Research Article

For reprint orders, please contact: reprints@futuremedicine.com

Future MICROBIOLOGY

Molecular epidemiology of nonpharyngeal group A streptococci isolates in northern Lebanon

Rayane Rafei[1], Malaik Hawli[1], Marwan Osman[1], Simon Khelissa[2], Tamara Salloum[3], Fouad Dabboussi[1], Sima Tokajian[3] & Monzer Hamze*[,1]

[1]Laboratoire Microbiologie Santé et Environnement (LMSE), Doctoral School for Science & Technology, Faculty of Public Health, Lebanese University, Tripoli, 1300, Lebanon

[2]Université de Lille, Centre National de la Recherche Scientifique, Institut National de Recherche pour l'Agriculture, l'Alimentation et l'Environnement, Centrale Lille, Unité Mixte de Recherche 8207 – Unité Matériaux et Transformations, Lille, 59000, France

[3]Department of Natural Sciences, School of Arts & Sciences, Lebanese American University, Byblos campus, Postal Box 36, Byblos, 1401, Lebanon

*Author for correspondence: Tel.: +961 621 3383; mhamze@monzerhamze.com

Aim: To characterize the epidemiology of group A *Streptococcus* (GAS) involved in nonpharyngeal infections sparingly addressed in Lebanon. **Materials & methods:** A collection of 63 nonpharyngeal GAS isolates recovered between 2010 and 2019 from northern Lebanon were analyzed through *emm* typing, virulence gene profiling, FCT typing and antibiotic susceptibility analysis. **Results & conclusion:** A total of 29 *emm* subtypes was detected, with *emm1* being the most dominant. A great intraclonal divergence driven by the loss and gain of superantigens or by the structural variability within the FCT regions was unraveled. The resistance rates for erythromycin and tetracycline were 8 and 20.6%, respectively. The 30-valent vaccine coverage was 76%. This study evidences the complexity of the neglected GAS pathogen in Lebanon.

First draft submitted: 2 April 2020; Accepted for publication: 20 October 2020; Published online: 25 November 2020

Keywords: 30-valent vaccine • *emm* typing • FCT • Lebanon • nonpharyngeal • *Streptococcus pyogenes* • superantigen

Although group A *Streptococcus* (GAS, *Streptococcus pyogenes*) is a common pathogen responsible for relatively mild superficial diseases, such as acute pharyngitis, it can take advantage from noninvasive infections as a portal either to reach deeper tissues and cause invasive infections (iGAS) as necrotizing fasciitis or to alter the immune system leading to poststreptococcal autoimmune sequelae as rheumatic heart disease [1]. Adding to this chorus of infections, GAS can produce a diverse array of virulence factors including among other superantigens (SAgs): three chromosomally encoded genes (*speG*, *speJ* and *smeZ*) and eight genes harbored by temperate phages (*speA*, *speC*, *speH*, *speI*, *speK*, *speL*, *speM* and *ssa*) [2] that can overstimulate the immune-inflammatory cascade and eventually lead to streptococcal toxic shock syndrome [3]. Additional gene-encoding virulence factors are gathered in a putative pathogenicity island known as FCT (the fibronectin-binding, collagen-binding, T antigen) region that gives rise to pilus-like appendages and/or extracellular matrix-binding proteins, and includes among others regulatory genes (*rofA* or *nra*), genes that encode the major backbone pilin protein (as *fctA*), ancillary pilin proteins (as *cpa, fctB*), and at least one sortase [3,4]. Besides, GAS can acquire resistance to antibiotics, most notable being that against macrolides and lincosamides. Indeed, macrolides are considered as paramount alternatives for penicillin-allergic patients while lincosamides are recommended together with ß-lactam for iGAS treatment [5]. In the past decades, GAS was reckoned as a dynamic pathogen with the ongoing reshaping of its population structure and appearance of new virulent strains linked to the current upsurge of iGAS infections [6]. Investigating the clonal relationships among GAS isolates continues to be the most convenient approach to elucidate the GAS epidemiology. One of the key epidemiological methods for tracking iGAS infections is to characterize the circulating *emm* types that can vary according to the geographical location and the clinical specimen type [7]. In Lebanon, three studies concerning the circulating *emm* types were published, the main sample type was pharyngeal, rarely from discharge [8,9] and

Future Medicine

perineal and vulvovaginal specimens [10]. Our study examined the epidemiological patterns of 63 nonpharyngeal GAS collected between 2010 and 2019 in two hospitals in northern Lebanon through *emm* typing, virulence gene profiling, FCT typing and antibiotic susceptibility analysis. The determination of *emm* types helped us to assess the effectiveness of the 30-valent vaccine, which is under clinical trial investigation, and is being developed using 30 pharyngitis-associated type-specific M antigens.

Materials & methods

GAS isolates

All nonpharyngeal clinical GAS isolates conserved at the Collection de Microbiologie Université Libanaise of the Laboratoire Microbiologie Santé Environnement (LMSE) were investigated. These isolates were collected between 2010 and 2019 mainly from two tertiary care hospitals in northern Lebanon during their routine analysis wherein no consent was required. They were identified by colony morphology, beta-hemolysis on blood agar, sensitivity to bacitracin, serogrouping using a latex agglutination test and, additionally, by matrix-assisted laser desorption ionization–time of flight (MALDI-TOF) technique using Vitek® MS machine (bioMérieux®, Marcy l'Etoile, France). Clinical isolates and patient records/information were anonymous and de-identified prior to analysis.

Antibiotic susceptibility testing

The susceptibility data of all isolates for standard streptococcal antibiotic list (erythromycin [15 µg], clindamycin [2 µg], penicillin G [1 unit] and tetracycline [30 µg]) were analyzed by the disc diffusion method following the 2019 European Committee on Antimicrobial Susceptibility Testing recommendations. A double-disc diffusion method (D-test) was also done according to European Committee on Antimicrobial Susceptibility Testing in order to examine the inducible character of clindamycin resistance and categorize the resistance patterns into M phenotype (isolate resistant to erythromycin and susceptible to clindamycin), and inducible or iMLSb (isolate resistant to erythromycin with any blunting in the shape of the clindamycin inhibition zone). The PCR of the *erm(A), erm(B)* and *mef(A)* genes was carried out as described previously on erythromycin-resistant GAS isolates (ERGAS) [11,12].

Epidemiological typing

PCR and sequencing of the 5′ region of the *emm* gene were done and interpreted following the protocol provided by 'CDC' [13]. The *emm* patterns and clusters were assigned based on *emm* types [14]. Multiplex PCR was used to profile 11 SAg genes [15]. *speL, speK, speM, speC, smeZ, speI* were tested by a multiplex PCR mix1 whereas *ssa, speA, speH, speG, speJ* were examined by multiplex PCR mix2 as previously described [15]. The presence of *speB* was analyzed separately by PCR using the same primers as elsewhere [15]. The virulence genes within the fibronectin- and collagen-binding T antigen (FCT) region were profiled using 14 different allele-specific primer pairs as outlined previously by Kratovac *et al.* [4]. The 'FCT-region variants' were defined according to the classification proposed by Kratovac *et al.* and Bencardino *et al.* [3,4]. Initially, the strains with profiles matching one FCT type were used as references for FCT-region variants and were assigned a specific code composed of their corresponding FCT type number (one to eight) followed by a capital letter A (ex: 1A, 2A). Strains showing an FCT region pattern differing only by one PCR product were recognized as a 'single FCT-region variant' and were labeled by a capital letter B (ex, 1B, 2B) after their corresponding reference FCT type number. The same approach was used to categorize the double and triple FCT-region variants being respectively designated as C and D after the corresponding reference FCT type number. The *sof* gene was also amplified following the same FCT typing scheme by PCR [4].

Vaccine coverage

emm types identified in this study were categorized based on the cross-opsonization experiments in animals that suggested that 30-valent vaccine protection may extend to noncognate *emm* types [16]. *emm* types included in the 30-valent vaccine were classified as 'vaccine *emm* types'. Nonvaccine *emm* types with more than 50% killing in the cross opsonization studies were classified as 'cross opsonization-positive'. Nonvaccine *emm* types that were untested in the cross opsonization experiments were stratified as 'cross opsonization unknown'. Vaccine coverage was defined as the proportion of all isolates in the region that was covered by the 30-valent M-protein-based GAS vaccine currently under clinical investigation (*emm* types *1*, *2*, *3.1*, *4*, *5.14*, *6.4*, *11*, *12*, *14.3*, *18*, *19*, *22*, *24*, *28*, *29*, *44*, *49*, *58*, *73*, *75*, *77*, *78*, *81*, *82*, *83.1*, *87*, *89*, *92*, *114* and *118*).

Statistical analysis

Statistical analyses were performed with GraphPad Prism 6.0 software (GraphPad Software, Inc., CA, USA). Univariate analyses of qualitative data are done using Fisher's exact test. The results are expressed in terms of p-values, odds-ratios (OR) and 95% CIs. All the statistical tests are bilateral with a risk of type I error of 5%.

Results

Bacterial isolates

A total of 63 nonpharyngeal GAS isolates were analyzed and were mainly collected from two hospitals Nini (40 isolates) and El Youssef Hospital Center (20 isolates) located respectively in Tripoli and Akkar. These isolates were recovered from ear (n = 38, 60.3%), wound and discharge (n = 15, 23.8%), blood of sepsis cases (n = 9, 14.3%) and vulvovaginal tract (n = 1, 1.6%). They were collected from 32 (50.8%) females and 31 (49.2%) males aged between 2.5 months and 93 years old. Of note, data related to age for five patients and collection hospital for three patients were missing.

Antibiotic susceptibility patterns & macrolide resistance mechanisms

The antibiotic susceptibility testing showed that all isolates were fully susceptible to penicillin G, and that five (8%) were ERGAS. Among these five ERGAS strains, two were also resistant to clindamycin and showed the constitutive macrolides, lincosamides, streptogramin b (cMLSb) phenotype where one harbored the *erm(B)* gene and another co-harbored the *erm(B)* and *mef(A)* genes. The three other ERGAS isolates expressed the M phenotype and only harbored the *mef(A)* gene. The iMLSb phenotype and *erm(A)* gene were not detected. Thirteen out of 63 isolates (20.6%) were resistant to tetracycline. Notably, all tetracycline-resistant isolates were erythromycin-susceptible.

emm typing: *emm* types, patterns, clusters & their association with specimen types & resistance patterns

All of the 63 GAS isolates were *emm* typable giving a total of 25 different *emm* types and 29 *emm* subtypes. Table 1 classifies the isolates according to their *emm* patterns, *emm* clusters and *emm* types with a special focus on the specimen type. Eight *emm* types accounted for more than 65% of the population and were distributed as follows: *emm1* (22.2%, 14/63), followed by *emm2* (11.1%, 7/63), *emm12* (8%, 5/63), *emm6* (6.4%, 4/63), *emm41* (4.7%, 3/63), *emm87* (4.7%, 3/63), *emm101* (4.7%, 3/63) and *emm165* (4.7%, 3/63). Seventeen different additional *emm* types were also recovered in percentages under four for each (*emm3.93, emm5.74, emm8, emm22, emm28, emm44, emm48, emm49, emm58.8, emm75, emm81, emm82, emm83, emm89, emm108, emm113* and *emm118*). Seven out of 25 *emm* types were identified in nine blood isolates (*emm1*, *emm12*, *emm49*, *emm89*, *emm101*, *emm165* and *emm113*).

Roughly, 46% (29/63) of the isolates belonged to *emm* pattern E, 39.7% (25/63) to *emm* pattern A–C and 14.3% (9/63) to *emm* pattern D. Ten *emm* clusters were identified with more than 75% of the isolates being associated with five clusters. Twenty-three *emm* types were grouped into eight large defined *emm* clusters, while the remaining (*emm6* and *emm5*) grouped as single clusters belonging to the clade Y. Variants of *emm* clusters E1-6 and A–C1–5 were dominant, accounting together for 79% (50/63) of the isolates. The majority of *emm* cluster E1–6 isolates belonged to the E4 cluster and those of *emm* cluster A–C1–5 to A–C3. All *emm* cluster D1–5 isolates were *emm* cluster D4. Six *emm* clusters comprised each a single *emm* type. While no *emm* types were significantly tied to a specific specimen, D pattern ($p = 0.0287$; OR = 5.5, IC95% [1.24–24.25]) and D4 cluster ($p = 0.015$; OR = 7.5, IC95% [1.53–36.68]) were found to be associated with wound infections. No association between *emm* type or pattern or cluster and time of isolation was detected.

The five ERGAS isolates were stratified into three *emm* types: *emm2* (1), *emm12* (3) and *emm113* (1) with two of the isolates (*emm12* and *emm113*) being recovered from blood. The two ERGAS isolates expressing the cMLSb phenotype belonged to *emm12*, while the three remaining isolates expressing the M phenotype were diverse and belonged to *emm2, emm12* and *emm113*. *emm* pattern A–C ($p = 0.02$; OR = 4.78, IC95% [1.21–18.84]) and cluster A–C3 ($p = 0.006$; OR = 17.03, IC95% [0.96–302.7]) were significantly associated with macrolide and tetracycline susceptibility. The 13 tetracycline-resistant isolates belonged to 10 *emm* types (*emm101* [3], *emm41* [2], *emm8* [1], *emm22* [1], *emm75* [1], *emm81* [1], *emm82* [1], *emm83* [1], *emm108* [1] and *emm165* [1]). Notably, a significant association was also found between tetracycline resistance and each of the *emm101* type ($p = 0.007$; OR = 33.67, IC95% [1.6–702.1]), pattern D ($p < 0.0001$; OR = 78.4, IC95% [8.06–761.9]) and cluster D4 ($p < 0.0001$; OR = 57.17, IC95% [5.96–548.3]).

Table 1. Classification of the isolates according to their *emm* patterns, *emm* clusters and *emm* types with a special focus on specimen types, and patient age.

emm patterns, nb (%)	*emm* clusters, nb (%)	*emm* types, nb (%)	*emm* subtypes‡	Specimen types				Age (years)			
				E	W	B	V	Infants (<1 years)	Children (1–<18)	Adult (<60)	Elderly (>60)
A–C (25)	A–C3 (14)	*emm1* (14)		9	2	2	1	1	10	1	2
	A–C4 (5)	*emm12*† (5)		3	1	1	0	1	2	-	-
	A–C5 (1)	*emm3* (1)	*emm3.93*	1	0	0	0	0	1	0	0
	Clade Y (4)	*emm6* (4)	*emm6.0* (1), *emm6.103* (2), *emm6.4* (1)	4	0	0	0	0	4	0	0
	Clade Y (1)	*emm5*	*emm5.74*	1	0	0	0	0	1	0	0
D (9)	D4 (8)	*emm41* (3)	*emm41.0* (1), *emm41.2* (2)	2	1	0	0	0	1	2	0
		emm83 (1)	*emm83.1*	0	1	0	0	0	0	1	0
		emm101 (3)		0	2	1	0	0	1	0	2
		emm108 (1)	*emm108.1*	0	1	0	0	0	0	1	0
	E6 (1)	*emm81* (1)		1	0	0	0	0	1	0	0
E (29)	E1 (3)	*emm165* (3)		1	0	2	0	0	1	1	1
	E3 (10)	*emm44* (1)		1	0	0	0	0	1	0	0
		emm49 (2)		0	1	1	0	0	0	2	0
		emm58 (1)	*emm58.8*	0	1	0	0	0	0	1	0
		emm82 (1)		1	0	0	0	0	1	0	0
		emm87† (3)		2	1	0	0	-	1	-	1
		emm113† (1)		0	0	1	0	-	-	-	-
		emm118 (1)		1	0	0	0	0	1	0	0
	E4 (13)	*emm2* (7)		6	1	0	0	0	7	0	0
		emm8 (2)		0	2	0	0	0	1	1	0
		emm22 (1)		1	0	0	0	0	1	0	0
		emm28 (1)		1	0	0	0	0	1	0	0
		emm89 (2)†		1	0	1	0	-	-	-	1
	E6 (3)	*emm48* (2)	*emm48.0* (1), *emm48.1* (1)	2	0	0	0	0	2	0	0
		emm75 (1)		0	1	0	0	0	1	0	0
Total (63 isolates)				38	15	9	1	2	39	10	7

Minus in the age column means that this case can contain 0 or a specific number other than 0 because the age information is not available for all isolates in the corresponding emm type.
†The age information was not available for all isolates having a particular *emm* type, so a symbol – was put instead of 0 in a particular age category.
‡The *emm* subtype was only indicated when the type contained subtypes other than those denoted by 0.
B: Blood; E: Ears, V: Vulvovaginal; W: Discharges and wound.

speB & superantigen-encoding genes profiling results & their association with *emm* types & specimens

Variations in the distribution and types of virulence factors were observed among GAS strains (Table 2). *speB* was detected in 92% (58/63) of the isolates, followed by *speG* 73% (46/63) and *speC* 65% (41/63). *speC* was mainly associated with isolates recovered from ear infections (p = 0.0309; OR = 3.49, IC95% [1.18–10.32]). The other SAg genes were distributed as follows: 39.7% (25/63) for *speJ*, 27% (17/63) for *speI*, 22.2% (14/63) for *speK*, 19% (12/63) for *ssa*, 15.8% (10/63) for *speA*, 11.1% (7/63) for *speL*, 9.5% (6/63) for *speH*, 3.1% (2/63) for *smeZ* and finally 1.6% (1/63) for *speM*.

Out of 12 virulence genes tested (*speB* and the 11 SAg-encoding genes), all isolates harbored at least one virulence determinant and 62 (98%)/63 isolates carried at least one SAg gene. Overall, 41 virulence-gene profiles were detected among the 63 isolates, with '*speB–speC*' (6.3%) being the predominant (Table 2). In addition, 11.1% of the isolates carried *speB* together with one SAg gene (*speB-speC*; *speB-speG*; *speB-speI*), 19% carried *speB* with two SAg genes, 28.6% co-harbored *speB* with three SAgs, 28.6% co-harbored *speB* with four SAg genes and 3.2%

Table 2. Genotypic diversity within the group A Streptococci isolates: correlation between virulence genes, *emm* types, FCT regions, antibiotic resistance profiles, specimens and *sof* gene.

nb of virulence genes	nb of virulence profiles	Virulence gene profiles (nb of isolates)	*emm* type (nb of isolates)	FCT regions type	*sof* gene	nb of different profiles† by merging data	Antibiotic resistance profile (nb of isolates)			Isolate source	Year of isolation
							ERY	CLN	TET		
1	1	*speB* (1)	*emm8* (1)	7D	+	1	-	-	-	W	2016
2	2	*speB-C* (4)	*emm2* (1)	5A	+	2	-	-	-	E	2013
			emm2 (1)	6A	+	3	-	-	-	E	2017
			emm48.1 (1)	5B	+	4	-	-	-	E	2017
			emm118 (1)	8C	+	5	-	-	-	E	2016
	3	*speB-G* (2)	*emm8* (1)	4C	+	6	-	-	1	W	2018
			emm113 (1)	5A	+	7	1	-	-	B	2010
	4	*speB-I* (1)	*emm82* (1)	7C	+	8	-	-	1	E	2015
	5	*speC-G* (2)	*emm1* (1)	5A	-	9	-	-	-	E	2015
			emm5.74 (1)	5B	-	10	-	-	-	E	2016
3	6	*speA-B-J* (2)	*emm1* (1)	1A	-	11	-	-	-	E	2017
			emm1 (1)	2A	+	12	-	-	-	E	2016
	7	*speB-C/ssa* (1)	*emm22* (1)	7C	+	13	-	-	1	E	2017
	8	*speB-C-G* (1)	*emm89* (1)	7A	+	14	-	-	-	E	2011
	9	*speB-C-K* (1)	*emm6.103* (1)	1A	-	15	-	-	-	E	2017
	10	*speB-G/ssa* (1)	*emm41.0* (1)	8C	-	16	-	-	-	E	2013
	11	*speB-G-K* (1)	*emm58.8* (1)	5A	+	17	-	-	-	W	2015
	12	*speB-H-J* (1)	*emm1* (1)	1A	-	18	-	-	-	E	2014
	13	*speB-J/ssa* (2)	*emm1* (1)	1A	-	19	-	-	-	W	2018
			emm108.1 (1)	1A	-	20	-	-	1	W	2017
	14	*speB-G-L* (2)	*emm2* (1)	5A	+	21	-	-	-	E	2017
			emm83.1 (1)	3A	-	22	-	-	1	W	2017
	15	*speC-G-J* (1)	*emm44* (1)	4C	+	23	-	-	-	E	2017
	16	*speG-I/ssa* (1)	*emm165* (1)	5A	+	24	-	-	1	B	2017
4	17	*speA-B-C-M* (1)	*emm49* (1)	7D	+	25	-	-	-	B	2013
	18	***speA-B-G-J* (2)**	***emm1* (2)**	**2A**	**-**	**26**	-	-	-	**E**	**2017 and 2018**
	19	*speB-C-G/ssa* (1)	*emm41.2* (1)	4C	-	27	-	-	1	E	2017
	20	*speB-C-G-J* (2)	*emm1* (1)	1A	-	28	-	-	-	E	2014
			emm28 (1)	7C	+	29	-	-	-	E	2016
	21	*speB-C-G-I* (2)	*emm3.93* (1)	X1	-	30	-	-	-	E	2019
			emm48 (1)	5A	+	31	-	-	-	E	2012

The bolded rows showed clusters of isolates with similar *emm* type, virulence gene profiles, FCT types and antibiotic resistance profiles.
†Number of different profiles obtained when merging the data of the following *loci* (emm, FCT regions, and virulence genes).
B: Blood; E: Ears; W: Discharges and wound; V: Vulvovaginal.

Table 2. Genotypic diversity within the group A Streptococci isolates: correlation between virulence genes, *emm* types, FCT regions, antibiotic resistance profiles, specimens and *sof* gene (cont.).

nb of virulence genes	nb of virulence profiles	Virulence gene profiles (nb of isolates)	*emm* type (nb of isolates)	FCT regions type	*sof* gene	nb of different profiles† by merging data	Antibiotic resistance profile (nb of isolates)			Isolate source	Year of isolation
							ERY	CLN	TET		
	22	*speB-C-G-K* (3)	*emm1* (1)	1A	-	32	-	-	-	B	2016
			emm2 (1)	6A	+	33	-	-	-	E	2016
			emm89 (1)	4C	+	34	-	-	-	B	2018
	23	*speB-C-I-K* (1)	*emm6* (1)	1A	+	35	-	-	-	E	2013
	24	*speB-C-L/smeZ* (1)	*emm165* (1)	3B	+	36	-	-	-	E	2017
	25	*speB-G-I-J* (3)	*emm12* (1)	7C	+	37	-	-	-	W	2014
			emm49 (1)	8D	+	38	-	-	-	W	2014
			emm101 (1)	3B	-	39	-	-	1	B	2018
	26	*speB-G-J-L* (2)	*emm41.2* (1)	8C	-	40	-	-	1	W	2015
			emm81 (1)	1A	-	41	-	-	1	E	2014
	27	*speC-G-I-K* (1)	*emm165* (1)	5B	+	42	-	-	-	B	2017
5	28	*speA-B-C-G-J* (3)	*emm1* (1)	2A	-	43	-	-	-	E	2017
			emm1 (1)	1A	-	44	-	-	-	V	2014
			emm87 (1)	X2	+	45	-	-	-	E	2011
	29	*speA-B-G-I-J* (1)	*emm1* (1)	2A	+	46	-	-	-	W	2016
	30	*speB-C-G-H-I* (2)	*emm12* (1)	7C	+	47	1	-	-	E	2015
			emm12 (1)	4A	+	48	1	1	-	B	2019
	31	*speB-C-H-I/ssa* (1)	*emm12* (1)	7A	+	49	-	-	-	E	2012
	32	*speB-C-G-H-J* (1)	*emm1* (1)	1A	-	50	-	-	-	E	2014
	33	*speB-C-G-J/ssa* (2)	*emm87* (1)	8B	+	51	-	-	-	W	2017
			emm87 (1)	4D	+	52	-	-	-	E	2016
	34	*speB-C-G-I-K* (1)	*emm6.4* (1)	5C	-	53	-	-	-	E	2017
	35	*speB-C-G-I-L* (2)	*emm2* (1)	6A	+	54	-	-	-	E	2017
			emm2 (1)	X3	+	55	-	-	-	E	2017
	36	*speB-C-G-I/ssa* (1)	*emm12* (1)	8D	+	56	1	1	-	E	2019
	37	***speB-C-G-J-K* (2)**	***emm101* (2)**	**3A**	**-**	**57**	**-**	**-**	**2**	**W**	**2018**
	38	*speB-C-G-K/smeZ* (1)	*emm75* (1)	5B	+	58	-	-	1	W	2018
	39	*speB-C-G-K/ssa* (1)	*emm2* (1)	6A	+	59	1	-	-	W	2017
6	40	*speA-B-C-G-J-K* (1)	*emm1* (1)	2A	-	60	-	-	-	B	2017
	41	*speB-C-G-H-K/ssa* (1)	*emm6.103* (1)	X4	-	61	-	-	-	E	2016

The bolded rows showed clusters of isolates with similar *emm* type, virulence gene profiles, FCT types and antibiotic resistance profiles.
†Number of different profiles obtained when merging the data of the following *loci* (emm, FCT regions, and virulence genes).
B: Blood; E: Ears; W: Discharges and wound; V: Vulvovaginal.

of isolates carried *speB* with five SAg genes. Although, 77.7% (7/9) of the blood isolates co-harbored four or more virulence genes, no significant association was found between the number of virulence factors and specimen types nor between virulence signature and specimen types. Interestingly, *speA* ($p < 0.0001$; OR = 31.33, IC95% [5.35–183.5]) and *speJ* ($p = 0.0001$; OR = 16.62, IC95% [3.26–84.48]) were significantly linked to *emm1*, while *speI* ($p = 0.0009$; OR = 40.92, IC95% [2.12–791.5]) with *emm12*. Additionally, *speH* was only detected in *emm* pattern A–C ($p = 0.0026$; OR = 25.67, IC95% [1.37–479.9]). However, no significant association was retrieved between *emm* types and the virulence factor signature nor between *emm* type and the number of virulence factors.

FCT typing & *sof* profiling

Roughly, a total of 30 different FCT regions were identified by the PCR-typing scheme. Out of the 30 detected FCT regions, 26 were either identical (seven regions, which are assigned by letter A) or nearly similar (19 regions, which are considered single, double or triple locus variants and assigned respectively by letter B, C and D) to previously described reference FCT regions. On the other hand, four types (denoted here as X1, X2, X3 and X4) differed completely from known FCT regions (Table 3).

The prevalence of FCT-A type (identical to previously described reference FCT-region) in this study was 54% (34/63). FCT-1, FCT-2 and FCT-6 were detected only as reference FCT regions. Overall, 39.7% (25/63) of the isolates had either a single locus FCT-region (FCT-B), double (FCT-C) or triple (FCT-D) variant (Table 3).

As shown in Table 3, the 14 *emm1* isolates contained FCT-1A (7/14), 2A (6/14) and 5A (1/14), while in *emm2* isolates the most common was FCT-6A (4/7) followed by FCT-5A (2/7) and unclassified FCT region (1/7). FCT-2A was exclusively associated with *emm1* ($p < 0.0001$; OR = 75.7, IC95% [3.89–1472]), as well as FCT-6A was only found in *emm2* ($p < 0.0001$; OR = 145.3, IC95% [6.44–3276]). Besides, FCT-1A was significantly associated with *emm1* ($p = 0.0013$; OR = 11.25, IC95% [2.6–48.65]) and *emm* pattern A–C ($p = 0.0045$; OR = 10.13, IC95% [1.96–52.29]). Significant associations were also revealed between *emm* pattern D and FCT-3A ($p = 0.0021$; OR = 58.69, IC95% [2.71–1269]), and *emm* pattern E and FCT-5A ($p = 0.0417$; OR = 8.61, IC95% [0.97–76.42]).

Noteworthy, out of five *emm12* isolates, three carried FCT-7A or closely related FCT-7 (FCT-7C). Isolates of *emm101* (three isolates) and *emm48* (two isolates) harbored closely related variants belonging respectively to FCT-3, and FCT-5. Out of the three *emm41* type, two harbored the same FCT-8C variant. Notably, all *emm* pattern E isolates (29/63) were positive for *sof* gene ($p < 0.0001$; OR = 183.9, IC95% [10.11–3345]) while all *emm* pattern D isolates ($p = 0.0001$; OR = 40.71, IC95% [2.24–740.4]) and 68% of *emm* pattern A–C isolates ($p = 0.0007$; OR = 6.85, IC95% [2.22–21.1]) were *sof*-negative. Regarding FCT types, FCT-1A was significantly associated with the absence of *sof* gene ($p = 0.0003$; OR = 22.5%, IC95% [2.65–91]).

Genetic diversity of GAS isolates based on *emm* types, virulence genes, FCT regions & antibiotic resistance patterns

Overall, when considering the complete set of studied genes (*emm* type, virulence genes and FCT) along with antibiotic resistance patterns, a total of 61 different profiles were categorized (Table 3). Among the 63 studied isolates, only two pairs belonging to *emm1* and *emm101* had the same genetic profile.

Vaccine coverage

Among the 25 identified *emm* types, 18 were categorized as 30-valent vaccine types, two were stratified as *emm* types with positive cross opsonization and five as *emm* types with unknown cross opsonization (Table 4). The 30-valent vaccine coverage was 76% when considering only the vaccine *emm* types, and 82.5% when also including the *emm* types with positive cross-opsonization during experiments done in animals [16].

Discussion

Our study is the first to examine a considerable collection of GAS isolates involved in nonpharyngeal infections in Lebanon. The high number of *emm* subtypes identified herein (29 *emm* subtypes) among the 63 isolates reflected a higher genetic heterogeneity than that reported in previous studies from Lebanon with the identification of 41 *emm* subtypes among 150 isolates [9], 33 among 103 isolates [8] and three among 17 isolates [10]. The higher detected diversity could be tied to: variations in the source of isolation (nonpharyngeal isolates in the present study vs primarily pharyngeal isolates in the Bahnan and Karaky studies [8,9], and genital isolates in the Mogielnicki study [10]), different sampling intervals [8–10], or different collection sites (Beirut in the three Lebanese studies

Table 3. FCT typing classification according to genetic variability by FCT-PCR typing and association of FCT types with *sof* gene, *emm* patterns, clusters and types.

Identified FCT types	*prtF1*	*Cpa*	*prtF2*	*sipA2*	*bridge*	*fctA*	*srtC2*	*fctB*	*srtB*	*rofA*	*nrA*	Total nb of isolates	*sof* (nb)	*emm* pattern (nb)	*emm* type (nb)	*emm cluster* (nb)
Previously described FCT (34 isolates)																
1A	+	-	-	-	-	-	-	-	+	+	-	11	*sof*-(10), *sof*+ (1)	A–C (9), D (2)	*emm1* (7), *emm6.103* (1), *emm6.0* (1), *emm108.1* (1), *emm81* (1)	A–C3 (7), clade Y (2), D4 (1), E6 (1)
2A	-	-	-	-	-	-	-	-	+	+	-	6	*sof*- (4), *sof*+ (2)	A–C	*emm1*	A–C3
3A	-	+	+	+	+	+	+	+	-	-	+	3	*sof*-	D	*emm101* (2), *emm83.1* (1)	D4
4A	+	+	+	+	+	+	+	+	+	+	-	1	*sof*+	A–C	*emm12*	A–C4
5A	+	-	-	-	-	-	-	-	-	+	-	7	*sof*+ (6), *sof*- (1)	E (6), A–C (1)	*emm2* (2), *emm1* (1), *emm48.0* (1), *emm58.8* (1), *emm113* (1), *emm165* (1)	E4 (2), A–C3 (1), E6 (1), E3 (2), E1 (1)
6A	-	-	-	-	-	-	-	-	-	+	-	4	*sof*+	E	*emm2*	E4
7A	+	-	+	+	+	+	+	+	+	+	-	2	*sof*+	A–C (1), E (1)	*emm12* (1) *emm89* (1)	A–C4 (1) E4 (1)
Single locus FCT variants (7 isolates)																
3B	-	+	+	+	+	+	+	+	+†	-	+	1	*sof*+	E	*emm165*	E1
	-	+	-†	+	+	+	+	+	-	-	+	1	*sof*-	D	*emm101*	D4
5B	+	-	+†	-	-	-	-	-	-	+	-	3	*sof*+ (2), sof- (1)	E (2), A–C (1)	*emm5.74* (1), *emm75* (1), *emm165* (1)	Clade Y (1), E6 (1), E1 (1)
	+	-	-	-	-	-	-	+†	-	+	-	1	*sof*+	E	*emm48.1*	E6
8B	+	+	+	+	+	+	+	-†	+	-	+	1	*sof*+	E	*emm87*	E3

†In cases or rows means a different locus or new FCT region respectively found by comparison with reference FCT variants. Nb for the isolate number.

Table 3. FCT typing classification according to genetic variability by FCT-PCR typing and association of FCT types with *sof* gene, *emm* patterns, clusters and types (cont.).

Identified FCT types	*prtF1*	*Cpa*	*prtF2*	*sipA2*	*bridge*	*fctA*	*srtC2*	*fctB*	*srtB*	*rofA*	*nrA*	Total nb of isolates	*sof* (nb)	*emm* pattern (nb)	*emm* type (nb)	*emm cluster* (nb)
Double locus FCT variants (13 isolates)																
4C	+	+	+	+	+	+	-†	+	-†	+	-	1	*sof*-	D	*emm41.2*	D4
	+	+	+	+	-†	-†	+	+	+	+	-	1	*sof*+	E	*emm44*	E3
	+	+	+	+	+	-†	+	-†	+	+	-	2	*sof*+	E	*emm8, emm89*	E4
5C	+	-	+†	-	-	-	-	+†	-	+	-	1	*sof*-	A–C	*emm6.4*	Clade Y
7C	+	-	+	+	-†	+	-†	+	+	+	-	2	*sof*+	A–C	*emm12*	A–C4
	+	-	+	+	+	-†	+	-†	+	+	-	2	*sof*+	E	*emm22, emm28*	E4
	+	-	+	+	-†	-†	+	+	+	+	-	1	*sof*+	E	*emm82*	E3
8C	+	+	+	+	-†	+	+	+	-†	-	+	2	*sof*-	D	*emm41.0, emm41.2*	D4
	+	+	+	+	+	+	+	-†	-†	-	+	1	*sof*+	E	*emm118*	E3
Triple locus FCT variants (5 isolates)																
4D	+	+	+	+	-†	-†	+	-†	+	+	-	1	*sof*+	E	*emm87*	E3
7D	+	-	+	+	-†	-†	+	-†	+	+	-	1	*sof*+	E	*emm8*	E4
	+	-	+	+	-†	-†	+	+	-†	+	-	1	*sof*+	E	*emm49*	E3
8D	+	-†	+	+	-†	-†	+	+	+	-	+	1	*sof*+	A–C	*emm12*	A–C4
	+	-†	+	+	-†	-†	+	+	+	-	+	1	*sof*+	E	*emm49*	E3
Newly described FCT variants (4 isolates)																
X1†	-	-	+	+	+	-	+	-	+	-	+	1	*sof*-	A–C	*emm3.93*	A–C5
X2†	+	-	+	+	-	-	-	-	+	+	-	1	*sof*+	E	*emm87*	E3
X3†	-	-	+	-	+	-	-	+	-	+	-	1	*sof*+	E	*emm2*	E4
X4†	+	-	+	-	+	-	+	-	-	+	-	1	*sof*-	A–C	*emm6.103*	CladeY

†In cases or rows means a different locus or new FCT region respectively found by comparison with reference FCT variants. Nb for the isolate number.

Table 4. Distribution of *emm* types identified in this study into 30-valent vaccine protection categories.

emm cluster	Vaccine *emm* types (Nb)	*emm* types with positive cross opsonization (Nb)	*emm* types with unknown cross opsonization (Nb)
A–C3	1 (14)		
A–C4	12 (5)		
A–C5	3 (1)		
Clade_Y	6 (4), 5 (1)		
D4	83 (1)		41 (3), 101 (3), 108 (1)
E1			165 (3)
E3	44 (1), 49 (2), 58 (1), 82 (1), 87 (3), 118 (1)		113 (1)
E4	2 (7), 22 (1), 28 (1), 89 (2)	8 (2)	
E6	75 (1), 81 (1)	48 (2)	

vs northern Lebanon in this study). A detailed comparison showed the persistence of seven *emm* types (*emm1*, *emm12*, *emm22*, *emm28*, *emm49*, *emm75* and *emm89*) detected at different occurrence rates with the remarkable absence of some of the previously reported *emm* types (*emm4*, *emm9*, *emm11*, *emm76*, *emm77* and *emm85*) [8–10].

Our findings were in line with other studies including one from Lebanon revealing that *emm1* was the dominant circulating type [8,17–19]. Noteworthy, the contemporary renaissance of iGAS infections has been linked to the dissemination of a hypervirulent *emm1* clone (the so-called M1T1) [20]. Most *emm* types in this study linked to isolates recovered from blood samples of sepsis cases were notorious for their invasiveness and their associations with invasive diseases worldwide [7,17,21]. *emm89* was previously associated with a 19% case fatality rate for iGAS infections in Toronto, Canada [19]. Other rare *emm* types were also detected herein, such as *emm113*, *emm165* and *emm101* [22–24]. Interestingly, *emm* types largely prevailing in other countries were either not detected or had low prevalence in this study, such as *emm11* common in Taiwan [25], and *emm77* in Finland [26]. Moreover, *emm3* one of the most common worldwide *emm* types being ranked fourth at the global level (after *emm*1, *emm12*, *emm28*) [7] as well as at the European and North American scale (after *emm1*, *emm28*, *emm89*) [17], it accounted only for a small percentage (one isolate) congruently with France [27]. Subtype 3.93 described in this study, was also identified in Scotland, England and Ireland [28]. *emm87* and *emm83* however, were remarkably associated with cellulitis according to a large scale European surveillance study of iGAS infections [29]. Interestingly, both *emm* types in this study were linked to skin being recovered from wounds and discharges.

The *emm* pattern scheme classifies M proteins into *emm*-patterns based on specific markers within the *emm* gene and *emm*-like flanking genes and serves as a potential marker to predict the preferred tissue site of GAS infection [30]. In our study, the *emm* pattern E strains (generalist strains), showing affinity to throat or skin, were the predominant ones followed by *emm* pattern A–C (regarded as throat specialists), and *emm* pattern D strains (regarded as skin specialists). Prevalence of pattern A–C in absence of isolates recovered from throat infections could be explained by the fact that ear being proximal to the throat, accounted for more than half (60.3%) of the isolates, and 72% of *emm* pattern A–C isolates. *emm* pattern A–C was also associated with iGAS [31], and with three out of our nine isolates recovered from the blood.

While the distribution of *emm* types could change significantly over time, the *emm* clusters, defined based on binding and structural properties of M proteins, could remain stable and be used together with *emm* types to follow GAS infections over time [19,31–33]. The 10 identified clusters in this study, were somewhat similar to the 11 *emm* clusters responsible for 99.8% of iGAS in the USA (E4, A–C3, E3, A–C4, E6, A–C5, E1, E2, D4, M6, M5) [31]. Consistent with our findings, A–C3 was the predominant *emm* cluster responsible for iGAS and streptococcal toxic shock syndrome in Alberta, Canada from 2003 to 2017 [19]. Besides monitoring GAS epidemiology, *emm* cluster typing can be useful to also predict tissue tropism [30]. Consistent with our findings regarding the association of *emm* cluster D4 (a plasminogen-binding *emm* cluster) with skin, a significant number of skin isolates were grouped in cluster D4 [16,30].

Resistance to erythromycin in this study was low (8%) and fits within the lower ends of the ERGAS rate reported by several studies from Lebanon (4.2–23%) [34,35]. Additionally, the predominance of M resistance phenotype over the cMLSb phenotype was consistent but not with all previous studies from Lebanon [8,9,35]. The low number of ERGAS strains identified in the present study could be linked to the limited number of strains recovered from throat swabs. However, a downtrend of ERGAS rate was reported in several European countries as well as in

Lebanon without any considerable reduction in macrolide usage [34]. The decline remains vague but could be correlated to the high fitness cost of ERGAS and the success of erythromycin-susceptible clones transmission [5,36]. *emm1*, the dominant type detected in this study and associated with life-threatening clinical features, retained susceptibility to all antistreptococcal agents including macrolides and tetracycline [36]. Interestingly, 60% (3/5) of the studied *emm12* isolates were resistant to erythromycin, while being susceptible to tetracycline. *emm12* clone showing co-resistance to tetracycline and macrolide was associated in 2011 with the upsurge of scarlet fever in Hong Kong, with resistance determinants being carried on integrative conjugative elements [37]. The lack of currently detected co-resistance could be attributed either to differences in circulating subclones or the loss of resistance to tetracycline associated with a high fitness cost. Resistance to tetracycline was 20.6% lower than what was previously reported in Lebanon (37%) [9] and Yemen (68.8%) [38] but higher than those in Germany (9.8%) [39]. To our best knowledge, we showed statistically significant previously untackled associations that require further validation and characterization, such as the finding that 88.9% of *emm* pattern D isolates were tetracycline resistant, with all the ones typed as *emm101* and *emm41.2* being within the resistant population.

In this study, all 11 SAgs genes were detected along with the predominance of *speG* (73%), *speC* (65%) and *speJ* (39.7%). These three SAgs also prevailed in iGAS isolates in Germany from 2009 to 2014 [40], although noteworthy that the distribution of virulence genes differed heavily between studies and can be time- and/or region dependent [22]. Studying virulence genes particularly those encoding SAgs has been widely used in examining disease associations and GAS typing [28,41]. Several studies suggested that SAgs were a better indicator than *emm* types in predicting disease manifestations [42]. *speH, speJ* and *speK* were associated with chronic skin lesions [40], *speH* and *speI* with septic arthritis and invasive wound infections [28,43], and *speA* and *speC* with streptococcal septic shock syndrome [36]. Although we found a significant overrepresentation of *speC* in isolates recovered from ear infections, we could not further elaborate on any correlation between invasive diseases and SAgs due to the absence of pharyngitis-linked isolates and lack of data related to the invasiveness of wound isolates. Although SAgs are diverse and dynamic with the continuous surge appearance of new combinations [2], phage-encoded SAgs showed some associations with certain *emm* types, indicating a presumptive link between phage preference and *emm* type. While the 14 isolates with *emm1* type were stratified in 11 virulence gene profiles (Table 2), our results were similar to other studies about the significant association of *speA* and *speJ* with *emm1* [9,41], and the lack of *speL* or *speM* from *emm1* isolates [41]. Interestingly, s*peC* occurred in 50% of our *emm1* isolates as reported elsewhere [41] and was infrequently associated with *speA*. Such association between *speA* and *speC* was found to be more frequent in isolates from patients with scarlet fever than in those from patients with pharyngitis [44]. Besides, *speI* was significantly herein linked to *emm12*, mirroring thus the situation in Western Norway where all *emm12* isolates either invasive or noninvasive ones harbored the *speI* [45]. Notably, *speH* was found exclusively in *emm* pattern A–C. Although some studies outlined an association between *emm12* and *speH* [29,42], those investigating linkage between *emm* clusters or patterns and SAgs are rare and urgently needed to draw global conclusions.

To date, nine distinct FCT variants have been identified [4]. However, the glut of unprecedented FCT regions detected in this study as well as in the Bencardino study [3] could imply either the versatility of GAS virulence mechanisms or the extent of genetic recombination events within this region triggered by the immune system pressure on the different surface-exposed GAS elements including those encoded by FCT-encoding genes. More detailed studies should be done to elucidate factors contributing to the observed FCT diversity. In this study, the three most frequent identified FCT regions were the FCT-1A (17.4%), FCT-5A (11.1%) and FCT-2A (9.5%). In contrast, Kratovac *et al.* showed that FCT-3 (32%), FCT-4 (27%) and FCT-5 (17%) were the most prevalent ones [4], while with Bencardino *et al.* it was FCT-5 (16%) and FCT-4 (10%) [3]. The observed FCT variability could be due to differences in the studied isolates, with a significant correlation being previously reported between FCT and several detected *emm* types or *emm* patterns, highlighting a possible FCT role in GAS tissue specificity [30]. While *emm* pattern E prevailed in our collection, the predominance of *emm1* interpreted the greater prevalence of FCT-1 (which is significantly associated with *emm1*) contrasting thus the initial findings of the preponderance of FCT-3 and FCT-4 [3,4]. Our findings re-emphasized the association between *emm* pattern A–C and FCT-1A, *emm* pattern D and FCT-3A, and *emm* pattern E and FCT-5A. Additionally, and as previously reported, FCT-2 was only detected in *emm1* (nearly in half of the *emm1* isolates), partly revealing the unusual success of this clone through its unique pilus structure [30]. FCT-6A, a rare FCT region, was first described in *emm2* isolates [4]. Congruently, we also found an exclusive association between FCT-6A and *emm2*. Additionally, we studied the presence and distribution of the *sof* gene. As expected, all *emm* pattern E strains had the *sof* gene while all the skin specialist (*emm* pattern D) and most of the throat specialist (*emm* pattern A–C) strains proved to be negative for the gene [30].

Combining the *emm* types, virulence genes and FCT types along with antibiotic resistance profiles, revealed the presence of 61 different patterns (or genotypes) among the isolates in this study, underlining thus a great degree of clonal heterogeneity that exceeds previous reports [3]. Bencardino *et al.* identified 83 profiles among 122 pharyngeal isolates using the same approach [3]. Intriguingly, this intraclonal divergence driven by the loss and gain of SAgs or by the structural variability within the FCT regions could contribute to the persistence of clones showing enhanced fitness or virulence [46].

With the lack of a licensed vaccine, strategies to develop one were essentially based on the M protein [34]. To maximize the expected coverage for the previously insufficiently tested vaccines, a new refined 30-valent M-protein-based vaccine was developed that induced bactericidal antibodies not only against all the 30 vaccine types but also against some nonvaccine types likely present within *emm* clusters [47]. In New Caledonia (a pacific region with high *emm* diversity), the theoretical protective coverage of the 30-valent vaccine increased from 24.2% after including the vaccine *emm* types to 76.1% after including the cross-opsonization *emm* types [32]. When *emm* types with positive cross-opsonization were not considered, the coverage rate of the 30-valent vaccine varied in the MENA region from 42 to 100% [34]. The 30-valent vaccine coverage of our study fits within the MENA range with being 76% without considering the cross-opsonization issue and reaching up to 82.5% when taking it into account. However, this value could further increase, as five *emm* types encompassing 11 isolates (of which four of nine blood isolates) had unknown cross-opsonization status. However, the D4 cluster encompasses many *emm* types with unknown cross-opsonization reactions [16], which urges the rapid experimental characterization of *emm* types for isolates linked to invasive and skin infections.

Conclusion & future perspective

In summary, this study evidences the diversity of the circulating GAS strain pool involved in nonpharyngeal infections in Lebanon. More studies are needed to understand and elucidate the molecular epidemiology of GAS. We need to have established surveillance mechanisms and programs for infectious diseases. Surveillance programs will help in better determining vaccine candidates and will provide tools to continuously monitor the emergence of new clones.

Summary points

***Streptococcus pyogenes*: overview**

- *Streptococcus pyogenes* (group A *Streptococcus* or GAS) is a common pathogen responsible for a large spectrum of clinical manifestations including severe invasive diseases.
- Of hallmark peculiarities of GAS is its outstanding ability to accumulate a wide array of gene-encoding virulence factors, to develop antibiotic resistance and to dynamically reshape its population structure.
- There is no licensed vaccine yet, and strategies to develop one were essentially based on the M protein as the 30-valent vaccine.

***Streptococcus pyogenes*: epidemiology of nonpharyngeal GAS in northern Lebanon**

- This study is the first to examine a considerable collection of GAS isolates (63 isolates) involved in nonpharyngeal infections in Lebanon.
- In line with international data, *emm1* was the predominant clone along with other *emm* types notorious for their invasiveness.
- The resistance rate to erythromycin was 8% with a slight predominance of M phenotype over the constitutive macrolides, lincosamides, streptogramin b phenotype.
- The resistance to tetracycline was 20.6% with a significant previously untackled association with *emm* pattern D.
- Profiling virulence-encoding genes demonstrated convoluted associations along with the omnipresence of *speB* followed by *speG* and *speC*.
- A wide variety of FCT regions (30 variants) was disclosed with the description of novel ones.
- Although clones were genetically heterogeneous, significant associations were revealed between *emm* types, patterns or clusters with either FCT types, *sof* gene or virulence genes.
- The 30-valent vaccine coverage was 76% when considering only the *emm* types included in the vaccine formulation, and 82.5% when also including the *emm* types with positive cross-opsonization.

Conclusion & perspective

- This study adds new diversity scores in the circulating GAS strain pool.
- More studies are badly needed to appraise at real-time the efficacy of the vaccine candidates and to surveil the potential emergence of new clones.

Author contributions

R Rafei, M Osman, F Dabboussi, S Tokajian, M Hamze was responsible for study conception and design; R Rafei, M Hawli, M Osman, T Salloum, S Khelissa performed the experiments; all authors participated in data analysis, and drafting and revision of the manuscript

Acknowledgments

The authors thank T Abdo, N Lazkani, C Khairallah, A Allouch, I Darwich and M Yehya for their excellent technical assistance.

Financial & competing interests disclosure

This study was supported by the Lebanese University. The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

No writing assistance was utilized in the production of this manuscript.

Ethical conduct of research

Ethical approval was not required as clinical isolates were collected and stored as part of routine clinical care.

References

Papers of special note have been highlighted as: • of interest; •• of considerable interest

1. Gherardi G, Petrelli D, Di Luca MC *et al.* Decline in macrolide resistance rates among *Streptococcus pyogenes* causing pharyngitis in children isolated in Italy. *Eur. J. Clin. Microbiol. Infect. Dis.* 34(9), 1797–1802 (2015).
2. Friães A, Pinto FR, SilvA–Costa C, Ramirez M, Melo-Cristino J. Superantigen gene complement of *Streptococcus pyogenes* – relationship with other typing methods and short-term stability. *Eur. J. Clin. Microbiol. Infect. Dis.* 32(1), 115–125 (2013).
3. Bencardino D, Luca MCD, Petrelli D, Prenna M, Vitali LA. High virulence-gene diversity in *Streptococcus pyogenes* isolated in central Italy. *Peer J.* 7, e6613 (2019).
4. Kratovac Z, Manoharan A, Luo F, Lizano S, Bessen DE. Population genetics and linkage analysis of *loci* within the FCT region of *Streptococcus pyogenes*. *J. Bacteriol.* 189(4), 1299–1310 (2007).

•• **Evidence of strong linkage between *loci* within FCT types and emm regions and/or patterns.**

5. SilvA-Costa C, Friães A, Ramirez M, Melo-Cristino J. Macrolide-resistant *Streptococcus pyogenes*: prevalence and treatment strategies. *Expert Rev. Anti Infect. Ther.* 13(5), 615–628 (2015).

•• **An interesting review that discusses macrolide resistance mechanisms in *Streptococcus pyogenes* and the epidemiology of resistant macrolide *Streptococcus pyogenes*.**

6. Barnett TC, Bowen AC, Carapetis JR. The fall and rise of group A *Streptococcus* diseases. *Epidemiol. Infect.* 147, 1–6 (2018).

•• **A useful review which highlights the prominent role of genetic changes in the emergence of new group A *Streptococcus* clones with increased potential to cause disease.**

7. Steer AC, Law I, Matatolu L, Beall BW, Carapetis JR. Global emm type distribution of group A streptococci: systematic review and implications for vaccine development. *Lancet Infect. Dis.* 9(10), 611–616 (2009).

•• **A systematic review which represents the global distribution of emm types of *Streptococcus pyogenes* and potential implications on vaccine development.**

8. Bahnan W, Hashwa F, Araj G, Tokajian S. emm typing, antibiotic resistance and PFGE analysis of *Streptococcus pyogenes* in Lebanon. *J. Med. Microbiol.* 60(Pt 1), 98–101 (2011).
9. Karaky NM, Araj GF, Tokajian ST. Molecular characterization of *Streptococcus pyogenes* group A isolates from a tertiary hospital in Lebanon. *J. Med. Microbiol.* 63(Pt 9), 1197–1204 (2014).
10. Mogielnicki NP, Schwartzman JD, Elliott JA. Perineal group A streptococcal disease in a pediatric practice. *Pediatrics* 106(2 Pt 1), 276–281 (2000).
11. El Ashkar S, Osman M, Rafei R *et al.* Molecular detection of genes responsible for macrolide resistance among *Streptococcus pneumoniae* isolated in north Lebanon. *J. Infect. Public Health.* 10(6), 745–748 (2017).
12. Pérez-Trallero E, Montes M, Orden B, Tamayo E, García-Arenzana JM, Marimón JM. Phenotypic and genotypic characterization of *Streptococcus pyogenes* isolates displaying the MLSB phenotype of macrolide resistance in Spain, 1999 to 2005. *Antimicrob. Agents Chemother.* 51(4), 1228–1233 (2007).
13. CDC. *Streptococcus* laboratory. http://www.cdc.gov/streplab/index.html
14. CDC. Assignment of emm type, subtype, predicted pattern and predicted cluster. https://pubmlst.org/spyogenes/info/Assignment_of_emm_type.pdf

15. Borek AL, Obszańska K, Hryniewicz W, Sitkiewicz I. Detection of *Streptococcus pyogenes* virulence factors by multiplex PCR. *Virulence* 3(6), 529–533 (2012).
16. Giffard PM, Tong SYC, Holt DC, Ralph AP, Currie BJ. Concerns for efficacy of a 30-valent M-protein-based *Streptococcus pyogenes* vaccine in regions with high rates of rheumatic heart disease. *PLoS Negl. Trop. Dis.* 13(7), e0007511 (2019).
17. Gherardi G, Vitali LA, Creti R. Prevalent emm types among Invasive GAS in Europe and North America since Year 2000. *Front. Public Health.* 6, 59 (2018).

• **Representation of dominant emm types involved in invasive *Streptococcus pyogenes* infections in Europe and North America.**

18. Nelson GE, Pondo T, Toews K-A *et al.* Epidemiology of invasive group A streptococcal infections in the United States, 2005–2012. *Clin. Infect. Dis.* 63(4), 478–486 (2016).
19. Tyrrell GJ, Fathima S, Kakulphimp J, Bell C. Increasing rates of invasive group A streptococcal disease in Alberta, Canada; 2003–2017. *Open Forum Infect. Dis.* 5(8), ofy177 (2018).
20. Cole JN, Barnett TC, Nizet V, Walker MJ. Molecular insight into invasive group A streptococcal disease. *Nat. Rev. Microbiol.* 9(10), 724–736 (2011).
21. Ikebe T, Endo M, Ueda Y *et al.* The genetic properties of *Streptococcus pyogenes* emm49 genotype strains recently emerged among severe invasive infections in Japan. *Jpn J. Infect. Dis.* 57(4), 187–188 (2004).
22. Friães A, Melo-Cristino J, Ramirez M. Portuguese Group for the Study of Streptococcal Infections. Changes in emm types and superantigen gene content of *Streptococcus pyogenes* causing invasive infections in Portugal. *Sci. Rep.* 9(1), 18051 (2019).
23. Plainvert C, Dinis M, Ravins M *et al.* Molecular epidemiology of sil locus in clinical *Streptococcus pyogenes* strains. *J. Clin. Microbiol.* 52(6), 2003–2010 (2014).
24. Athey TBT, Teatero S, Sieswerda LE *et al.* High incidence of invasive group A *Streptococcus* disease caused by strains of uncommon emm types in Thunder Bay, Ontario, Canada. *J. Clin. Microbiol.* 54(1), 83–92 (2016).
25. Chiang-Ni C, Zheng P-X, Wang S-Y *et al.* Epidemiology analysis of *Streptococcus pyogenes* in a Hospital in Southern Taiwan by use of the updated emm cluster typing system. *J. Clin. Microbiol.* 54(1), 157–162 (2016).
26. Vähäkuopus S, Vuento R, Siljander T, Syrjänen J, Vuopio J. Distribution of emm types in invasive and non-invasive group A and G streptococci. *Eur. J. Clin. Microbiol. Infect. Dis.* 31(6), 1251–1256 (2012).
27. Lepoutre A, Doloy A, Bidet P *et al.* Epidemiology of invasive *Streptococcus pyogenes* infections in France in 2007. *J. Clin. Microbiol.* 49(12), 4094–4100 (2011).
28. Meehan M, Murchan S, Gavin PJ, Drew RJ, Cunney R. Epidemiology of an upsurge of invasive group A streptococcal infections in Ireland, 2012–2015. *J. Infect.* 77(3), 183–190 (2018).
29. Luca-Harari B, Darenberg J, Neal S *et al.* Clinical and microbiological characteristics of severe *Streptococcus pyogenes* disease in Europe. *J. Clin. Microbiol.* 47(4), 1155–1165 (2009).
30. Bessen DE, Smeesters PR, Beall BW. Molecular epidemiology, ecology, and evolution of group A Streptococci. *Microbiol. Spectr.* 6(5), CPP3-0009-2018 (2018).

•• **An eloquent review which addresses the ecology, different facets of epidemiology and molecular mechanisms of genetic change at the long- and short-term of *Streptococcus pyogenes*.**

31. Smeesters PR, Laho D, Beall B, Steer AC, Van Beneden CA. Seasonal, geographic, and temporal trends of emm clusters associated with invasive group A streptococcal infections in US multistate surveillance. *Clin. Infect. Dis.* 64(5), 694–695 (2017).
32. Baroux N, D'Ortenzio E, Amédéo N *et al.* The emm-cluster typing system for group A *Streptococcus* identifies epidemiologic similarities across the Pacific region. *Clin. Infect. Dis.* 59(7), e84–92 (2014).
33. Koutouzi F, Tsakris A, Chatzichristou P *et al. Streptococcus pyogenes* emm types and clusters during a 7-year period (2007 to 2013) in pharyngeal and nonpharyngeal pediatric isolates. *J. Clin. Microbiol.* 53(7), 2015–2021 (2015).
34. Rafei R, Hawli M, Osman M, Dabboussi F, Hamze M. Distribution of emm types and macrolide resistance determinants among group A streptococci in the Middle East and North Africa region. *J. Glob. Antimicrob. Resist.* 22, 334–348 (2020).

• **Useful review that garners the circulating emm types and discusses the macrolide resistance epidemiology in *Streptococcus pyogenes* in the MENA region.**

35. Hamze M, Osman M, Mallat H, Achkar M. Prevalence and antibiotic susceptibility of ear pathogens isolated from patients in Tripoli, north of Lebanon. *The International Arabic Journal of Antimicrobial Agents* 7(1), (2017).
36. Plainvert C, Doloy A, Loubinoux J *et al.* Invasive group A streptococcal infections in adults, France (2006–2010). *Clin. Microbiol. Infect.* 18(7), 702–710 (2012).
37. Davies MR, Holden MT, Coupland P *et al.* Emergence of scarlet fever *Streptococcus pyogenes* emm12 clones in Hong Kong is associated with toxin acquisition and multidrug resistance. *Nat. Genet.* 47(1), 84–87 (2015).
38. AL Ameri AG, AL Kolaibe MA. Identification and genotypic analysis of *Streptococcus pyogenes* isolated from pharyngitis and tonsillitis infected children in IBB city in Yemen. *J. Microbiol. Antimicrob.* 7(3), 21–27 (2015).

39. Imöhl M, van der Linden M. Antimicrobial susceptibility of invasive *Streptococcus pyogenes* isolates in Germany during 2003–2013. *PLoS ONE* 10(9), e0137313 (2015).
40. Imöhl M, Fitzner C, Perniciaro S, van der Linden M. Epidemiology and distribution of 10 superantigens among invasive *Streptococcus pyogenes* disease in Germany from 2009 to 2014. *PLoS ONE* 12(7), e0180757 (2017).
41. Commons R, Rogers S, Gooding T *et al.* Superantigen genes in group A streptococcal isolates and their relationship with emm types. *J. Med. Microbiol.* 57(Pt 10), 1238–1246 (2008).
42. Rantala S, Vähäkuopus S, Siljander T *et al. Streptococcus pyogenes* bacteraemia, emm types and superantigen profiles. *Eur. J. Clin. Microbiol. Infect. Dis.* 31(5), 859–865 (2012).
43. Strus M, Heczko PB, Golińska E *et al.* The virulence factors of group A streptococcus strains isolated from invasive and non-invasive infections in Polish and German centres, 2009–2011. *Eur. J. Clin. Microbiol. Infect. Dis.* 36(9), 1643–1649 (2017).
44. Tyler SD, Johnson WM, Huang JC *et al.* Streptococcal erythrogenic toxin genes: detection by polymerase chain reaction and association with disease in strains isolated in Canada from 1940 to 1991. *J. Clin. Microbiol.* 30(12), 3127–3131 (1992).
45. Kittang BR, Skrede S, Langeland N, Haanshuus CG, Mylvaganam H. emm gene diversity, superantigen gene profiles and presence of SlaA among clinical isolates of group A, C and G streptococci from western Norway. *Eur. J. Clin. Microbiol. Infect. Dis.* 30(3), 423–433 (2011).
46. Friães A, Lopes JP, Melo-Cristino J, Ramirez M. Changes in *Streptococcus pyogenes* causing invasive disease in Portugal: evidence for superantigen gene loss and acquisition. *Int. J. Med. Microbiol.* 303(8), 505–513 (2013).
47. Frost HR, Laho D, Sanderson-Smith ML *et al.* Immune cross-opsonization within emm clusters following group A *Streptococcus* skin infection: broadening the scope of type-specific immunity. *Clin. Infect. Dis.* 65(9), 1523–1531 (2017).