

Multilocus genotyping of *Streptococcus pyogenes* isolated from different clinical sources in Al-Diwaniyah city

Huda Refaat Ibraheem¹, Ghaidaa Jihadi Mohammed¹

Department of Biology, College of Science, University of Al-Qadisiyah, Al-Diwaniayh,

Iraq.

Abstract

Background: *Streptococcus pyogenes*, often known as group A beta-hemolytic Streptococcus (GAS), is one of the most bacterial pathogens that widely distributed and mostly infects the superficial tissues such as upper respiratory tract and skin, leading to a variety of invasive diseases.

Aim :Multilicus genotyping of Streptococcus pyogenes isolated from clinical sources.

Method :The study included the collection of (125) samples from different clinical cases in Al-Diwaniyah Teaching hospital and burns hospital in Al-Diwaniyah city during the period from 15-8-2022 to 1-12-2023.The bacterial isolates were identified by traditional biochemical tests and confirmed by vitek-2 system.The internal fragments of seven housekeeping genes (*gki, gtr, murI, mutS, recP, xpt,* and *yqiL*) were amplified with primers and sequenced by Multilocus sequence typing.

Result :A 20 isolates were diagnosed as *S.pyogenes*; include one isolate from throat swab and 19 isolates from tonsils, all the isolates were having seven housekeeping genes with rate(100%) as detected by PCR.Also, the present resultsidentified two ST that included MLST-STs 49 and 150 were common in study isolates, Single locus variations were seen in isolate 4 .Whereas,other four isolates not show any variation.

Conclusion :The current study showed most *Streptococcuspyogenes* isolated from different clinical samples have a high prevalence of housekeeping gene such as,glucose kinase, glutamine transporter protien, glutamate rasemase, transketolase, xanthine phosphoribosyltransferase and acetyle-CoA acetyltransferase which have a major role in the pathogenicity of bacteria and the effectiveness of MLST scheme for *Streptococcus pyogenes* strains supplies a preferable genotyping tool for comprehension the phylogeny of strains.

Key words:Clinical samples, *Streptococcus pyogenes*, Housekeeping genes, Multilocus sequence typing.

Introduction:

Streptococcus pyogenes (S.pyogenes) is a species of Gram-positive, aerotolerant bacteria in the genus streptococcus. These bacteria are extracellular, and made up of non-motile and nonsporingcocci that tend to link in chains. They are clinically important for humans, as they are infrequent, but usually pathogenic, part of the skin microbiota that can cause Group A streptococcal infection(Kimberlinet al., 2015). In microbiology research, genotyping is an effective method for monitoring bacterial strains (Ochoa-Díazet al., 2018). Many genotypic techniques have been sophisticated for typing of *S.pyogenes*; like ribotyping (Thenmozhi*et al.*, 2010), pulse field gel electrophoresis (Vitaliet al., 2015), random amplified polymorphic DNA (RAPD) analysis (Rahimi-Moghaddam et al., 2019), Vir/emm typing (Frost et al., 2020) and multilocus sequence typing(MLST) (Pérez-Losadaet al., 2018) are being used worldwide for typing of GAS isolates. In general, MLST typing involves the amplification of seven loci of housekeeping genes(glucose kinase (gki), glutamine transporter protein (gtr), glutamate racemase (murI), DNA mismatch repair protein (mutS), transketolase (recP), xanthine phosphoribosyltransferase (xpt), and acetyl-coenzyme A acetyltransferase (yqiL) that code for enzymes necessary for the viability of the cell by PCR, followed by DNA sequencing of the resulting PCR fragments. Specific DNA sequences are then matched to allelic profiles. A single nucleotide variation at any of these loci defines a different allele and informs the sequence type(Riddle et al., 2010). Housekeeping genetic sequences are used for analysis because they are present in every organism and their products serve vital functions in the cell.

Further, mutations within them are largely believed to be selectively neutral. It would also help to identify the nature and magnitude of development of new strains in a bacterial population. MLST have been used for epidemiological typing of a variety of pathogenic microbes (Maiden et al., 2013).

Multilocus sequence typing (MLST) is a comparatively new method for bacterial molecular typing. MLST is a nucleotide sequence-depended method that is well appropriate towards differentiating the genetic connection between the organisms of a bacterial species .The sequence data generated for several neutral housekeeping loci by MLST are unambiguous, electronically portable, and easily queried via the internet, which is a significant advantage over gel-based methods(Rajkumariet al., 2017). So, the aim of this study isto typing Streptococcus pyogenes isolated from different clinical sources by multilocus genes.

Methods

Samples collection and Identification of Bacterial Isolates

A 125 clinical samples collected from Al-Diwaniyah Teaching hospital and burns hospital in Al-Diwaniyah city during the period from 15-8-2022 to 1-12-2023 and from various clinical sources including; 35 throat swab, 55 tonsils, 10 urine and 25 burns swabs. All samples were culture on Azide blood agar and selective streptococcus agar; incubated at 37°C for 24 hours. The isolates were diagnosis based on the standard biochemical methods and then subjected to VITEK 2 system for confirmed detection of S.pyogenes.

Molecular detection of housekeeping genes

DNA Extraction of *S.s pyogenes*:

Bacterial genomic DNA was extracted from *S.pyogenes* isolates that they were activated by culturing them in brain heart infusion broth for twenty-four hours at 37 degrees Celsius. DNA was extracted by using (Presto[™] Mini gDNA Bacteria KitGeneaid, Taiwan).and done according to company instructions. A NanoDrop(Thermo Scientific NanoDrop, USA) was used to measure the concentration of DNA for both quality and quantity.Seven unique polymerase chain reactions were performed to determine the presence of housekeeping genes (gki,gtr,murI,mutS,recP,xpt,yqiL).The PCR primers for MLST housekeeping gene of S.pyogenes were designed according to (Enright et al., 2001). These primers were provided by Scientific Resercher. Co. Ltd in Iraq. Table 1 describe the primer sequences and the size of the amplicons. PCR product was analyzed in a 1.5 percent agarose gel hold 3µl of ethidium bromide stain in TBE buffer, and a UV Transilluminator(Wised,Korea) was used to visualize the PCR products.

Primer	Seque	Product Size	
gki gene	F	GGCATTGGAATGGGATCACC	498hn
	R	TCTCCTGCTGCTGACAC	1900
gtr gene	F	GAGGTTGTGGTGATTATTGG	450bp
	R	GCAAAGCCCATTTCATGAGTC	
murI gene	F	TGCTGACTCAAAATGTTAAAATGATTG	4201
	R	GATGATAATTCACCGTTAATGTCAAAATAG	4380p

Table (1): PCR primers of the Housekeeping gene with their nucleotide sequence and product size

mutS gene	F	GAAGAGTCATCTAGTTTAGAATACGAT	405h.a	
	R	AGAGAGTTGTCACTTGCGCGTTTGATTGCT		
recP gene	F	GCAAATTCTGGACACCCAGG	450hr	
	R CTTTCACAAGGATATGTTGCC		4390p	
<i>xpt</i> gene	F	TTACTTGAAGAACGCATCTTA	4501-	
	R ATGAGGTCACTTCAATGCCC		4300p	
yqiL-gene	F	TGCAACAGTATGGACTGACCAGAGAACAAGA TGC	4201-	
	R CAAGGTCTCGTGAAACCGCTAAAGCCTGAG		4370p	

DNA sequence method:

The DNA sequencing method was performed for study the MLST genotyping based 7 housekeeping genes in five Streptococcus pyogenesisolates. The PCR products were sent to Macrogen Company in Korea in ice bag by DHL for performed the DNA sequencing by Applied Biosystems DNA sequencing system.

The DNA sequencing analysis was conducted by using Molecular Evolutionary Genetics Analysis version 6.0. (Mega 6.0) and Multiple sequence alignment analysis of the partial metE and metF gene based ClustalW alignment analysis and The evolutionary distances were computed using the Maximum Composite Likelihood method by phylogenetic tree UPGMA method.

The homology sequence identity and mutation analysis by using NCBI BLAST analysis.

Finally, the MLST genotyping based 7 housekeeping genes were submitted into NCBI-Genbank data base for get Genbank accession numbers.

Statistical Analysis

Data were collected, summarized, analyzed and presented using statistical package for social sciences (SPSS) version 26 and Microsoft Office Excel 2010.Chi-square test was used to study association between any two categorical variables. The level of significance was considered at P-value of 0.05 (Daniel 2009).

Results:

Identification of Streptococcus pyogenes

Twenty isolates were diagnosed as *S.pyogenes* from (125) clinical specimensthat collected from patients in Hospitals of Al-Diwanyiah province, by traditional biochemical tests and Vitek 2 system; include one isolate from throat swab and 19 isolates from tonsils, as shown in table (2).

Characteristic	Total number	S. pyogenes	Р
Clinical samples			
Throat Swab, n (%)	35 (28 %)	1 (2.8%)	0.001

Tonsils, n (%)	55 (44 %)	19 (34.54%)	¥ HS
Urine, <i>n</i> (%)	10 (8 %)	0	
Burns, <i>n</i> (%)	25 (20 %)	0	
Total	125(100)	20(16%)	

0.05 <*n*: number of cases; ¥: Chi-square test; S: significant at P

Revealing of housekeeping genes for *Streptococcus pyogenes*

The presence of the seven housekeeping genes(*gki*, *gtr*, *murI*, *mutS*, *recP*,*xpt* and *yqiL*)in the 20 *Streptococcus pyogenes* isolated from clinical samples ;was detected in all 20 isolates (100%) ,and the results of gel electrophoresis are shown in fig. 1 to fig. 7.



Figure (1):Agarose gel electrophoresis image that showed the PCR product analysis of *gki* gene in *S.pyrogens* isolates. Where Marker ladder (100-2000bp), Lane (1-20) showed positive isolates for*gki* gene at 498bp PCR product size.



Figure (2):Agarose gel electrophoresis image that showed the PCR product analysis of *gtr* gene in *S.pyrogens* isolates. Where Marker ladder (100-2000bp), Lane (1-20) showed positive isolatesfor *gtr* gene at 450bp PCR product size.



Figure (3):Agarose gel electrophoresis image that showed the PCR product analysis of *murI* gene in *S.pyrogens* isolates. Where Marker ladder (100-2000bp), Lane (1-20) showed positive isolates for *murI* gene at 438bp PCR product size.



Figure (4):Agarose gel electrophoresis image that showed the PCR product analysis of *mutS* gene in *S.pyrogens* isolates. Where Marker ladder (100-2000bp), Lane (1-20) showed positive isolates for *mutS* gene at 405bp PCR product size.



Figure (5):Agarose gel electrophoresis image that showed the PCR product analysis of *recP* gene in *S.pyrogens* isolates. Where Marker ladder (100-2000bp), Lane (1-20) showed positive isolates for *recP* gene at 459bp PCR product size.



Figure (6):Agarose gel electrophoresis image that showed the PCR product analysis of *xpt* gene in *S.pyrogens* isolates. Where Marker ladder (100-2000bp), Lane (1-20) showed positive isolates for *xpt* gene at 450bp PCR product size.



Figure (7):Agarose gel electrophoresis image that showed the PCR product analysis of *yqiL* gene in *S.pyrogens* isolates. Where Marker ladder (100-2000bp), Lane (1-20) showed positive isolates for *yqiL* gene at 439bp PCR product size. **Sequencing analysis of Housekeepinggenes:**

BioGecko



The results of the PCR product of seven housekeeping genesforfive *S. pyogenes* isolatesthat were multidrug resistance selected and subjected to partial sequencing and blasted in National Center for Biotechnology Information (NCBI) against standard strains of *S. pyogenes*. These strains obtained the official accession numbers in the NCBI-Gene Bank for *S. pyogenes* isolates from (1 to 7) as show in figure(8) to figure (14) and table(3), which recorded by NCBI.



Figure (8): Phylogenetic tree analysis based MLST housekeeping *gki gene* partial sequence in local *S.pyogenes* IQ. isolates that used for confirmative genetic identification of MLST typing analysis. The phylogenetic tree was constructed using Neighbor-Joining tree method in (MEGA 6.0 version). The local *S.pyogenes*IQ. isolates IQ-No.1-IQ-No.5 were showed closed related to Pub-MLST data base *S.pyogenes* MLST-ST11 at total genetic changes (0.0020%).



Figure (9): Phylogenetic tree analysis based MLST housekeeping *gtr* gene partial sequence in local *S.pyogenes* IQ. isolates that used for confirmative genetic identification of MLST typing analysis. The phylogenetic tree was constructed using Neighbor-Joining tree method in

Vol 12 Issue 03 2023 ISSN NO: 2230-5807

(MEGA 6.0 version). The local *S.pyogenes* IQ. isolates IQ-No.1-IQ-No.5 were showed closed related to Pub-MLST data base *S.pyogenes* MLST-ST2 at total genetic changes (0.0010%).



Figure (10): Phylogenetic tree analysis based MLST housekeeping *murI* gene partial sequence in local *S.pyogenes* IQ. isolates that used for confirmative genetic identification of MLST typing analysis. The phylogenetic tree was constructed using Neighbor-Joining tree method in (MEGA 6.0 version). The local *S.pyogenes* IQ. isolates IQ-No.1-IQ-No.5 were showed closed related to Pub-MLST data base *S.pyogenes* MLST-ST1 at total genetic changes (0.0010%).



Figure (11): Phylogenetic tree analysis based MLST housekeeping *mutS* gene partial sequence in local *S.pyogenes* IQ. isolates that used for confirmative genetic identification of MLST typing analysis. The phylogenetic tree was constructed using Neighbor-Joining tree method in (MEGA 6.0 version). The local *S.pyogenes* IQ. isolates IQ-No.1-IQ-No.5 were showed closed related to Pub-MLST data base *S.pyogenes* MLST-ST3 at total genetic changes (0.0020%).



Figure (12): Phylogenetic tree analysis based MLST housekeeping *recP* gene partial sequence in local *S.pyogenes* IQ. isolates that used for confirmative genetic identification of MLST typing analysis. The phylogenetic tree was constructed using Neighbor-Joining tree method in (MEGA 6.0 version). The local *S.pyogenes* IQ. isolates IQ-No.1, No.2, No.3 and -IQ-No.5 were showed closed related to Pub-MLST data base *S.pyogenes* MLST-ST12, and the local *S.pyogenes* IQ-No.4 isolate was showed closed related to Pub-MLST data base *S.pyogenes* MLST-ST50 at total genetic changes (0.01%).



Figure (13): Phylogenetic tree analysis based MLST housekeeping *xpt* gene partial sequence in local *S.pyogenes* IQ. isolates that used for confirmative genetic identification of MLST typing analysis. The phylogenetic tree was constructed using Neighbor-Joining tree method in (MEGA 6.0 version). The local *S.pyogenes* IQ. isolates IQ-No.1, No.2, No.3 and -IQ-No.5 were showed closed related to Pub-MLST data base *S.pyogenes* MLST-ST3, and the local *S.pyogenes* IQ. isolate IQ-No.8 was showed closed related to Pub-MLST data base *S.pyogenes* MLST-ST50 at total genetic changes (0.00300-0.00050%).



Figure (14): Phylogenetic tree analysis based MLST housekeeping *yqiL* gene partial sequence in local *S.pyogenes* IQ. isolates that used for confirmative genetic identification of MLST typing analysis. The phylogenetic tree was constructed using Neighbor-Joining tree method in (MEGA 6.0 version). The local *S.pyogenes* IQ. isolates IQ-No.1-IQ-No.5 were showed closed related to Pub-MLST data base *S.pyogenes* MLST-ST7 at total genetic changes (0.0010%). **Table (3).The Pub-MLST Typing analysis for local** *Streptococcus pyogenes* IQ. isolates

related Pub-WILST ST typing								
local Streptococcus pyogenes	Streptococcus pyogenes MLST genes Pub-MLST Typing allele mutation analysis					Pub-MLST MLST ST genotype		
	gki	Gtr	murI	mutS	recP	xpt	yqiL	ST-49
IQ.No.1	11	2	1	3	12	3	7	ST-49
IQ.No.2	11	2	1	3	12	3	7	ST-49
IQ.No.3	11	2	1	3	12	3	7	ST-49
IQ.No.4	11	2	1	3	50	8	7	ST-150
IQ.No.5	11	2	1	3	12	3	7	ST-49

A Journal for New Zealand Herpetology

Discussion:

Multilocus sequence typing (MLST) is a method to study genetic variation of bacterial species by studying allelic variations in housekeeping genes (Rajkumari*et al.*, 2017). It is a very useful nucleotide sequencebased method to study genetic relationships between bacterial species (Maiden, 2006). In case of *S. pyogenes*, this is done through amplification of internal fragments of the seven housekeeping genes that code for enzymes necessary for the viability of the cell (McGregor *et al.*, 2004).

MLST detects changes at DNA level that cannot be inferred from the phenotype, such as serotyping or multilocus enzyme electrophoresis (MLEE) (Achtman*et al.*, 2012).

Seven housekeeping loci were chosen for the characterization of *S. pyogenes* isolates by MLST and for determining their population genetic structure. The nucleotide sequence was determined for an internal portion of about 400 to 500 bp of each gene (Bessen, 2009).

The present results showed that sevenhousekeeping genes was observed in all *S. pyogenes* isolates, 20 (100.0%). The continued expansion of global sampling strategies has revealed that not all seven target genes are carried by all *S. pyogenes* strains, with some sub-populations lacking the target genes for *yqiL* and *xpt* (Davies *et al.*, 2019). Strain variations among different isolates were studied with respect to their alleles in the housekeeping genes, table (3). The present resultswere identified two ST among 20 isolates that included MLST-STs 49 and 150 were common in studied isolates. Single locus variations were seen in isolate 4 with respect to **Transketolase**. Other four isolates not show any variation. Such heterogeneity is a unique feature of *Streptococcus pyogenes* isolates from developing countries (Steer *et al.*, 2016). The relative contributions of point mutation and recombination to the initial stages of clonal diversification can be assessed from MLST data, by identifying those STs that are very closely related, differing at only one of the seven MLST loci (single-locus variants).

MLST is based on the sequence of housekeeping genes that result in each strain having a distinct numerical allelic profile, which is abbreviated to a unique identifier: the sequence type (ST). The relatedness between two strains can then be inferred by the differences between allelic profiles (Rajkumariet al., 2020). For a more comprehensive analysis of the possible patterns of evolutionary descent, a set of rules were proposed and implemented in the eBURST algorithm. These rules allow the division of a data set into several clusters of related strains, dubbed clonal complexes, by implementing a simple model of clonal expansion and diversification. Within each clonal complex, the rules identify which links between STs correspond to the most probable pattern of descent (Francisco et al., 2009). The present results show the MLST data into non-overlapping groups of STs with a user defined level of similarity in their allelic profiles. The most stringent definition of an eBURST group, where all STs assigned to the same group must share alleles at least six of the seven MLST loci with at least one other ST in the group, identifies clusters of closely related genotypes that are considered to be descended from the same founder and that are defined as clonal complexes (Janßenet al., 2015). Two STs were identified among 20 isolates showing that MLST is a good indicator of strain variation among S. pyogenes strains. The allelic change in the isolates result from either mutation or recombination was based on the following assumptions. If there are multiple (more than one) nucleotide differences among the alleles at the locus that differs among SLVs, a recombination event is assumed to have occurred, because the probability of multiple independent point mutations at one locus with none at any of the other six loci is low. If there is only a single nucleotide difference, assignment of the variant as the result of mutation or recombination is more complicated (Yu et al., 2012). The best estimate at present is that recombination changes alleles of housekeeping loci at least 1.4 times more commonly than point mutation (Rajkumariet al., 2020).

Conclusion :The current study showed most *Streptococcus pyogenes* isolated from clinical samples have a high prevalence of housekeeping gene such as,glucose kinase, glutamine

transporter protein, glutamate rasemase, transketolase, xanthine phosphoribosyltransferase and acetyle-CoA acetyltransferase which have a major role in the pathogenicity of bacteria and the effectiveness of MLST scheme for *Streptococcus pyogenes* strains supplies a preferable genotyping tool for comprehension the phylogeny of strains.

References:

- Kimberlin, D.W.; Brady, M.T.; Jackson, M.A.(2015) Group A Streptococcal Infections; American Academy of Pediatrics: Itasca, IL, USA, pp. 732–744.
- Ochoa-Díaz, M.M., Daza-Giovannetty, S. and Gómez-Camargo, D., 2018. Bacterial genotyping methods: from the basics to modern. Host-Pathogen Interactions, pp.13-20.
- Thenmozhi, R., Balaji, K., Kanagavel, M. and KaruthaPandian, S., 2010. Development of species-specific primers for detection of Streptococcus pyogenes from throat swabs. FEMS microbiology letters, 306(2), pp.110-116.
- Vitali, L.A., Gherardi, G. and Petrelli, D., 2015. Pulsed field gel electrophoresis of group A streptococci. In Pulse Field Gel Electrophoresis (pp. 129-138). Humana Press, New York, NY.
- Rahimi-Moghaddam, A., Salmanzadeh-Ahrabi, S., Tahereh, T., Seifali, M. and Pourramezan, Z., 2019. Genotyping of Streptococcus Pyogenes Isolates using Optimized RAPD-PCR Protocol. Biological Journal of Microorganism, 8(32), pp.131-138.
- Frost, H.R., Davies, M.R., Velusamy, S., Delforge, V., Erhart, A., Darboe, S., Steer, A., Walker, M.J., Beall, B., Botteaux, A. and Smeesters, P.R., 2020. Updated emm-typing protocol for Streptococcus pyogenes. Clinical Microbiology and Infection, 26(7), pp.946-e5.
- Riddle, D.J., Bessen, D.E. and Caparon, M.G., 2010. Variation in Streptococcus pyogenes NAD+ glycohydrolase is associated with tissue tropism. Journal of bacteriology, 192(14), pp.3735-3746.
- Pérez-Losada, M., Arenas, M., & Castro-Nallar, E. (2018). Microbial sequence typing in the genomic era. Infection, genetics and evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases, 63, 346–359.
- Maiden, M.C., Van Rensburg, M.J.J., Bray, J.E., Earle, S.G., Ford, S.A., Jolley, K.A. and McCarthy, N.D., 2013. MLST revisited: the gene-by-gene approach to bacterial genomics. Nature Reviews Microbiology, 11(10), pp.728-736.
- Rajkumari, R., Jose, J.M. and Brahmadathan, K.N., 2017. Genetic diversity and allelic variation in south Indian isolates of group A streptococci causing invasive disease. Indian Journal of Medical Microbiology, 35(4), pp.575-579.
- Daniel W. W. (2009). Biostatistics: A Foundation for Analysis in the Health Sciences, John Wiley & Sons New York.
- Maiden M.C., (2006). Multilocus sequence typing of bacteria. Annu Rev Microbiol 2006;60:561-88.
- McGregor K.F., Spratt B.G., Kalia A., Bennett A., Bilek N., Beall B., et al., (2004). Multilocus sequence typing of Streptococcus pyogenes representing most known emm types and distinctions among subpopulation genetic structures. J Bacteriol.186:4285-94.
- Achtman M., Wain J., Weill F.X., Nair S., Zhou Z., Sangal V., et al., (2012). Multilocus sequence typing as a replacement for serotyping in Salmonella enterica. PLoSPathog. 8(6):e1002776.
- Bessen D.E. (2009). Population biology of the human restricted pathogen, Streptococcus pyogenes. Infect. Genet. Evol. 9:581-593.

- Davies M.R., McIntyre L., Mutreja A., Lacey J.A., Lees J.A., et al., (2019). Atlas of group A streptococcal vaccine candidates compiled using large-scale comparative genomics. Nat. Genet. 51,
- Steer A.C., Carapetis J.R., Dale J.B., Fraser J.D., et al., (2016).
- Francisco A.P., Bugalho M., Ramirez M. et al., (2009). Global optimal eBURST analysis of multilocus typing data using a graphic matroid approach. BMC Bioinformatics 10, 152.
- Janßen T., Voss M., Kühl M. et al., (2015). A combinational approach of multilocus sequence typing and other molecular typing methods in unravelling the epidemiology of Erysipelothrixrhusiopathiae strains from poultry and mammals. Vet Res 46, 84.
- Rajkumari, J., Chakraborty, S., &Pandey, P. (2020). Distinctive features gleaned from the comparative genomes analysis of clinical and non-clinical isolates of Klebsiellapneumoniae. Bioinformation, 16(3), 256–268.
- Yu S., Fearnhead P., Holland B.R., Biggs P., Maiden M., French N., (2012). Estimating the relative roles of recombination and point mutation in the generation of single locus variants in Campylobacter jejuni and Campylobacter coli. J MolEvol. 2012 Jun;74(5-6):273-80.