JOURNAL OF CLINICAL MICROBIOLOGY, Mar. 2005, p. 1433–1436 Vol. 43, No. 3
0095-1137/05/$08.00+0 doi:10.1128/JCM.43.3.1433–1436.2005

Survey of *emm*-Like Gene Sequences from Pharyngeal Isolates of Group C and Group G Streptococci Collected in Spain

Sebastián Albertí,[1]* Cesar García-Rey,[2] María Isabel García-Laorden,[1] Rafael Dal-Ré,[2] Juan García-de-Lomas,[3] and the Spanish Surveillance Group for Respiratory Pathogens†

Institut Universitari d'Investigacions en Ciències de la Salut and Departament de Biologia, Universitat de les Illes Balears, Palma de Mallorca,[1] Departamento Médico, GlaxoSmithKline, Tres Cantos,[2] and Servicio de Microbiología, Hospital Clínico Universitario y Facultad de Medicina, Instituto Valenciano de Microbiología, Valencia,[3] Spain

Received 30 July 2004/Returned for modification 15 October 2004/Accepted 20 October 2004

We conducted a nationwide surveillance of the variable 5′ *emm*-like (M-like protein gene) sequences from 214 pharyngeal group C and group G streptococci. Almost 75% of the isolates exhibited *emm* or *emm*-like sequences previously described. We identified six new 5′ *emm*-like regions, and almost 23% of the isolates were nontypeable. Five *emm*-like sequences accounted for more than 50% of the isolates in group C and group G, suggesting horizontal gene transfer between strains of different species.

Group C beta-hemolytic streptococci (GCS) and group G beta-hemolytic streptococci (GGS) form a heterogeneous collection of microorganisms, which, unlike other beta-hemolytic streptococci, such as group A streptococci (GAS), can be isolated from either human or animal specimens (10). In humans, GCS and GGS may cause pharyngitis and skin and soft tissue infections, although not as frequently as GAS (6, 9, 18).

Human-associated strains of GGS and GCS may present virulence factors similar to those expressed by GAS, including the major virulence factor of these important pathogens, the M protein (5, 7, 12, 16).

M proteins of GAS exhibit a substantial polymorphism due to the high degree of heterogeneity among the N termini of diverse M proteins, leading to more than 80 different serological M or OF types. The variable sequences of surface-exposed amino termini of M proteins may be deduced by sequencing the 5′ M protein-encoding gene, *emm*. Thus, since the *emm* gene's possible polymorphism was first reported, in 1995 (4), its gene sequence analysis has proven quite useful for routine M typing of GAS and many epidemiological studies have applied this methodology for monitoring GAS strain diversity (2, 3, 8). M proteins of GAS, GCS, and GGS recovered from human infections have similar biological, immunochemical, and structural features (5–7, 12, 16). Collins et al. reported the first complete sequence of an M-like protein gene of a human GGS strain (7). Later, Smirnov et al. published the sequence of a GGS gene encoding an immunoglobulin G-binding protein (17). More recently, Bisno et al. reported that *emm* gene sequences from GCS pharyngeal isolates were highly homologous to those found in GGS (6). The proteins encoded by these genes exhibited features similar to those present in the GAS M proteins, but they differed significantly in the N terminal region. These data suggest that, like GAS, GGS and GCS can exhibit various M-like proteins encoded by *emm*-like genes (*emmL*). Based on these observations, Schnitzler et al. extended the earlier studies and demonstrated that the *emmL* gene from 38 epidemiologically unrelated GGS strains isolated from human specimens displayed a substantial diversity, suggesting that this polymorphism could provide the basis for a molecular typing of GGS (15).

In the present work, we analyzed a larger number of GCS and GGS isolates for the variability of the 5′ region of the *emmL* to study the M-like type distribution of these isolates in Spain.

Bacterial strains. GCS and GGS isolates included in this study belong to the SAUCE II (November 1998 to October 1999) surveillance collection and were isolated from patients with acute pharyngitis in 17 different hospitals selected on the basis of population and geographic location in Spain (13). Among 2,292 beta-hemolytic streptococcus isolates collected during this period, 155 (6.8%) and 59 (2.6%) were GCS and GGS, respectively. Isolates were tested initially for beta-hemolysis by using 5% (vol:vol) defibrinated horse blood agar and serotyped by using specific C and G group antisera (Remel, Lenexa, Kans.). Colony size (smaller or larger than 0.5 mm in diameter), Voges-Proskauer test results (11), and production of β-D-glucuronidase (20) were determined to differentiate between *Streptococcus equi* subsp. *zooepidemicus* and *Streptococcus dysgalactiae* subsp. *equisimilis* from the "*Streptococcus milleri*" group (*S. anginosus*, *S. intermedius*, or *S. constellatus*).

Further testing to differentiate among species included acid production from trehalose and sorbitol to differentiate *S. equi* subsp. *zooepidemicus* from *S. dysgalactiae* subsp. *equi* and production of β-D-fucosidase, β-D-glucosidase, and/or α-D-glucosidase by using 4-methyllumbelliferyl-linked fluorogenic substrates in microtitration plates as previously described (20) to differentiate among *S. anginosus*, *S. intermedius*, and *S. constellatus.*

Differentiation criteria are summarized as follows. *S. equi* subsp. *zooepidemicus* or *S. dysgalactiae* subsp. *equisimilis* were presumed for those isolates forming minute colonies (less than

* Corresponding author. Mailing address: Unidad de Investigación, Edificio D, 1a planta, Hospital Universitario Son Dureta, Andrea Doria, 55, Palma de Mallorca 07014, Spain. Phone: 34-971-175334. Fax: 34-971-173184. E-mail: salberti@hsd.es.

† Contributing members of the Spanish Surveillance Group for Respiratory Pathogens are listed in Acknowledgments.

TABLE 1. *emm* or *emm*-like sequences of pharyngeal GCS and GGS isolated in Spain in 1998-1999[a]

emm or *emmL* type	No. of isolates of:								Total (%)
	S. constellatus		*S. anginosus*		*S. dysgalactiae* subsp. *equisimilis*		*S. intermedius*		
	C	G	C	G	C	G	C	G	
emm12	1								1 (0.47)
emm28D	4		5		10	1	1		21 (9.81)
emm3							1		1 (0.47)
emm9			1		2				3 (1.40)
emmC1139	3		4		7	1	4		19 (8.87)
emmC839	2		10	1	19	3			35 (16.35)
emmLG480	1				1	5			7 (3.27)
emmLG485		1			2	5			8 (3.74)
emmLG494			1			2			3 (1.40)
emmLG643						1			1 (0.47)
emmLG935		1				4			5 (2.34)
fcrV	3	1	3	3	4	17			31 (14.48)
C74a					1	1			2 (0.93)
emm73					1				1 (0.47)
mlc36	3		3		9	1			16 (7.48)
emmLC1900					2	1			3 (1.40)
emmLC1901					1				1 (0.47)
emmLC1902					1				1 (0.47)
emmLC1903					1				1 (0.47)
emmLC1904	2								2 (0.93)
emmLC1905						1			1 (0.47)
st1400c	1								1 (0.47)
NT	9		22	4	9	5	1		50 (23.36)
Subtotal	29	3	49	8	70	48	7	0	
Total for species		32		57		118		7	214

[a] C, GCS; G, GGS; NT, nontypeable.

0.5 mm in diameter) and confirmed only when positive β-D-glucuronidase and negative Voges-Proskauer test results were obtained. Further differentiation between both species was based on acid production from trehalose and/or sorbitol (*S. equi* subsp. *zooepidemicus* is trehalose negative or variable and sorbitol positive, while *S. dysgalactiae* subsp. *equisimilis* is trehalose positive and sorbitol negative). The *S. anginosus*, *S. intermedius*, and *S. constellatus* group was presumed when isolates formed colonies larger than 0.5 mm in diameter and confirmed by negative β-D-glucuronidase and positive Voges-Proskauer tests. Further species differentiation among these three species was done according to the production of β-D-fucosidase, β-D-glucosidase, and/or α-D-glucosidase (*S. anginosus*, β-D-glucosidase positive and β-D-fucosidase negative; *S. intermedius*, β-D-fucosidase and α-D-glucosidase positive; *S. constellatus*, β-D-fucosidase negative and α-D-glucosidase positive).

Among the GCS isolates, we identified 29 as *S. constellatus*, 70 as *S. dysgalactiae* (subsp. *equisimilis*), 7 as *S. intermedius*, and 49 as *S. anginosus*. Almost 82% (48) of the GGS isolates were *S. dysgalactiae* (subsp. *equisimilis*), and only 8 and 3 isolates were identified as *S. anginosus* and *S. constellatus*, respectively.

***emm*-like gene typing.** The *emmL* gene types of GCS and GGS isolates were determined by amplification and sequencing of the *emmL* gene with primers G1F and G1R (15). Lysates of the GCS and GGS isolates were prepared with mutanolysin as we previously described for GAS (1). The sizes of the amplicons ranged from about 0.8 kb to approximately 1.6 kb, a size range similar to that found for GAS *emm* genes (2). PCR products were sequenced with primer G1F and a dye terminator mix (Perkin-Elmer, Applied Biosystems, Madrid, Spain) and were subjected to automated sequence analysis on a 377 DNA sequencer (Perkin-Elmer, Applied Biosystems). DNA sequences were subjected to homology searches against the bacterial DNA database with BLASTN. Sequences were given the GenBank *emm* or *emmL* designation following the criteria previously described (4, 15).

Prevalent *emmL* genes. Results of the survey of the *emmL* gene sequences are shown in Table 1. Overall, 155 (72.43%) of 214 GCS and GGS pharyngeal isolates included in our study had 5′ *emm* sequences that were ≥95% identical to the first 120 bases of one of the *emm* or *emm*-like genes deposited in GenBank. For most of these sequences, this high level of identity actually extended to 200 to 450 bases without diminishing. For some sequences, we found more than one *emm* or *emm*-like gene in the GenBank with a homology of between 95 and 100% in the 5′ *emm* region. In these cases, we gave the *emm* designation with the highest homology. Nine of the 214 isolates had an undocumented *emm* or *emmL* gene sequence. All of the sequences were less than 90% identical over the first 160 bases to sequences deposited in GenBank. The sequences were designated *emmLC1900*, *emmLC1901*, *emmLC1902*, *emmLC1903*, *emmLC1904*, and *emmLC1905* and have been deposited in the EMBL data library.

We were not able to amplify the *emm* or *emmL* gene from 50 (23.36%) isolates (nontypeable isolates) with primer pair G1F-G1R or GASM1-GASM2 (4). Although it has been reported that all human-associated GGS strains harbor *emmL* genes, it

may be possible that nontypeable isolates do not contain this gene, as occurs with those strains recovered from animal sources (16). In fact, Podbielski et al. (14), utilizing PCR assays and genomic fingerprinting, detected *emm* gene in 75% of 28 GCS strains isolated from human specimens. The author's findings are in agreement with our results, since 76.63% of our isolates harbor *emm* or *emmL* sequences that were amplified by PCR. Interestingly, more than half of the nontypeable isolates were *S. anginosus*, and almost half of the isolates of this species were nontypeable. These results suggest that the nontypeable isolates express another protein(s) or use a different mechanism to replace the M protein since this protein is a critical virulence factor causing pharyngitis and resists opsonophagocytosis in GAS and in many GCS and GGS strains. Alternatively, it may be possible that the G1F-G1R and GASM1-GASM2 annealing regions are not conserved in the nontypeable isolates and as a consequence we failed to amplify these genes.

emmC839, *fcrV*, *emm28D*, *emmC1139*, and *mlc36* were the most prevalent *emm* sequences among the GCS and GGS isolated in Spain. They accounted for 16.35, 14.48, 9.81, 8.87, and 7.48%, respectively, of the isolates. These results indicate that there are ongoing horizontal gene transfers between the species of these two groups of beta-hemolytic streptococci. The remaining isolates represented less than 50% of the total and were associated with 17 different *emm* or *emmL* types. As occurred in the survey of *emm* gene sequences from pharyngeal GAS isolates conducted in Spain (2), only five *emm* types account for the majority of the isolates, suggesting that the M-like proteins encoded by these genes are more advantageous for the colonization of the human pharynx by GCS and GGS. However, the most prevalent GAS *emm* types (*emm1*, *emm3*, *emm12*, and *emm9*) were not represented among the most frequently observed *emm* or *emmL* types for GCS and GGS. By contrast, two GCS and five GGS strains exhibited *emm* gene sequences (*emm12*, *emm3*, *emm9*, and *st1400c*) that were identical over the first 160 bases to those reported for GAS, suggesting a horizontal gene transfer between strains of these species of beta-hemolytic streptococci. This transference could be responsible for the development of gene mosaics and the evolution of *emm* genes in beta-hemolytic streptococci (19).

In conclusion, using a large number of unrelated GCS and GGS pharyngeal isolates, we have confirmed the existence of an *emmL* gene polymorphism in these isolates and observed that this *emmL* gene diversity provides the basis for an *emmL* typing technique. The application of this typing technique to nationwide multicenter surveillance has revealed the existence of five prevalent types in Spain and has suggested that there is a horizontal transfer of *emm* genes between beta-hemolytic streptococci.

Nucleotide sequence accession numbers. The sequences described in this study were designated *emmLC1900*, *emmLC1901*, *emmLC1902*, *emmLC1903*, *emmLC1904*, and *emmLC1905* and were deposited in the EMBL data library with the accession numbers AY686726, AY686727, AY686728, AY686729, AY686730, and AY686731, respectively.

We thank C. Vidal and C. Santos for DNA sequencing.

This study was supported by a grant from GlaxoSmithKline, Madrid, Spain. M.I.G.-L. was supported by Instituto de Salud Carlos III through a fellowship from Red Respira (RTIC C03/11).

Participants in the Spanish Surveillance Group for Respiratory Pathogens include the following: F. Marco and T. Jiménez de Anta, Hospital Clinic i Provincial, Barcelona; C. Fernández-Mazarrasa, Hospital Marqués de Valdecilla, Santander; C. García-Riestra, B. Regueiro, A. Jato, and M. Prieto, Hospital Clínico Universitario, Santiago de Compostela; M. Casal and A. Ibarra, Hospital Reina Sofía Cordova; M. de la Rosa, Hospital Virgen de las Nieves, Granada; C. de la Torre and A. Gené, Hospital San Joan de Deu, Barcelona; C. Rubio and C. García, Hospital Lozano Blesa, Zaragoza; E. Perea and L. Martínez, Hospital Virgen de la Macarena, Seville; J. Nogueira, Hospital Dr. Peset, Valencia; E. Pérez-Trallero and J. Larruskain, Hospital Donostia, San Sebastián; J. A. García-Rodríguez and I. Trujillano, Hospital Clínico Universitario, Salamanca; J. Ruiz and E. Simarro, Hospital Virgen de la Arrixaca, Murcia; E. Bouza and E. Cercenado, Hospital Gregorio Marañón, Madrid; A. M. Martín and F. Cañas, Hospital Insular, Las Palmas; J. Barrón and L. López, Hospital de Cruces, Bilbao; A. García, S. García, and M. Güeni, Hospital La Paz, Madrid; D. Romero and M. González, Hospital Nuestra Señora de Alarcos, Ciudad Real; A. Fenoll and J. Casal, Instituto Carlos III, Madrid; J. J. Granizo, Fundación Jiménez Díaz, Madrid; C. Gimeno and E. Esteban; Instituto Valenciano de Microbiología, Valencia; L. Aguilar, Universidad Complutense, Madrid; R. Dal-Ré, GlaxoSmithKline S.A., Madrid.

REFERENCES

1. **Albertí, S., C. D. Ashbaugh, and M. R. Wessels.** 1998. Structure of the *has* operon promoter and regulation of hyaluronic acid capsule expression in group A *Streptococcus.* Mol. Microbiol. **28:**343–353.
2. **Albertí, S., C. García-Rey, M. A. Domínguez, L. Aguilar, E. Cercenado, M. Gobernado, A. García-Perea, and Spanish Surveillance Group for Respiratory Pathogens.** 2003. Survey of *emm* gene sequences from pharyngeal *Streptococcus pyogenes* isolates collected in Spain and their relationship with erythromycin susceptibility. J. Clin. Microbiol. **41:**2385–2390.
3. **Beall, B., R. Facklam, T. Hoenes, and B. Schwartz.** 1997. Survey of *emm* gene sequences and T-antigen types from systemic *Streptococcus pyogenes* infection isolates collected in San Francisco, California; Atlanta, Georgia; and Connecticut in 1994 and 1995. J. Clin. Microbiol. **35:**1231–1235.
4. **Beall, B., R. Facklam, and T. Thompson.** 1995. Sequencing *emm*-specific polymerase chain reaction products for routine and accurate typing of group A streptococci. J. Clin. Microbiol. **34:**953–958.
5. **Bisno, A. L., R. E. Campo, and C. M. Collins.** 1994. M proteins of group G streptococci; structural and functional studies, p. 225–227. *In* A. Totolian (ed.), Pathogenic streptococci: present and future. Lancer Publications, St. Petersburg, Russia.
6. **Bisno, A. L., C. M. Collins, and J. C. Turner.** 1996. M proteins of group C streptococci isolated from patients with acute pharyngitis. J. Clin. Microbiol. **34:**2511–2515.
7. **Collins, C. M., A. Kimura, and A. L. Bisno.** 1992. Group G streptococcal M protein exhibits structural features analogous to those of class I M protein of group A streptococci. Infect. Immun. **60:**3689–3696.
8. **Dicuonzo, G., G. Gherardi, G. Lorino, S. Angeletti, M. De Cesaris, E. Fiscarelli, E. D. Bessen, and B. Beall.** 2001. Group A streptococcal genotypes from pediatric throat isolates in Rome, Italy. J. Clin. Microbiol. **39:** 1687–1690.
9. **Efstratiou, A.** 1989. Outbreaks of human infections caused by pyogenic streptococci of Lancefield groups C and G. J. Med. Microbiol. **29:**207–219.
10. **Facklam, R.** 2002. What happened to the streptococci: overview of taxonomic and nomenclature changes. Clin. Microbiol. Rev. **15:**613–630.
11. **Facklam, R. R., and J. A. Washington.** 1991. *Streptococcus* and related catalase-negative gram-positive cocci, p. 248. *In* A. Balows, W. J. Hausler, Jr., K. L. Herrmann, H. D. Isenberg, and H. J. Shadomy (ed.), Manual of Clinical Microbiology, 5th ed. American Society for Microbiology, Washington, D.C.
12. **Jones, K. F., and V. A. Fischetti.** 1987. Biological and immunochemical identity of M protein on group G streptococci with M protein on group A streptococci. Infect. Immun. **55:**502–506.
13. **Pérez-Trallero, E., C. Fernández-Mazarrasa, C. García-Rey, E. Bouza, L. Aguilar, J. García de Lomas, F. Baquero, and the Spanish Surveillance Group for Respiratory Pathogens.** 2001. Antimicrobial susceptibilities of 1,684 *Streptococcus pneumoniae* and 2,039 *Streptococcus pyogenes* isolates and their ecological relationships: results of a 1-year (1998–1999) multicenter surveillance study in Spain. Antimicrob. Agents Chemother. **45:**3334–3340.
14. **Podbielski, A., M. Mignon, J. Weber-Heynemann, R. Schnitzler, R. Lut-**

ticken, and A. Kaufhold. 1994. Characterization of groups C (GCS) and G (GGS) streptococcal M protein (*emm*) genes, p. 234–236. *In* A. Totolian (ed.), Pathogenic streptococci: present and future. Lancer Publications, St. Petersburg. Russia.
15. **Schnitzler, N., A. Podbielski, G. Baumgarten, M. Mignon, and A. Kaufhold.** 1995. M or M-like protein gene polymorphisms in human group G streptococci. J. Clin. Microbiol. **33:**356–363.
16. **Simpson, W. J., J. C. Robbins, and P. P. Cleary.** 1992. Evidence for group A-related M protein genes in human but not animal-associated group G streptococcal pathogens. Microb. Pathog. **3:**339–350.
17. **Smirnov, O. Y., A. I. Denesyuk, M. V. Zakharov, V. M. Abramov, and V. P. Zav′yalov.** 1992. Protein V, a novel type-II IgG receptor from *Streptococcus* sp.: sequence, homologies and putative Fc-binding site. Gene **20:**127–132.
18. **Stryker, W. S., D. W. Fraser, and R. R. Facklam.** 1992. Foodborne outbreak of group G streptococcal pharyngitis. Am. J. Epidemiol. **116:**533–540.
19. **Whatmore, A. M., and M. A. Kehoe.** 1994. Horizontal gene transfer in the evolution of group A streptococcal *emm*-like genes: gene mosaics and variation in Vir regulons. Mol. Microbiol. **11:**363–374.
20. **Whiley, R. A., H. Fraser, J. M. Hardie, and D. Beighton.** 1990. Phenotypic differentiation of *Streptococcus intermedius*, *Streptococcus constellatus*, and *Streptococcus anginosus* strains within the "*Streptococcus milleri* group." J. Clin. Microbiol. **28:**1407–1501.