

## Mobile genetic elements provide evidence for a bovine origin of clonal complex 17 of *Streptococcus agalactiae*

Geneviève Héry-Arnaud, Guillaume Bruant, Philippe Lanotte, Stella Brun, Bertrand Picard, Agnès Rosenau, Nathalie Van Ver Mee-Marquet, Pascal Rainard, Roland Quentin, Laurent Mereghetti

► **To cite this version:**

Geneviève Héry-Arnaud, Guillaume Bruant, Philippe Lanotte, Stella Brun, Bertrand Picard, et al.. Mobile genetic elements provide evidence for a bovine origin of clonal complex 17 of *Streptococcus agalactiae*. *Applied and Environmental Microbiology*, American Society for Microbiology, 2007, 73 (14), pp.4668-4672. <10.1128/AEM.02604-06>. <hal-00557613>

**HAL Id: hal-00557613**

**<http://hal.univ-brest.fr/hal-00557613>**

Submitted on 21 Jan 2011

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1 Mobile genetic elements provide evidence for a bovine origin of clonal complex 17 of  
2 *Streptococcus agalactiae*

3

4 Geneviève Héry-Arnaud,<sup>1</sup> Guillaume Bruant,<sup>1,a</sup> Philippe Lanotte,<sup>1,2</sup> Stella Brun,<sup>1</sup>  
5 Bertrand Picard,<sup>3</sup> Agnès Rosenau,<sup>1</sup> Nathalie van der Mee-Marquet,<sup>1,2</sup> Pascal Rainard,<sup>4</sup>  
6 Roland Quentin,<sup>1,2</sup> and Laurent Mereghetti<sup>1,2\*</sup>

7

8 Université François-Rabelais, IFR 136, Faculté de Médecine, EA 3854 "Bactéries et  
9 risque materno-fœtal", Tours, France <sup>1</sup>; Centre Hospitalier Universitaire, Tours,  
10 France<sup>2</sup>; Service de Microbiologie, Hôpital Avicenne, Assistance Publique-Hôpitaux de  
11 Paris, Bobigny, France<sup>3</sup>; Institut National de la Recherche Agronomique, UR1282,  
12 Infectiologie Animale et Santé Publique, Nouzilly, France <sup>3</sup>

13

14 \* Corresponding author: EA 3854 "Bactéries et risque materno-fœtal", Laboratoire de  
15 Bactériologie, Faculté de Médecine de Tours, 2 bd Tonnellé, 37032 Tours, France  
16 laurent.mereghetti@med.univ-tours.fr

17

18 <sup>a</sup> Present affiliation: Groupe de recherche sur les maladies infectieuses du porc, Faculté  
19 de Médecine Vétérinaire, Université de Montréal, Saint-Hyacinthe, Québec.

20

21 This work was presented in part at the XVI<sup>th</sup> Lancefield International Symposium on  
22 Streptococci and Streptococcal Diseases, Cairns, 25-29 September 2005.

23           We sought an explanation for epidemiological changes in *Streptococcus*  
24 *agalactiae* infections, by investigating the link between ecological niches of the  
25 bacterium, by determining the prevalence of 11 mobile genetic elements. The  
26 prevalence of nine of these elements differed significantly according to the human or  
27 bovine origin of the isolate. Correlating this distribution with the phylogeny obtained by  
28 multilocus sequence analysis, we observed that human isolates harboring GBSi1, a clear  
29 marker of the bovine niche, clustered in clonal complex 17. Our results are thus  
30 consistent with the emergence of this virulent human clone from a bovine ancestor.

31           Epidemiological changes in the pattern of *Streptococcus agalactiae* infections  
32 over the last six decades, from predominantly bovine infections (13), to a leading cause  
33 of neonatal infections (25, 26) and, more recently, a cause of invasive disease in elderly  
34 adults (6), requires explanation. Genomic lineages grouping the most invasive isolates  
35 have been described by multilocus enzyme electrophoresis (17, 23), by pulsed-field gel  
36 electrophoresis (1, 24) and more recently by the analysis of the polymorphism in the  
37 allelic profiles of housekeeping genes (11), but are insufficient to account for the  
38 emergence of *S. agalactiae* from such diverse ecosystems. A link between the bovine  
39 and human niches has been proposed (8, 19), but remains a matter of debate. Indeed, it  
40 was recently suggested that human and bovine *S. agalactiae* isolates are largely  
41 unrelated (3) and represent distinct populations or subtypes (5, 27). Conversely, another  
42 study provided evidence to suggest that the major hyperinvasive ST-17 clone implicated  
43 in neonatal invasive diseases had arisen from bovine lineage ST-61 (2), whereas others  
44 suggested that the genome of the common ancestor was closer to that of human ST-17  
45 strains than to that of the ST-61 strains of bovine origin (4). We hypothesized that  
46 mobile genetic elements (MGEs), such as insertion sequences (ISs) or group II introns,  
47 which drive bacterial evolution (12, 30), might provide insight into changes in the  
48 epidemiology of *S. agalactiae* infections.

49           We tested this hypothesis by analyzing a collection of 98 epidemiologically  
50 unrelated isolates, mostly collected in France (3 isolates from Germany and 1 from  
51 Malaysia) between 1966 and 2003: 63 human isolates from colonized individuals, adults  
52 with invasive diseases and neonates with meningitis (52 of which were previously typed  
53 by MLST and studied for the prevalence of IS1548, IS861, IS1381, ISSa4 and GBSi1)  
54 (10), and 35 bovine isolates obtained from cows with clinical evidence of udder  
55 infection.

56 The prevalence of 11 MGEs [one group II intron (GBSi1), and ten IS elements  
57 corresponding to six insertion sequences (*IS1548*, *IS861*, *IS1381*, *ISSa4*, *ISSag1* and  
58 *ISSag2*), and four transposase genes corresponding to remnants of IS elements (*ISSag9*,  
59 *sag0448*, *sag1241*, and *sag0610*, relating to the identification of the ORFs in strains  
60 A909 and 2605V/R) (28, 29)], was evaluated by PCR. *IS1548*, *IS861*, *IS1381*, *ISSa4*,  
61 *ISSag1*, *ISSag2* and *GBSi1* were detected as previously described (10). Other  
62 transposase genes were detected with the primers listed in Table 1. The prevalence of all  
63 the MGEs except *IS861* and *ISSag2* differed significantly as a function of the original  
64 reservoir (Fig. 1). This was particularly true for *IS1381*, *IS1548*, *ISSag1* and *ISSag9*,  
65 which were mostly present in human isolates, and for *GBSi1* and *ISSa4*, which were  
66 mostly present in bovine isolates (Fig. 1).

67 Multilocus sequence analysis, carried out with the standard multilocus sequence  
68 typing (MLST) scheme described by Jones *et al.* (11), resolved the 98 isolates into 47  
69 STs. A phylogenetic tree was generated, based on the nucleotide sequences of the  
70 supergene obtained by concatenation of the seven loci for the 98 isolates, using MEGA  
71 software version 3.1 ([www.megasoftware.net/](http://www.megasoftware.net/)) with the neighbor joining (NJ) method  
72 (14). Clonal complexes (CCs) (isolates sharing six or seven identical alleles) were  
73 defined by eBURST analysis (<http://eburst.mlst.net/>) (7). Seven main NJ phylogenetic  
74 divisions were identified, closely matching the seven main CCs identified by eBURST  
75 analysis (Fig. 2): subCC19 (n=21), subCC2 (n=11), CC10 (n=12), CC7 (n=6), CC17  
76 (n=13), CC67 (n=15) and CC23 (n=10). Three of these CCs were “host-specific”: CC17  
77 and subCC19 consisted almost exclusively of human isolates (100% and 95% of the  
78 isolates, respectively), whereas CC67 was almost exclusively composed of bovine  
79 isolates (93.8% of the isolates). Moreover, the NJ dendrogram shows that human

80 lineage CC17 arose from bovine lineage CC67 (Fig. 2), consistent with the assertions of  
81 Bisharat *et al.* (2).

82 A factorial analysis of correspondence (FAC) was performed on the whole data  
83 set, with the 11 MGEs as active variables and the three main clonal complexes (CC19,  
84 CC17 and CC67) and the bovine or human origins of the strains as illustrative variables.  
85 Computations were performed with SPAD.N software (Centre International de  
86 Statistiques et Informatique Appliquées, St Mandé, France), as previously described  
87 (22). The projections of the active and illustrative variables and of the strains on the  
88 F1/F2 plane accounting for 40.97 % of the total variance (Fig. 3), separated two groups  
89 of variables and of strains. GBSi1 and ISSa4 (and to a lesser degree *sag0448* and  
90 *sag1241*), bovine origin, CC17 and CC67 and 94.3% of bovine isolates were projected  
91 onto the negative values of F1. IS1548, IS1381, ISSag9 and ISSag1 (and to a lesser  
92 degree *sag0610*), human origin, CC19 and 79.4% of human isolates were projected onto  
93 the positive values of F1. These findings confirm those of Luan *et al.* in that the two  
94 major human *S. agalactiae* lineages are each marked by one MGE: subCC19 (defined as  
95 CC19 by Luan *et al.*) (16), is characterized by IS1548, whereas CC17 is characterized  
96 by GBSi1. All but one of the human isolates projected with the bovine isolates and  
97 variable CC67 onto the negative values of F1 belonged to CC17, highlighting the  
98 stronger linkage of CC17 with the bovine lineage CC67 than with any of the other  
99 human lineages (Fig. 3). Moreover, our data also suggest that the presence of GBSi1 in  
100 the bovine reservoir predates its presence in the human reservoir because (1) all our  
101 bovine isolates harbored GBSi1 whereas this element was present in less than one third  
102 of human isolates (almost all of which belonged to CC17) and (2) the bovine isolates  
103 had higher copy numbers and more complex patterns for GBSi1 than the human isolates  
104 (data not shown), consistent with an increase in IS pattern complexity over time (18).

105 MGEs therefore provide evidence for a bovine ancestor of the human clonal lineage  
106 CC17, all the isolates of which harbor GBSi1. This statement is consistent with the  
107 proposed evolutionary scheme of MGE acquisition, in which GBSi1 is thought to have  
108 been acquired recently by human isolates (10).

109       The particular distribution of MGEs among bovine and human isolates (Fig. 3) is  
110 also consistent with a physical barrier reducing the exchange of genetic material  
111 between the bovine and human reservoirs. The human isolate H14, which occupies a  
112 prominent position on the NJ dendrogram, between the bovine CC67 and the human  
113 CC17 (Fig. 2), displayed two particular features. First, the copy number of GBSi1 in  
114 this isolate was much higher than that in other isolates (human isolates had a median of  
115 3 GBSi1, whereas H14 harbored 12 copies — data not shown). Parkhill *et al.* suggested  
116 that IS expansion occurs during evolutionary bottlenecks (20, 21). They argued that in  
117 *Yersinia*, IS expansion coincided with a change in the ecological niche of the species,  
118 from commensal organism of the gut to systemic pathogen (21). Second, isolate H14  
119 harbored both IS1548 and GBSi1. In our study, as in previous studies (9, 15), very few  
120 isolates harbored both GBSi1 and IS1548 (see Fig. 2). Granlund *et al.* explained this  
121 mutually exclusive distribution by the presence of a closed insertion site for the two  
122 MGEs on the bacterial chromosome (9), but this distribution could also be explained by  
123 IS1548 and GBSi1 originating from two different ecological niches. Thus, human clone  
124 CC17 seems to have evolved from the bovine lineage containing GBSi1, and has  
125 retained the GBSi1 genetic marker from its bovine donor.

126       In conclusion, our results provide at least a partial explanation of the emergence  
127 of *S. agalactiae* as a major cause of neonatal meningitis, despite its original  
128 identification as the causal agent of bovine mastitis. Assuming that CC17 evolved from  
129 a bovine ancestor, as shown both by MLST analysis and by the presence of the bovine

130 marker GBSi1, the isolates of this clonal complex may have genetic features different  
131 from those of other *S. agalactiae* isolates that have adapted to their human hosts over a  
132 longer period of time. These features probably concern genes not present in the core  
133 genome, as recently defined by Tettelin *et al.*, but that have been acquired by horizontal  
134 exchange due to the environmental gene pool available in other ecosystems (28).  
135 Indeed, we recently reported a link between prophage DNA fragments and *S. agalactiae*  
136 isolates from neonatal cases of meningitis, suggesting the presence of additional genetic  
137 material in these strains (10).

138

139

#### **ACKNOWLEDGMENTS**

140 We would like to thank Nicolas Jones, who manages the MLST website database,  
141 for assigning new allele numbers and new sequence type numbers. We also thank  
142 Stéphanie Gouriou for her help in using the SPAD.N software.



143

## REFERENCES

- 144 1. **Bidet, P., N. Brahim, C. Chalas, Y. Aujard, and E. Bingen.** 2003. Molecular  
145 characterization of serotype III group B-streptococcus isolates causing neonatal  
146 meningitis. *J Infect Dis* **188**:1132-7.
- 147 2. **Bisharat, N., D. W. Crook, J. Leigh, R. M. Harding, P. N. Ward, T. J.**  
148 **Coffey, M. C. Maiden, T. Peto, and N. Jones.** 2004. Hyperinvasive neonatal  
149 group B streptococcus has arisen from a bovine ancestor. *J Clin Microbiol*  
150 **42**:2161-7.
- 151 3. **Bohnsack, J. F., A. A. Whiting, G. Martinez, N. Jones, E. E. Adderson, S.**  
152 **Detrick, A. J. Blaschke-Bonkowsky, N. Bisharat, and M. Gottschalk.** 2004.  
153 Serotype III *Streptococcus agalactiae* from bovine milk and human neonatal  
154 infections. *Emerg Infect Dis* **10**:1412-9.
- 155 4. **Brochet, M., E. Couve, M. Zouine, T. Vallaey, C. Rusniok, M. C. Lamy, C.**  
156 **Buchrieser, P. Trieu-Cuot, F. Kunst, C. Poyart, and P. Glaser.** 2006.  
157 Genomic diversity and evolution within the species *Streptococcus agalactiae*.  
158 *Microbes Infect* **8**:1227-43.
- 159 5. **Dogan, B., Y. H. Schukken, C. Santisteban, and K. J. Boor.** 2005.  
160 Distribution of serotypes and antimicrobial resistance genes among  
161 *Streptococcus agalactiae* isolates from bovine and human hosts. *J Clin*  
162 *Microbiol* **43**:5899-5906.
- 163 6. **Edwards, M. S., M. A. Rench, D. L. Palazzi, and C. J. Baker.** 2005. Group B  
164 streptococcal colonization and serotype-specific immunity in healthy elderly  
165 persons. *Clin Infect Dis* **40**:352-7.
- 166 7. **Feil, E. J., B. C. Li, D. M. Aanensen, W. P. Hanage, and B. G. Spratt.** 2004.  
167 eBURST: inferring patterns of evolutionary descent among clusters of related

- 168 bacterial genotypes from multilocus sequence typing data. J Bacteriol **186**:1518-  
169 30.
- 170 8. **Finch, L. A., and D. R. Martin.** 1984. Human and bovine group B streptococci:  
171 two distinct populations. J Appl Bacteriol **57**:273-8.
- 172 9. **Granlund, M., F. Michel, and M. Norgren.** 2001. Mutually exclusive  
173 distribution of IS1548 and GBSi1, an active group II intron identified in human  
174 isolates of group B streptococci. J Bacteriol **183**:2560-9.
- 175 10. **Héry-Arnaud, G., G. Bruant, P. Lanotte, S. Brun, A. Rosenau, N. van der**  
176 **Mee-Marquet, R. Quentin, and L. Mereghetti.** 2005. Acquisition of insertion  
177 sequences and the GBSi1 intron by *Streptococcus agalactiae* isolates correlates  
178 with the evolution of the species. J Bacteriol **187**:6248-52.
- 179 11. **Jones, N., J. F. Bohnsack, S. Takahashi, K. A. Oliver, M. S. Chan, F. Kunst,**  
180 **P. Glaser, C. Rusniok, D. W. Crook, R. M. Harding, N. Bisharat, and B. G.**  
181 **Spratt.** 2003. Multilocus sequence typing system for group B streptococcus. J  
182 Clin Microbiol **41**:2530-6.
- 183 12. **Kazazian, H. H., Jr.** 2004. Mobile elements: drivers of genome evolution.  
184 Science **303**:1626-32.
- 185 13. **Keefe, G. P., I. R. Dohoo, and E. Spangler.** 1997. Herd prevalence and  
186 incidence of *Streptococcus agalactiae* in the dairy industry of Prince Edward  
187 Island. J Dairy Sci **80**:464-70.
- 188 14. **Kumar, S., K. Tamura, and M. Nei.** 2004. MEGA3: Integrated software for  
189 Molecular Evolutionary Genetics Analysis and sequence alignment. Brief  
190 Bioinform **5**:150-63.

- 191 15. **Luan, S. L., M. Granlund, and M. Norgren.** 2003. An inserted DNA fragment  
192 with plasmid features is uniquely associated with the presence of the GBSi1  
193 group II intron in *Streptococcus agalactiae*. *Gene* **312**:305-12.
- 194 16. **Luan, S. L., M. Granlund, M. Sellin, T. Lagergard, B. G. Spratt, and M.**  
195 **Norgren.** 2005. Multilocus sequence typing of Swedish invasive group B  
196 streptococcus isolates indicates a neonatally associated genetic lineage and  
197 capsule switching. *J Clin Microbiol* **43**:3727-33.
- 198 17. **Musser, J. M., S. J. Mattingly, R. Quentin, A. Goudeau, and R. K. Selander.**  
199 1989. Identification of a high-virulence clone of type III *Streptococcus*  
200 *agalactiae* (group B streptococcus) causing invasive neonatal disease. *Proc Natl*  
201 *Acad Sci USA* **86**:4731-5.
- 202 18. **Naas, T., M. Blot, W. M. Fitch, and W. Arber.** 1995. Dynamics of IS-related  
203 genetic rearrangements in resting *Escherichia coli* K-12. *Mol Biol Evol* **12**:198-  
204 207.
- 205 19. **Norcross, N. L., and N. Oliver.** 1976. The distribution and characterization of  
206 group B streptococci in New York State. *Cornell Vet* **66**:242-8.
- 207 20. **Parkhill, J., M. Sebahia, A. Preston, L. D. Murphy, N. Thomson, D. E.**  
208 **Harris, M. T. Holden, C. M. Churcher, S. D. Bentley, K. L. Mungall, A. M.**  
209 **Cerdeno-Tarraga, L. Temple, K. James, B. Harris, M. A. Quail, M.**  
210 **Achtman, R. Atkin, S. Baker, D. Basham, N. Bason, I. Cherevach, T.**  
211 **Chillingworth, M. Collins, A. Cronin, P. Davis, J. Doggett, T. Feltwell, A.**  
212 **Goble, N. Hamlin, H. Hauser, S. Holroyd, K. Jagels, S. Leather, S. Moule,**  
213 **H. Norberczak, S. O'Neil, D. Ormond, C. Price, E. Rabinowitsch, S.**  
214 **Rutter, M. Sanders, D. Saunders, K. Seeger, S. Sharp, M. Simmonds, J.**  
215 **Skelton, R. Squares, S. Squares, K. Stevens, L. Unwin, S. Whitehead, B. G.**

- 216 **Barrell, and D. J. Maskell.** 2003. Comparative analysis of the genome  
217 sequences of *Bordetella pertussis*, *Bordetella parapertussis* and *Bordetella*  
218 *bronchiseptica*. *Nat Genet* **35**:32-40.
- 219 21. **Parkhill, J., B. W. Wren, N. R. Thomson, R. W. Titball, M. T. Holden, M.**  
220 **B. Prentice, M. Sebaihia, K. D. James, C. Churcher, K. L. Mungall, S.**  
221 **Baker, D. Basham, S. D. Bentley, K. Brooks, A. M. Cerdeno-Tarraga, T.**  
222 **Chillingworth, A. Cronin, R. M. Davies, P. Davis, G. Dougan, T. Feltwell, N.**  
223 **Hamlin, S. Holroyd, K. Jagels, A. V. Karlyshev, S. Leather, S. Moule, P. C.**  
224 **Oyston, M. Quail, K. Rutherford, M. Simmonds, J. Skelton, K. Stevens, S.**  
225 **Whitehead, and B. G. Barrell.** 2001. Genome sequence of *Yersinia pestis*, the  
226 causative agent of plague. *Nature* **413**:523-7.
- 227 22. **Picard, B., J. S. Garcia, S. Gouriou, P. Duriez, N. Brahimi, E. Bingen, J.**  
228 **Elion, and E. Denamur.** 1999. The link between phylogeny and virulence in  
229 *Escherichia coli* extraintestinal infection. *Infect Immun* **67**:546-53.
- 230 23. **Quentin, R., H. Huet, F. S. Wang, P. Geslin, A. Goudeau, and R. K.**  
231 **Selander.** 1995. Characterization of *Streptococcus agalactiae* strains by  
232 multilocus enzyme genotype and serotype: identification of multiple virulent  
233 clone families that cause invasive neonatal disease. *J Clin Microbiol* **33**:2576-  
234 81.
- 235 24. **Rolland, K., C. Marois, V. Siquier, B. Cattier, and R. Quentin.** 1999.  
236 Genetic features of *Streptococcus agalactiae* strains causing severe neonatal  
237 infections, as revealed by pulsed-field gel electrophoresis and *hylB* gene  
238 analysis. *J Clin Microbiol* **37**:1892-8.
- 239 25. **Schrag, S. J.** 2004. The past and future of perinatal group B streptococcal  
240 disease prevention. *Clin Infect Dis* **39**:1136-8.

- 241 26. **Schuchat, A.** 2001. Group B streptococcal disease: from trials and tribulations  
242 to triumph and trepidation. *Clin Infect Dis* **33**:751-6.
- 243 27. **Sukhnanand, S., B. Dogan, M. O. Ayodele, R. N. Zadoks, M. P. Craver, N.**  
244 **B. Dumas, Y. H. Schukken, K. J. Boor, and M. Wiedmann.** 2005. Molecular  
245 subtyping and characterization of bovine and human *Streptococcus agalactiae*  
246 isolates. *J Clin Microbiol* **43**:1177-86.
- 247 28. **Tettelin, H., V. Massignani, M. J. Cieslewicz, C. Donati, D. Medini, N. L.**  
248 **Ward, S. V. Angiuoli, J. Crabtree, A. L. Jones, A. S. Durkin, R. T. Deboy,**  
249 **T. M. Davidsen, M. Mora, M. Scarselli, I. Margarit y Ros, J. D. Peterson, C.**  
250 **R. Hauser, J. P. Sundaram, W. C. Nelson, R. Madupu, L. M. Brinkac, R. J.**  
251 **Dodson, M. J. Rosovitz, S. A. Sullivan, S. C. Daugherty, D. H. Haft, J.**  
252 **Selengut, M. L. Gwinn, L. Zhou, N. Zafar, H. Khouri, D. Radune, G.**  
253 **Dimitrov, K. Watkins, K. J. O'Connor, S. Smith, T. R. Utterback, O. White,**  
254 **C. E. Rubens, G. Grandi, L. C. Madoff, D. L. Kasper, J. L. Telford, M. R.**  
255 **Wessels, R. Rappuoli, and C. M. Fraser.** 2005. Genome analysis of multiple  
256 pathogenic isolates of *Streptococcus agalactiae*: implications for the microbial  
257 "pan-genome". *Proc Natl Acad Sci USA* **102**:13950-5.
- 258 29. **Tettelin, H., V. Massignani, M. J. Cieslewicz, J. A. Eisen, S. Peterson, M. R.**  
259 **Wessels, I. T. Paulsen, K. E. Nelson, I. Margarit, T. D. Read, L. C. Madoff,**  
260 **A. M. Wolf, M. J. Beanan, L. M. Brinkac, S. C. Daugherty, R. T. DeBoy, A.**  
261 **S. Durkin, J. F. Kolonay, R. Madupu, M. R. Lewis, D. Radune, N. B.**  
262 **Fedorova, D. Scanlan, H. Khouri, S. Mulligan, H. A. Carty, R. T. Cline, S.**  
263 **E. Van Aken, J. Gill, M. Scarselli, M. Mora, E. T. Iacobini, C. Brettoni, G.**  
264 **Galli, M. Mariani, F. Vegni, D. Maione, D. Rinaudo, R. Rappuoli, J. L.**  
265 **Telford, D. L. Kasper, G. Grandi, and C. M. Fraser.** 2002. Complete genome

266 sequence and comparative genomic analysis of an emerging human pathogen,  
267 serotype V *Streptococcus agalactiae*. Proc Natl Acad Sci USA **99**:12391-6.  
268 30. **Toro, N.** 2003. Bacteria and Archaea Group II introns: additional mobile genetic  
269 elements in the environment. Environ Microbiol **5**:143-51.  
270

271 TABLE 1. Primers used for the amplification of transposase genes from the  
 272 *S. agalactiae* isolates.

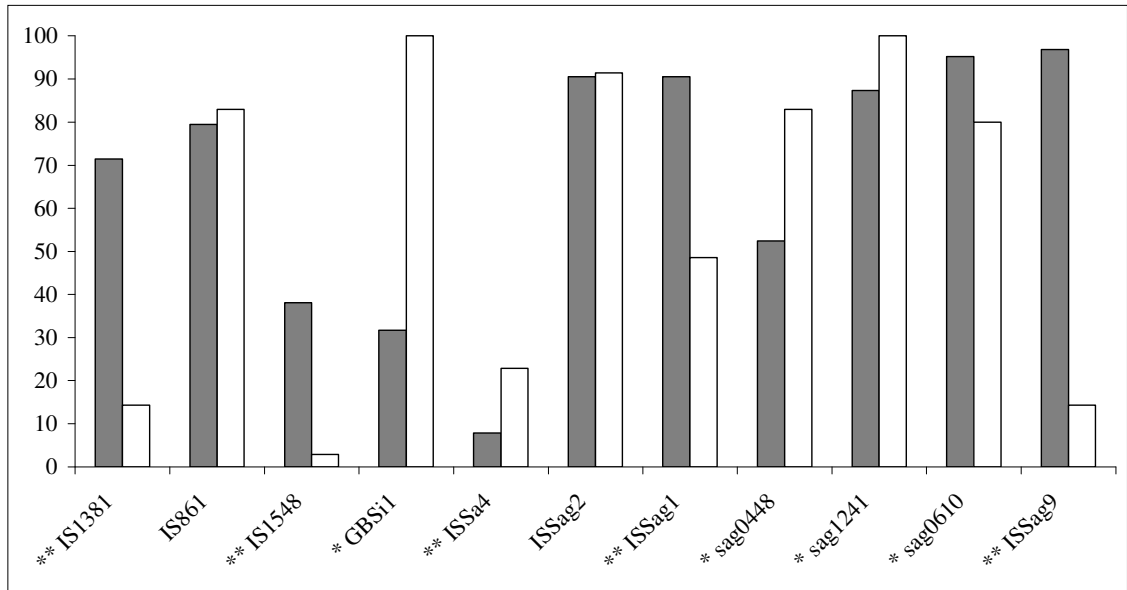
273

Target	Genbank accession no. <sup>a</sup>	Forward primer (5'-3')	Reverse primer (5'-3')	Tm (°C)	Amplicon (bp)
ISSag9	NC_004116	gaccgtaatgacatccaatctgg	ggcttgaaatggacacgggt	50	600
<i>sag0448</i>	NC_004116	ccacagaattacttaacttcttagc	cgaccaaccttattgaatcact	50	994
<i>sag1241</i> ( <i>orfA</i> )	NC_004116	ctcttgtaacctttatcaaactgg	cctaaaaaagcaagtgccca	50	218
<i>sag0610</i>	NC_004116	tactccacattagctaaagaacagg	ggcctgatttaaccgagatt	50	394

274

275 <sup>a</sup> NC\_004116 refers to strain 2603 V/R (29)

276 FIG. 1. Prevalence of the mobile genetic elements within the human (filled bars) and the  
 277 bovine (open bars) *S. agalactiae* isolates. Significant differences (without correction for  
 278 bias due to multiple comparisons) are indicated by one asterisk for p values below 0.05  
 279 and two asterisks for p values below 0.001.  
 280



281

282



283 FIG. 2. Relatedness of the 98 isolates of *S. agalactiae* analyzed by the NJ method, using  
284 MEGA v3.1 software and the Kimura 2-parameter mutation model of genetic distance.  
285 The nucleotide differences between the supergene sequences obtained in multilocus  
286 sequence analysis, as described by Jones *et al.* (11), were used to infer phylogenetic  
287 relationships between isolates. We used eBURST v3 software to group the 98 isolates  
288 into clonal complexes (CC). For each isolate, sequence type (ST),  $\beta$ -hemolysis, serotype  
289 and presence/absence of the 11 mobile genetic elements (MGEs) are indicated. Clonal  
290 complexes are also indicated on the right of the figure (several isolates were not  
291 attributed a CC number by the Start software and are considered “singletons”).

292

293

294

295

296

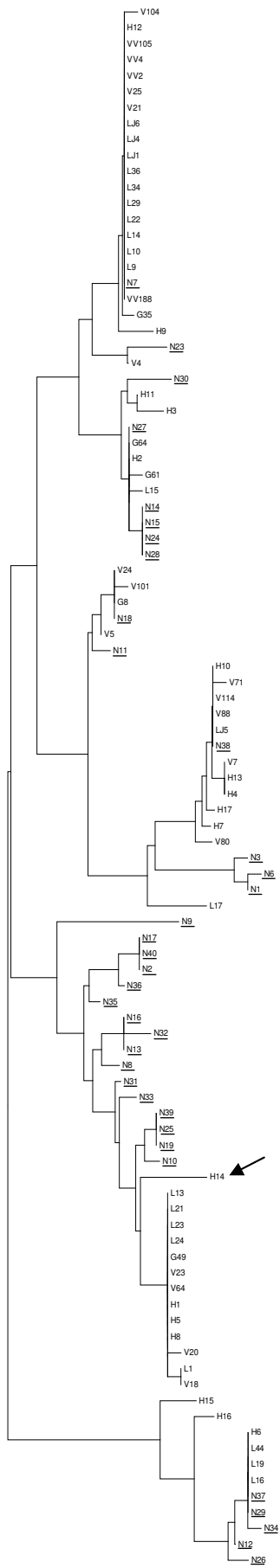
297 The particular position of isolate H14 is indicated by a black arrowhead.

298 The human isolates are named L or LJ (meningocerebral fluid), H (adult), V or VV  
299 (vaginal carriage) or G (gastric fluid); the bovine isolates are named N (bovine isolates  
300 are underlined).

301 NT: non typable

302 FIG. 3. Factorial analysis of correspondence of 98 *S. agalactiae* isolates based on the  
303 mobile genetic element (MGE) data. Eleven active variables [the MGEs (▲) *IS1381*,  
304 *IS861*, *IS1548*, *GBSi1*, *ISSa4*, *ISSag2*, *ISSag1*, *sag0448*, *sag1241*, *sag0610* and  
305 *ISSag9*], and five illustrative variables [the main clonal complexes CC19 (◆), CC17 (◆)  
306 and CC67 (◆), and the human (●) and bovine (■) origins of the isolates], and the 98  
307 isolates are projected onto the F1/F2 plane. This plane, obtained by computation, is  
308 defined by the two principal axes of the analysis; the first axis, F1, explains most of the  
309 variance, and the second axis, F2 (orthogonal to F1), explains most of the variance not  
310 explained by F1. Together, F1 and F2 accounted for 40.97% of the total variance.  
311 Isolates of CC17 (n=13) are shown in red, isolates of CC67 (n=15) in green, and isolates  
312 of subCC19 (n=21) in pink.  
313





Isolate	Sequence Type	β-haemolysis	Serotype	MGEs										
				GBS1	IS1381	IS861	IS1548	ISSa4	ISSag2	ISSag1	sag0448	sag1241	sag0610	ISSag9
V104	ST 28	+	II	+	+	+	-	-	+	+	+	+	+	+
H12	ST 19	+	III	-	+	+	+	-	-	+	+	+	+	+
VV105	ST 19	+	III	-	+	+	+	-	-	+	+	+	+	+
VV4	ST 19	+	III	-	+	+	+	-	-	+	+	+	+	+
VV2	ST 19	+	III	-	+	+	+	-	-	+	+	+	+	+
V25	ST 19	+	NT	+	+	+	-	-	+	-	+	+	+	+
V21	ST 19	+	II	+	+	+	-	-	+	-	+	+	+	+
LJ6	ST 19	+	III	-	+	+	+	-	-	+	+	+	+	+
LJ4	ST 19	+	III	-	+	+	+	-	-	+	+	+	+	+
LJ1	ST 19	+	III	-	+	+	+	-	-	+	+	+	+	+
L36	ST 19	+	III	-	+	+	+	-	-	+	+	+	+	+
L34	ST 19	+	III	-	+	+	+	-	-	+	+	+	+	+
L29	ST 19	+	III	-	+	+	+	-	-	+	+	+	+	+
L22	ST 19	+	III	-	+	+	+	-	-	+	+	+	+	+
L14	ST 19	+	III	-	+	+	+	-	-	+	+	+	+	+
L10	ST 19	+	III	-	+	+	+	-	-	+	+	+	+	+
L9	ST 19	+	III	-	+	+	+	-	-	+	+	+	+	+
NZ	ST 19	+	III	+	+	+	+	-	-	+	+	+	+	-
VV188	ST 19	+	III	-	+	+	+	-	-	+	+	+	+	+
G35	ST 197	+	NT	+	+	+	+	-	-	+	+	-	+	+
H9	ST 131	+	V	+	+	+	+	-	-	+	+	+	+	+
N23	ST 226	+	II	+	-	-	-	-	-	+	+	+	+	+
V4	ST 366	+	Ia	-	-	-	+	-	-	+	+	+	+	-
N30	ST 261	+	IV	-	-	-	-	-	-	+	+	+	+	-
H11	ST 1	+	NT	-	-	-	-	-	-	+	+	+	+	+
H3	ST 186	+	NT	-	+	-	-	-	-	+	+	-	+	+
N27	ST 2	-	II	+	+	+	-	+	+	+	+	+	+	+
G64	ST 2	+	II	+	+	+	-	+	+	+	+	+	+	+
H2	ST 2	+	NT	+	+	-	-	-	-	+	+	-	+	+
G61	ST 202	+	V	-	-	-	-	-	-	+	+	+	+	+
L15	ST 196	+	IV	-	+	-	-	-	-	+	+	+	+	+
N15	ST 250	+	NT	+	+	-	-	-	-	+	+	+	+	-
N14	ST 250	+	NT	+	+	-	-	-	-	+	+	+	+	-
N24	ST 357	+	NT	+	+	+	+	+	+	+	+	+	+	-
N28	ST 250	+	NT	+	+	+	-	-	-	+	+	+	+	+
V24	ST 7	+	IV	-	+	+	-	-	-	+	+	+	+	+
V101	ST 41	+	V	-	+	+	-	-	-	+	+	+	+	+
G8	ST 7	+	Ia	+	+	+	-	-	-	+	+	+	+	+
N18	ST 7	+	Ia	+	+	+	-	-	-	+	+	+	+	+
V5	ST 6	+	Ib	+	+	-	-	-	-	+	+	+	+	+
N11	ST 354	+	Ib	+	+	+	-	-	-	+	+	+	+	+
H10	ST 10	+	Ib	+	+	+	-	-	-	+	+	+	+	+
V71	ST 8	+	Ib	-	+	+	-	-	-	+	+	+	+	+
V114	ST 10	+	Ib	-	+	+	+	-	-	+	+	+	+	+
V88	ST 10	+	II	-	+	+	-	-	-	+	+	+	+	-
LJ5	ST 10	+	Ib	-	+	+	-	-	-	+	+	+	+	+
N38	ST 10	+	Ib	+	-	+	-	-	-	+	+	+	+	-
V7	ST 12	+	Ib	-	+	+	-	-	-	+	+	+	+	+
H13	ST 12	+	Ib	-	+	+	-	-	-	+	+	+	+	+
H4	ST 12	+	Ib	-	+	+	-	-	-	+	+	+	+	+
H17	ST 200	+	Ib	-	+	+	-	-	-	+	+	+	+	+
H7	ST 10	+	II	-	+	+	+	-	-	+	+	+	+	+
V80	ST 12	+	II	-	+	+	+	+	+	+	+	+	+	+
N3	ST 952	-	NT	+	-	-	-	-	-	+	+	+	+	+
N6	ST 353	-	NT	+	-	+	-	-	-	+	+	+	+	-
N1	ST 350	+	NT	+	-	+	-	-	-	+	+	+	+	-
L17	ST 195	+	III	+	+	+	-	-	-	+	+	+	+	+
N9	ST 26	+	V	+	+	+	-	-	-	+	+	+	+	+
N17	ST 351	-	II	+	-	+	-	-	-	+	+	+	+	-
N40	ST 351	-	NT	+	-	+	-	-	-	+	+	+	+	-
N2	ST 351	-	II	+	-	+	-	-	-	+	+	+	+	-
N36	ST 365	+	II	+	-	+	-	-	-	+	+	+	+	-
N35	ST 364	-	NT	+	-	+	-	-	-	+	+	+	+	-
N16	ST 67	+	NT	+	+	-	-	-	-	+	+	+	+	-
N32	ST 85	+	NT	+	+	-	-	-	-	+	+	+	+	-
N13	ST 67	+	NT	+	+	-	-	-	-	+	+	+	+	-
N8	ST 64	-	II	+	-	+	-	-	-	+	+	+	+	-
N31	ST 361	-	III	+	-	+	-	-	-	+	+	+	+	-
N33	ST 362	-	III	+	-	+	-	-	-	+	+	+	+	-
N39	ST 61	+	NT	+	-	+	-	-	-	+	+	+	+	-
N25	ST 61	+	III	+	+	-	-	-	-	+	+	+	+	-
N19	ST 61	+	NT	+	+	-	-	-	-	+	+	+	+	-
N10	ST 63	+	II	+	+	+	-	-	-	+	+	+	+	-
H14	ST 22	+	II	+	+	+	-	-	-	+	+	+	+	-
L13	ST 17	+	III	+	+	+	-	-	-	+	+	+	+	+
L21	ST 17	+	III	+	+	+	-	-	-	+	+	+	+	+
L23	ST 17	+	III	+	+	+	-	-	-	+	+	+	+	+
L24	ST 17	+	III	+	+	+	-	-	-	+	+	+	+	+
G49	ST 17	+	III	+	+	+	-	-	-	+	+	+	+	+
V23	ST 17	+	NT	+	+	+	+	-	-	+	+	+	+	+
V64	ST 17	+	III	+	+	+	-	-	-	+	+	+	+	+
H1	ST 17	+	III	+	+	+	-	-	-	+	+	+	+	+
H5	ST 17	+	III	+	+	+	-	-	-	+	+	+	+	+
H8	ST 17	+	III	+	+	+	-	-	-	+	+	+	+	+
V20	ST 367	+	III	+	-	+	-	-	-	+	+	+	+	+
L1	ST 201	+	III	+	+	+	-	-	-	+	+	+	+	-
V18	ST 201	+	III	+	+	+	-	-	-	+	+	+	+	-
H15	ST 198	+	III	-	-	-	-	-	-	+	+	+	+	+
H16	ST 199	+	III	-	-	-	-	-	-	+	+	+	+	+
H6	ST 23	+	III	-	-	-	-	-	-	+	+	+	+	+
L44	ST 23	+	Ia	+	+	+	-	-	-	+	+	+	+	+
L19	ST 23	+	Ia	+	+	-	-	-	-	+	+	+	+	+
L16	ST 23	+	III	-	-	-	-	-	-	+	+	+	+	+
N37	ST 23	+	III	+	+	+	-	-	-	+	+	+	+	-
N29	ST 23	+	III	+	-	+	-	-	-	+	+	+	+	-
N34	ST 363	+	III	+	-	+	-	-	-	+	+	+	+	-
N12	ST 355	+	III	+	+	+	-	-	-	+	+	+	+	+
N26	ST 358	+	III	+	-	-	-	-	-	+	+	+	+	+

subCC19

subCC2

CC7

CC10

CC67

CC17

CC23

0.001

Axis F2

