

## Identification of *Streptococcus agalactiae* isolates from various phylogenetic lineages by matrix-assisted laser desorption ionization-time of flight mass spectrometry

Marie-Frédérique Lartigue, Geneviève Héry-Arnaud, Eve Haguenoer, Anne-Sophie Domelier, Pierre-Olivier Schmit, Nathalie Van Der Mee-Marquet, Philippe Lanotte, Laurent Mereghetti, Markus Kostrzewa, Roland Quentin

### ► To cite this version:

Marie-Frédérique Lartigue, Geneviève Héry-Arnaud, Eve Haguenoer, Anne-Sophie Domelier, Pierre-Olivier Schmit, et al. Identification of *Streptococcus agalactiae* isolates from various phylogenetic lineages by matrix-assisted laser desorption ionization-time of flight mass spectrometry. *Journal of Clinical Microbiology, American Society for Microbiology*, 2009, 47 (7), pp.2284-2287. <10.1128/JCM.00175-09>. <hal-00557578>

**HAL Id: hal-00557578**

**<http://hal.univ-brest.fr/hal-00557578>**

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 **JCM00175-09 - Revised Manuscript.**

2 **Identification of *Streptococcus agalactiae* isolates from various phylogenetic**  
3 **lineages by Matrix-Assisted Laser Desorption/Ionization -Time of Flight Mass**  
4 **Spectrometry**

5 Marie-Frédérique Lartigue<sup>1,2</sup>, Geneviève Héry-Arnaud<sup>1</sup>, Eve Haguenaer<sup>1</sup>, Anne-  
6 Sophie Domelier<sup>1,2</sup>, Pierre-Olivier Schmit<sup>3</sup>, Nathalie van der Mee-Marquet<sup>1,2</sup>, Philippe  
7 Lanotte<sup>1</sup>, Laurent Mereghetti<sup>1</sup>, Markus Kostrzewa<sup>4</sup>, and Roland Quentin<sup>1,2</sup>.

8

9 <sup>1</sup> Université François-Rabelais de Tours, UFR de Médecine, EA 3854 « Bactéries et  
10 risque materno-foetal », Institut Fédératif de Recherche 136 « Agents Transmissibles et  
11 Infectiologie », 37032 Tours Cedex, France

12 <sup>2</sup> Service de Bactériologie et Hygiène Hospitalière, Hôpital Trousseau, CHRU de Tours,  
13 37044 Tours, France.

14 <sup>3</sup> Bruker Daltonique, 34 rue de l'Industrie, 67166 Wissembourg Cedex, France

15 <sup>4</sup> Bruker Daltonik GmbH, Permoser Str. 15, D-04318 Leipzig, Germany

16

17 **Corresponding author. Mailing address :**

18 Marie-Frédérique Lartigue

19 Service de Bactériologie-Hygiène

20 CHRU Trousseau

21 37044 Tours cedex 9

22 France.

23 Phone : 33 2 47 47 81 13

24 Fax : 33 2 47 47 85 30

25 E-mail : lartigue@med.univ-tours.fr

26 Abstract

27 Variations in proteins related to bacterial diversity may affect species  
28 identification performed using matrix-assisted laser desorption/ionization-time of flight  
29 mass spectrometry (MALDI-TOF MS). Using this method, we identified 110  
30 *Streptococcus agalactiae* isolates characterized by serotyping and MLST. Serotype III  
31 and ST23 strains expressed the widest variations in molecular weights of putative  
32 “species-identifying” biomarker ions. Recognition of the diversity of MALDI patterns  
33 observed in strains that represent all major intraspecies lineages assists in the  
34 constitution of an optimal reference database.

35 *Streptococcus agalactiae* is the main cause of neonatal infections and has  
36 emerged as an increasingly frequent pathogen in non-pregnant humans (15). Several  
37 studies have identified diversity in *S. agalactiae* species, first by serotyping, and then  
38 extensively by multilocus enzyme electrophoresis (MLEE) (12, 13). A large number of  
39 different Sequence Types (STs) distributed over several major phylogenetic lineages or  
40 clonal complexes (CC) have recently been identified by multilocus sequence typing  
41 (MLST) (<http://pubmlst.org/sagalactiae/>). STs 17 and 19 account for the majority of  
42 cases of *S. agalactiae* meningitis in infants (7). CCs 1, 12, 17, 19 and 23 are mostly  
43 associated with infections in adults (7).

44 Matrix-assisted laser desorption/ionization-time of flight mass spectrometry  
45 (MALDI-TOF MS) has emerged as a new technique for species identification. By  
46 measuring the exact molecular masses of many peptides and small proteins, it is  
47 possible to determine the species within a few minutes, whether the analysis is started  
48 with whole cells, cell lysates, or crude bacterial extracts (4, 6, 9, 11). Nevertheless, wide  
49 variations in protein expression have been reported, specifically according to the  
50 distribution of strains in various phylogenetic lineages that make up many species. For  
51 example, we have shown such variations in the expression of metabolic enzymes (12,  
52 13), in catabolic functions (3), in the expression of surface-exposed bacterial proteins  
53 involved in the adhesion and/or invasion of host cells (14). In addition, the diversity of  
54 the rRNA gene region found in *S. agalactiae* species (2) may induce variations in the  
55 expression of the ribosomal proteins detected by MALDI-TOF MS, a method based  
56 mainly on the detection of ribosomal protein fractions of bacteria (1, 4, 11).

57 The aim of this study was therefore to determine whether variations in protein  
58 expression related to the phylogenetic position of strains affect the results obtained by  
59 MALDI-TOF MS when used to identify *S. agalactiae* isolates. One hundred and ten

60 strains were selected from an epidemiologically unrelated national collection (10, 13)  
61 isolated from the vagina, the anatomical site at which the genetic diversity of strains is  
62 the highest. As usually performed in medical laboratories, strains were plated on 5%  
63 sheep's blood agar (laboratoire bioMérieux, Marcy l'Etoile, France). Serotyping was  
64 performed with a Pastorex rapid latex agglutination test (Bio-Rad, Hercules, Calif.) and  
65 by a previously reported PCR serotype identification method (8). Forty-five isolates  
66 were from serotype III (40.9% of isolates), 17 from serotype II (15.5%), 16 from  
67 serotype V (14.5%), 15 from serotype Ia (13.6%), 7 from serotype Ib (6.4%), and 7  
68 from serotype IV (6.4%). Three isolates were not typeable (2.7%). MLST analysis,  
69 carried out with the standard MLST scheme (7), identified 38 STs for the 110 strains  
70 (Fig. 1). The relationships between STs were defined by UPGMA analysis and  
71 represented as a tree generated from allelic profile data using Phylodendron  
72 (<http://pubmlst.org/sagalactiae/>) (Fig. 1). By adding together data from serotyping and  
73 MLST, 52 patterns were obtained that represent the wide diversity of the *S. agalactiae*  
74 population studied and the major STs and CCs (Fig. 1).

75 For MALDI-TOF-MS analysis, cell extracts were prepared using ten colonies for  
76 each preparation. Samples were prepared according to the microorganism profiling  
77 « ethanol/formic acid extraction » procedure as recommended by the manufacturer, with  
78 minor modifications as recently described (1). After drying and addition of a chemical  
79 matrix, the samples were analyzed by MALDI-TOF MS on a Bruker Ultraflex  
80 TOF/TOF III in positive linear mode (Bruker Daltonique, Wissembourg, France). The  
81 spectra were recorded as recently described (1). For each spectrum, 500 laser shots were  
82 collected and analyzed. For automated data analysis, raw spectra were processed using  
83 the MALDI Biotyper 1.1 software (Bruker Daltonique, France) with default settings. To  
84 identify bacteria, the peak lists generated were used for matches against the initial

85 MALDI Biotyper reference library (*S. agalactiae* ATCC 27956 THL and *S. agalactiae*  
86 DSM 2134) directly using the integrated pattern-matching algorithm of the software (1).  
87 In a typical analysis of *S. agalactiae* strains by MALDI-TOF MS, 70 prominent ion  
88 peaks were noted in the spectra in the region between 2,000 and 20,000 Da, the highest-  
89 intensity peaks being consistently in the range of 3,000 to 10,000 Da. The log(score) of  
90 the MALDI Biotyper pattern matching algorithm is calculated according to the log of  
91 the product of three factors: the matches of the unknown spectrum against the reference  
92 spectrum in the database (main spectrum), the reverse matches of the main spectrum  
93 with the unknown spectrum, and the correlation of relative intensities of the unknown  
94 spectrum and the main spectrum. The product has a maximum value of 1000, leading to  
95 a maximum log(score) of 3. Differences in the distribution of the strains in log (score)  
96 groups were tested by the Chi<sup>2</sup> test (STAT-ITCF software, Paris). A p value < 0.05 was  
97 considered to indicate statistical significance.

98 On this basis, the log (score) values obtained by MALDI-TOF MS correctly  
99 identified all 110 *S. agalactiae* isolates at the species level [log (score),  $\geq 2.0$ ]; 86 of the  
100 110 (78%) being identified with excellent scores [log (score),  $\geq 2.3$ ] (Fig. 1, Table 1).  
101 Nevertheless, significant variations in the log (score) were observed according to  
102 serotypes and major STs. Indeed, excellent scores [log (score)  $\geq 2.3$ ] were obtained  
103 significantly less frequently (64%) for serotype III strains than for the other strains (73  
104 to 100%) (Table 1) ( $p = 0.0027$ ). Similarly, excellent scores [log (score),  $\geq 2.3$ ] were  
105 obtained less frequently for ST23 strains (43%) than for the other major STs (77 to  
106 88%) (Table 1) (Chi<sup>2</sup> test,  $p = 0.025$ ).

107 To test the reproducibility of the method, five strains representative of the five  
108 main STs (ST1, ST10, ST17, ST19, and ST23) were grown in quintuplicate and  
109 analyzed. The mass spectrum was measured five times for each replicate of each of the

110 five representative strains of main STs (125 measurements). The log(score) calculated  
111 using the mass spectra data allowed determination of the coefficient of variation,  
112  $CV=\delta/\mu$  in which  $\delta$  is the standard deviation and  $\mu$  is the average. The CV for intra runs  
113 (5 measurements for each replicate) varied from 0.006 to 0.025 and the highest CV for  
114 inter runs (5 different replicas) was 0.027. Moreover, reproducibility tests were also  
115 performed by cultivation of bacteria on different culture media (5% sheep blood agar,  
116 CPS ID 2<sup>®</sup>, chocolate PVX agar (bioMérieux), Granada<sup>®</sup> (Biolys, Taluyers, France),  
117 and Mueller Hinton agar). A strain of GBS ST17 was grown in triplicate on each  
118 medium. All the log(score) were above 2.3. The low values of the intra runs and inter  
119 runs CV and the low variations in the scores obtained using various culture media  
120 indicate that the reproducibility of the method was high. Thus the variations found  
121 between populations of various serotypes or STs could not be related to the procedure  
122 used for preparing the samples.

123 This study confirmed variations in expression of proteins according to the  
124 distribution of strains in serotypes and in various phylogenetic lineages that make up *S.*  
125 *agalactiae* species. Strains of serotype III and of the phylogenetic lineage ST23  
126 expressed greater variations in the molecular weights of putative “species-identifying”  
127 biomarker ions than strains of other serotypes and STs. The ability to produce an intense  
128 peak in the protein pattern correlates with the ionization efficiency, combined with the  
129 protein quantity used. The variations observed for serotype III strains may be explained  
130 by the greater genetic diversity of this serotype population compared to the other  
131 serotype populations, as shown in Figure 1 and as previously described (7, 12, 13). As  
132 shown previously, CC23 (that mainly contains ST23 strains) is genetically a highly  
133 divergent clone in the species (Fig. 1) (5, 7, 16). This characteristic may explain why  
134 ST23 strains express differences in the nature or quantities of proteins when compared

135 to other major ST populations that are genetically more closely related. These variations  
136 did not affect the ability of MALDI-TOF MS combined with the MALDI Biotyper  
137 software method to identify *S. agalactiae* at a species level; log (score) values were  $\geq 2.0$   
138 for all strains. Nevertheless, to optimize the ability of MALDI-TOF MS to identify *S.*  
139 *agalactiae* strains whatever their phylogenetic origin, we selected five strains to  
140 constitute a new reference database on the basis of the variations observed for each  
141 lineage. These strains were retained because they generated the best representative  
142 spectra for each major intraspecies lineage (ST1, ST10, ST17, ST19, and ST23). The  
143 log (score) values obtained by using the new reference database identified 109 of the  
144 110 *S. agalactiae* isolates (>99%) with a higher log (score) ( $\geq 2.3$ ) (Fig. 1, Table 1).

145 In conclusion, genetic diversity between bacterial species affects the diversity of  
146 MALDI patterns observed when using MALDI-TOF MS. Consequently, the most  
147 effective performance for identification of bacterial strains at the species level was  
148 obtained by using a reference database designed, not by selecting one or two strains  
149 randomly, but by choosing them from the major phylogenetic lineages that represent the  
150 species studied.

151

152 Conceived and designed the experiments: MFL, RQ. Performed the experiments: MFL,  
153 GH, EH, ASD, NVDM, PL, LM (serotyping and MLST), MFL, POS (MALDI-TOF  
154 MS). Analyzed the data: MFL, MK, RQ. Wrote the paper: MFL, RQ.

155

156

157 **Acknowledgments.** This study was supported by the Ministère de l'Enseignement  
158 Supérieur et de la Recherche and by the Centre Hospitalier Universitaire de Tours.



159 **REFERENCES**

160

- 161 1. **Barbuddhe, S.B., Maier, T., Schwarz, G., Kostrzewa, M., Hof, H., Domann,**  
162 **E., Chakraborty, T., and T. Hain.** 2008. Rapid Identification and Typing of  
163 *Listeria* species using Matrix Assisted Laser Desorption Ionization-Time of  
164 Flight Mass Spectrometry. *Appl. Environ. Microbiol.* **74**:5402-5407.
- 165 2. **Chatellier, S., Huet, H., Kenzi, S., Rosenau, A., Geslin, P., and R. Quentin.**  
166 1996. Genetic diversity of rRNA operons of unrelated *Streptococcus agalactiae*  
167 strains isolated from cerebrospinal fluid of neonates suffering from meningitis.  
168 *J. Clin. Microbiol.* **34**:2741-2747.
- 169 3. **Domelier, A.S., van der Mee-Marquet, N., Grandet, A., Mereghetti, L.,**  
170 **Rosenau, A., and R. Quentin.** 2006. Loss of catabolic function in  
171 *Streptococcus agalactiae* strains and its association with neonatal meningitis. *J.*  
172 *Clin. Microbiol.* **44**:3245-3250.
- 173 4. **Fenselau, C., and P.A. Demirev.** 2001. Characterization of intact  
174 microorganisms by MALDI mass spectrometry. *Mass. Spectrom. Rev.* **20**:157-  
175 171.
- 176 5. **Héry-Arnaud, G., Bruant, G., Lanotte, P., Brun, S., Picard, B., Rosenau, A.,**  
177 **van der Mee-Marquet, N., Rainard, P., Quentin, R., and L. Mereghetti.**  
178 2007. Mobile genetic elements provide evidence for a bovine origin of clonal  
179 complex 17 of *Streptococcus agalactiae*. *Appl. Environ. Microbiol.* **73**:4668-  
180 4672.
- 181 6. **Holland, R.D., Wilkes, J.G., Rafii, F., Sutherland, J.B., Persons, C.C.,**  
182 **Voorhees, K.J., and J.O.J. Lay.** 1996. Rapid identification of intact whole  
183 bacteria based on spectral patterns using matrix-assisted laser

- 184 desorption/ionization with time-of-flight mass spectrometry. Rapid. Commun.  
185 Mass. Spectrom. **10**:1227-1232.
- 186 7. **Jones, N., Bohnsack, J.F., Takahashi, S., Oliver, K.A., Chan, M.S., Kunst,**  
187 **F., Glaser, P., Rusniok, C., Crook, D.W., Harding, R.M., Bisharat, N., and**  
188 **B.G. Spratt.** 2003. Multilocus sequence typing system for group B  
189 *Streptococcus*. J. Clin. Microbiol. **41**:2530-2536.
- 190 8. **Kong, F., Gowan, S., Martin, D., James, G., and G.L. Gilbert.** 2002.  
191 Serotype identification of group B streptococci by PCR and sequencing. J. Clin.  
192 Microbiol. **40**: 216-226.
- 193 9. **Krishnamurthy, T., Ross, P.L., and U. Rajamani.** 1996. Detection of  
194 pathogenic and non-pathogenic bacteria by matrix-assisted laser  
195 desorption/ionization time-of-flight mass spectrometry. Rapid. Commun. Mass.  
196 Spectrom. **10**:883-888.
- 197 10. **Loulergue, J., Couhé, C., Grasmick, C., Laudat., P., and R. Quentin.** 2004.  
198 Sensibilité aux antibiotiques des souches de streptocoque du groupe B de  
199 portage vaginal isolées en France, 2003. B.E.H. **18**:69-70.
- 200 11. **Mellmann, A., Cloud, J., Maier, T., Keckevoet, U., Ramminger, I., Iwen, P.,**  
201 **Dunn, J., Hall, G., Wilson, D., Lasala, P., Kostrzewa, M., and D. Harmsen.**  
202 2008. Evaluation of matrix-assisted laser desorption ionization-time-of-flight  
203 mass spectrometry in comparison to 16S rRNA gene sequencing for species  
204 identification of nonfermenting bacteria. J. Clin. Microbiol. **46**:1946-1954.
- 205 12. **Musser, J.M., Mattingly, S.J., Quentin, R., Goudeau, A., and R.K. Selander.**  
206 1989. Identification of a high-virulence clone of type III *Streptococcus*  
207 *agalactiae* (group B *Streptococcus*) causing invasive neonatal disease. Proc.  
208 Natl. Acad. Sci. U. S. A. **86**: 4731-4735.

- 209 13. **Quentin, R., Huet, H., Wang, F.S., Geslin, P., Goudeau, A., and R.K.**  
210 **Selander.** 1995. Characterization of *Streptococcus agalactiae* strains by  
211 multilocus enzyme genotype and serotype: identification of multiple virulent  
212 clone families that cause invasive neonatal disease. J. Clin. Microbiol. **33**:2576-  
213 2581.
- 214 14. **Rosenau, A., Martins, K., Amor, S., Gannier, F., Lanotte, P., van der Mee-**  
215 **Marquet, N., Mereghetti, L., and R. Quentin.** 2007. Evaluation of the ability  
216 of *Streptococcus agalactiae* strains isolated from genital and neonatal  
217 specimens to bind to human fibrinogen and correlation with characteristics of  
218 the *fbxA* and *fbxB* genes. Infect. Immun. **75**:1310-1317.
- 219 15. **Schuchat, A.** 1998. Epidemiology of group B streptococcal disease in the  
220 United States: shifting paradigms. Clin. Microbiol. Rev. **11**:497-513.
- 221 16. **Van der Mee-Marquet N, Fourny L, Arnault L, Domelier AS, Salloum M,**  
222 **Lartigue MF, Quentin R.** 2008. Molecular characterization of human-  
223 colonizing *Streptococcus agalactiae* strains isolated from throat, skin, anal  
224 margin, and genital body sites. J. Clin. Microbiol. **46**:2906-2911

225 **Figure Legend**

226

227 Figure 1. MLST UPGMA tree of 110 *S. agalactiae* strains used to test MALDI-TOF  
228 MS for identification of *S. agalactiae* strains and log (score) obtained for each strain  
229 analyzed.

230 <sup>a</sup> The MALDI-TOF MS analysis calculates the log (score) that allows classification of  
231 the results of identification in four categories: i) a log (score)  $\geq 2.3$  indicates a highly  
232 probable level of species identification, ii) a log (score)  $\leq 2.299$  and  $\geq 2.0$  indicates a  
233 highly probable level of genus identification and probable species identification, iii) a  
234 log (score)  $\leq 1.999$  and  $\geq 1.7$  indicates a probable genus identification and iv) a log  
235 (score)  $\leq 1.699$  indicates non-reliable identification. Whatever the phylogenetic origin of  
236 the strain tested, no log (score)  $\leq 1.999$  was found.

237 <sup>b</sup> Initial MALDI Biotyper reference library consisted of main spectra of *S. agalactiae*  
238 ATCC 27956 THL and *S. agalactiae* DSM 2134.

239 <sup>c</sup> New reference library consisted of the 5 newly created mass spectra for *S. agalactiae*  
240 strains representative of each major intraspecies lineage (ST1, ST10, ST17, ST19, and  
241 ST23).

242 Table 1. Identification scores produced by the pattern-matching algorithm against the  
 243 initial MALDI Biotyper reference library (main spectra of *S. agalactiae* ATCC 27956  
 244 THL and *S. agalactiae* DSM 2134) and against a new reference library (mass spectra of  
 245 five *S. agalactiae* strains selected to represent each major intraspecies lineage ST1,  
 246 ST10, ST17, ST19, and ST2) for 110 *S. agalactiae* strains, according to serotypes and  
 247 major intraspecies STs.

	Log(score)			
	Initial reference library		New reference library	
<i>S. agalactiae</i>	$\geq 2.3$	$< 2.3 - \geq 2.0$	$\geq 2.3$	$< 2.3 - \geq 2.0$
strains (N)	N (%)	N (%)	N (%)	N (%)
All (110)	86 (78)	24 (22)	109 (99)	1 (1)
Serotype Ia (15)	11 (73)	4 (27)	15 (100)	0
Serotype Ib (7)	7 (100)	0	7 (100)	0
Serotype II (17)	15 (88)	2 (12)	17 (100)	0
Serotype III (45)	29 (64)	16 (36)	44 (98)	1 (2)
Serotype IV (7)	7 (100)	0	7 (100)	0
Serotype V (16)	15 (94)	1 (6)	16 (100)	0
ST1 (16)	13 (81)	3 (19)	16 (100)	0
ST10 (8)	7 (88)	1 (12)	8 (100)	0
ST17 (15)	12 (80)	3 (20)	15 (100)	0
ST19 (13)	10 (77)	3 (23)	13 (100)	0
ST23 (7)	3 (43)	4 (57)	7 (100)	0

248

249

250

