11-2016

The Prevalence of Streptococcus Pyogenes and its emm Gene Types among School Children in AI Ain, UAE

Fatima Hammad Ali Al Shamisi

United Arab Emirates University

College of Science

Department of Biology

THE PREVALENCE OF *STREPTOCOCCUS PYOGENES* AND ITS *EMM* GENE TYPES AMONG SCHOOL CHILDREN IN AL AIN, UAE

Fatima Hammad Mohammed Ali Al Shamisi

This thesis is submitted in partial fulfillment of the requirements for the degree of Master of Science in Molecular Biology and Biotechnology

Under the Supervision of Dr. Abdulmajeed Alkhajeh

November 2016

Declaration of Original Work

I, Fatima Hammad Mohammed Ali Al Shamisi, the undersigned, a graduate student at the United Arab Emirates University (UAEU), and the author of this thesis entitled *"The Prevalence of Streptococcus Pyogenes and its emm Gene Types among School Children in Al Ain, UAE"*, hereby, solemnly declare that this thesis is my own original research work that has been done and prepared by me under the supervision of Dr. Abdulmajeed Alkhajeh in the College of Science at UAEU. This work has not previously been presented or published, or formed the basis for the award of any academic degree, diploma or a similar title at this or any other university. Any materials borrowed from other sources (whether published or unpublished) and relied upon or included in my thesis have been properly cited and acknowledged in accordance with appropriate academic conventions. I further declare that there is no potential conflict of interest with respect to the research, data collection, authorship, presentation and/or publication of this thesis.

Student's Signature: _____ Date: 20/12/2016

Advisory Committee

1) Advisor: Dr. Abdulmajeed Alkhajeh

Title: Associate Professor

Department of Biology

College of Science

2) Co-advisor: Dr. Rabah Iratni

Title: Associate Professor

Department of Biology

College of Science

Approval of the Master Thesis

This Master Thesis is approved by the following Examining Committee Members:

1) Advisor (Committee Chair): Dr. Abdulmajeed Alkhajeh

Title: Associate Professor

Department of Biology

College of Science

Signature _____ Date 23/11/2016

2) Member: Dr. Khaled Abbas El Tarabily

Title: Associate Professor

Department of Biology

College of science

Signature Khaled El Tarabily Date 23/11/2016

4) Member (External Examiner): Dr. Nihar Ranjan Dash

Title: Assistant professor

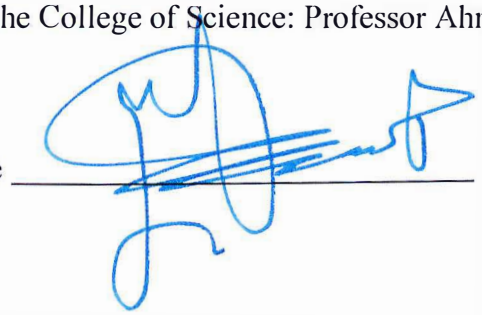
Department of Clinical Sciences

Institution: College Of Medicine-University Of Sharjah

Signature _____ Date 23/11/2016

This Master Thesis is accepted by:

Dean of the College of Science: Professor Ahmed Murad

Signature  Date 20/12/2016

Dean of the College of the Graduate Studies: Professor Nagi T. Wakim

Signature  Date 20/12/2016

Copy 6 of 6

Abstract

Streptococcus pyogenes, also known as Group A streptococcus (GAS), causes a wide variety of diseases in children and adults, and it frequently infects the pharynx and skin; pharyngitis and tonsillitis are one of the commonest diseases caused by this pathogen. GAS evaded by many virulence factors, however, the most significant part is the M protein, which plays a major role in its pathogenicity. Over 225 M protein types have been identified so far by the *emm* typing technique. Identification and *emm* typing of GAS is an essential part of epidemiological and pathogenic studies of streptococcal diseases, as different regions of the world have different common *emm* types. This diversity and the different distribution of the *emm* types among regions cause an obstacle in developing an effective vaccine. Currently, there are many studies on a 26 polyvalent vaccine targets a certain common *emm* types in the high-income countries, such as the USA, Canada, and Europe while this vaccine has low coverage in the low-income countries and the pacific. As there are not enough information about the distribution of GAS and diversity of the *emm* types in the UAE and the region, it would be difficult to evaluate the efficacy and the coverage of the proposed 26 polyvalent vaccine for the UAE and the region. The main objective of this thesis is to establish the carriage rate of Group A Streptococcal throat colonization among healthy school children in Al Ain, to identify the common *emm* types from the positive GAS isolates and to compare our findings with the most common *emm* types used to develop the 26-polyvalent vaccine. Total of 500 throat swabs samples were collected from 500 school children aged from 5-10 years old (250 male students and 250 Female students) from primary schools from Al Ain city in the UAE. The samples were collected during the month of October 2015 and

November 2015, and cultured on blood agar using the standard microbiological techniques for GAS identification. Only10% (50/500) of the school children carried GAS in their throats and a total of 50 GAS isolates were analyzed and *emm* typed using the sequence *emm* typing technique. A total of 7 different *emm* types were detected from the 50 GAS isolates, where *emm3* was the most common which accounted for 24 (48%) followed by *emm12* and *emm1*, which represented 10 (20%) and 4 (8%) respectively. These findings indicate that only 23% of *emm* types from Al Ain is covered by the 26-poly vaccine which considered a poor coverage in contrast to the 90% coverage in the western region.

Keywords: GAS, *emm* types, *emm* sequence typing, M protein, and pharyngitis.

Title and Abstract (in Arabic)

مدى انتشار البكتيريا السبحية المقيحة و أنواع جين ال emm الشائعه لهذه البكتيريا في مدينة العين في دولة الإمارات العربية المتحدة

الملخص

البكتيريا السبحية المقيحة، المعروفة ايضا باسم (GAS) تسبب العديد من الأمراض في الأطفال و البالغين على حد سواء، ولكن هذه البكتيريا غالبا ما تصيب البلعوم و الجلد. أحد أكثر الأمراض الشائعة التي تسببها هذه البكتيريا هو التهاب اللوزتين أو البلعوم. تمتلك هذه البكتيريا الكثير من العوامل التي تجعلها تسبب المرض و تساهم في انتشارها لكن أحد أهم هذه العوامل هو بروتين M. هناك حوالي 255 نوع من هذا البروتين و هو يعتبر الاساس للتقنية التي تعتمد على ايجاد سلسلة ال DNA المعطية لهذا البروتين او ما يسمى بالingemm typ. ايجاد سلسلة ال DNA لهذا البروتين في هذه البكتيريا يساعد في فهم علم الأوبئة و طريقة انتشارها وكيفية تسببها للمرض. تختلف الانواع السائده لهذه البروتينات (emm types) باختلاف المناطق الجغرافية حيث أن تعدد انواع هذه البروتينات (emm types) و اختلاف انتشار انواعها السائده باختلاف المناطق الجغرافيه يسبب عائق في خلق لقاح ضد هذه البكتيريا. حاليا، هناك العديد من الدراسات و البحوث على فعالية اللقاح المتعدد 26 الذي يستهدف الانواع السائدة في امريكا وكندا و اوروبا فقط و لكن لا يغطي الانواع السائده في منطقتنا او في دول اخرى كالدول الناميه كافريقيا و منطقة المحيط الهادئ. و بما انه لاتوجد المعلومات الكافية عن أنواع بروتين إم (emm types) في دولة الامارات والمنطقة عموما فمن الصعب تقييم فعالية هذا اللقاح و نسبة تغطيته للأنواع السائده في منطقتنا.

الاهداف الرئيسية من هذه الدراسة هي معرفة نسبة وجود هذه البكتيريا بدون اعراض في منطقة الحلق (carraige rate) لطلاب المرحله الابتدائية في مدينة العين في دولة

الاماراتالعبية المتحدة. و ايضا تحديد الانواع السائدة لجينات البروتين M (emm types) من العينات الموجبة لهذه البكتيريا و مقارنة النتائج التي سنحصل عليها مع الانواع السائدة التي تغطي اللق اح المتعدد 26 (polyvalent vaccine 26). تم اخذ 500 عينة من الحلق من 500 طالب متوسط اعمارهم من 5-الى 10 سنوات (250 ذكر و 250 انثى) من مدارس ابتدائية في مدينة العين. تم جمع العينات خلال شهر اكتوبر و نوفمبر لسنة 2015 ثم زرعت في اطباق اجار الدم باستخدام تقنيات علم الاحياء المجهريه لبكتيريا GAS.

فقط 10% (50 من ضمن 500 عينه) من طلاب المدارس حملوا هذه البكتيريا في منطقة الحلق. 50 عينة موجبة للبكتيريا تم تحليلها و ايجاد سلسلة ال DNA لبروتين M (emm types) باستخدام تقنيات البيولوجيا الجزيئية. تم ايجاد سبعة انواع مختلفه لجين بروتين M (emm types) من الخمسين عينه لبكتيريا السبحية المتقيحة (GAS). واكثر هذه الانواع شيوعا هو نوع emm3 بنسبة 48% (24 عينة) يليها emm12 ثم emm1 و التي شكلت ما نسبته 20% و 8% على التوالي. هذه النتائج توضح أن اللقاح يغطي 23% من أنواع جينات البروتين M (emm types) في مدينة العين، هذه النسبة تعتبر نسبة قليله مقارنة بنسبة فعالية هذا اللقاح في الدول الغربية و التي تشكل 90% من الأنواع السائدة هناك.

مفاهيم البحث الرئيسية: DNA، انواع بروتين ام، تسلسل ام، التهاب الحلق

Acknowledgements

Thanks to the merciful Allah for His guidance and the countless gifts He has offered me. Thanks to my loving mother for everything, thanks to my sisters and my encouraging family. Also a big thank you sent up to the sky to the rested soul of Bangei Mendoza, who helped me all through the way to samples collection. Many thanks to Al Tumoh and Shamma Bint Mohammed schools and their principles for allowing me to collect samples from their students and facilitating my task. A special thanks to Al Tumoh principl, Mrs. Wafaa, for her kindness and hospitality. I would like to express my deep gratitude to my supervisor, the Associate Professor Dr. Abdulmajeed Alkhajeh for his continuous support, collegiality, and mentorship throughout this project.

I would like to extend my thanks to all of my professors in the program of molecular biology and biotechnology.

Dedication

To my beloved mother who always keep me in her prayers, to my sister Mariam who's been supporting me all through the way, to my whole loving family, and to my little angel Talah who fills my life with sunshine and my heart with love

Table of Contents

List of Tables

List of Figures

List of Charts

List of Abbreviations

ARF	Acute Rheumatic Fever
GAS	Group A Streptococcus
NF	Necrotizing Fasciitis
RHD	Rheumatic Heart Disease
STSS	Streptococcal Toxic Shock Syndrome

Chapter 1: Introduction

1.1 *Streptococcus pyogenes*

Streptococcus pyogenes, commonly known as group A streptococcus (GAS) is a gram-positive bacteria that appear to be arranged in chains under the microscope (Figure 1). This prevalent and versatile bacterium is responsible for a broad spectrum of human diseases, ranging from superficial skin infections to an occasionally lethal syndromes leading to millions of dollars in healthcare expenditure (Cunningham, 2003). GAS spreads by respiratory secretions and fomites through asymptomatic carriers where the upper respiratory tract is the site where mostly the infections take place (Patterson, 1996). GAS has several powerful virulence factors that allow it to escape the innate immune responses of the host (Cole *et al*., 2011; Olsen *el at*., 2009; Olsen and Musser, 2010). It is characterized by being covered with a hyaluronic acid capsule (Kendall *et al*., 1941) and forming a β-haemolysis type when it is streaked on blood agar (Figure 2). In addition to that, GAS responds to a range of host environmental conditions to keep its metabolic activities intact (Kreikemeyer *et al*., 2003).

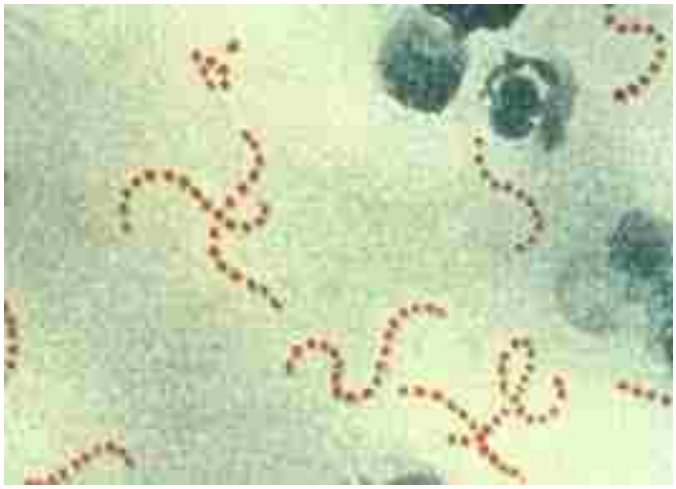


Figure 1: *Streptococcus pyogenes* (GAS)

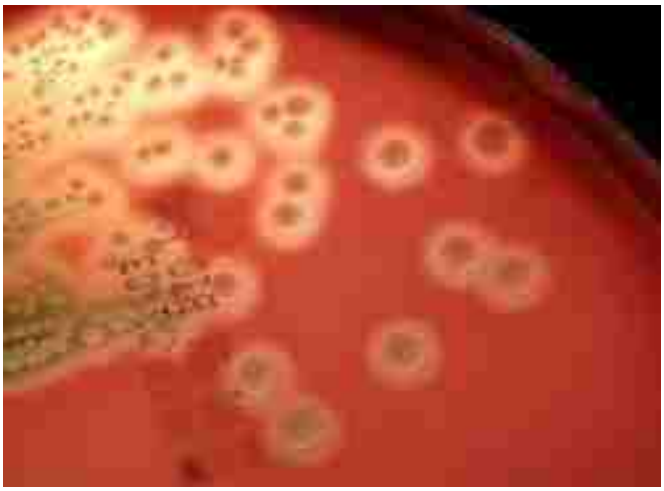


Figure 2: *Streptococcus pyogenes causes* β-haemolysis on blood agar

Consequently, these properties of GAS help the bacteria to cause a wide range of localized and invasive diseases, such as pharyngitis, impetigo, cellulitis, necrotizing fasciitis, and toxic shock syndromes (Aziz and Kotb, 2008). Despite of all the advances in hygiene procedures and methods of prevention, GAS remains a very successful bacterium that is responsible for more than half a million deaths globally each year (Carapetis *et al*., 2005) due to pharyngitis, impetigo, scarlet fever, cellulitis and erysipelas and invasive infections such as necrotizing fasciitis, myositis and streptococcal toxic shock syndrome (Ferretti *et al*., 2016). Patients may also develop immune-mediated sequelae as post streptococcal infections such as acute rheumatic fever and acute glomerulonephritis (Ferretti *et al*., 2016). Recent populations indicate that there are at least 517,000 deaths each year due to sever GAS diseases as previously mentioned (e.g., acute rheumatic fever), rheumatic heart diseases, acute post-streptococcal glomerulonephritis (APSGN), and invasive infections (Carapetis *et al*., 2005).

1.2 GAS infections

Historically, GAS diseases have been known for centuries, and the first description of streptococcal infection is attributed to Theodor Billroth, an Austin surgeon, in 1874, where he described the bacteria in cases of erysipelas and wound infections (Billroth, 1874; Billroth, 1877).

Streptococcus pyogenes, the ubiquitous organism, can cause a variety of both suppurative as well as non-suppurative post-infectious sequelae. Many of these infections were associated with high mortality rates. Table 1 lists the diseases under each category.

Suppurative	Non-suppurative post-infectious sequelae
Pharyngitis	Acute rheumatic fever (ARF)
Impetigo	Rheumatic heart disease (RHD)
Pneumonia	Acute glomerulonephritis
Necrotizing fasciitis	
Cellulitis	
Streptococcal bacteremia	
Osteomyelitis	
Otitis media	
Sinusitis	
Meningitis or brain abscess	

Table 1: Examples of Suppurative and Non-suppurative post-infectious sequelae

GAS infections can be also subdivided into four categories according to the type of tissue the bacteria attack and the way it is mediated, these subdivisions are superficial, deep, toxin mediated and immunologically mediated diseases (Wannamaker, 1970). Superficial Diseases include pharyngitis, impetigo and

infections of soft tissues. Deep infections include necrotizing fasciitis, meningitis and infections of deep soft tissues. Toxic shock like syndrome is toxin mediated while rheumatic fever and glomerulonephritis are immunologically mediated.

1.2.1 Pharyngitis

Pharyngitis (Figure 3), or commonly known as sore throat or strep throat, is the most common manifestation of infection with *Streptococcus pyogenes* (GAS) (Wessels, 2016). Infection with this bacterium is diagnosed in 20 to 40% of pharyngitis cases in children and in 5 to 15% in adults (Ebell *et al*., 2000; Shaikh *et al*., 2010). The peak incidence of GAS pharyngitis occurs in children aged 5 to 15 years old (Danchin *et al*., 2007). Pharyngitis is more common to occur during winter and spring and can be transmitted by direct or indirect contact with an infected person via large respiratory secretions droplets, or through contaminated objects and food (Wessels, 2016).

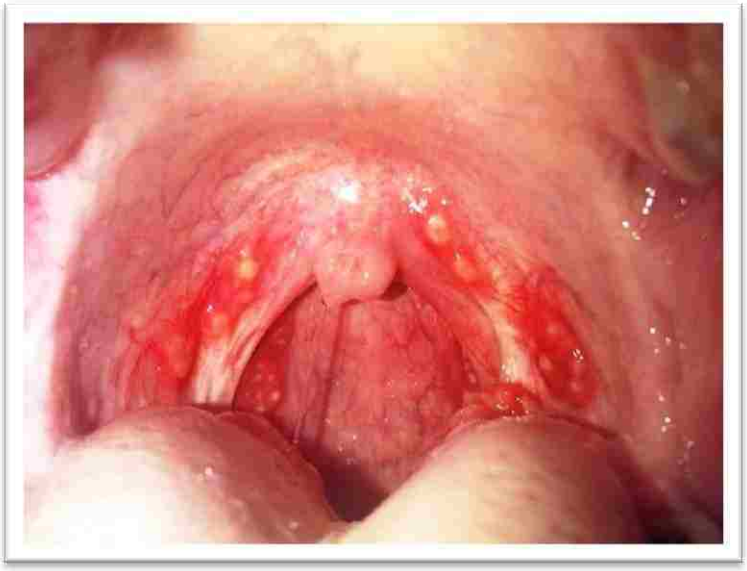


Figure 3: Infection with pharyngitis

Non-infectious auto-inflammatory complications of pharyngitis can include acute rheumatic fever and post-streptococcal glomerulonephritis, both of which are thought to result from immune responses to streptococcal infection (Wessels, 2016).

1.2.2 Acute Rheumatic Fever (ARF)

ARF and rheumatic heart disease are estimated to affect over 20 million people globally; they lead to cause cardiovascular death during the first 5 decades of life in developing countries (Bisno, 1991) and occur commonly in children of 5-15 age range. ARF, which is the precursor to RHD, results from an abnormal autoimmune response to GAS infection in a genetically susceptible host, it affects the heart, joints, brain, and subcutaneous tissue (Marijon *et al*., 2012).

1.2.3 Impetigo

Impetigo is group A streptococcal superficial infection of the outer layers of skin. It commonly affects young children (Mahe and Hay, 2005). The estimations show that more than 162 million children can be affected by impetigo at any time (Bowen *et al*., 2015). The burden of disease is highest in the low-income and developing countries. This infection is caused by invasion of the epidermis by GAS colonizing the skin following minor trauma. Autoinoculation is common and the infection is highly transmissible (Bowen *et al*., 2005). Factors that play a role in frequent transmission of the diseases in epidemic areas include hot and humid climatic conditions, poor access to water and crowded areas (Mahe and Hay, 2005).

1.2.4 Necrotizing Fasciitis (NF)

Necrotizing fasciitis (Streptococcal Gangrene) is an infection of the deeper subcutaneous tissues and fascia; it is characterized by extensive and rapidly spreading necrosis (gangrene) of the skin and underlying structures (Stevens and

Bryant, 2016). Streptococcal gangrene begins at a site of an operative incision or occurs due to curtain medical conditions (Kihiczak *et al*., 2006). In the current era, it has been estimate that the mortality rate of NF is approximately 70-80% despite the lack of antibiotics, IV fluids, ventilators, and dialysis (Stevens and Bryant, 2016).

1.2.5 Streptococcal Toxic Shock Syndrome (STSS)

Streptococcal Toxic Shock Syndrome is a fatal streptococcal toxin-mediated infection that is associated with the sudden onset of shock and organ failure (Stevens and Bryant, 2016, p570). The common ways of entry for the bacteria include the vagina, pharynx, mucosa, and skin (Stevens *et al*., 1989). In some cases, surgical procedures could lead to STSS and it is rarely occurs secondary to streptococcal pharyngitis (Herold, 1990; Bradley *et al*., 1991; Chapnick *et al*., 1992). Several studies have revealed that a high prevalence of M-types 1 or 3 strains among throat isolates may indicate an increased incidence of STSS in a community (Kiska *et al*., 1997; Holm *et al*., 1992; Sellers *et al*., 1996).

1.3 GAS virulence factors

GAS has a wide range of antigenic factors that causes the large number of streptococcal infections. These virulence factors that are engaged in both adhering and phagocytosis inhibition: (1) M protein (phagocytosis inhibition and adherence), (2) Protein F (Fibronectin-binding protein) (adherence), (3) lipoteichoic acid (adherence), (4) hyaluronic acid capsule (phagocytosis inhibition), (5) invasions such as streptokinase, streptodornase (DNase B), hyaluronidase, and streptolysins, (6) exotoxins, such as pyrogenic (erythrogenic) toxin which is responsible for scarlet fever's rash as well as systemic toxic shock syndrome (Beachey and Courtney, 1987).

1.3.1 The M protein

The cell wall of the GAS carries an important and characteristic antigenic part called 'M protein' which was identified by Rebecca Lancefield around 90 years ago (Lancefield, 1928). M protein is one, and most significant, among various antigenic structure on the cell wall of the GAS. All of the molecules on the bacterial surface have specifically evolved to enable survival in the human environment like the throat, blood, or skin, specifically; the ability of GAS to persist in infected tissues can be mainly attributed to the cell surface M protein (Fischetti, 2016). This molecule gives GAS the ability to resist phagocytosis by polymorphonuclear leukocytes in the absence of type-specific antibodies (Fischetti, 2016). Results from other studies showed that the M-knockout mutant of the M protein gene (*emm* gene) could not survive in human blood that contains phagocytes (Perez-Casal *et al*., 1992), therefore, resistance to GAS infection is attributed to the presence of type-specific M protein antibodies (Lancefield, 1959; Lancefield, 1962).

Structurally, M molecules are proteins that are attached to cell surface of the bacteria and composed of two polypeptide chains that form an alpha-helical coiled-coil configuration (Fischetti *et al*., 1990), the M protein molecules are arranged in a hair-like structure under the electron microscope as figure 4 shows (Oehmcke *et al*., 2010). Swanson, *et al.* (Swanson *et al*., 1969) published the first EM (electron micrographs) of the M protein (Figure 5), they are composed of a C-terminal region, which is the region where the M protein is attached to the cell wall of the bacteria, and it is the conserved region of the molecule, The carboxy-terminal region is hydrophobic and membrane-spanning domain that is rich in proline/glycine (Oehmcke *et al*., 2010). The M protein is also composed of an N-terminal region (amino-terminal), which is the variable region and type specific and it's the one that

accounts for its pathogenicity, the N-terminal region consists of 4 repeat regions or blocks (Figure 6) that differs in size and amino acids form a strain to another (Oehmcke *et al*., 2010). The structure of the M6 protein (Figure 6) was studied (Hollingshead, *et al.*, 1986) and showed that repeat of the As is composed of 14 amino acids, the B repeats consists 25 amino acids while the C repeats have 42 amino acids and the D repeats have small number of amino acids that showed some homology among each others (Ferretti *et al.*, 2016). The combination among these repeats in the M protein accounts for the ability of it to be variable and for the bacteria to escape the immune recognition (Fischetti *et al*., 1985; Fischetti *et al*., 1986; Hollingshead *et al*., 1987).

Figure 4: M protein molecules under electron microscope
appears as green hair-like structure

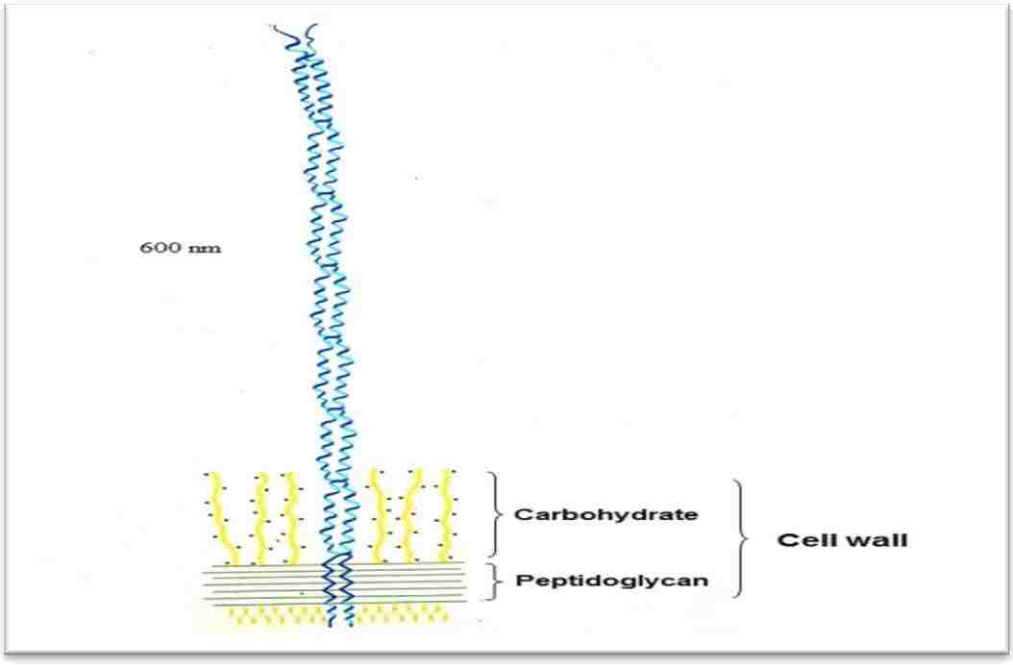


Figure 5: An electron micrograph of the M protein shows the structure of the M protein on the surface of the GAS

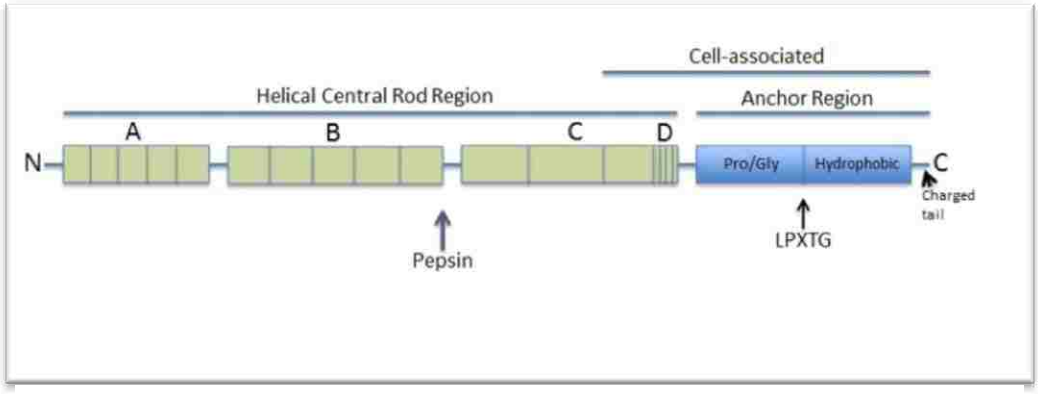


Figure 6: M6 protein sequence shows the 4 repeat blocks of the N-terminal region

1.4 Classical M- typing

Since 1919, from Lancefield work and her aim to understand the protective immunity against GAS diseases, immunity against it was believed to be type-specific; Lancefield developed a serological typing system that is based on antigen-antibody reactions; the extracted anti-phagocytic M protein surface fibrils and their type

specific antisera, the antisera against the M protein was produced by injecting rabbits with whole-cell streptococcal vaccine (Lancefield, 1928). Acceptable antisera should hold specific-precipitin antibodies and type specific antibodies that require enhancing the phagocytosis of the strain used to immunize the rabbit (Lancefield, 1957; Johnson *et al*., 1996). "The precipitin antibodies are made specific by absorption of the serum with streptococcal cells to remove the carbohydrate group antibodies and any cross reactive precipitin antibody to heterologous M-type strains. Each rabbit antiserum is tested for reaction with antigens of all known M types" (Facklam *et al*., 1999).

At the time, more than 80 distinct M types were identified (Lancefield, 1962) and more than 120 M protein validated to date (Facklam *et al*., 1999). The variable sequences of the N-terminal end of the M protein provide the basis of the classic M typing scheme (Gillespie and Hawkey, 2016). The problem with the classical typing method of the M protein is that it's time and money consuming due to the unavailability and cost of the typing reagents as well as the difficulties in their preparation and maintenance. Therefore, the classical serological way has been replaced with a molecular method that depends on sequencing the 5' end of the N-terminal region of the M protein.

1.5 M protein gene (*emm*) typing

The advances in the molecular biology field have led to the establishment of the *emm* gene typing. The *emm* gene encodes for the M protein and the variability within the N-terminal sequence of the *emm* gene accounts for the different *emm* types among the GAS strains, and recent studies shows that there are more than 234 *emm* types exist (Ferretti *et al*., 2016).

The system of *emm* typing is based on sequencing the 5’ end of *emm* gene sequence, which corresponds to the N-terminus end of the M protein. Between 160 and 660 nucleotides bases are sequenced in the 5’ end of the *emm* gene and matched with the *emm* genotype databases in the CDC (Introduction to *emm* typing, 2008) to determine the emm type) (Gillespie and Hawkey, 2016). Two highly conserved primers are used to amplify a large portion of the *emm* gene, primer 1 and primer 2. The variable sequence of the *emm* gene that encodes for the M sera-specificity lies adjacent to one of the amplifying sequences of the primers, hence, allowing for direct sequencing (Introduction to *emm* typing, 2008). If two strains of GAS share equal or greater than 95% identity in the N-terminus of the M protein then they are regarded to be of the same *emm* type (Beall *et al.*, 1996; Facklam *et al*. 2002).

The reference strains types of Dr. Lancefield was *emm* gene sequenced (5’ end *emm* sequencing) and submitted to the center for disease control (CDC). The Lancefield referenced strains are from M-type M1 to M51 and from M-type M52 to M81 (Beall *et al*., 1996; Facklam *et al*. 2002).

It was found that the *emm* sequencing results of most of the Lancefield referenced M-types matched the sequences in GenBank that was previously submitted by other researchers. These M-types are: Types M1, M2, M3, M4, M5, M6, M8, M9, M11, M12, M14, M15, M17, M18, M19, M22, M23, M24, M25, M26, M27, M28, M29, M30, M31, M33, M36, M37, M39, M41, M43, M44, M46, M47, M48, M49, M50, M51, M52, M53, M54, M55, M56, M57, M58, M59, M60, M61, M62, M63, M64, M65, M66, M72, M73, M74, M75, M76, M77, M78, M80, M81. Each has accession number in GenBank. However, M-types M13, M67, M68 and M79 were found to have different 5' emm sequences from those submitted to GenBank by previous investigators (Whatmore, *et al*, 1994). For unknown reason, the *emm* sequences of

M-type M67 and 68 in GenBank did not match those from CDC. For M-type M79, the *emm* sequence in CDC matched the one in GenBank but in a different label, which is M-type M80 (Beall *et al*., 1996; Facklam *et al*., 2002).

Large epidemiological studies of pharyngitis and invasive disease have been done using *emm* sequence typing, particularly in the USA, Canada, and Europe (Tanz *et al*., 2005). The *emm* typing has added a huge change as a reliable epidemiological tool for GAS epidemiology and subdivision; it also has the potential to classify isolates that was difficult to type by serological methods. It is also able to determine the non-typeable and weakly antigenic isolates (Beall *et al*., 1996; Facklam *et al*., 2002).

1.6 GAS epidemiology

The epidemiology and distribution of the major GAS diseases varies among different regions of the globe; despite the declined incidence of many diseases in the developed countries, but pharyngitis and invasive disease are GAS diseases of greatest public health importance in high-income countries (Beall *et al*., 2009) and the majority of GAS-associated deaths are due to the clinical manifestations associated with invasive disease (Efstratiou and Lamagni, 2016). On the other hand, regions of low-income and poor infrastructure suffer a high burden of GAS diseases like acute rheumatic fever, rheumatic heart disease (RHD), invasive disease, and acute post-streptococcal glomerulonephritis that causes millions of death annually (Carapetis *et al*., 2005). However, The RHD remains a concern in both developed and developing countries (Efstratiou and Lamagni, 2016).

GAS infections can be observed in persons of any age, but the prevalence of infection is usually higher in children more than adults because of the multiple exposures in schools or nurseries and due to their immunity compared to adults

(Efstratiou and Lamagni, 2016). The prevalence of pharyngeal infection is highest in children older than three years but in neonates it is uncommon due to the protective transplacentally acquired immunity (Martin *et al*., 2004).

For invasive diseases, which have been always the main focus of the epidemiological studies along with RHD and toxin-mediated diseases, contemporary data show that infection incidence of over 2 to 4/100,000 population in developed countries In developing countries higher rates has been observed and ranges from 12 to 83/100,000 population (Steer *et al*., 2012). Other data shows that the global burden of invasive GAS diseases is surprisingly high, with at least 663,000 new cases as well as 163,000 deaths observed annually (Carapetis *et al*., 2005).

For superficial diseases such as pharyngitis and tonsillitis, it is estimated that over 600 million cases of symptomatic GAS pharyngitis occur each year among people aged over 4 and around 550 million of these cases occur in developing countries (Gerber *et al*., 2009). In Europe, they account for around 800 consultations per 10,000 patients annually with all the economic impact of days missed from work or school (Gerber *et al*., 2009)**.** In developed countries, around 15% of school aged children develop a case of pharyngitis (generally called strep throat) each year whereas in developing countries the number is five to ten times more (Carapetis *et al*., 2005). In United States, 11 million people get infected with pharyngitis caused by GAS each year. 15-30% of the pharyngitis cases in children are caused by GAS while many cases are also caused by viral infections (Choby, 2009). Scarlet fever incidence has fallen considerably over the last century, but recent studies show increased incidence of outbreaks in some developed countries despite the fact it is no longer a life-threating disease (Efstratiou and Lamagni, 2016). Rheumatic heart

disease (RHD), which follows the infection of pharyngitis in some cases, accounts for the greatest global burden of GAS disease, 15 million cases of RHD and 349,000 deaths occur worldwide each year. This disease has a notable effect on the health of children and young adults in low and middle-income countries causing 95% of disease burden, which is a very high percentage. Other studies show that there are 2.4 million of affected children in developing countries (Steer *et al*., 2009).

1.7 GAS diseases and their associated M types

There is a significant association of some *emm* types with a particular disease manifestation (Table 2). Fore example, M types M2, M4, M6, M12, and M44/M61 are found to be associated with superficial disease, specifically, the common GAS M protein serotypes that are associated with pharyngitis are M1, M3, M4, M5, M6, M12, M14, M18, M19 and M24 (Stollerman, 2001). Types M33, M41, M42, M52, M53, M70 are associated with causing impetigo (Cunningham, 2000). Additionally, skin infections are found to be associated with M2, M49, M57, M59, M60, and M61 (Cunningham, 2000). GAS M types that are associated with the non-suppurative diseases are generally M1, M3, M5, M14, M18, M19 and M24 (Stollerman, 2001). Acute rheumatic fever, which is a non-suppurative streptococcal diseases is shown to be associated with M1, M4, M12, M49, M55, M57, M60 types (Parks *et al*., 2012), whereas acute poststreptococcal glomerulonephritis is caused by being infected with types M1, M4, M12, M49, M55, M57, M60 (Bisno, 2010). On the other hand, types M1, M3 and M28 are associated with invasive diseases such as Necrotizing fasciitis (Olsen and Musser, 2010).

Disease	Associated M types
Superficial	

Pharyngitis, scarlet fever	1, 3, 5, 6, 12, 14, 17, 19, 24
Impetigo	33, 41, 42, 52, 53, 70
Invasive	
Necrotizing Fasciitis	1, 3, 28
Streptococcal toxic shock Syndrome	1, 3
Puerperal sepsis	28
Post streptococcal infections	
Acute rheumatic fever, Rheumatic heart disease	1, 3, 5, 6, 11, 12, 14, 17, 18, 19, 24, 27, 29, 30, 32, 41
Acute post-streptococcal glomerulonephritis	1, 4, 12, 49, 55, 57, 60

Table 2: Major Group A Streptococcal infections and the associated M types

1.8 *emm* types distribution

There is a notable difference in *emm* type distribution among the different region. A systematic review was conducted (Steer *et al*., 2009) to show the global distribution of *emm* types and the most common types in each region. A total of 205 *emm* types were recorded and the most common *emm* type was found to be *emm1*, which accounted for 18.3% of all isolates in the study, second most common *emm* type was *emm12*, which accounted for 11.1% of the isolates, *emm28* (8.5%), *emm3* (6.9%), and *emm4* (6.9%). For the distribution among the different parts of the glop, the *emm* type distribution between the high-income countries, Asia, the Middle East, and Latin America, by contrast with the distribution in Africa and the Pacific region were obviously similar. Figure 7 shows a comparison between the proportions of the 25 most common *emm* types in high-income countries, Africa, and the Pacific.

In high-income countries, 25 *emm* types accounted for 90.3% of all isolates and *emm1* was the most common type. Where in Africa, 26 *emm* types accounted for only 62.5% of all isolates. Where as in the pacific region, 26 *emm* types accounted for 61.8% of all isolates. As *emm1* and *emm12* were the most common *emm* types in high-income countries, Asia, and Latin America, while the same types (*emm1* and *emm12*) came as second and third most common *emm* types in the Middle East.

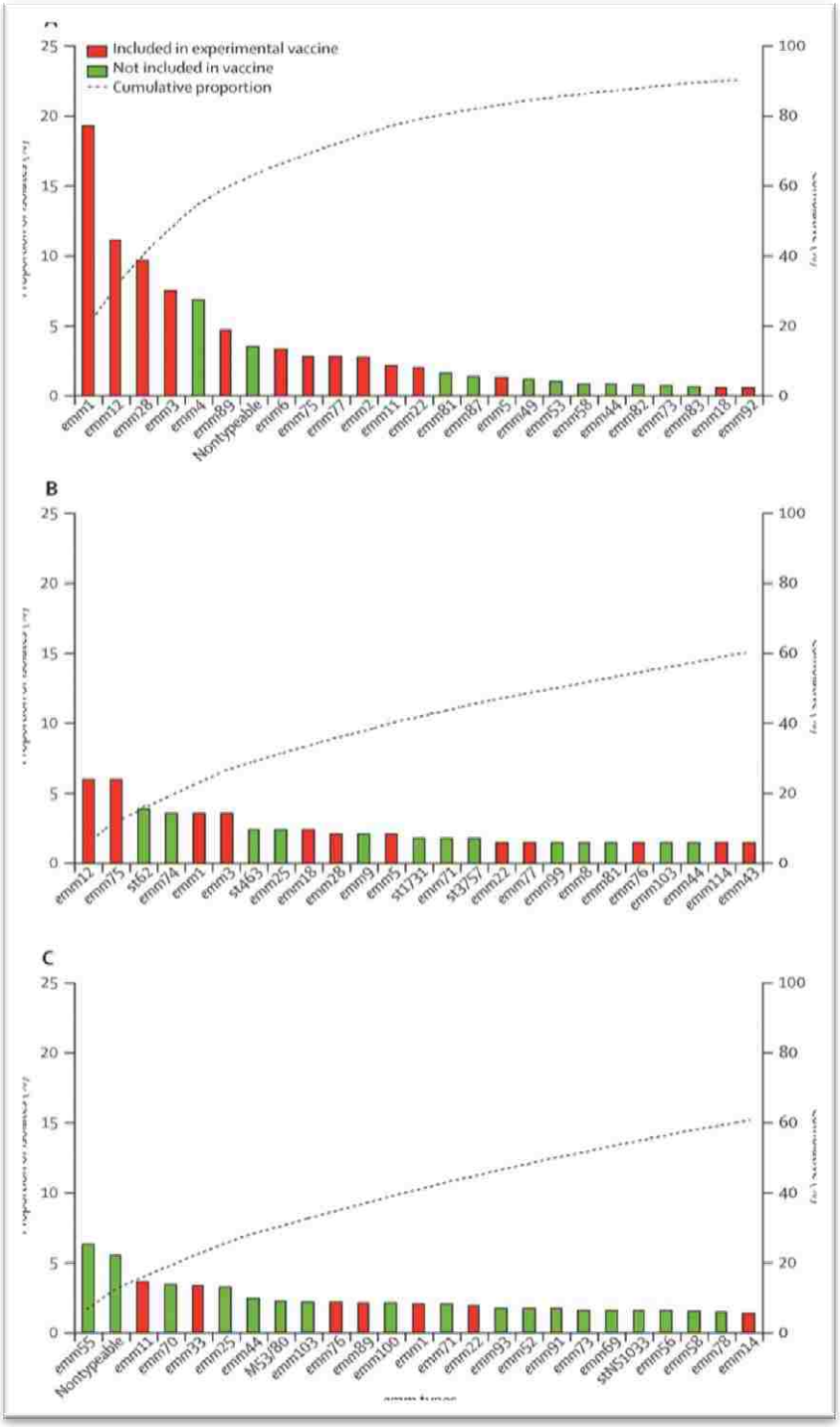


Figure 7: The most common 25 *emm* types in high-income countries (A), Africa (B), and the Pacific region (C)

Regarding *emm* diversity, the greatest diversity was found in Africa as the cumulative curve in figure 7 shows. The Middle East comes fourth after Asia and Latin America, respectively, as indicated by calculation of Simpson's index (Table 3). (Steer *et al*., 2009).

Diversity of *emm* types by global region	
	Simpson's index of diversity (% [95% CI])
Africa	98·1%
Asia	88·7%
Latin America	93·2%
Middle East	93·0%
Pacific region	97·9%
High-income countries	92·1%
Combined	92·8%

Table 3: Diversity of *emm* types by global region [Table from (Steer, et al., 2009)]

1.9 GAS vaccines

Creating an effective vaccine targeting the group A streptococcal diseases has been always the aim for the researchers and public health authorities. GAS vaccines can be either an M protein–based or non–M protein based. Since the M protein is the major virulence determinant of the GAS, the vaccines that have been under clinical trials are the N-terminal M protein based vaccines such as 26-valent and 30-valent vaccines, and the conserved M protein vaccines such as the J8 vaccine and the StreptInCor vaccine (Steer, *et al*., 2009). There are also a variety of other GAS vaccine candidates that are at different stages of development but we will focus on the 26 and 30 polyvalent vaccines.

1.9.1 26-polyvalent vaccine

The 26 Multi-valent vaccine is composed of fused recombinant protein peptides from the N-terminal region of M proteins that are encoded by their corresponding *emm* types of GAS (McNeil, *et al*., 2005). These 26 M types that are covered by this vaccine are M1, M1.2, M2, M3, M5, M6, M11, M12, M13, M14, M18, M19, M22, M24, M28, M29, M33, M43, M59, M75, M76, M77, M89, M92, M101, and M114 (Hu *et al*., 2002) which are the most common types in the high-income countries, such as the USA, Canada, and Europe with a coverage percentage of 72%. However, the vaccine has a poor coverage (only 24%) in the low-income countries such as Africa and the pacific and an intermediate coverage in Asia and the Middle East (Steer *et al*., 2009).

1.9.2 30-valent Vaccines

The 30-valent vaccine is a reformulated form of the 26-valent vaccine to increase its coverage the United States, Canada and Europe as well as developing countries (Steer *et al*., 2009). The M types that are represented in this vaccine and were not contained in the 26-valent vaccine are: M4, M29, M73, M58, M44, M78, M118, M82, M83, M81, M87, and M49, while the serotypes that are represented in the 26-valent vaccine and are not contained in the 30-valent vaccine are: M1.2, M43, M13, M59, M33, M101, and M76. The vaccine serotypes covers 98% of all cases of pharyngitis in the United States and Canada and 90% of invasive disease in the United States, in Europe, it accounts for 78% of invasive diseases (Dale *el al*., 2011).

1.10 GAS *emm* typing studies in the UAE

The first GAS typing study in the United Arab Emirates of M serotypes was conducted by Dr. Abdulmajeed Alkhajeh (published under Ameen) in 1997 and

showed that out of 100 isolates of GAS from school children only 76% were serologically typable and felled in to 14 different M types. The study showed that the commonest M types were M1, M6, M2, M22, M28 and M75 which consisted of 57% of all the isolates followed by M3, M4, M5, M60, M58/75, M12, M76, M33, and M58 which consisted of only 19 % of all the tested GAS strains. (Ameen *et al*., 1997).

Other resent study in UAE by Alfaresi 2010 published in BMC Research Notes, which demonstrated the heterogeneity of GAS population. In his study a total of 38 clinical GAS isolates were analyzed, including 35 isolates from throat and 3 from skin, among the 38 isolates, a total of 25 different *emm* types were detected, there were no dominant *emm* types detected (Alfaresi, 2010). Eight different *emm* types (st3211.0, st0721.0, *emm89*, *emm75*, st4695.0, st75.0, stIL62.0, *emm78*.3) made up 55.4% of the isolates. The most common type, which is st3211.0, compromised only 10.5% of the total isolates. The study found that some common *emm* types in the Western world are totally absent in the collection of that study, furthermore, type *emm89*, which is frequent in Canada, was found to be frequent among the isolates (Alfaresi, 2010).

More recent study by Dr. Abdulmajeed Alkhajeh (Unpublished study) showed that out of 60 clinical GAS isolates typed for *emm* gene only 4 *emm* types are represented in the 26-polyvalent vaccine, which indicates that only 15.4% of the local isolates can be covered by the vaccine. Thus the 26-polyvalent vaccine which is basically designed for the Western countries populations will be ineffective in preventing GAS infections in the UAE population.

1.11 Objectives

1. To establish the carriage rate of Group A Streptococcal throat colonization among healthy school children in Al-Ain, UAE.
2. To identify the common GAS *emm* types from the positive isolates.
3. To compare our findings of *emm* types with the most common types which are used to develop the 26-polyvalent vaccine.

Chapter 2: Materials and Methods

2.1 Materials

- Gram Stain Kit (BioLab, Australia).
- Sterile throat swabs (Puritan Opti-Tranzr®, USA)
- Loops (Celltreat®, USA)
- Cryo tubes (CryoTubes®, USA).
- Cultural Media 5% Sheep Blood Agar. (Manufactured in association with Becton Dickinson In UAE)
- Nutrient broth media (20% glycerol, PanGulf®, UAE).
- Bacitracin differentiation disc (HardyDisk™, USA).
- SLIDEX® STREPTO Plus latex agglutination Kit (BioMérieux®, France).
- Bijoux bottles (Sterillin™, USA).
- Liquid glycerol.
- Eppendorf pipette.
- Pipette tips.

2.2 Samples collection

Throat swabs samples were collected from 500 school children aged from 5-10 years old in Al Ain; 250 male students from Al Tumoh primary School and 250 Female students from Shamma Bint Mohammed primary school, in Al Ain city in the UAE. The samples were collected during the month of October 2015 and November 2015.

2.3 Area of study

This study was conducted in United Arab Emirates and samples were collected from the city of Al Ain in the United Arab Emirates (Figure 8).

Figure 8: Map of United Arab Emirates

2.4 Samples preparation and identification

2.4.1 Culture

The collected throat swab samples were cultured on sheep blood agar plates immediately, and Bacitracin disc was added and then incubated for 24 hours at 37 C^{o}. GAS colonies with a β-haemolysis or complete lysis of blood cells around the colonies (Figure 8) were sub-cultured on fresh blood agar plates to get pure colonies of the suspected group A *streptococcus*, then the colonies of the suspected GAS were picked for Lancefield grouping and for GAS confirmation.

Figure 9: β-haemolysis (clear zone) around GAS colonies

2.4.2 GAS confirmation by Lancefield grouping

SLIDEX STREPTO Plus latex agglutination Kit was used to group the β-haemolytic samples. The GAS positive samples were then taken to the microbiology laboratory in Tawam hospitals for further comparing and confirmation. The samples were confirmed to be group A streptococcal by its sensitivity to bacitracin disc and by using the Lancefield grouping kit SLIDEX Strepto plus.

2.4.3 Samples preservation

Nutrient broth media with 20% glycerol was used to preserve the GAS positive samples and stored at -20 C^{o} prior sending to a reference laboratory for *emm* typing.

2.5 *emm* typing

emm typing was performed according to the protocol described by the CDC (http://www.cdc.gov/ncidod/biotech/strep/protocol_emm-type.htm) in the group A streptococci reference laboratory at the Faculty of Medicine at the University of Lisbon in Portugal.

Chapter 3: Results

3.1 GAS identification

Out of the 500 throat swabs samples collected from the from the school children, only 65 samples showed a β-haemolytic colonies, where 50 samples confirmed to be GAS, while the other 16 β-haemolytic samples turned out to be; 11 isolates were of group G and 5 isolates were of group C.

3.2 GAS carriage rate among children

Our finding showed that the GAS carriage rate is 10% among school children, where out of 500 screened samples, only 50 GAS isolates were detected and the rate is slightly higher in boys than girls, 27 and 23, respectively.

3.3 *emm* type distribution

Our *emm* typing results showed that only 7 different *emm* types were detected from our 50 GAS isolates (Table 4). The most common type was *emm 3* with 24 isolates, which represents 48% of the total of GAS isolates (Chart1). The Second common type was *emm 12 with* 10 isolates, which represent 20%, where *emm* types 1, 22 and 89 accounts for 24% of the total GAS isolates with 4 GAS isolates of each (8%). The least common types were *emm29* and *emm4* with 2 samples of each, which both accounted for 8% of the GAS isolates.

No.	*emm* type	Frequency (out of 50 GAS)	Percentage%
1	***emm3***	24	48%
2	***emm12***	10	20%
3	***emm1***	4	8%
4	***emm22***	4	8%
5	***emm89***	4	8%
6	***emm29***	2	4%
7	***emm4***	2	4%
Total		50 GAS isolates	

Table 4: *emm* types distribution, their frequency out of the 50 GAS isolates

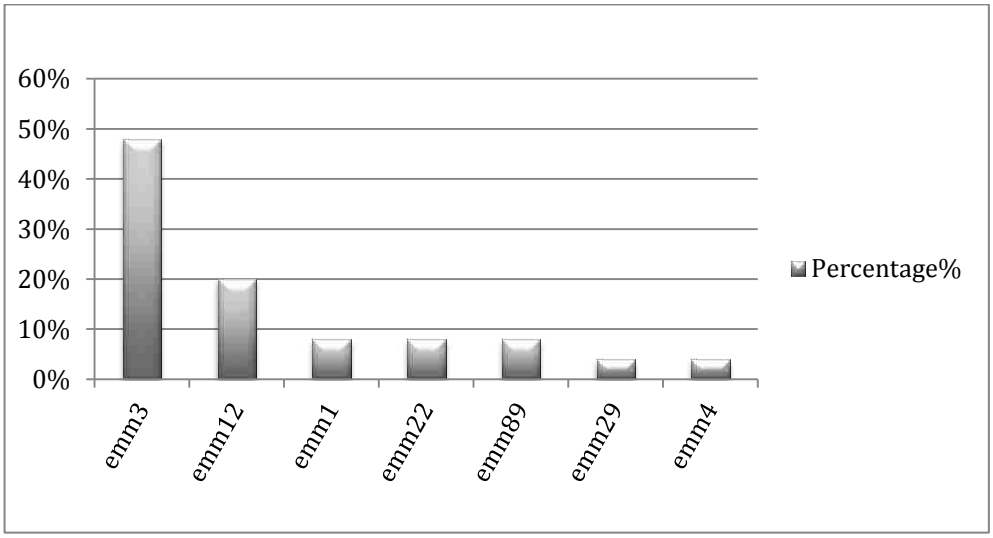


Chart 1: The 7 *emm* types and their percentage from the most common to the least

Chapter 4: Discussion

4.1 Carriage rate of GAS among children

The pharyngeal carriage rates of GAS among healthy school children vary with seasons and geographical location, and the peak incidence of spread of this bacterium is during winter and early spring (December-February). In the present study, the throat samples were collected during the period time between October and November 2015 and the carriage rate among the 500 healthy school children was found to be 10%. Previous studies in the UAE conducted by Dr. Abdulmajeed Alkhajeh in 1995 showed a carriage rate of 35.4% among 1000 school children in Al Ain, which is considered to be high, compared to our findings (Ameen *et al*., 1997). In a follow up study by Ameen, the carriage rate among the same cohort of children a year later showed to be 26% (Ameen *et al*., 1998).

In a study conducted in Ethiopia, the carriage rate for GAS was 9.7% (91/937) (Abdissa, *et al*, 2006) which is similar to our findings. In Nepalese school children, GAS was isolated from 10.9% (38/350) of the screened children, which is almost similar to our findings as well (Dumer *et al*., 2009). Other study conducted in Brazil, 2194 school children demonstrated 11% carriage rate, which is approximately similar to the carriage rate found in the present study (Tartof *et al*., 2010). Although the current GAS carriage rate is lower compared to the findings of GAS studies 20 years ago in the UAE, however, our 10% carriage is similar to many recent studies conducted around the glob as shown previously.

4.2 *emm* type distribution

Table 4 represents the different *emm* types of the 50 typable GAS isolates from school children in Al Ain and their frequencies. The litters "*emm"* refers to the sequence types that fall under the standard GAS strains, while "st" refers to the new sequence types that were discovered later (Facklam *et al.*, 2002), however, no new strains were detected from our study and all of the *emm* types belong to the standard GAS strains. The number of distinct *emm* types found in the present study was 7 types out of the 50 GAS isolates, which indicates that the genetic heterogeneity of the 50 GAS isolates from the school children in Al Ain is not considerably high. The most common type is *emm3*, 24 isolates were reported which represents 48% out of isolates. Type *emm12* is the second most common type with 10 isolates out of the 50 (20%). Type *emm1*, *emm22* and *emm89* occur in an equal frequency, which is 4 isolates out of the 50 and that account for only 4% of the isolates for each. The least common types are *emm29* and *emm4* that represent 4% for each type (Table 4). Five out of the seven GAS *emm* types (*emm*3, *emm*12, *emm*1, *emm*29 and *emm*4) found in the present study is known to be associated with major streptococcal infections such as pharyngitis, NF, ARF, STSS, RHD and glomerulonephritis.

In comparison with other previous local studies, the diversity of *emm* types in the present study from the 50 isolates is considered to be low, where a study conducted by Dr. Mubarak Alfarsi, 25 different *emm* types were identified from only 38 GAS isolates, which is considered to be a very heterogeneous (Alfaresi, 2010), however, his isolates were obtained from clinical samples, and the most common type showed to be *emm*89 and st4696.0 which is not matching with our most common types.

However, the heterogeneity of *emm* types that our results demonstrated is similar to

the heterogeneity of a study conducted by (Ameen *et al*., 1997) where 14 M types patterns were determined out of 100 GAS isolates, however, serotyping techniques were used instead of the *emm* gene sequencing technique. For the most common types, the study of (Ameen *et al*., 1997) reported that M1, M6, M2, M22, M28 and M75 were the most common types which is different that our results.

The differences in the findings among the local studies indicate that distribution of *emm* types varies over time and within different geographic region (Su *et al*, 2009). Similarly, in the United States, the most common types were *emm*1, *emm*28, *emm*12, *emm*3, and *emm*11 during the period from 1995 to 1999 (O'Brien *et al*., 2002) while from 2000 to 2004, *emm1*, *emm3* and *emm12* dominated and *emm28* was no longer among the dominant types (O'Loughlin, 2007). In Athens, the distribution of *emm* types in two different periods of times was detected (from 2003 to 2006 and 2007 to 2013), during the first period, *emm* types *emm*12, *emm*4, *emm*5, and *emm*61 were dominant but decreased significantly during the second period between 2007 and 2013. The prevalence of *emm* *emm*89, *emm*75, and *emm*11 increased notably between 2007 and 2013 and the diversity of the *emm* types shows an incline during the latter period (Koutouzi, 2015).

In a similar study in Ethiopia in which 82 GAS isolates were collected from healthy school children, a high diversity of GAS was reported with 43 different *emm* types among the 82 GAS isolates (Abdissa *et al.*, 2006).

Another similar and very recent study in Iran, where 25 GAS isolates were obtained from swapping the throats of 1000 children showed that only 3 *emm* types were found and the *emm3* type is the most common one which is similar to our findings, however, our carriage rate is three folds higher than their rate. (Khosravi *et al.*,

2016).

In a Brazilian study mentioned earlier, 61 distinct *emm* types out of 238 GAS isolates were identified with *emm12* and *emm*1 being the most common types (Tartof *et al*., 2010), which is similar to our common *emm* types.

Another study conducted in India, Bangalore, also showed a very high heterogeneity among the GAS isolates, where out of 60 GAS isolates, 35 *emm* types were detected with again *emm12* being the most common type which fell second in our present study (Gowda *et al*., 2011).

Overall, these findings share some similarities with other *emm* typing studies of other regions, which were conducted in other countries and involved the same or more numbers of GAS isolates.

4.3 *emm* types and 26-valent vaccine coverage

Clinical trials of a 26-valent M protein-based vaccine have been always in progress in the USA and Canada. The vaccine underwent a phase I/II clinical trial in human adult volunteers and was reported to be safe and immunogenic (McNeil *et al*., 2005). However, the number of M type strains causing infections in the developing countries, particularly the US, is somehow limited so the potential efficiency of the vaccine coverage in other regions will not be the same. In the US the vaccine covers around 90% of invasive GAS diseases. The 26-valent vaccine covers the following *emm* types *emm*1, *emm*1.2, *emm*2, *emm*3, *emm*5, *emm*6, *emm*11, *emm*12, *emm*14, *emm*18, *emm*19, *emm*22, *emm*24, *emm*28, *emm*29, *emm*33, *emm*43, *emm*59, *emm*75, *emm*76, *emm*77, *emm*89, *emm*92, *emm*94, *emm*101, and *emm*114 (Hu *et al*., 2002).

In our present study, 6 *emm* types were found to be covered by the 26 polyvalent

vaccine, these *emm* types are *emm1*, *emm3*, *emm12*, *emm22*, *emm29* and *emm89*, so, the percentage of coverage by the vaccine is only 23%, which is a poor coverage compared to the vaccine coverage in the high-income countries such as USA and Canada (Table5).

Although this finding shows a poor coverage of the vaccine among school children, we have to keep in mind that these are not a clinical samples and that our results shows carriage of *emm* types among healthy children and does not represent the clinical isolates from different sites of infections which might give us a total different picture, and this was demonstrated by Dr. Abdulmajeed Alkhajeh and Alfaresi in previous studies in the UAE). However, we believe that these findings will add to the previous GAS studies finding in the UAE, and will contribute in drawing the real picture of GAS *emm* typing which might help the health authorities in making the right decision to introduce such vaccine.

No.	*emm* types included in 26-polyvalent vaccine	*emm* types of present study covered by 26-polyvalaent vaccine
1	1	(4 isolates)
2	1.2	
3	2	
4	3	(24 isolates)
5	5	
6	6	
7	11	
8	12	(10 isolates)
9	13	
10	14	
11	18	
12	19	
13	22	(4 isolates)
14	24	
15	28	
16	29	(2 isolates)
17	33	
18	43	
19	59	
20	75	

21	76	
22	77	
23	89	(4 isolates)
24	92	
25	101	
26	114	
Percentage of coverage%		**6 *emm* types are covered by the polyvalent vaccine (23% coverage)**

Table 5: *emm* types of the present study that are covered by the 26-polyvalent vaccine

If we combine the findings of the previous local studies on GAS M typing in the UAE, the percentage of vaccine coverage will increase but still doesn't exceed 56%, which means the vaccine will cover only half of the patients. From the study of (Ameen *et al*., 1997), 8 M types (in a total of 49 isolates) were found to be covered by the 26-polyvalent vaccine. In (Alfaresi, 2010) study, 4 *emm* types (in 7 isolates) were reported to be covered by the. Our study reported 44 isolates distributed among 6 *emm* types as mentioned previously. Calculating the total isolates from the 3 studies, it gives totally102 (56%) isolates that are included in the 26-polyvalent vaccine (Table 6). In addition, all the studies covered only 10 *emm* types, which are *emm1, emm2, emm3, 4mm12, emm22, emm28, emm29, emm33, emm75, and emm89* as highlighted in Table 6. They represent 38% of the 26-polyvaccine types.

Then, the 56% coverage of the vaccine is not enough; this means 44% will not be covered by the vaccine, which was expected to cover 90% of invasive GAS isolates in the USA. Further studies with more number of GAS Isolates from different regions of the country will help reaching the real mapping of GAS *emm* typing in the UAE which will help measuring the efficiency of applying the 26-polvalent vaccine.

No.	*emm* types covered by 26-poly vaccine	Present study 2016 (50 isolates)	Alfaresi, 2010) (38 isolates)	(Ameen, at el, 1997) (100 isolates)	Total of isolates from the 3 studies=183
1	***emm1***	YES (4/50)	YES (1/38)	YES (17/100)	22
2	***emm1.2***				
3	***emm2***			YES (8/100)	8
4	***emm3***	YES (24/50)		YES (3/100)	27
5	***emm5***				
6	***emm6***				
7	***emm11***				
8	***emm12***	YES (10/50)		YES (1/100)	11
9	***emm13***				
10	***emm14***				
11	***emm18***				
12	***emm19***				
13	***emm22***	YES (4/50)	YES (1/38)	YES (5/100)	10
14	***emm24***				
15	***emm28***			YES (8/100)	8
16	***emm29***	YES (2/50)			2
17	***emm33***			YES (2/100)	2
18	***emm43***				
19	***emm59***				
20	***emm75***		YES (2/38)	YES (5/100)	7
21	***emm76***				
22	***emm77***				
23	***emm89***	YES 2 (2/50)	YES (3/38)		5
24	***emm92***				
25	***emm101***				
26	***emm114***				
Total of *emm* types are covered by the 26-polyvalent vaccine		**6 *emm* types (23% coverage)**	**4 *emm* types (15.4% coverage)**	**9 *emm* types (34.6% coverage)**	**11 *emm* types** **42% coverage of 26-polyvalent vaccine** **(11/26*100)**

Table 6: Total vaccine coverage of the GAS *emm* types from 3 studies in the UAE, 1997, 2010 and 2016

Chapter 5: Conclusion

The aim of this study was to determine the carriage rate of the GAS carriers among school children and to survey the genetic diversity of group A streptococcal isolates in the UAE using the *emm* gene sequence technique in order to compare our findings of *emm* types with the most common types which are used to develop the 26-polyvalent vaccine. Our GAS isolates were obtained from the throats of healthy school children aged from 6-10 years old, and they showed a carriage rate of 10%. The findings demonstrated also moderate heterogeneity nature of the GAS population among school children in Al Ain during the period when the samples were being collected. According to our findings, the vaccine will theoretically cover 23% of the studied subjects, which is a very poor coverage; only 6 *emm* types were to be covered by the 26-polyvalent vaccine. On the contrary, the previous studies in the UAE showed similar findings that also the vaccine couldn't be theoretically effective if it were to be applied in our regions as it doesn't cover most of the *emm* types and isolates in those studies.

Deciding whether the 26-polyvalent vaccine is effective or ineffective for our region is still challenging due to the lack of comprehensive information and the actual picture of GAS *emm* types and their distribution in our region. To better understand the dynamics of GAS epidemiology and the efficiency of the 26-polyvalent vaccine, further surveillance very much needed to be conducted in the region to help finding the real molecular epidemiology of this bacterium in order to be able to introduce the vaccine and control it or even design our own polyvalent vaccine.

References

Abdissa A, Asrat D, Kronvall G, et al. High Diversity of Group A Streptococcal *emm* Types among Healthy Schoolchildren in Ethiopia. *Clinical Infectious Diseases.* 2006; 42: 1362-1367.

Alfaresi MS. Group A streptococcal genotypes from throat and skin isolates in the United Arab Emirates. *Alfaresi BMC Research Notes*. 2010; 3:94

Ameen AS, Dawson KP, Mustafa N, Nsanse H. Group A streptococcal contact within schools over a 1-year period *Clinical Microbiology infection*. 1998; 4: 659-661

Ameen AS, Dawson K, Mustafa N, Nsanze H, Othman S. The prevalence of Group A Streptococcal throats carriage in Al Ain, UAE. *Annals of Tropical Paediatrics.* 1997; 16:123-127.

Aziz RK, Kotb M. Rise and persistence of global M1T1 clone of Streptococcus pyogenes. *Emerging Infectious Diseases*. 2008;14(10):1511–1517.

Beachey EH, Courtney HS. Bacterial adherence: The attachment of group A streptococci to mucosal surfaces. *Clinical Infectious Diseases.* 1987;S475-S481.

Beall B, Facklam R, Thompson T. Sequencing *emm*-specific PCR products for routine and accurate typing of group A streptococci. *Journal of Clinical Microbiology*. 1996; 34, 953–958

Beall BW, Carapetis JR, Law I, Matatolu L, Steer SA. Global *emm* type distribution of group A streptococci: Systematic review and implications for vaccine development. *Lancet Infectious Diseases.* 2009; 9:611-16

Billroth, T. Lectures on surgical pathology and therapeutics. London: The New Sydenham Society. 1877.

Billroth, T. Untersuchungen über die Vegetationsformen von Coccobacteria septica und der Antheil, welchen sie an der Entstehung und Verbreitung der accidentellen Wundkrankheiten haben. Berlin: G. Reimer. 1874.

Bisno AL. Group A streptococcal infections acute rheumatic fever. *New England Journal of Medinice.* 1991; 325:783-93.

Bisno AL. Nonsuppurative poststreptococcal sequelae: rheumatic fever and glomerulonephritis, p 2611–2622. In Mandel GL, Bennett JE, Dolan R (ed), Mandell, Douglas, and Bennett's principles and practice of infectious diseases, 7th ed. 2010. Livingstone Elsevier, Philadelphia, PA.

Bowen AC, Carapetis JR, Yeoh DK. Impetigo and scabies disease burden and modern treatment strategies. *Journal of Infection.* 2016; 72

Bowen AC, Mahe A, Hay RJ, Andrews RM, Steer AC, Tong SY, et al. The global epidemiology of impetigo: a systematic review of the population prevalence of impetigo and pyoderma. *PLoS One.* 2015; 10

Bradley JS, Schlievert PM, Peterson BM. Toxic shock-like syndrome: a complication of strep throat. *The Pediatric Infectious Disease Journal.* 1991;10(10):790.

Carapetis JR, Steer A. C., Mulholland E. K., Weber M. The global burden of group A streptococcal diseases. *The Lancet Infectious Diseases.* 2005;5(11):685–694.

Chapnick EK, Gradon DJ, Lutwick LI, Kim J, Levi M, Kim MH, et al. Streptococcal toxic shock syndrome due to noninvasive pharyngitis. *Clinical Infectious Diseases.* 1992;14(5):1074–1077.

Choby BA. Diagnosis and treatment of streptococcal pharyngitis. *American Family Physician.* 2009; 79(5):383–390.

Cole JN, Barnett TC, Nizet V, Walker MJ. Molecular insight into invasive group A streptococcal disease. *Nature Reviews Microbiology*. 2011;9:724–736.

Cunningham MW. Pathogenesis of group A streptococcal infections *Clinical Microbiology Review. 2000;* **13**: 470–511.

Cunningham MW. *Patho*genesis of group A streptococcal infections. *Clinical Microbiology Reviews. 2003; 13:470-511.*

Dale JB, Penfound TA, Chiang EY, Walton WJ. New 30-Valent M Protein-Based Vaccine Evokes Cross-Opsonic Antibodies Against Non-Vaccine Serotypes of Group A Streptococci. *Vaccine*. 2011;29(46):8175-8178.

Danchin MH, Rogers S, Kelpie L, Selvaraj G, Curtis N, Carlin JB, et al. Burden of acute sore throat and group A streptococcal pharyngitis in school-aged children and their families in Australia. *Pediatrics*. 2007; 120(5):950–957.

Dumre SP1, Sapkota K, Adhikari N, Acharya D, et al. Asymptomatic throat carriage rate and antimicrobial resistance pattern of Streptococcus pyogenes in Nepalese school children. *Kathmandu Univ Med J (KUMJ).* 2009; 7(28):392-6.

Ebell MH, Smith MA, Barry HC, Ives K, Carey M. The rational clinical examination. Does this patient have strep throat? *JAMA.* 2000; 284(22):2912–

2918.

Efstratiou A, Lamagni T. *Epidemiology of Streptococcus pyogenes. 2016 Feb 10. In: Ferretti JJ, Stevens DL, Fischetti VA, editors*. Streptococcus pyogenes: Basic Biology to Clinical Manifestations [Internet]. Oklahoma City (OK): University of Oklahoma Health Sciences Center; 2016-.[Cited: 25th September]. 465-470 P. Available from: http://www.ncbi.nlm.nih.gov/books/NBK343616/

Facklam R, Beall B, Efstratiou A, Fischetti V, Johnson D, Kaplan E, et al. *Emm* typing and validation of provisional M types for group A streptococci. *Emerging Infectious Diseases*. 1999; 5(2):247–253. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2640698/

Facklam RF, Martin DR, Lovgren M et al. Extension of the Lancefield classification for group A streptococci by addition of 22 new M protein gene sequence types from clinical isolates: *emm103* to *emm124*. *Clinical Infectious Diseases*. 2002; 34:28–38.

Ferretti JJ, Stevens DL, Fischetti VA, editors. *Streptococcus pyogenes: Basic Biology to Clinical Manifestations [Internet]*. Oklahoma City (OK): University of Oklahoma Health Sciences Center; 2016. [Cited: 2nd September]. 384 P. Available from: http://www.ncbi.nlm.nih.gov/books/NBK333417/.

Fischetti VA, Jarymowycz M, Jones KF, Scott J. R. Streptococcal M protein size mutants occur at high frequency within a single strain. *The Journal of Experimental Medicine.* 1986;164(4):971–980. Available from: http://www.ncbi.nlm.nih.gov/pubmed/3760782

Fischetti VA, Jones KF, Scott JR. Size variation of the M protein in group A streptococci. *The Journal of Experimental Medicine*. 1985;161(6):1384–1401. Available from: http://www.ncbi.nlm.nih.gov/pubmed/2409199

Fischetti VA, Pancholi V, Schneewind O. Conservation of a hexapeptide sequence in the anchor region of surface proteins from gram-positive cocci. *Mol Microbiol.* 1990;4:1603–5.

Fischetti VA. M Protein and Other Surface Proteins on Streptococci. 2016 Feb 10. In: Ferretti JJ, Stevens DL, Fischetti VA, editors. Streptococcus pyogenes: Basic Biology to Clinical Manifestations [Internet]. Oklahoma City (OK): University of Oklahoma Health Sciences Center; 2016-. Available from: http://www.ncbi.nlm.nih.gov/books/NBK333431/.

Ferretti JJ, Stevens DL, Fischetti VA, editors. Streptococcus pyogenes: Basic Biology to Clinical Manifestations [Internet]. Oklahoma City (OK):

University of Oklahoma Health Sciences Center; 2016-. Available from: http://www.ncbi.nlm.nih.gov/books/NBK333431/.

Fox KL, Born MW, Cohen MA. Fulminant infection and toxic shock syndrome caused by *Streptococcus pyogenes*. *The Journal of Emergency Medicine.* 2002;22(4):357–366.

Gerber MA, Baltimore CRS, Eaton CB, et al. Prevention of rheumatic fever and diagnosis and treatment of acute streptococcal pharyngitis. *Circulation*. 2009;119:1541–1551.

Gillespie SH, Hawkey PM, editors. *Principles and Practice of Clinical Bacteriology second Edition [internet]*. England: Royal Free and University College Medical School, London, UK and University of Birmingham, Birmingham, UK; 2006. [cited: 16th September]. Chapter: β-haemolytic streptococci. 12 P. Available from: https://www.researchgate.net/publication/292173603_Principles_and_Practice_of_Clinical_Bacteriology_Second_Edition.

Gowda KL, Melbin JJ, Patil SA, Bindu SR, Sanjay MK, Shivannavar CT, Brahmadathan KN. Prevalence of *emm* types of Group A streptococci recovered from school children and hospital patients in Bangalore City, India. *World J Microbiol Biotechnol*. 2011; 27:319–323.

Herold AH. Group A beta-hemolytic streptococcal toxic shock from a mild pharyngitis. *the Journal of Family Practice.* 1990; 31(5): 549–551.

Hollingshead SK, Fischetti VA, Scott JR. Complete nucleotide sequence of type 6 M protein of the group A streptococcus: repetitive structure and membrane anchor. *The Journal of Biological Chemistry.* 1986;261(4):1677–1686. Available from: http://www.ncbi.nlm.nih.gov/pubmed/3511046.

Hollingshead SK, Fischetti VA, Scott JR. Size variation in group A streptococcal M protein is generated by homologous recombination between intragenic repeats. *Molecular and General Genetics.* 1987;207(2-3):196–203. http://www.ncbi.nlm.nih.gov/pubmed/3039291

Holm SE, Norrby A, Bergholm AM, Norgren M. Aspects of pathogenesis of serious group A streptococcal infections in Sweden, 1988-1989. *The Journal of Infectious Diseases*. 1992;166(1):31–37.

Hu MC, Walls MA, Stroop SD, Reddish MA, et al. Immunogenicity of a 26-valent group a streptococcal vaccine. *Infectious Immunology.* 2002;70: 2171–2177.

Introduction to *emm* typing: M protein gene (*emm*) typing *Streptococcus pyogenes.* Centers for disease prevention and control. Retrieved from: http://www.cdc.gov/streplab/m-proteingene-typing.html. Updated: February

8, 2008.

Johnson DR, Sramek J, Kaplan EL, Bicova R, Havlicek J, Havlickova H, et al. Laboratory diagnosis of group A streptococcal infections. Geneva: *World Health Organization*; 1996.

Kendall FE, Heidelberger M, Dawson MN. *A serologically inactive polysaccharide elaborated by mucoid strains of group A hemolytic streptococcus*. *J Biol Chem*. 1941;118:61–69

Khosravi AD, Ebrahimifard N, Shamsizadeh A, Shoja S. Isolation of *Streptococcus pyogenes* from children with pharyngitis and *emm* type analysis. *Journal of the Chinese Medical Association*. 2016; 276**e**280

Kihiczak GG, Schwartz RA, Kapila R. Necrotizing fasciitis: a deadly infection. *Journal of European Academy Dermatology and Venereology.* 2006; 4: 365-369.

Kiska DL, Thiede B, Caracciolo J, Jordan M, Johnson D, Kaplan EL, et al. Invasive group A streptococcal infections in North Carolina: Epidemiology, clinical features, and genetic and serotype analysis of causative organisms. *The Journal of Infectious Diseases.* 1997;176(4):992–1000.

Koutouzi F, Tsakris A, Chatzichristou P, Koutouzis E, et al. *Streptococcus pyogenes emm* types and clusters during a 7-year period (2007 to 2013) in pharyngeal and nonpharyngeal pediatric isolates. Journal of Clinical Microbiology. 2015; 53:2015–2021. doi:10.1128/JCM.00301-15.

Kreikemeyer B, McIver KS, Podbielski A. Virulence factor regulation and regulatory networks in *Streptococcus pyogenes* and their impact on pathogen-host interactions. *Trends Microbiol.* 2003*;* 11, 224–232.

Lancefield RC. Current knowledge of the type-specific M antigens of group A streptococci. *The Journal of Immunology*. 1962;89(3):307–313. Available from: http://www.ncbi.nlm.nih.gov/pubmed/14461914.

Lancefield RC. Differentiation of group A streptococci with a common R antigen into three serological types, with special reference to the bactericidal test. *Journal of Experimental Medicine*. 1957;106:525-44.

Lancefield RC. Persistence of type-specific antibodies in man following infection with group A streptococci. *The Journal of Experimental Medicine*. 1959;110(2):271–292. Available from: http://www.ncbi.nlm.nih.gov/pubmed/13673139.

Lancefield RC. The antigenic complex of Streptococcus hemolyticus. I

Demonstration of a type-specific substance in extracts of Streptococcus hemolyticus. *The Journal of Experimental Medicine*. 1928;47(1):91–103. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19869404.

Mahe A, Hay RJ. Epidemiology and management of common skin diseases in children in developing countries. Geneva:World Health Organisation; 2005.

Marijon E, Mirabel M, Celermajer DS, Jouven X. Rheumatic heart disease. *Lancet*. 2012;379:953–964.

Martin JM, Green M, Barbadora KA, Wald ER. Group A streptococci among school-aged children: clinical characteristics and the carrier state. *Pediatrics*. 2004;114(5):1212–1219.

McNeil SA, Halperin SA, Langley JM, et al. Safety and immunogenicity of 26-valent group A streptococcus vaccine in healthy adult volunteers. *Clinical Infectious Diseases*. 2005;41:1114–1122.

O'Brien KL, Beall B, Barrett NL, Cieslak PR, et al. Epidemiology of invasive group A streptococcus disease in the United States, 1995-1999. *Clinical Infectious Diseases.* 2002. 35:268-276.

O'Loughlin RE, Roberson A, Cieslak PR, Lynfield R, et al. The epidemiology of invasive group A streptococcal infection and potential vaccine implications: United States, 2000-2004. *Clinical Infectious Diseases*. 2007. 45853-862.

Oehmcke S, Shannon O, Mörgelin M, Herwald H. Streptococcal M proteins and their role as virulence determinants. *Clinica Chimica Acta. 2010; 1172–1180*

Olsen RJ, Musser JM. Molecular pathogenesis of necrotizing fasciitis. *Annual Review of Pathology*. 2010;5:1–31.

Olsen RJ, Musser JM. Molecular pathogenesis of necrotizing fasciitis. *Annu. Rev. Pathol.* 2010; 5:1–31. http://dx.doi.org/10.1146/annurev -pathol-121808-102135.

Olsen RJ, Shelburne SA, Musser JM. Molecular mechanism underlying group A streptococcal pathogenesis. *Cellular Microbiology*. 2009;11(1):1–12.

Parks T, Smeesters PR, Steer AC. Streptococcal skin infection and rheumatic heart disease. *Current Opinions in Infectious Diseases*. 2012; 25:145–153. http://dx.doi.org/10.1097/QCO.0b013e3283511d27.

Patterson MJ. *Streptococcus*. In: Baron S, editor. *Medical Microbiology [internet]*. 4th edition. Galveston (TX): University of Texas Medical Branch at Galveston; 1996. Chapter 13. Available from:

http://www.ncbi.nlm.nih.gov/books/NBK7611/?report=classic.

Perez-Casal J, Caparon MG, Scott JR. Introduction of the *emm6* gene into an *emm*-deleted strain of Streptococcus pyogenes restores its ability to resist phagocytosis. *Research in Microbiology.* 1992;143(6):549–558. Available from: http://www.ncbi.nlm.nih.gov/pubmed/1475516

Schlievert PM, Assimacopoulos AP, Cleary PP. Severe invasive group A streptococcal disease: clinical description and mechanisms of pathogenesis. *The Journal of Laboratory and Clinical Medicine*. 1996;127(1):13–22.

Sellers BJ, Woods ML, Morris SE, Saffle JR. Necrotizing group A streptococcal infections associated with streptococcal toxic shock syndrome. *American Journal of Surgery.* 1996;172(5):523–527.

Shaikh N, Leonard E, Martin JM. Prevalence of streptococcal pharyngitis and streptococcal carriage in children: a meta-analysis. *Pediatrics.* 2010;126(3):e557–e564.

Sharkawy A, Low DE, Saginur R, Gregson D, Schwartz B, Jessamine P, et al. Severe group a streptococcal soft-tissue infections in Ontario: 1992-1996. *Clinical Infectious Diseases*. 2002;34(4):454–460.

Steer AC, Batzloff MR, Mulholland K, et al. Group A streptococcal vaccines: facts versus fantasy. *Current Opinions in Infectious Diseases*. 2009;22:544–552.

Steer AC, Lamagni T, Curtis N, Carapetis JR. Invasive group a streptococcal disease: epidemiology, pathogenesis and management. *Drugs.* 2012;72(9):1213–1227.

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*. 2009;9(10): 611–616.

Stevens DL, Bryant AE. Severe Group A Streptococcal Infections. 2016 Feb 10. In: Ferretti JJ, Stevens DL, Fischetti VA, editors. Streptococcus pyogenes: Basic Biology to Clinical Manifestations [Internet]. Oklahoma City (OK): University of Oklahoma Health Sciences Center; 2016-. Available from: https://www.ncbi.nlm.nih.gov/books/NBK333425/

Stevens DL, Tanner MH, Winship J, Swarts R, Reis KM, Schlievert PM, et al. Reappearance of scarlet fever toxin A among streptococci in the Rocky Mountain West: Severe group A streptococcal infections associated with a toxic shock-like syndrome. *The New England Journal of Medicine.* 1989;321(1):1–7

Stollerman GH. Rheumatic fever in the 21st century. *Clinical Infectious Diseases.*

2001; 33: 806-814.

Su Y-F, Wang S-M, Lin Y-L, et al. Changing Epidemiology of *Streptococcus pyogenes emm* Types and Associated Invasive and Noninvasive Infections in Southern Taiwan . *Journal of Clinical Microbiology*. 2009;47(8):2658-2661. doi:10.1128/JCM.01078-09.

Swanson J, Hsu KC, Gotschlich EC. Electron microscopic studies on streptococci. I. M antigen. *The Journal of Experimental Medicine.* 1969;130(5):1063–1091. Available from: http://www.ncbi.nlm.nih.gov/pubmed/5347694

Tanz RR, Shulman ST, Kabat W, et al. Five-year group A streptococcal pharyngitis serotype surveillance in North America, 2000-2005. *Proceedings of the XVIth Lancefield International Symposium on Streptococci and streptococcal Diseases, Palm Cove, Australia; Sept 25-29, 2005: 3-33.*

Tartof SY, Reis JN, Andrade AN, Ramos RT, Reis MG, Riley LW. Factors associated with Group A Streptococcus *emm* type diversification in a large urban setting in Brazil: a cross-sectional study. *BMC Infectious Diseases*. 2010;10:327. doi:10.1186/1471-2334-10-327.

Wessels MR. Pharyngitis and Scarlet Fever. 2016 Feb 10 [Updated 2016 Mar 25]. In: Ferretti JJ, Stevens DL, Fischetti VA, editors. Streptococcus pyogenes: Basic Biology to Clinical Manifestations [Internet]. Oklahoma City (OK): University of Oklahoma Health Sciences Center; 2016-. Available from: https://www.ncbi.nlm.nih.gov/books/NBK333418/.

Whatmore AM, Kapur V, Sullivan DJ, et al. Non-congruent relationships between variation in *emm* gene sequences and the population genetic structure of group A streptococci. *Mol Microbiol* 1994;14:619-631

Appendix

Protocol for *emm* typing

A. Lysate preparation

(Note that one can often simply boil very fresh growth with good results for DNA extracts).

1. A fair amount of fresh growth was picked up with a loop (perhaps half of a standard loop-full). Resuspended in 300 ul 0.85% NaCl.
2. Heated at 70C for 15 minutes.
3. Samples were spanned down full speed for 2 min in microfuge and pipeted out the supernatant.
4. The pellet were resuspended in 50 ul TE (10mM Tris, 1mM EDTA, pH8), 10 ul mutanolysin (3,000 units/ml), and 2 ul hyaluronidase (30 mg/ml, Sigma H-3506; 300-750 units/mg).).
5. Then incubated in 37C, for 30 min.
6. Heated at 100C for 10 min.
7. Immediately were proceeded with PCR or stored lysates at -20C until use.

Occasionally, isolates can be hard to amplify. In these cases the lysate preparation should be altered as follows:

1. A fresh plate were prepared, follow the procedure above.
2. Heated for 15 min at 70C.
3. Samples were centrifuged at low speed (5000rpm) for 2 min; then carefully discarded the supernatant, 500ul of TE were added, vortexed, centrifuged again at 5000rpm for 2 minutes, then supernatant were discarded and again 500ul of TE buffer were added, vortexed, and spanned at full speed 14000rpm for 2 minutes.

4. 50ul of TE buffer were added and 10ul of mutanolysin (and if colonies look very mucoid, 6 uL of hyaluronidase). With usual lysate preparation protocol were proceeded after this step.

B. PCR

Master mix is prepared with this ratio of components:

- 10 uL 10X buffer containing 15mM MgCl2 (commercial, eg. Applied Biosystems B07332)
- 2.0 ul of dNTP mixture (10mM)
- 2.0 ul of each primer 1 and primer 2 (70 picomole/ul)
- primer 1: tatt (c/g) gcttagaaaattaa
- primer 2: gcaagttcttcagcttgttt
- 0.5 ul Taq (3U/ul)
- 82 ul dH20

(Note: Primers 1 and 2 are more dependable with a wider variety of strains than **other primers that were tested).**

20 ul PCR reactions are prapared:

1. Lysate were spanned down full speed for 1 min.
2. For 1 sample on ice aliquot 19.5 ul master mix.
3. No more than 0.5 ul lysate supernatant was added.
4. Using the following program the lysate were placed in cycler. After the initial sample temperature (94C) is reached, the samples were placed in thermocycler at 94C for 1 min.

Then the following 10 X was done:

- 94C: 15s.
- 46.5C: 30s.
- 72C: 1 min 15s.

Followed by the following 20X:

- 94C: 15s.
- 46.5C: 30s.
- 72C: 1 min 15s with a 10 sec increment for each of the subsequent 19 cycles.
- 72C 10 min., then stored at 4C.

PCR products were stored at -20C until use.

C. Sequencing

1. Sequencing template were prepared from 2-11 ul aliquot of PCR prep to be sequenced with ExoSAP-IT as described by USB corporation.
2. BigDye v.1.1 sequencing mixture were diluted, 1 part BigDye to 5 parts of provided buffer.
3. PCR product and sequencing primer *emm* seq2 (tattcgcttagaaaattaaaaacagg) were used exactly as described in the Applied Biosystems BigDye protocol. For sequencing (we used Perkin Elmer Gene Amp PCR 2400 or 2700) we used the following cycling parameters:

    - 25X
    - 96C - 10 sec
    - 55C - 5 sec
    - 60C - 4 min
    - 4C - storage

Sequencing reactions were stored at -20C. Sequencing reactions were purified on Quiagen columns, dryed, and dissolved in formamide-EDTA exactly as described by manufacturer (in Big Dye sequencing handbook) and stored at -20C. Alternatively, we have recently started purifying sequencing reactions with good results using Agencourt's paramagnetic bead technology as described in CleanSEQ Reaction Clean-Up.

***emm* type/previous designation**	**GenBank accession number**	**[1, 2]Countries where isolated** (Partial Listing)	**Closest N-terminal M protein sequence match[3] (% identity)**
emm103/st2034	U74320	**PNG**, Bra, Egy, Mal, Nep, NZ, US	M87 (66%)
emm104/st2035	AF056300	**PNG**, Egy, Mal, Nep, NZ, US	M66 (72%)
emm105/st4529	AF060227	**Mal**, Nep, NZ, US	M5 (45%)
emm106/st4532	AF077666	**Mal**,Egy, Iran, Nep	M27G (71%)
emm107/st4264	AF163686	**Mal**, NZ	M25 (52%)
emm108/st4547	AF052426	**Mal**, Bra, Egy, Ira, NZ	M70 (84%)
emm109/st3018	AF077667	**Mal,**Egy, NZ	M28(74%)
emm110/st4935	U92492	**Ind**, Bul, NZ, Rus, US	M13 (60%)
emm111/st4973	AF128960	**Ind**, Bra, Nep, US	M80 (40%)
emm112/stCmuk16	AF091806	**Thi**, Bra, Rus, US	M27L/77 (59%)

emm113/st2267	AF078068	**NZ**, Thai, Chi	M13 (50%)
emm114/st2967	U50338	**US**, Can, Gam, NZ, PNG	M73 (80%)
emm115/st2980	AF028712	**US**, Bra, Rus	M36 (64%)
emm116/st2370	AF156180	**US**, Nep, NZ	M52 (60%)
emm117/st436	AF058801	**US**	M13 (59%)
emm118/st448	AF058802	**US**, Bra, Egy, Nep, NZ	M49 (79%)
emm119/st3365	AF083874	**US,** Br, Nep	M52 (59%)
emm120/st1135	AF296181	**Egy**	M56 (78%)
emm121/st1161	AF296182	**Egy**	M64 (64%)
emm122/st1432	AF222860	**Egy**, Rus, Nep	M18 (40%)
emm123/st6949	AF213451	**Arg**, NZ	M80 (68%)
emm124/st1160	AF149048 and AF018178	**Egy**, Mal, NZ	M2 (82%

PNG: Papua New Guinea, **Egy:** Egypt, **Mal:** Malaysia, **Nep:** Nepal, **NZ:** New Zealand, **US:** United States, **Bra:** Brazil, **Ira:** Iran, **Rus:** Russian, **Thi:** Thailand, **Chi:** China, **Arg:** Argentina.