

**Table 6. Clusters of genes derived from phylogenetic profiling of *S. agalactiae* strains based on comparative genome hybridizations**

Clusters of five or more genes (see Fig. 5) are listed.

**Cluster 1**

SAG0230	conserved hypothetical protein
SAG0231	hypothetical protein
SAG0232	hypothetical protein
SAG0233	hypothetical protein
SAG0234	hypothetical protein
SAG0235	hypothetical protein

**Cluster 2**

SAG0222	conserved domain protein
SAG0223	conserved hypothetical protein, fusion
SAG0225	hypothetical protein
SAG0226	recombination protein
SAG0227	hypothetical protein
SAG0228	conserved hypothetical protein
SAG0229	conserved hypothetical protein

**Cluster 3**

SAG0634 hypothetical protein  
SAG0635 acid phosphatase, class B  
SAG0636 conserved hypothetical protein  
SAG0638 cell wall surface anchor family protein, interruption-N  
SAG0640 transposase OrfA, IS3 family  
SAG0642 hypothetical protein  
SAG0643 chaperonin, 33 kDa, degenerate  
SAG0644 transcriptional regulator, AraC family  
SAG0645 cell wall surface anchor family protein  
SAG0646 cell wall surface anchor family protein  
SAG0647 sortase family protein  
SAG0648 sortase family protein  
SAG0649 cell wall surface anchor family protein, putative  
SAG0650 sortase family protein  
SAG0651 protein of unknown function

**Cluster 4**

SAG1898 PTS system, IID component  
SAG1899 PTS system, IIC component  
SAG1900 PTS system, IIB component  
SAG1901 glucuronyl hydrolase  
SAG1902 PTS system, IIA component  
SAG1905 conserved hypothetical protein  
SAG1906 carbohydrate kinase, PfkB family

**Cluster 5**

SAG0247 hypothetical protein  
SAG0248 hypothetical protein  
SAG0249 hypothetical protein  
SAG0674 hypothetical protein  
SAG0675 putative secreted protein  
SAG0676 proteinase, putative  
SAG0677 hypothetical protein  
SAG0680 protein of unknown function  
SAG0681 conserved domain protein  
SAG0684 ABC transporter, ATP-binding protein  
SAG1698 conserved hypothetical protein

**Cluster 6**

SAG0261 IS1381, transposase OrfB  
SAG0262 IS1381, transposase OrfA  
SAG0965 IS1381, transposase OrfA  
SAG0966 IS1381, transposase OrfB  
SAG2002 IS1381, transposase OrfB

**Cluster 7**

SAG1027 conserved hypothetical protein  
SAG1028 hypothetical protein  
SAG1029 hypothetical protein  
SAG1030 protein of unknown function  
SAG1031 conserved domain protein  
SAG1032 conserved hypothetical protein

**Cluster 8**

SAG1253 transposase, ISL3 family  
SAG1254 mercuric reductase  
SAG1255 mercuric resistance operon regulatory protein MerR  
SAG2022 transposase, ISL3 family  
SAG2023 mercuric reductase  
SAG2024 mercuric resistance operon regulatory protein MerR

**Cluster 9**

SAG1993 site-specific recombinase, phage integrase family  
SAG1994 conserved hypothetical protein  
SAG1995 hypothetical protein  
SAG1996 cell wall surface anchor family protein, putative  
SAG1997 hypothetical protein  
SAG1998 hypothetical protein  
SAG2000 membrane protein, putative  
SAG2001 conjugal transfer protein, interruption-C  
SAG2007 conserved hypothetical protein  
SAG2008 conserved hypothetical protein  
SAG2009 conserved hypothetical protein  
SAG2010 hypothetical protein  
SAG2011 conserved hypothetical protein  
SAG2012 hypothetical protein  
SAG2016 hypothetical protein  
SAG2017 transcriptional regulator, Cro/CI family  
SAG2025 Mn<sup>2+</sup>/Fe<sup>2+</sup> transporter, NRAMP family

**Cluster 10**

SAG1039 conserved hypothetical protein  
SAG1447 conserved hypothetical protein  
SAG1448 glycosyl transferase, group 1 family protein  
SAG1449 preprotein translocase SecA subunit, putative  
SAG1450 conserved domain protein  
SAG1452 conserved hypothetical protein  
SAG1453 preprotein translocase SecY family protein  
SAG1454 glycosyl transferase, putative  
SAG1455 glycosyl transferase, group 2 family protein  
SAG1456 glycosyl transferase, family 8, degenerate

SAG1459 glycosyl transferase family 8  
SAG1460 glycosyl transferase, family 8  
SAG1461 conserved hypothetical protein  
SAG1462 cell wall surface anchor family protein  
SAG1463 transcriptional regulator, RofA family, authentic point mutation  
SAG1469 conserved hypothetical protein  
SAG1471 conserved hypothetical protein  
SAG1933 PTS system, IIC component, putative

#### **Cluster 11**

SAG0009 hypothetical protein  
SAG0120 hypothetical protein  
SAG0157 deoxyribonuclease-related protein, degenerate  
SAG0186 hypothetical protein  
SAG0216 hypothetical protein  
SAG0236 hypothetical protein  
SAG0307 hypothetical protein  
SAG0308 ABC transporter, ATP-binding protein  
SAG0311 DNA-binding response regulator, authentic point mutation  
SAG0518 peptide chain release factor 2, programmed frameshift  
SAG0553 hypothetical protein  
SAG0555 prophage LambdaSa1, antirepressor, putative  
SAG0564 conserved hypothetical protein  
SAG0579 conserved hypothetical protein  
SAG0580 conserved hypothetical protein, truncation  
SAG0611 transposase, degenerate  
SAG0637 transcriptional regulator, TetR family, putative, authentic frameshift  
SAG0641 Tn5252, Orf 10 protein, degenerate  
SAG0652 Tn5252, Orf 28 protein, degenerate  
SAG0655 conserved hypothetical protein  
SAG0678 endopeptidase O, degenerate  
SAG0683 transmembrane protein Vexp3, putative, degenerate  
SAG0855 glycogen biosynthesis protein GlgD, authentic frameshift  
SAG0898 hypothetical protein  
SAG0899 hypothetical protein  
SAG0901 hypothetical protein  
SAG0902 hypothetical protein  
SAG0903 hypothetical protein  
SAG0917 Tn916, hypothetical protein  
SAG0920 Tn916, hypothetical protein  
SAG0922 Tn916, hypothetical protein  
SAG0924 Tn916, tetM leader peptide  
SAG0928 Tn916, hypothetical protein, authentic frameshift  
SAG0936 Tn916, hypothetical protein  
SAG0943 hypothetical protein  
SAG0972 conserved hypothetical protein, authentic frameshift

SAG1023 hypothetical protein  
SAG1080 hypothetical protein  
SAG1123 hypothetical protein  
SAG1129 hypothetical protein  
SAG1136 conserved hypothetical protein  
SAG1217 conserved hypothetical protein, authentic frameshift  
SAG1231 transposase OrfB, IS3 family, degenerate  
SAG1242 transposase OrfB, IS3 family, truncation  
SAG1309 hypothetical protein  
SAG1331 R5 protein  
SAG1437 hypothetical protein  
SAG1445 MutT/nudix family protein, authentic frameshift  
SAG1484 ribosomal protein L33  
SAG1493 hypothetical protein  
SAG1539 hypothetical protein  
SAG1543 conserved hypothetical protein, authentic frameshift  
SAG1560 hypothetical protein  
SAG1568 phosphoserine aminotransferase, authentic frameshift  
SAG1570 conserved hypothetical protein  
SAG1601 conserved hypothetical protein  
SAG1644 hypothetical protein  
SAG1646 hypothetical protein  
SAG1699 hypothetical protein  
SAG1705 peptidase, M24 family, authentic point mutation  
SAG1708 hypothetical protein  
SAG1857 prophage LambdaSa2, HNH endonuclease family protein  
SAG1864 hypothetical protein  
SAG1868 hypothetical protein  
SAG1869 prophage LambdaSa2, type II DNA modification methyltransferase,  
putative  
SAG1872 hypothetical protein  
SAG1874 hypothetical protein  
SAG1876 prophage LambdaSa2, HNH endonuclease family protein  
SAG1878 conserved domain protein  
SAG1881 hypothetical protein  
SAG1883 conserved hypothetical protein  
SAG1886 hypothetical protein  
SAG1903 hypothetical protein  
SAG1937 streptococcal histidine triad family protein, degenerate  
SAG1971 hypothetical protein  
SAG1979 membrane protein, putative  
SAG1980 ABC transporter, ATP-binding protein  
SAG1981 hypothetical protein  
SAG1982 transcriptional regulator, Cro/CI family  
SAG1983 conserved hypothetical protein  
SAG1984 conserved hypothetical protein TIGR00730

SAG1985 hypothetical protein  
SAG1991 transcriptional regulator, Cro/CI family  
SAG1992 protein of unknown function  
SAG1999 hypothetical protein  
SAG2004 conjugal transfer protein, interruption-N  
SAG2039 conserved hypothetical protein  
SAG2044 hypothetical protein  
SAG2052 hypothetical protein  
SAG2065 ribosomal protein L33  
SAG2094 competence/damage-inducible protein CinA, authentic frameshift  
SAG2099 hypothetical protein

#### **Cluster 12**

SAG1164 glycosyl transferase CpsJ(V)  
SAG1165 glycosyl transferase CpsO(V)  
SAG1166 glycosyl transferase CpsN(V)  
SAG1167 polysaccharide biosynthesis protein CpsM(V)  
SAG1168 polysaccharide biosynthesis protein cpsH(V)

#### **Cluster 13**

SAG0581 conserved hypothetical protein  
SAG0582 conserved hypothetical protein  
SAG0583 conserved hypothetical protein  
SAG0585 conserved hypothetical protein  
SAG0586 conserved hypothetical protein  
SAG0587 prophage LambdaSa1, structural protein, putative  
SAG0588 conserved hypothetical protein  
SAG0589 conserved hypothetical protein  
SAG0590 conserved hypothetical protein  
SAG0591 conserved hypothetical protein  
SAG0593 prophage LambdaSa1, structural protein  
SAG0594 conserved hypothetical protein  
SAG0595 conserved hypothetical protein  
SAG0596 prophage LambdaSa1, pblA protein, internal deletion

#### **Cluster 14**

SAG0915 Tn916, transposase  
SAG0918 Tn916, hypothetical protein  
SAG0919 Tn916, hypothetical protein  
SAG0921 Tn916, transcriptional regulator, putative  
SAG0925 Tn916, hypothetical protein  
SAG0926 Tn916, NLP/P60 family protein  
SAG0927 membrane protein, putative  
SAG0929 Tn916, hypothetical protein  
SAG0930 Tn916, hypothetical protein  
SAG0931 Tn916, hypothetical protein

SAG0932 Tn916, transcriptional regulator, putative  
SAG0933 Tn916, FtsK/SpoIIIE family protein  
SAG0934 Tn916, hypothetical protein  
SAG0935 Tn916, hypothetical protein  
SAG0937 ABC transporter, ATP-binding protein, authentic frameshift

### **Cluster 15**

SAG1835 conserved hypothetical protein  
SAG1837 prophage LambdaSa2, lysin, putative  
SAG1839 conserved hypothetical protein  
SAG1840 hypothetical protein  
SAG1842 prophage LambdaSa2, PblB, putative  
SAG1843 conserved hypothetical protein  
SAG1844 conserved hypothetical protein  
SAG1849 hypothetical protein  
SAG1851 conserved domain protein  
SAG1852 conserved domain protein  
SAG1853 prophage LambdaSa2, protease, putative  
SAG1854 conserved hypothetical protein  
SAG1855 prophage LambdaSa2, terminase large subunit, putative  
SAG1856 hypothetical protein  
SAG1858 hypothetical protein  
SAG1859 prophage LambdaSa2, site-specific recombinase, phage integrase family  
SAG1860 conserved hypothetical protein  
SAG1861 prophage LambdaSa2, transcriptional regulator, Cro/CI family  
SAG1862 hypothetical protein  
SAG1863 prophage LambdaSa2, single-strand binding protein  
SAG1865 conserved hypothetical protein  
SAG1866 conserved hypothetical protein  
SAG1867 conserved hypothetical protein  
SAG1870 prophage LambdaSa2, DNA replication protein DnaC, putative  
SAG1871 prophage LambdaSa2, bacteriophage replication protein/hypothetical protein, truncation/fusion  
SAG1873 prophage LambdaSa2, replicative DNA helicase  
SAG1877 prophage LambdaSa2, antirepressor protein, putative  
SAG1879 hypothetical protein  
SAG1882 prophage LambdaSa2, repressor protein, putative  
SAG1884 hypothetical protein  
SAG1885 prophage LambdaSa2, site-specific recombinase, phage integrase family

### **Cluster 16**

SAG1247 site-specific recombinase, phage integrase family  
SAG1250 Tn5252, relaxase  
SAG1251 Tn5252, Orf 9 protein  
SAG1252 Tn5252, Orf 10 protein  
SAG1256 IS861, transposase OrfB, truncation

SAG1257 cation-transporting ATPase, E1-E2 family  
SAG1258 cadmium efflux system accessory protein  
SAG1259 conserved hypothetical protein  
SAG1260 hypothetical protein  
SAG1261 conserved hypothetical protein  
SAG1262 cation-transporting ATPase, E1-E2 family  
SAG1263 conserved domain protein, authentic frameshift  
SAG1264 transcriptional repressor CopY, putative  
SAG1265 cadmium resistance transporter, putative  
SAG1266 hypothetical protein  
SAG1267 hypothetical protein  
SAG1268 repressor protein, putative  
SAG1270 ImpB/MucB/SamB family protein  
SAG1271 conserved hypothetical protein  
SAG1272 conserved hypothetical protein  
SAG1273 conserved hypothetical protein  
SAG1274 conserved hypothetical protein  
SAG1276 conserved hypothetical protein  
SAG1277 hypothetical protein  
SAG1278 hypothetical protein  
SAG1279 conserved domain protein  
SAG1280 SNF2 family protein  
SAG1281 hypothetical protein  
SAG1283 agglutinin receptor  
SAG1284 abortive infection protein AbiGI  
SAG1285 abortive infection protein AbiGII  
SAG1286 Tn5252, Orf28  
SAG1287 Tn5252, Orf26  
SAG1288 Tn5252, Orf25, degenerate  
SAG1289 Tn5252, Orf23  
SAG1290 hypothetical protein  
SAG1291 Tn5252, Orf 21 protein, internal deletion  
SAG1292 hypothetical protein  
SAG1293 protease, putative  
SAG1294 conserved hypothetical protein  
SAG1295 conserved hypothetical protein  
SAG1296 conserved hypothetical protein  
SAG1297 C-5 cytosine-specific DNA methylase  
SAG1299 conserved hypothetical protein  
SAG1304 hypothetical protein