSX family                       |
| H36B.ORF00604  | 509 aa | phage portal protein, SPP1 family                                 |
| H36B.ORF00605  | 480 aa | phage putative head morphogenesis protein, SPP1 gp7 family domain |
| protein        |        |   |
| H36B.ORF00606  | 67 aa  | conserved hypothetical protein                                    |
| H36B.ORF00607  | 189 aa | conserved hypothetical protein                                    |
| H36B.ORF00608  | 298 aa | conserved hypothetical protein                                    |
| H36B.ORF00609  | 118 aa | conserved hypothetical protein                                    |
| H36B.ORF00610  | 108 aa | phage protein, HK97 gp10 family                                   |

|                |         |  |
|----------------|---------|--|
| H36B.ORF00611  | 119 aa  | conserved hypothetical protein                                   |
| H36B.ORF00612  | 210 aa  | conserved hypothetical protein                                   |
| H36B.ORF00613  | 151 aa  | conserved hypothetical protein                                   |
| H36B.ORF00614  | 76 aa   | conserved hypothetical protein                                   |
| H36B.ORF00616  | 1408 aa | reticulocyte binding protein                                     |
| H36B.ORFA00587 | 67 aa   | hypothetical protein   |
| H36B.ORF00617  | 280 aa  | conserved hypothetical protein                                   |
| H36B.ORF00618  | 1134 aa | reticulocyte binding protein                                     |
| H36B.ORFA00591 | 35 aa   | hypothetical protein   |
| H36B.ORF00619  | 817 aa  | prophage LambdaSa1, N-acetylmuramoyl-L-alanine amidase, family 4 |
| H36B.ORF00620  | 50 aa   | hypothetical phage protein-related protein                       |
| H36B.ORF00621  | 138 aa  | conserved hypothetical protein                                   |
| H36B.ORF00622  | 74 aa   | hypothetical protein   |
| H36B.ORF00623  | 98 aa   | conserved hypothetical protein                                   |
| H36B.ORF00624  | 111 aa  | phage holin, LL-H family   |
| H36B.ORF00625  | 248 aa  | endolysin domain protein   |
| H36B.ORF00626  | 37 aa   | hypothetical protein   |
| H36B.ORF00627  | 83 aa   | conserved hypothetical protein                                   |
| H36B.ORF00628  | 69 aa   | cI-like repressor  |
| H36B.ORF00629  | 274 aa  | hypothetical protein   |
| H36B.ORF00630  | 181 aa  | pyrogenic exotoxin C precursor, phage associated, putative       |
| H36B.ORF00631  | 58 aa   | putative exotoxin precursor                                      |
| H36B.ORF00632  | 59 aa   | conserved hypothetical protein                                   |

## REGION 7.6 [1 sequence(s)]:

H36B [608015-616671]

|               |        |   |
|---------------|--------|---|
| H36B.ORF00687 | 38 aa  | A/G-specific adenine glycosylase                                  |
| H36B.ORF00688 | 95 aa  | ribosomal protein S6  |
| H36B.ORF00689 | 163 aa | single-strand binding protein                                     |
| H36B.ORF00690 | 439 aa | prophage LambdaSa1, reverse transcriptase/maturase family protein |
| H36B.ORF00691 | 79 aa  | ribosomal protein S18   |
| H36B.ORF00692 | 314 aa | magnesium transporter, CorA family                                |
| H36B.ORF00693 | 223 aa | conserved hypothetical protein                                    |
| H36B.ORF00694 | 226 aa | excinuclease ABC, A subunit                                       |
| H36B.ORF00695 | 175 aa | sensor histidine kinase VncS                                      |
| H36B.ORF00696 | 186 aa | transposase OrfB, IS3 family, truncation                          |
| H36B.ORF00697 | 66 aa  | CsbD-like superfamily   |

## REGION 7.7 [1 sequence(s)]:

H36B [1133255-1140047]

|               |       |                             |
|---------------|-------|-----------------------------|
| H36B.ORF01241 | 91 aa | Unknown                     |
| H36B.ORF01242 | 55 aa | repressor protein, putative |
| H36B.ORF01243 | 52 aa | repressor protein, putative |

|                |        |  |
|----------------|--------|--|
| H36B.ORF01244  | 208 aa | prophage pi3 protein 59 [imported]                             |
| H36B.ORF01245  | 263 aa | prophage LambdaSa2, site-specific recombinase, phage integrase |
| family         |        |  |
| H36B.ORF01246  | 358 aa | ribosomal protein S1   |
| H36B.ORFA01211 | 32 aa  | hypothetical protein   |
| H36B.ORF01247  | 76 aa  | conserved hypothetical protein                                 |
| H36B.ORF01248  | 89 aa  | branched-chain amino acid aminotransferase                     |
| H36B.ORF01249  | 232 aa | branched-chain amino acid aminotransferase                     |

## REGION 7.8 [1 sequence(s)]:

H36B [1910893-1923194]

|                |        |  |
|----------------|--------|--|
| H36B.ORF02058  | 169 aa | Unknown                                    |
| H36B.ORF02059  | 92 aa  | conserved hypothetical protein             |
| H36B.ORF02060  | 128 aa | protein of unknown function, putative      |
| H36B.ORF02061  | 190 aa | Unknown                                    |
| H36B.ORF02062  | 110 aa | Unknown                                    |
| H36B.ORF02063  | 42 aa  | Unknown                                    |
| H36B.ORF02064  | 96 aa  | Unknown                                    |
| H36B.ORF02065  | 48 aa  | hypothetical protein                       |
| H36B.ORF02066  | 102 aa | response regulator                         |
| H36B.ORF02067  | 171 aa | sugar ABC transporter, ATP-binding protein |
| H36B.ORF02068  | 181 aa | GTP pyrophosphokinase family protein       |
| H36B.ORF02069  | 46 aa  | Holliday junction DNA helicase RuvA        |
| H36B.ORF02070  | 129 aa | Transposase (IS4 family)                   |
| H36B.ORF02071  | 127 aa | degenerate transposase (orf1) [imported]   |
| H36B.ORF02072  | 45 aa  | hypothetical protein                       |
| H36B.ORF02073  | 205 aa | ABC transporter, ATP-binding protein       |
| H36B.ORF02074  | 36 aa  | conserved hypothetical protein             |
| H36B.ORF02075  | 204 aa | ABC transporter, ATP-binding protein       |
| H36B.ORFA02003 | 28 aa  | hypothetical protein                       |
| H36B.ORFA02004 | 47 aa  | conserved hypothetical protein             |

## REGION 7.9 [1 sequence(s)]:

H36B [1990820-1998923]

|               |        |  |
|---------------|--------|--|
| H36B.ORF02142 | 419 aa | competence/damage-inducible protein CinA                   |
| H36B.ORF02143 | 112 aa | DNA-3-methyladenine glycosylase I                          |
| H36B.ORF02144 | 127 aa | degenerate transposase (orf1) [imported]                   |
| H36B.ORF02145 | 129 aa | Transposase (IS4 family)                                   |
| H36B.ORF02146 | 217 aa | response regulator [imported]                              |
| H36B.ORF02147 | 423 aa | sensor histidine kinase [imported]                         |
| H36B.ORF02148 | 211 aa | Isoprenylcysteine carboxyl methyltransferase (ICMT) family |
| H36B.ORF02149 | 308 aa | iga fc receptor precursor                                  |
| H36B.ORF02150 | 58 aa  | DNA mismatch repair protein HexA                           |

|               |        |                                  |
|---------------|--------|----------------------------------|
| H36B.ORF02151 | 141 aa | DNA mismatch repair protein HexA |
| H36B.ORF02152 | 113 aa | histidyl-tRNA synthetase         |
| H36B.ORF02153 | 60 aa  | ribosomal protein L32            |

## REGION 7.10 [1 sequence(s)]:

H36B [2131055-2229117]

|                |        |  |
|----------------|--------|--|
| H36B.ORF02300  | 137 aa | ABC transporter, ATP-binding protein                             |
| H36B.ORF02301  | 290 aa | transketolase  |
| H36B.ORF02302  | 214 aa | transaldolase, putative  |
| H36B.ORF02303  | 229 aa | ribulose-phosphate 3-epimerase                                   |
| H36B.ORF02304  | 367 aa | PTS system, IIC component [imported]                             |
| H36B.ORF02305  | 149 aa | PTS system, IIA component [imported]                             |
| H36B.ORF02306  | 103 aa | PTS system, IIB component [imported]                             |
| H36B.ORF02307  | 629 aa | MW0309, putative   |
| H36B.ORF02308  | 102 aa | transcription regulator, putative                                |
| H36B.ORF02309  | 149 aa | probable bud operon transcriptional regulator , fragment         |
| H36B.ORF02310  | 116 aa | oxidoreductase, short chain dehydrogenase/reductase family       |
| H36B.ORF02311  | 245 aa | oxidoreductase, aldo/keto reductase family                       |
| H36B.ORF02312  | 48 aa  | conserved hypothetical protein                                   |
| H36B.ORF02313  | 176 aa | Glutamate-rich protein grpB                                      |
| H36B.ORF02314  | 230 aa | phenazine biosynthesis protein PhzF family                       |
| H36B.ORF02315  | 116 aa | acetyltransferase, GNAT family family                            |
| H36B.ORF02316  | 69 aa  | conserved hypothetical protein                                   |
| H36B.ORF02317  | 124 aa | protein of unknown function                                      |
| H36B.ORF02318  | 41 aa  | hypothetical protein   |
| H36B.ORF02319  | 127 aa | protein of unknown function                                      |
| H36B.ORF02320  | 34 aa  | hypothetical protein   |
| H36B.ORF02321  | 114 aa | Unknown  |
| H36B.ORF02322  | 102 aa | Unknown  |
| H36B.ORF02323  | 84 aa  | DNA-damage-inducible protein J, putative                         |
| H36B.ORFA02244 | 48 aa  | conserved hypothetical protein                                   |
| H36B.ORF02324  | 79 aa  | Unknown  |
| H36B.ORF02325  | 133 aa | Unknown  |
| H36B.ORF02326  | 261 aa | c protein alpha-antigen precursor                                |
| H36B.ORF02327  | 258 aa | prophage LambdaSa1, N-acetylmuramoyl-L-alanine amidase, family 4 |
| H36B.ORF02328  | 47 aa  | c protein alpha-antigen precursor                                |
| H36B.ORF02329  | 396 aa | Unknown  |
| H36B.ORF02330  | 124 aa | exonuclease, putative  |
| H36B.ORF02331  | 171 aa | Unknown  |
| H36B.ORF02332  | 263 aa | fibrinogen-binding protein                                       |
| H36B.ORF02333  | 341 aa | aspartate-semialdehyde dehydrogenase                             |
| H36B.ORF02334  | 112 aa | Unknown  |
| H36B.ORF02335  | 172 aa | Unknown  |

|                |        |  |
|----------------|--------|--|
| H36B.ORF02336  | 393 aa | site-specific recombinase, phage integrase family                            |
| H36B.ORF02337  | 77 aa  | Unknown-related protein  |
| H36B.ORF02338  | 207 aa | Plasmid replication protein  |
| H36B.ORF02339  | 56 aa  | phosphoribosylaminoimidazolecarboxamide formyltransferase/IMP cyclohydrolase |
| H36B.ORFA02259 | 52 aa  | hypothetical protein   |
| H36B.ORFA02261 | 80 aa  | hypothetical protein   |
| H36B.ORFA02260 | 47 aa  | permease   |
| H36B.ORF02340  | 212 aa | BRO family, N-terminal domain protein  |
| H36B.ORF02341  | 57 aa  | conserved hypothetical protein   |
| H36B.ORFA02264 | 50 aa  | hypothetical protein   |
| H36B.ORF02342  | 154 aa | c protein alpha-antigen precursor  |
| H36B.ORF02343  | 50 aa  | c protein alpha-antigen precursor  |
| H36B.ORF02344  | 148 aa | amidase family protein   |
| H36B.ORF02345  | 236 aa | c protein alpha-antigen precursor  |
| H36B.ORF02346  | 176 aa | translation initiation factor IF-3   |
| H36B.ORF02347  | 66 aa  | ribosomal protein L35  |
| H36B.ORF02348  | 119 aa | ribosomal protein L20  |
| H36B.ORF02350  | 91 aa  | DNA polymerase I   |
| H36B.ORF02349  | 86 aa  | CoA-binding domain protein   |
| H36B.ORFA02274 | 44 aa  | hypothetical protein   |
| H36B.ORF02352  | 264 aa | DNA polymerase delta large chain   |
| H36B.ORF02356  | 97 aa  | cell wall surface anchor family protein                                      |
| H36B.ORF02367  | 209 aa | GDP dissociation inhibitor isoform 1; GDI-1                                  |
| H36B.ORFA02288 | 48 aa  | diarrheal toxin  |
| H36B.ORF02357  | 166 aa | c protein alpha-antigen precursor  |
| H36B.ORF02363  | 298 aa | sensor histidine kinase BlpH, probable [imported]                            |
| H36B.ORFA02283 | 33 aa  | conserved hypothetical protein   |
| H36B.ORF02364  | 153 aa | Unknown  |
| H36B.ORF02365  | 39 aa  | phosphoribosylformylglycinamidine synthase, putative                         |
| H36B.ORF02366  | 236 aa | condensin subunit  |
| H36B.ORF02359  | 46 aa  | SNF2 family protein  |
| H36B.ORF02361  | 224 aa | putative nucleoporin   |
| H36B.ORF02358  | 79 aa  | oligopeptide ABC transporter, ATP-binding protein                            |
| H36B.ORF02362  | 207 aa | FtsK/SpoIIIE family protein  |
| H36B.ORFA02296 | 55 aa  | conserved protein  |
| H36B.ORF02369  | 101 aa | Unknown  |
| H36B.ORF02371  | 170 aa | modulator of drug activity (mda66)   |
| H36B.ORF02377  | 77 aa  | hypothetical protein   |
| H36B.ORF02378  | 115 aa | type IIS restriction enzyme R and M protein (ECO57IR), putative              |
| H36B.ORFA02310 | 139 aa | undecaprenol kinase  |
| H36B.ORFA02312 | 30 aa  | threonine-tRNA ligase Cj0206   |
| H36B.ORF02383  | 109 aa | conserved hypothetical protein   |

|                |        |   |
|----------------|--------|---|
| H36B.ORFA02316 | 35 aa  | conserved hypothetical protein                |
| H36B.ORF02384  | 101 aa | Na+/H+ exchanger family protein               |
| H36B.ORFA02318 | 35 aa  | hypothetical protein                          |
| H36B.ORFA02319 | 37 aa  | ABC transporter, ATP-binding protein          |
| H36B.ORF02385  | 60 aa  | lipoprotein, putative                         |
| H36B.ORF02386  | 81 aa  | serine O-acetyltransferase                    |
| H36B.ORF02387  | 72 aa  | hypothetical protein                          |
| H36B.ORF02388  | 84 aa  | hypothetical protein                          |
| H36B.ORF02389  | 48 aa  | hypothetical protein                          |
| H36B.ORF02391  | 71 aa  | ABC transporter, ATP-binding protein          |
| H36B.ORF02392  | 67 aa  | membrane protein, putative                    |
| H36B.ORFA02327 | 41 aa  | hypothetical protein                          |
| H36B.ORF02393  | 159 aa | conserved hypothetical protein                |
| H36B.ORF02394  | 212 aa | aminotransferase, class I                     |
| H36B.ORFA02330 | 201 aa | hypothetical protein                          |
| H36B.ORFA02333 | 45 aa  | conserved hypothetical protein                |
| H36B.ORF02396  | 168 aa | hypothetical protein                          |
| H36B.ORF02398  | 64 aa  | conserved hypothetical protein                |
| H36B.ORF02397  | 79 aa  | hypothetical protein                          |
| H36B.ORF02401  | 115 aa | cell division protein FtsH                    |
| H36B.ORF02402  | 175 aa | endopeptidase O                               |
| H36B.ORF02403  | 35 aa  | conserved hypothetical protein                |
| H36B.ORFA02338 | 74 aa  | F27J15.23                                     |
| H36B.ORFA02341 | 41 aa  | hypothetical protein                          |
| H36B.ORF02405  | 37 aa  | phosphoribosylamine--glycine ligase           |
| H36B.ORFA02342 | 51 aa  | Tn5252, Orf23                                 |
| H36B.ORF02406  | 72 aa  | ABC transporter, ATP-binding/permease protein |
| H36B.ORF02407  | 97 aa  | RNA methyltransferase, TrmH family, group 3   |
| H36B.ORFA02345 | 41 aa  | conserved hypothetical protein                |
| H36B.ORF02408  | 56 aa  | GTP-binding protein                           |
| H36B.ORF02409  | 62 aa  | GTP-binding protein                           |
| H36B.ORFA02348 | 47 aa  | conserved hypothetical protein                |
| H36B.ORF02415  | 42 aa  | conserved hypothetical protein                |
| H36B.ORFA02356 | 65 aa  | Oxidoreductase NAD-binding domain protein     |
| H36B.ORF02413  | 48 aa  | ABC transporter, ATP-binding protein          |
| H36B.ORF02414  | 38 aa  | IS1381, transposase OrfA [imported]           |
| H36B.ORF02422  | 102 aa | membrane protein, putative                    |
| H36B.ORF02417  | 116 aa | c protein alpha-antigen precursor             |
| H36B.ORF02416  | 257 aa | ammonium transporter family protein           |
| H36B.ORF02421  | 60 aa  | transcription-repair coupling factor          |
| H36B.ORFA02367 | 50 aa  | transcriptiorepair coupling factor mfd        |
| H36B.ORF02419  | 133 aa | chromosome segregation SMC protein            |
| H36B.ORF02420  | 40 aa  | hypothetical protein                          |

|                |        |  |
|----------------|--------|--|
| H36B.ORF02418  | 114 aa | c protein alpha-antigen precursor                  |
| H36B.ORFA02366 | 47 aa  | ABC transporter permease protein                   |
| H36B.ORFA02352 | 52 aa  | sugar ABC transporter, permease protein            |
| H36B.ORF02411  | 236 aa | iga fc receptor precursor                          |
| H36B.ORFA02350 | 49 aa  | type I restriction-modification system, M subunit  |
| H36B.ORF02412  | 287 aa | c protein alpha-antigen precursor                  |
| H36B.ORFA02360 | 44 aa  | hypothetical protein                               |
| H36B.ORFA02361 | 48 aa  | hypothetical protein                               |
| H36B.ORFA02351 | 46 aa  | conserved hypothetical protein                     |
| H36B.ORF02423  | 181 aa | hypothetical protein                               |
| H36B.ORF02424  | 190 aa | surface protein Rib                                |
| H36B.ORF02426  | 43 aa  | Stp1   |
| H36B.ORFA02372 | 73 aa  | conserved hypothetical protein                     |
| H36B.ORF02427  | 53 aa  | virulence factor MviM, putative                    |
| H36B.ORFA02375 | 88 aa  | hydrogen peroxide-inducible protein hic-5          |
| H36B.ORF02428  | 46 aa  | T3B23.2/T3B23.2                                    |
| H36B.ORF02429  | 122 aa | AT3g55020/T15C9_20                                 |
| H36B.ORFA02377 | 46 aa  | conserved hypothetical protein                     |
| H36B.ORF02430  | 157 aa | C protein alpha antigen                            |
| H36B.ORFA02379 | 43 aa  | conserved hypothetical protein                     |
| H36B.ORFA02380 | 57 aa  | amino acid ABC transporter, ATP-binding protein    |
| H36B.ORFA02381 | 41 aa  | endolysin  |
| H36B.ORF02432  | 118 aa | conserved hypothetical protein                     |
| H36B.ORFA02383 | 47 aa  | Resolvase  |
| H36B.ORF02433  | 62 aa  | methyltransferase (ssoIM)                          |
| H36B.ORF02434  | 137 aa | hypothetical protein                               |
| H36B.ORF02435  | 112 aa | ribonuclease G                                     |
| H36B.ORF02436  | 55 aa  | Fe-S oxidoreductases family 2                      |
| H36B.ORF02437  | 38 aa  | hypothetical protein                               |
| H36B.ORF02438  | 51 aa  | hypothetical protein                               |
| H36B.ORF02440  | 207 aa | mercuric reductase                                 |
| H36B.ORF02444  | 189 aa | transposon tn1546 resolvase                        |
| H36B.ORF02445  | 48 aa  | hypothetical protein                               |
| H36B.ORF02446  | 169 aa | tRNA delta(2)-isopentenylpyrophosphate transferase |
| H36B.ORF02447  | 253 aa | hypothetical protein                               |
| H36B.ORF02448  | 257 aa | hypothetical protein                               |
| H36B.ORF02449  | 263 aa | transposase, Tn3 family, putative                  |
| H36B.ORF02454  | 187 aa | transposase, IS4 family, putative                  |
| H36B.ORF02455  | 78 aa  | hypothetical protein                               |
| H36B.ORF02456  | 203 aa | resolvase  |
| H36B.ORF02457  | 218 aa | c protein alpha-antigen precursor                  |

REGION 8.1 [1 sequence(s)]:

## 18RS21 [25397-35846]

|                 |        |              |
|-----------------|--------|--------------|
| 18RS21.ORF00023 | 112 aa | PcsB protein |
| 18RS21.ORF00024 | 31 aa  | PcsB protein |

## REGION 8.2 [1 sequence(s)]:

## 18RS21 [1873333-2164750]

|                 |        |   |
|-----------------|--------|---|
| 18RS21.ORF01961 | 134 aa | aspartate kinase family protein                               |
| 18RS21.ORF01962 | 48 aa  | exodeoxyribonuclease VII, small subunit                       |
| 18RS21.ORF01963 | 154 aa | exodeoxyribonuclease VII, large subunit                       |
| 18RS21.ORF01964 | 63 aa  | 82 kDa heat shock protein HSP82                               |
| 18RS21.ORF01965 | 227 aa | phage nucleotide-binding protein                              |
| 18RS21.ORF01966 | 454 aa | 51.5 kDa protein  |
| 18RS21.ORF01967 | 160 aa | 18K protein GP18C   |
| 18RS21.ORF01968 | 517 aa | bacteriophage resistance protein                              |
| 18RS21.ORF01969 | 289 aa | hypothetical phage protein                                    |
| 18RS21.ORF01970 | 93 aa  | gp106   |
| 18RS21.ORF01971 | 52 aa  | hypothetical protein  |
| 18RS21.ORF01972 | 127 aa | phage conserved hypothetical protein TIGR01671                |
| 18RS21.ORF01973 | 112 aa | hypothetical protein  |
| 18RS21.ORF01974 | 313 aa | ABC transporter, permease protein SP1919                      |
| 18RS21.ORF01975 | 167 aa | phosphoribosylaminoimidazole carboxylase, ATPase subunit      |
| 18RS21.ORF01976 | 41 aa  | conserved hypothetical protein                                |
| 18RS21.ORF01977 | 52 aa  | immunoreactive protein Se23.5                                 |
| 18RS21.ORF01978 | 149 aa | sugar ABC transporter, ATP-binding protein SP0846             |
| 18RS21.ORF01979 | 188 aa | ABC transporter, permease protein                             |
| 18RS21.ORF01980 | 255 aa | tagatose-6-phosphate kinase                                   |
| 18RS21.ORF01981 | 319 aa | pli0047   |
| 18RS21.ORF01982 | 281 aa | phosphoribosylformylglycinamide cyclo-ligase                  |
| 18RS21.ORF01983 | 63 aa  | amidophosphoribosyltransferase                                |
| 18RS21.ORF01984 | 41 aa  | hypothetical protein  |
| 18RS21.ORF01985 | 214 aa | D-xylulose 5-phosphate/D-fructose 6-phosphate phosphoketolase |
| 18RS21.ORF01987 | 51 aa  | xylulose-5-phosphate/fructose-6-phosphate phosphoketolase     |
| 18RS21.ORF01986 | 278 aa | polysaccharide deacetylase family protein                     |
| 18RS21.ORF01988 | 64 aa  | amino acid ABC transporter, ATP-binding protein               |
| 18RS21.ORF01989 | 129 aa | iron compound ABC transporter, ATP-binding protein            |
| 18RS21.ORF01990 | 210 aa | iron compound ABC transporter, substrate-binding protein      |
| 18RS21.ORF01991 | 71 aa  | transcriptional regulator, Cro/CI family                      |
| 18RS21.ORF01992 | 74 aa  | conserved hypothetical protein                                |
| 18RS21.ORF01993 | 96 aa  | hypothetical phage protein                                    |
| 18RS21.ORF01994 | 44 aa  | conserved hypothetical protein                                |
| 18RS21.ORF01995 | 125 aa | conserved hypothetical protein                                |
| 18RS21.ORF01996 | 175 aa | conserved hypothetical protein                                |
| 18RS21.ORF01997 | 53 aa  | hypothetical protein  |

|                 |        |   |
|-----------------|--------|---|
| 18RS21.ORF01998 | 131 aa | hypothetical phage protein                                      |
| 18RS21.ORF01999 | 63 aa  | transcription regulator   |
| 18RS21.ORF02000 | 224 aa | hypothetical protein  |
| 18RS21.ORF02001 | 52 aa  | conserved hypothetical protein                                  |
| 18RS21.ORF02003 | 65 aa  | hypothetical protein  |
| 18RS21.ORF02002 | 42 aa  | hypothetical protein  |
| 18RS21.ORF02004 | 30 aa  | DNA gyrase, A subunit   |
| 18RS21.ORF02005 | 252 aa | DNA gyrase, A subunit   |
| 18RS21.ORF02006 | 248 aa | biotin--acetyl-CoA-carboxylase ligase                           |
| 18RS21.ORF02007 | 38 aa  | conserved hypothetical protein                                  |
| 18RS21.ORF02008 | 222 aa | conserved hypothetical protein TIGR00046                        |
| 18RS21.ORF02009 | 287 aa | amidophosphoribosyltransferase                                  |
| 18RS21.ORF02010 | 182 aa | transcription regulator   |
| 18RS21.ORF02011 | 98 aa  | conserved hypothetical protein                                  |
| 18RS21.ORF02012 | 32 aa  | hypothetical protein  |
| 18RS21.ORF02013 | 136 aa | conserved hypothetical protein                                  |
| 18RS21.ORF02014 | 151 aa | GtrA family protein   |
| 18RS21.ORF02015 | 166 aa | 50S ribosomal protein   |
| 18RS21.ORF02016 | 30 aa  | hypothetical protein  |
| 18RS21.ORF02017 | 40 aa  | branched-chain amino acid aminotransferase                      |
| 18RS21.ORF02018 | 76 aa  | conserved hypothetical protein                                  |
| 18RS21.ORF02019 | 466 aa | 30S ribosomal protein S1  |
| 18RS21.ORF02020 | 358 aa | integrase   |
| 18RS21.ORF02021 | 228 aa | conserved hypothetical protein                                  |
| 18RS21.ORF02022 | 62 aa  | conserved hypothetical protein                                  |
| 18RS21.ORF02023 | 36 aa  | hypothetical protein  |
| 18RS21.ORF02024 | 548 aa | putative phage-related tail protein                             |
| 18RS21.ORF02025 | 110 aa | cysteinyl-tRNA synthetase                                       |
| 18RS21.ORF02026 | 137 aa | glutamine ABC transporter ATP-binding protein glnQ              |
| 18RS21.ORF02027 | 113 aa | amino acid ABC transporter, amino acid-binding/permease protein |
| 18RS21.ORF02028 | 213 aa | ribokinase  |
| 18RS21.ORF02029 | 171 aa | endoribonuclease L-PSP family protein                           |
| 18RS21.ORF02030 | 129 aa | methyl-accepting chemotaxis protein                             |
| 18RS21.ORF02031 | 279 aa | amidophosphoribosyltransferase                                  |
| 18RS21.ORF02032 | 286 aa | phosphoglucomutase  |
| 18RS21.ORF02033 | 111 aa | polyprenyl synthetase family protein                            |
| 18RS21.ORF02034 | 93 aa  | ABC transporter, ATP-binding protein CydC                       |
| 18RS21.ORF02035 | 122 aa | acetyltransferase, GNAT family                                  |
| 18RS21.ORF02036 | 182 aa | transcriptional regulator, MerR family                          |
| 18RS21.ORF02037 | 223 aa | exodeoxyribonuclease VII, large subunit                         |
| 18RS21.ORF02038 | 43 aa  | protein of unknown function                                     |
| 18RS21.ORF02039 | 59 aa  | Oligopeptide transport ATP-binding protein oppF                 |
| 18RS21.ORF02040 | 66 aa  | CsbD-like family  |

|                            |         |   |
|----------------------------|---------|---|
| 18RS21.ORF02041<br>protein | 298 aa  | ATPase, histidine kinase-, DNA gyrase B-, and HSP90-like domain protein |
| 18RS21.ORF02042            | 43 aa   | hypothetical protein  |
| 18RS21.ORF02043            | 447 aa  | 49.7 kDa protein  |
| 18RS21.ORF02044            | 75 aa   | holin, phage phi LC3 family   |
| 18RS21.ORF02045            | 42 aa   | conserved hypothetical protein  |
| 18RS21.ORF02046            | 130 aa  | conserved hypothetical protein  |
| 18RS21.ORF02047            | 674 aa  | prophage LambdaSa1, minor structural protein, putative                  |
| 18RS21.ORF02048            | 1275 aa | prophage LambdaSa1, N-acetylmuramoyl-L-alanine amidase, family 4        |
| 18RS21.ORF02049            | 502 aa  | prophage LambdaSa1, minor structural protein, putative                  |
| 18RS21.ORF02050            | 415 aa  | putative tail protein, phage associated                                 |
| 18RS21.ORF02051            | 160 aa  | Oligopeptide transport ATP-binding protein oppF                         |
| 18RS21.ORF02052            | 180 aa  | cI repressor  |
| 18RS21.ORF02053            | 86 aa   | H <sup>+</sup> -transporting ATP synthase a chain                       |
| 18RS21.ORF02054            | 150 aa  | ATP synthase F0, B subunit  |
| 18RS21.ORF02055            | 253 aa  | Arginyl-tRNA synthetase (Arginine--tRNA ligase) (ArgRS)                 |
| 18RS21.ORF02056            | 31 aa   | hypothetical protein  |
| 18RS21.ORF02057            | 63 aa   | Oligopeptide transport ATP-binding protein oppF                         |
| 18RS21.ORF02058            | 390 aa  | mercury(II) reductase [similarity]                                      |
| 18RS21.ORF02059            | 130 aa  | Hg(II)-responsive transcriptional regulator                             |
| 18RS21.ORF02060            | 30 aa   | conserved hypothetical protein  |
| 18RS21.ORF02061            | 97 aa   | transcription regulator   |
| 18RS21.ORF02062            | 45 aa   | HAM1 protein homolog  |
| 18RS21.ORF02063            | 151 aa  | HAM1 protein homolog  |
| 18RS21.ORF02064            | 109 aa  | Glutamate racemase  |
| 18RS21.ORF02065            | 326 aa  | exonuclease RexB  |
| 18RS21.ORF02066            | 318 aa  | putative tail protein, phage associated                                 |
| 18RS21.ORF02067            | 201 aa  | enoyl-acyl carrier protein(ACP) reductase.                              |
| 18RS21.ORF02068            | 38 aa   | enoyl-(acyl-carrier-protein) reductase                                  |
| 18RS21.ORF02069            | 96 aa   | conserved hypothetical protein  |
| 18RS21.ORF02070            | 162 aa  | conserved hypothetical protein  |
| 18RS21.ORF02071            | 39 aa   | conserved hypothetical protein  |
| 18RS21.ORF02072            | 123 aa  | gp31  |
| 18RS21.ORF02073            | 204 aa  | conserved hypothetical protein  |
| 18RS21.ORF02074            | 62 aa   | heavy-metal transporting P-type ATPase CAC3655                          |
| 18RS21.ORF02075            | 55 aa   | copper-translocating P-type ATPase                                      |
| 18RS21.ORF02076            | 171 aa  | transcriptional repressor CopY  |
| 18RS21.ORF02078            | 106 aa  | transketolase   |
| 18RS21.ORF02077            | 151 aa  | transketolase   |
| 18RS21.ORF02079            | 40 aa   | hypothetical protein  |
| 18RS21.ORF02080            | 89 aa   | conserved hypothetical protein  |
| 18RS21.ORF02081            | 142 aa  | Phenylalanyl-tRNA synthetase beta chain(Phenylalanine--tRNA             |

## ligase beta chain) (PheRS)

|                 |        |  |
|-----------------|--------|--|
| 18RS21.ORF02082 | 238 aa | conserved hypothetical protein   |
| 18RS21.ORF02083 | 150 aa | B. subtilis YabF protein homolog lin0226   |
| 18RS21.ORF02084 | 30 aa  | B. subtilis YabF protein homolog lin0226   |
| 18RS21.ORF02085 | 111 aa | Aspartyl/glutamyl-tRNA(Asn/Gln) amidotransferase subunit B(EC 6.3.5.-) (Asp/Glu-ADT subunit B) |
| 18RS21.ORF02086 | 35 aa  | Glutamyl-tRNA(Gln) amidotransferase subunit A (Glu-ADTsubunit A)                               |
| 18RS21.ORF02087 | 204 aa | CAAX amino terminal protease family family   |
| 18RS21.ORF02088 | 30 aa  | hypothetical protein   |
| 18RS21.ORF02089 | 148 aa | transcriptional repressor CopY   |
| 18RS21.ORF02090 | 124 aa | copper-translocating P-type ATPase   |
| 18RS21.ORF02091 | 144 aa | gp31   |
| 18RS21.ORF02092 | 88 aa  | conserved hypothetical protein   |
| 18RS21.ORF02093 | 140 aa | conserved hypothetical protein   |
| 18RS21.ORF02094 | 112 aa | hypothetical protein   |
| 18RS21.ORF02095 | 59 aa  | phage protein homolog lin0091  |

## REGION 8.3 [1 sequence(s)]:

18RS21 [2173883-2212314]

|                 |        |   |
|-----------------|--------|---|
| 18RS21.ORF02106 | 79 aa  | cI repressor                                    |
| 18RS21.ORF02107 | 40 aa  | hypothetical protein                            |
| 18RS21.ORF02108 | 112 aa | amino acid ABC transporter, ATP-binding protein |
| 18RS21.ORF02109 | 87 aa  | DNA polymerase III gamma/tau subunits           |
| 18RS21.ORF02110 | 180 aa | putative major tail protein                     |
| 18RS21.ORF02111 | 100 aa | conserved hypothetical protein                  |
| 18RS21.ORF02112 | 770 aa | putative tail protein, phage assocaited         |
| 18RS21.ORF02113 | 35 aa  | conserved hypothetical protein                  |
| 18RS21.ORF02114 | 100 aa | conserved hypothetical protein                  |
| 18RS21.ORF02115 | 201 aa | putative major tail protein                     |
| 18RS21.ORF02116 | 126 aa | conserved hypothetical protein                  |
| 18RS21.ORF02117 | 123 aa | conserved hypothetical protein                  |
| 18RS21.ORF02118 | 49 aa  | conserved hypothetical protein                  |
| 18RS21.ORF02119 | 97 aa  | conserved hypothetical protein                  |
| 18RS21.ORF02120 | 386 aa | 77ORF006  |
| 18RS21.ORF02121 | 243 aa | phage protein                                   |
| 18RS21.ORF02122 | 382 aa | phage portal protein, HK97 family               |
| 18RS21.ORF02123 | 499 aa | phage terminase, large subunit, putative        |
| 18RS21.ORF02124 | 118 aa | conserved hypothetical protein                  |
| 18RS21.ORF02125 | 73 aa  | 77ORF040  |
| 18RS21.ORF02126 | 113 aa | hypothetical phage protein-related protein      |
| 18RS21.ORF02127 | 125 aa | hypothetical phage protein                      |
| 18RS21.ORF02128 | 71 aa  | gp31  |

|                 |        |   |
|-----------------|--------|---|
| 18RS21.ORF02129 | 140 aa | integral membrane protein                         |
| 18RS21.ORF02130 | 327 aa | inosine-uridine preferring nucleoside hydrolase   |
| 18RS21.ORF02131 | 37 aa  | hypothetical protein                              |
| 18RS21.ORF02132 | 464 aa | prophage Lp3 protein 8, helicase                  |
| 18RS21.ORF02133 | 285 aa | DNA primase domain protein                        |
| 18RS21.ORF02134 | 90 aa  | prophage ps1 protein 10                           |
| 18RS21.ORF02135 | 63 aa  | conserved hypothetical protein                    |
| 18RS21.ORF02136 | 110 aa | conserved hypothetical protein                    |
| 18RS21.ORF02137 | 150 aa | hypothetical protein                              |
| 18RS21.ORF02138 | 212 aa | BRO family, N-terminal domain protein             |
| 18RS21.ORF02139 | 200 aa | BRO family, N-terminal domain protein             |
| 18RS21.ORF02140 | 46 aa  | transcription regulator, probable                 |
| 18RS21.ORF02141 | 41 aa  | UDP-N-acetylglucosamine 1-carboxyvinyltransferase |
| 18RS21.ORF02142 | 223 aa | thiamine-phosphate pyrophosphorylase              |
| 18RS21.ORF02143 | 219 aa | hydroxyethylthiazole kinase                       |
| 18RS21.ORF02144 | 36 aa  | hypothetical protein                              |
| 18RS21.ORF02145 | 388 aa | Integrase core domain protein                     |
| 18RS21.ORF02146 | 384 aa | seryl-tRNA synthetase                             |

The gene content of the 69 genomic islands larger than 5 kb that differ between strains (as determined in Fig. 1) is listed. The entire content of each island in each strain appears even if the genomic island varies in size from strain to strain. ORF length and annotation are also reported.