

Research article

Open Access

M protein typing of Thai group A streptococcal isolates by PCR-Restriction fragment length polymorphism analysis

Nonglak Yoonim*¹, Colleen Olive³, Chulabhorn Pruksachatkunakorn², Michael F Good³ and Sumalee Pruksakorn¹

Address: ¹Department of Microbiology, Faculty of Medicine, Chiang Mai University, Chiang Mai, Thailand, ²Department of Pediatrics, Faculty of Medicine, Chiang Mai University, Chiang Mai, Thailand and ³Queensland Institute of Medical Research, 300 Herston Road, Herston, Brisbane, QLD, Australia

Email: Nonglak Yoonim* - amm_aon@yahoo.com; Colleen Olive - Colleen.Olive@qimr.edu.au; Chulabhorn Pruksachatkunakorn - cpruksac@mail.med.cmu.ac.th; Michael F Good - michaelG@qimr.edu.au; Sumalee Pruksakorn - spruksak@mail.med.cmu.ac.th

* Corresponding author

Published: 16 October 2005

Received: 08 June 2005

BMC Microbiology 2005, 5:63 doi:10.1186/1471-2180-5-63

Accepted: 16 October 2005

This article is available from: <http://www.biomedcentral.com/1471-2180/5/63>

© 2005 Yoonim et al; licensee BioMed Central Ltd.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

Background: Group A streptococcal (GAS) infections can lead to the development of severe post-infectious sequelae, such as rheumatic fever (RF) and rheumatic heart disease (RHD). RF and RHD are a major health concern in developing countries, and in indigenous populations of developed nations. The majority of GAS isolates are M protein-nontypeable (MNT) by standard serotyping. However, GAS typing is a necessary tool in the epidemiologically analysis of GAS and provides useful information for vaccine development. Although DNA sequencing is the most conclusive method for M protein typing, this is not a feasible approach especially in developing countries. To overcome this problem, we have developed a polymerase chain reaction-restriction fragment length polymorphism (PCR-RFLP)-based assay for molecular typing the M protein gene (*emm*) of GAS.

Results: Using one pair of primers, 13 known GAS M types showed one to four bands of PCR products and after digestion with *Alu* I, they gave different RFLP patterns. Of 106 GAS isolates examined from the normal Thai population and from patients with GAS-associated complications including RHD, 95 isolates gave RFLP patterns that corresponded to the 13 known M types. Only 11 isolates gave RFLP patterns that differed from the 13 known M types. These were then analyzed by DNA sequencing and six additional M types were identified. In addition, we found that M93 GAS was the most common M type in the population studied, and is consistent with a previous study of Thai GAS isolates.

Conclusion: PCR-RFLP analysis has the potential for the rapid screening of different GAS M types and is therefore considerably advantageous as an alternative M typing approach in developing countries in which GAS is endemic.

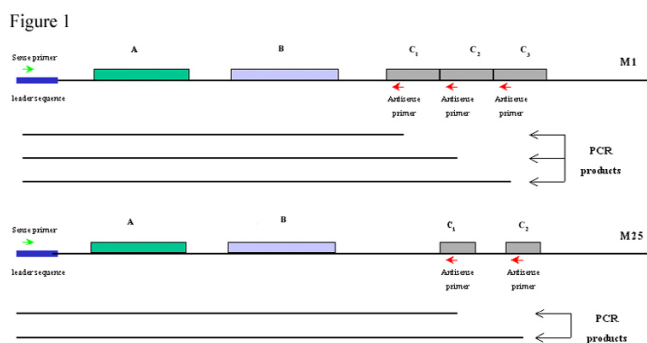


Figure 1
Diagrammatic representation of the *emmI* and *emm25* genes of M1 and M25 proteins, respectively, showing the different A-, B- and C-repeat regions encoding the M protein. The location of the sense and antisense primers used for PCR amplification of the *emm* gene are indicated as well as the different *emm* gene amplicons (PCR products) that arise depending upon the number of C-repeat regions within the *emm* gene.

Background

Streptococcus pyogenes or group A streptococcus (GAS) causes a number of clinical manifestations and diseases including sore throat, pyoderma, necrotizing fasciitis, toxic-shock syndrome, and the post-infectious sequelae – rheumatic fever (RF) and rheumatic heart disease (RHD) [1]. RF and RHD are a major health concern worldwide but especially in indigenous communities within developed countries, and in populations of developing countries [2]. The GAS surface M protein is known to prevent opsonophagocytosis and is a major virulence factor in GAS infection [1]. The N-terminal region of the M protein is highly variable between different GAS strains and contains a type-specific moiety, the antigenic variation of which forms the basis for the classical M protein serological typing of GAS. However, there are disadvantages with this method of typing including ambiguities in the results, the emergence of new M types, and the high rates of M protein-nontypeable (MNT) strains [3] largely due to the unavailability of specific typing antisera. As a consequence, there has been a surge of interest in the development of alternative methods for M typing utilizing molecular technologies.

Several methods have been developed for GAS typing such as enzyme electrophoretic polymorphism [4-6], genomic typing methods such as RAPD [7], 16s rDNA typing [8], RFLP analysis [9-12], vir typing [13,14], DNA hybridization using N-terminal sequences of the M protein gene (*emm*) as oligonucleotide probes [15,16], polymerase chain reaction (PCR)-enzyme linked immu-

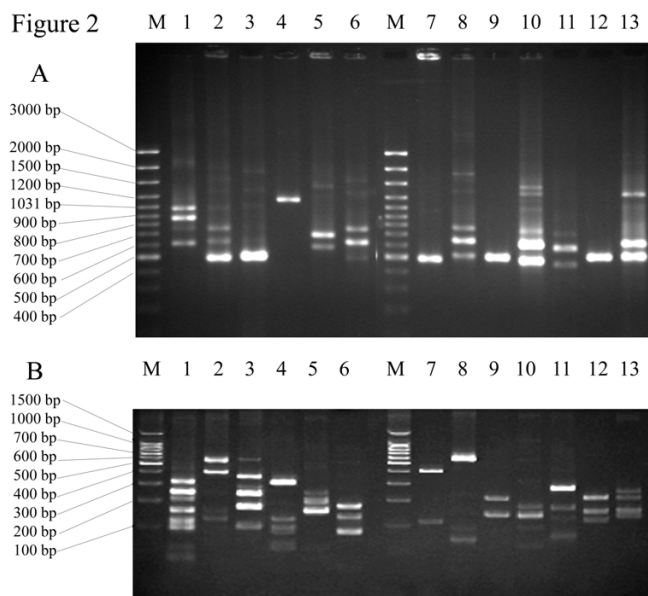


Figure 2
Agarose gel electrophoresis of *emm* gene amplicon PCR products (A), and polyacrylamide gel electrophoresis RFLP patterns using *Alu I* restriction enzyme (B). The size differences of PCR products and RFLP patterns of M1, M11, M49, M12, M25, M48, M112, st11014, M44/61, M93, M81, M104 and M66 are shown in lanes 1-13, respectively. Lane M represents a 100 bp ladder.

nosorbant assay [17], and PCR M typing using type-specific oligonucleotide primers for PCR amplification of the N-terminal region of the *emm* gene [18]. PCR-RFLP analysis which utilizes PCR to amplify the *emm* gene amplicons encoding the M protein prior to digestion with restriction endonucleases has been used for specific molecular M typing methods [19-22], as well as multilocus sequence typing (MLST) [23]. N-terminal sequencing of the M protein gene, however, is the most conclusive method for typing of GAS [24,25], This method of typing is not an option in most laboratories in developing countries worldwide, due to limited resources, and therefore an alternative approach is required for the identification of GAS types. The purpose of this study was therefore to analyze the N-terminal regions of the *emm* gene of Thai GAS isolates using PCR-RFLP analysis. PCR products were digested with an appropriate restriction enzyme and the fragments analyzed by polyacrylamide gel electrophoresis (PAGE).

Results and discussion

The N- and C-terminal region of the *emm* gene was amplified by PCR from all of the 119 GAS isolates. The amplicons consisted of one to four bands which varied from approximately 450 to 1200 bp depending on the M types

Table 1: PCR products and *Alu* I restriction fragment sizes and distribution in the Thai GAS isolates studied.

M type	Size of PCR product (bp)	Size of restriction fragment (bp)	No. (%) of isolates from PCR-RFLP analysis	M type	Size of PCR product (bp)	Size of restriction fragment (bp)	No. (%) of isolates from PCR-RFLP analysis		
M1	1071	354	16 (15.09%)	M49	450	336	2 (1.89%)		
	892	257				252			
	744	218				196			
	592	147				135			
		110				89			
		87							
		47				Emm58.5		370	1 (0.94%)
		37						185	
		29						206	1 (0.94%)
	M4	747				314		2 (1.89%)	M66
609		174	123						
476		57	123						
M11	710	593	8 (7.55%)	M81	649	322	7 (6.60%)		
	592	434				185			
	450	147				127			
M12	1121	117	3 (2.83%)	M89	404	350	1 (0.94%)		
		349				77			
		116							
		93				M93		254	22 (20.75%)
		78						227	
M25	649	51	10 (9.43%)	M102	420	168	1 (0.94%)		
		247				130			
		203				93			
		148				53			
		51							
M33	812	37	3 (2.83%)	M104	450	332	3 (2.83%)		
		347				92			
		168				210			
M44/61	450	94	11 (10.38%)	M109	509	146	3 (2.83%)		
		74				118			
		212				386			
M48	710	133	2 (1.89%)	M112	450	168	5 (4.72%)		
		162				435			
		121				104			
M49	450	81	5 (4.72%)	st11014	744	596	5 (4.72%)		
						620			
						493		87	
						29			

and GAS strains (Fig. 2A). However, some M types had similar sizes of their *emm* amplicons, but gave distinct bands after digestion with the restriction enzyme, *Alu* I. This revealed RFLP patterns consisting of two to nine distinct fragments, which ranged from approximately 25 to 700 bp depending on the variability of DNA sequences in the *emm* amplicons of different M types (Fig. 2B). The RFLP patterns of 95 unknown GAS M types corresponded to 13 known M types (Table 1, Fig. 2B). The M types of all isolates represented by the PCR-RFLP patterns were confirmed by DNA sequencing analysis. All isolates were sequenced independently of the PCR-RFLP analysis which was carried out blinded. The two sets of data were then compared. Only 11 isolates gave patterns that did not cor-

responded to the 13 known M types. The isolates that could not be identified by PCR-RFLP analysis were also analyzed by DNA sequencing. Six M types that differed from the 13 known M types were obtained from the DNA sequencing analysis (M4, M33, *emm*58.5, M102, M109 and M89). Although these M protein DNA sequences were $\geq 95\%$ similar to the data in GenBank, we found both point mutations and deletions in these isolates. M89 has a single-base substitution (T→C) that resulted in a new *Alu* I restriction site whereas M4 and M102 have 21-base and 33-base deletions, respectively. Alignment of the amino acid sequences of M4 and M102 with reference strains in GenBank, demonstrated $\geq 95\%$ similarity (Fig. 3).

Figure 3

```

(a)
Emm4      AVAVAVL GAGF ANQTEVKA AEIKKPOADS AUNWPKE YNALLKENEELKVEREKYL
U768      AVAVAVL GAGF ANQTEVKA AEIKKPOADS AUNWPKE YNALLKENEELKVEREKYL
*****
Emm4      SYADDKEKDPQYRALMGENQDLRKRGGQYODKIEELEKERKERKOEQOEQLEROYQIEADK
U768      SYADDKEKDPQYRALMGENQDLRKRGGQYODKIEELEKERKERKOEQOEQLEROYQIEADK
*****
Emm4      HYOEQOKKHQOEQQLEAEKOKLAKDKQISDASRQGLSRDLEASRAAKKVEADL AALTA
U768      HYOEQOKKHQOEQQLEAEKOKLAKDKQISDASRQGLSRDLEASRAAKK-----ELEA
*****
Emm4      EHQLKEDKQISEA
U768      EHQLKEDKQISEA
*****;****;

(b)
Emm102    SASVAVAVAVL GAGF ANQTTVKADNPSSVPVKKAAELVDKIKELEEGREELNLDLDKVK
P354      GTASVAVAAVL GAGF ANQTTVKADNPSSVPVKKAAELVDKIKELEEGREELN-----
;****;
emm102    EEHKDLKVKKEHKHKEKLEKKAEDVERHYLRQLDQDHKEQOERQKNLEERQSQRE
P354      -----DLKVKKEHKHKEKLEKKAEDVERHYLRQLDQDHKEQOERQKNLEERQSQRE
*****
emm102    VEKRYOEQLQKQOLETEKQISEA
P354      VEKRYOEQLQKQOLETEKQISEA
*****
    
```

Figure 3
 Alignment of amino acid sequences from U768 with M4 (*emm4* gene) and P354 with M102 (*emm102* gene) in GenBank.

For decades serotyping has been the method of choice for GAS M typing. However, serotyping is time consuming and it is often difficult to produce high-titer M type-specific antisera, and therefore this technique is limited to a few laboratories in the world. To overcome these limitations, genomic typing methods have been developed [7-12], as well as specific M typing methods [15-18]. RFLP analysis of the *emm* gene revealed different RFLP patterns among GAS isolates with the same M type [19-22]. Therefore, it could potentially be used to differentiate among isolates with the same M type, but which are not clonal, and may originate from different geographical locations or populations.

In this study, we applied PCR-RFLP analysis for GAS M typing. PCR products derived from the amplification of the N- and C-terminal region of the *emm* gene and digested with the restriction enzyme, *Alu I*, produced different RFLP patterns among different M types. Ninety five clinical isolates could be compared with known M types by their RFLP patterns. Only 11 isolates had novel RFLP patterns and required sequencing to confirm the M types. Our results showed that M93 GAS was the most common GAS strain isolated from the population studied with 22 isolates (20.75%) having this type, and is consistent with a previous study of Thai GAS isolates [3]. Other M types that represented more than 10% of the Thai GAS isolates were M1 and M44/61, whereas M66, *emm58.5*, M89 and M102 were quite rare (0.94%). Furthermore, the specificity of the PCR-RFLP analysis was confirmed by sequencing of the N-terminal region of the *emm* gene. The majority of isolates were sequenced for each M type. M4, M33, M48, *emm58.5*, M66, M81, M89, M102, M109, M112 and st11014 were sequenced in all isolates. M1, M11, M12, M25, M44, M49, M93 and M104 were

sequenced in 8 of 16, 7 of 8, 2 of 3, 3 of 10, 3 of 11, 1 of 2, 11 of 22 and 2 of 3 isolates, respectively. All isolates were sequenced independently of the PCR-RFLP analysis which was carried out blinded. The two sets of data were then compared. The results obtained from the PCR-RFLP analysis were in agreement with the sequencing data.

In comparison with other genotyping methods based on RFLP analysis [8-10,19-22], the PCR-RFLP analysis used here has some technical advantages. Only one pair of PCR primers is required for the amplification of all of the GAS isolates, and the digested PCR products can be discriminated using standard PAGE which is easy to perform, interpret and is less time consuming. In addition, compared with sequencing analysis [24,25], the PCR-RFLP analysis is technically less demanding and more economical. Therefore, with this simple and rapid protocol, GAS typing can be used in any laboratory in which PCR is routinely used, and is particularly useful for typing large numbers of GAS isolates. However, our findings are based on a relatively small number of *emm* types isolated in a relatively limited geographic region over a relatively short period of time. Greater geographic and temporal diversity may result in greater clonal diversity thus complicating interpretation of RFLP patterns. Therefore, we suggest that each laboratory should make their own reference RFLP patterns from M types that circulate in their region. However, there are several problems that may occur such as point mutation and deletion. These result in the strains that can not be identified. In this case, it should then be typed by sequence analysis and added these variants into their reference RFLP patterns. In addition, to confirm validation of their results, sequencing a subset of their analyzed isolated should be done periodically.

Conclusion

PCR-RFLP analysis is a rapid, economical and practical genotyping method for determining the M type of GAS. It is therefore particularly suited as a method of choice in developing countries in which GAS is endemic and the majority of GAS isolates are M protein-nontypeable by conventional serological typing.

Methods

Bacteria

One hundred and six strains of GAS isolated from the normal Thai population, patients with sore throat, RHD or impetigo from 1985, 1990, 1995, 2000, 2003 and 2004, and 13 known GAS M serotypes, were included in the study.

DNA isolation and PCR

DNA was isolated from GAS based on the method previously described [3]. The sense primer (CAGTATTCGCT-TAGAAAATTAATA) was derived from the conserved

leader sequence of *emm* gene [26]. The antisense primer (CCCTTACGGCTTGC TTCTGA) was derived from the C-repeat region of the *emm* gene [3]. The PCR conditions were as follows: denaturation at 94°C for 30 s, annealing at 45°C for 30 s, and extension at 72°C for 2 min for 35 cycles. These primers were also used for sequencing analysis. Using these primers, up to four *emm* gene amplicons can be generated for a particular GAS strain following PCR amplification depending on the number of C-repeat regions in the M protein (Fig. 1).

PCR-RFLP analysis

PCR-RFLP analysis can be used to type different *emm* genes. PCR amplification and digestion with appropriate restriction enzymes is expected to give different RFLP patterns among GAS. Based on the *emm* gene sequences in the GenBank database and analysis using the restriction mapper program [27], we chose the restriction enzyme, *Alu* I, because almost all sequences contained at least one *Alu* I site in different positions in their sequences. Therefore, digestion with *Alu* I is likely to give different RFLP patterns among GAS M types. Following PCR, the products (*emm* amplicons) were analyzed by electrophoresis on a 1% agarose gel in TBE buffer. The sizes of PCR products were compared with the 100 bp standard marker and recorded following staining with ethidium bromide. PCR products were then partially purified by ethanol precipitation prior to quantitation and digestion with the *Alu* I restriction enzyme according to the manufacturer's specifications. Digested PCR products were separated on a 15% polyacrylamide gel in TBE buffer and compared with the 100 bp standard marker following staining with ethidium bromide. A log-linear standard curve was initially generated in the Excel program using the known marker sizes versus gel running distance. The equation from the standard curve was then used to estimate the sizes of experimental bands.

DNA sequencing analysis

PCR products were sequenced using the ABI Dye Terminator Cycle Sequencing Ready Reaction Kit according to the manufacturer's instructions (The Perkin-Elmer Corporation) and analyzed using an ABI 310 automated sequencer (The Perkin-Elmer Corporation). DNA sequences were transferred to the DNASIS program for sequence comparison between isolates. The BLAST 2 program (NCBI) was used to determine the levels of DNA homology with published sequences in the GenBank database.

Authors' contributions

NY carried out the microbiologic experiments, performed the molecular genetic analysis, participated in the sequence alignment, interpretation of data and drafted the manuscript. CO contributed to the editing of the man-

uscript. CP provided clinical specimens and clinical support. MFG provided PCR primers and molecular reagents. SP conceived of the study, and participated in its design and coordination. All authors read and approved the final manuscript.

Acknowledgements

This study was supported by the Royal Golden Jubilee Ph.D. Program, the Thailand Research Fund, grant No. PHD/00100/2541.

References

- Fischetti VA: **Streptococcal M protein: molecular design and biological behavior.** *Clin Microbiol Rev* 1989, **2**:285-314.
- Bisno AL: **Group A streptococcal infections and acute rheumatic fever.** *N Engl J Med* 1991, **325**:783-793.
- Pruksakorn S, Sittisombut N, Phornphutkul C, Pruksachatkunakorn C, Good MF, Brandt E: **Epidemiological analysis of non-M-typeable group A Streptococcus isolates from a Thai population in northern Thailand.** *J Clin Microbiol* 2000, **38**:1250-1254.
- Musser JM, Gray BM, Schlievert PM, Pichichero ME: **Streptococcus pyogenes pharyngitis: characterization of strains by multilocus enzyme genotype, M and T protein serotype, and pyrogenic exotoxin gene probing.** *J Clin Microbiol* 1992, **30**(3):600-603.
- Haase AM, Melder A, Mathews JD, Kemp DJ, Adams M: **Clonal diversity of Streptococcus pyogenes within some M-types revealed by multilocus enzyme electrophoresis.** *Epidemiol Infect* 1994, **113**(3):455-62.
- Bert F, Picard B, Lambert-Zechovsky N, Goulet P: **Identification and typing of pyogenic streptococci by enzyme electrophoretic polymorphism.** *J Med Microbiol* 1995, **42**(6):442-51.
- Seppala H, Qiushui H, Osterblad M, Huovinen P: **Typing of group A streptococci by random amplified polymorphic DNA analysis.** *J Clin Microbiol* 1994, **32**(8):1945-1948.
- Stanley J, Linton D, Desai M, Efstratiou A, George R: **Molecular subtyping of prevalent M serotypes of Streptococcus pyogenes causing invasive disease.** *J Clin Microbiol* 1995, **33**(11):2850-2855.
- Patric Cleary P, Kaplan EL, Livdahl C, Skjold S: **DNA fingerprints of Streptococcus pyogenes are M type specific.** *J Infect Dis* 1988, **158**(6):1317-1323.
- Bingen E, Denamur E, Lambert-Zechovsky N, Boissinot C, Brahimi N, Aujard Y, Blot P, Elion J: **Mother-to-infant vertical transmission and cross-colonization of Streptococcus pyogenes confirmed by DNA restriction fragment length polymorphism analysis.** *J Infect Dis* 1992, **165**:147-50.
- Desai M, Tanna A, Wall R, Efstratiou A, George R, Stanley J: **Fluorescent amplified-fragment length polymorphism analysis of an outbreak of group A streptococcal invasive disease.** *J Clin Microbiol* 1998, **36**(11):3133-3137.
- Desai M, Efstratiou A, George R, Stanley J: **High-resolution genotyping of Streptococcus pyogenes serotype M1 isolates by fluorescent amplified-fragment length polymorphism analysis.** *J Clin Microbiol* 1999, **37**(6):1948-1952.
- Gardiner D, Hartas J, Currie B, Mathews JD, Kemp DJ, Sriprakash KS: **Vir typing: a long-PCR typing method for group A streptococci.** *PCR Methods Appl* 1995, **4**:288-293.
- Gardiner DL, Sriprakash KS: **Molecular epidemiology of impetiginous group A streptococcal infections in aboriginal communities of northern Australia.** *J Clin Microbiol* 1996, **34**:1448-1452.
- Kauffhold A, Podbielski A, Baumgarten G, Blokpoel M, Top J, Schouls L: **Rapid typing of group A streptococci by the use of DNA amplification and non-radioactive allele-specific oligonucleotide probes.** *FEMS Microbiol Lett* 1994, **119**:19-25.
- Penney TJ, Martin DR, Williams LC, de Malmanche SA, Bergquist PL: **A single emm gene-specific oligonucleotide probe does not recognise all members of the Streptococcus pyogenes M type I.** *FEMS Microbiol Lett* 1995, **130**(2-3):145-9.
- Saunders NA, Hallas G, Gaworzewska ET, Metherell L, Efstratiou A, Hookey JV, George RC: **PCR-enzyme-linked immunosorbent assay and sequencing as an alternative to serology for M-antigen typing of Streptococcus pyogenes.** *J Clin Microbiol* 1997, **35**:2689-2691.

18. Vitali LA, Zampaloni C, Prenna M, Ripa S: **PCR m typing: a new method for rapid typing of group a streptococci.** *J Clin Microbiol* 2002, **40**:679-681.
19. Stanley J, Desai M, Xerry J, Tanna A, Efstratiou A, George R: **High-resolution genotyping elucidates the epidemiology of group A streptococcus outbreaks.** *J Infect Dis* 1996, **174**:500-506.
20. Dicuonzo G, Gherardi G, Lorino G, Angeletti S, De Cesaris M, Fiscarelli E, Bessen DE, Beall B: **Group A streptococcal genotypes from pediatric throat isolates in Rome, Italy.** *J Clin Microbiol* 2001, **39**:1687-1690.
21. Perea-Mejia LM, Inzunza-Montiel AE, Cravioto A: **Molecular characterization of group A Streptococcus strains isolated during a scarlet fever outbreak.** *J Clin Microbiol* 2002, **40**:278-280.
22. Beall B, Facklam RR, Elliott JA, Franklin AR, Hoenes T, Jackson D, Laclaire L, Thomson T, Viswanathan R: **Streptococcal emm types associated with T-agglutination types and the use of conserved emm gene restriction fragment patterns for subtyping group A streptococci.** *J Clin Microbiol* 2002, **40**:278-280.
23. Enright MC, Spratt BG, Kalia A, Cross JH, Bessen D: **Multilocus sequence typing of Streptococcus pyogenes and the relationships between emm type and clone.** *Infection and Immunity* 2001, **69**(4):2416-2427.
24. Beall B, Facklam R, Thompson T: **Sequencing emm-specific PCR products for routine and accurate typing of group A streptococci.** *J Clin Microbiol* 1996, **34**:953-958.
25. Brandt CM, Spellerberg B, Honscha M, Truong ND, Hoeverner B, Luticken R: **Typing of Streptococcus pyogenes strains isolated from throat infections in the region of Aachen, Germany.** *Infection* 2001, **29**:163-165.
26. Manjula BN, Khandke KM, Fairwell T, Relf WA, Sriprakash KS: **Hep-tad motifs within the distal subdomain of the coiled-coil rod region of M protein from rheumatic fever and nephritis associated serotypes of group A streptococci are distinct from each other: nucleotide sequence of the M57 gene and relation of the deduced amino acid sequence to other M proteins.** *J Protein Chem* 1991, **10**:369-384.
27. **RestrictionMapper** [<http://www.restrictionmapper.org>]

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:
http://www.biomedcentral.com/info/publishing_adv.asp

