

Analysis of genome variation in Simple sequence repeat (SSR) of meningococcal isolates

Mohammad Abdul Rahmman Al-Maeni

Dept. of Biology, College of Science University of Baghdad, Baghdad, Iraq

*Corresponding Author: Mohammad Abdul Rahmman Al-Maeni

Email: mhogene@yahoo.com

Abstracts

Alteration of repeat tract length within the SSR of phase variable genes may enhance the persistence of isolates within their host for a long time (a period of months) (Alamro et al., 2014). Alamro et al. (2014) showed there was trend towards selection for OFF state or low expression for five phase variable genes (*opc*, *hmr*, *nadA*, *nalP*, *hpuAB*) in three carriers (V54, V124 and V64) within strains belonging into CC174 and CC167 in first, second and third time points. He concluded that the selection for low or OFF state helped *N. meningitidis* to persist for a long time (Alamro et al., 2014). The current study aimed to detect the alteration in the repeat tracts of the same five variable genes within the previous three carriers (V54, V124, and V64) but for other strains belonging to CC22, CC269, and CC198 at the fourth time point. There was also a trend towards selection for an OFF state or low expression for three genes which are (*opc*, *hpuAb*, *nalP*) with 2/3 (66.6%) carriers while *hmr* gene showed ON state in all carriers (100%). This indicates that antibodies formed against (*opc*, *HpuAb*, *nalP*) genes in strains belonging to CC174 and CC167 in first, second and third time points were able to enhance immunization against the isolates in the current study for CC269 and CC22 resulting in the selection for the OFF state or low expression. Conversely, antibodies formed against *hmr* gene in the previous time points on isolates within the CC167 may not show immunization against *hmr* gene in the isolates of current study CC22, CC198 and CC269 and so the gene stayed in the ON state. Interestingly, the *nadA* gene was missing from isolates under the current study. These results may reflect important aspects in the vaccination program especially as *nadA* is one of the component of Bexsero vaccine.

Keywords: Phase variation, SSR, *Neisseria meningitidis*.

Introduction

Phase variation is one of the adaptive strategies of pathogenic and commensal bacteria. This process is characterized by hyper-mutation of DNA sequences or hypervariable methylation in particular regions of the genome sequence and in a reversible manner. Phase variation mechanisms include slipped strand mispairing, site-specific recombination, homologous recombination and epigenetic modification (Bayliss et al., 2008). The major mechanism utilised in *Neisseria* is slipped strand mispairing and so this will be the focus of this section. Slipped strand mispairing of simple sequence repeats (SSR) during DNA replication leads to reversible changes in repeat number for tracts located in the ORF or the promoter region of a gene. Thus, the location of an SSR in an open reading frame (ORF) may lead to a frame shift mutation, whereas alterations in an SSR in a promoter may cause a change in the distance between different components of a promoter. Therefore, phase variation can lead to an abnormal or missing product or a level of expression that is higher or lower than the normal one (Metrucchio et al., 2009). Studies of the mutation rates and patterns of mutation of SSR provide an insight into the amount of genetic variation generated by these repetitive sequences (Bayliss et al., 2008). Mutability of the SSR is influenced by cis-acting factors such as repeat length and trans-acting factors such as DNA replication and repair factors (Moxon et al., 2006; Bayliss, 2009; Bayliss et al., 2001).

Many outer membrane proteins of *N. meningitidis* undergo phase variation and this may facilitate escape of the immune system. Bayliss et al. (2008) have shown that alterations in repeat tracts of *lgtG* were observed when *N. meningitidis* was subject to selection by the bactericidal

activity of a monoclonal antibody specific for a phosphoethanolamine-containing epitope of the LPS. The *LgtG* product controlled addition of a glucose blocking addition of phosphoethanolamine to this specific position in LPS. These results indicated that phase variation mediates changes in the expression of *lgtG* resulting in variants that could escape an immune response and hence adaptation to a stress condition.

Some experiments were designed to investigate alteration in SSR in five phase variable genes during a persistent carriage of three meningococcal strains in three carriers (a CC22, in V54, N428), (a CC198, in V64, N436) and (a CC269, in V124, N419). Fifteen isolates from 2008-2009 were investigated in this study. The data for the first, second and third time points were generated by (Alamro et al., 2014). The aim of this study was to visualize how the SSR of phase variable genes in the fourth time point altered as these carriers were colonized by different strains at this time point.

Materials and Methods

Bacterial isolates and growth conditions

The alteration of SSR in five phase variable genes during persistent carriage was examined using DNA extracts provided by Dr. C. D. Bayliss.

Study Primers

The primers used for the purpose of repeat tract analyses were either taken from (Alamro et al., 2014; Tauseef et al., 2011) or designed using Clone manager 9 software (Table 1).

Table 1: List and sequences of primers used in the current study

NMB1390: Wildtype Forward Primer	ACCCGCCAAACTGATGAC	for SAP method detection NMB1390
NMB1390: Mutant Forward Primer	ACCCGCCAAACTGATGAT	
NMB1390: Common Reverse Primer	TGTCGAGAGCCGAGCCG	
NMB1390: Forward primer	TACGCCGAATAAACAAGCCG	For sequencing NMB1390
NMB1390: Reverse primer	GAACGGCGCTAAGGGCAA	
NMB0329: Wildtype Reverse Primer	CAGCTCGAACACAGCCTAC	for SAP method detection NMB0329
NMB0329: Mutant Reverse Primer	CAGCTCGAACACAGCCTAA	
NMB0329: Common Forward Primer	ATGAGCGTAGGTTTGCTGAG	
NMB0329: Forward primer	ATGAGCGTAGGTTTGCTGAG	For sequencing NMB0329
NMB0329: Reverse primer	GATTAACCTGCCGGGCATC	
NMB0329: Reverse primer	CGCTGATGGGCATAAUCTC	

Name	Sequence	Comment
<i>Hpa</i> 4 - vic for	ATGCGATGAAATACAAAGCCC	Fluorescent labelled and non-fluorescent forward primers to amplify repeat tract in <i>Hpa</i> 4.
<i>Hpa</i> 4- 350 rev	GGATGAAAGGGCGTATTGCGC	
<i>Hpa</i> 4 C for	ATGCGATGAAATACAAAGCCC	
<i>nod</i> 4 for fam	TCGACGTCCTCGATTACGAAGGC	Fluorescent labelled and non-fluorescent forward primers to amplify repeat tract in <i>nod</i> 4.
<i>nod</i> 4 rep. rev	TGGCTGTGGTCAGTACTTTGGATGG	
<i>nod</i> 4 for	TCGACGTCCTCGATTACGAAGGC	
<i>Opc</i> for fam	GAGAATAACAATTCGTTGTA	Fluorescent labelled and non-fluorescent forward primers to amplify repeat tract in <i>opc</i> .
<i>Opc</i> rep rev	CTCATTAGCGGTTTGAAGCTCTTGTCAG	
<i>Opc</i> for	GAGAATAACAATTCGTTGTA	
<i>nal</i> PF1	GTTGCAACAACACTTTCGCCTGC	Fluorescent labelled and non-fluorescent forward primers to amplify repeat tract in <i>nal</i> P
<i>nal</i> PR1 for sequencing	GCAGGTTGTCGTTGCTCATCCAGG	
<i>nal</i> PR2 for sequencing	CAGGCGCTCCCTCCGCATATACG	
<i>nal</i> PF2-FAM	AAATGTGCAAAGACAGAAGCATGC	Fluorescent labelled and non-fluorescent forward primers to amplify repeat tract in <i>hmb</i> R
<i>hmb</i> R-RF3 for sequencing	TGCCAACCTCTTTTACGAATGG	
<i>hmb</i> R-RF4 for sequencing	GCTACTGAACACGTCGTTC	
<i>hmb</i> R-for 2	CGGCATTAGGTCAAAATCCC	
<i>hmb</i> R-Rev 7	GCCGAAGGATCCAATTGGG	

Polymerase chain reaction (PCR)

The target DNA sequences were amplified by PCR with the following components, 1 µl of a 1:10 dilution of DNA (~15 ng/µl) as template, 0.2 µM final concentration of forward and reverse primers (2µM stock concentration), 1 µl of a 10x PCR buffer (10x KAPA Taq buffer A with 15mM MgCl₂), 0.4 µl of 25mM MgCl₂, 0.25 µl of 10mM dNTPs, 0.1 µl of a 5U per µl stock of Taq DNA polymerase (KAPA Biosystems) and 5.25 µl of sterile distilled water. The reaction conditions for each cycle were:-95 °C for 30 seconds, an annealing step (52-60) of 60 seconds and an elongation step at 72 °C for 1 minutes.

A-tailing

An A-tailing step was used when PCR products were subject to GeneScan. The mixture contains 10 µl PCR reaction and A-tailing mix solution that is 0.4 µl PCR buffer, 0.05 µl Taq, DNA polymerase and 3.55 µl dH₂O, and then incubated for 45 minutes at 72°C.

Gene Scan

The variation in the length of PCR products spanning the repeat tracts of five genes was measured using GeneScan. Repeat tracts were amplified with specific primers, one of which was labeled with a fluorescent dye. The GeneScan components were:-0.5 µl of the 1:10 dilution of the PCR products (1:10 in sterile distilled water), 9.25 µl of formamide and 0.25 µl of DNA size standard GS500 LIZ (Life Technologies). Samples were analysed on an ABI3730 machine by the Protein and Nucleic Acids laboratory (PNAACL). Finally, Peak scanner software v1.0 (Applied Biosystems) and Microsoft Excel were used to analyze the sequencer data.

DNA Sequencing

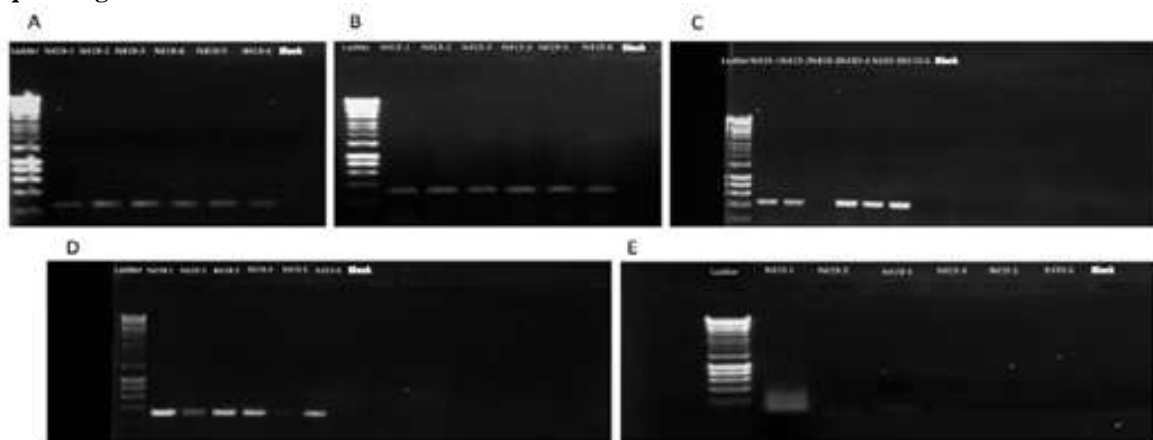


Fig. 1: Amplification of SSRs for five phase variable genes in six isolates. Panel A: *opc*, panel B: *hpuA* gene, panel C: *hmbr* gene, panel D: *nalP* gene, panel E: *nadA* gene.

The GeneScan and Peak Scanner Software were used to estimate the size of fluorescent PCR products. In general, there was one major and one minor peak. The peak with the highest signal was considered the main peak when the ratio between primary and secondary peaks was more than 1.2. On the other hands, if the ratio was between primary and secondary peaks were less than 1.2, the GeneScan technique for particular samples was repeated (Fig. 2).

The reactions for target DNA sequences were set up with the following components, 0.5 µl of DNA PCR product or 1 µl of plasmid DNA, 4 µl of sequencing mix (1:8 Big Dye v3.1 , 1:5 5x Sequencing buffer), 1 µl of primer required (forward or reverse), and water to make the total volume of 10 µl. The reaction conditions of each cycle were 96 °C for 30 second, 50°C for 15 second and 60 °C for 4 minutes. Reactions were analysed on an ABI 3730 DNA Sequencer (ABI, Applied Biosystems) at PNAACL.

Agarose gel Electrophoresis

PCR products were analyzed on gels containing 1% agarose (Seakem LE Agarose, Cambrex), 1x TAE buffer (40mM Tris acetate, 1mM EDTA, pH8.2), and 0.5 Mg/ml ethidium bromide. The DNA samples were mixed with 6x loading Dye (0.25% bromophenol blue). A DNA standard marker, Hyper ladder 1kb (Bioline Reagents Ltd.) was used to calculate the size of PCR fragments. Gel visualization was performed using a transilluminator gel documentation system (Syngene).

Results and Discussion

GeneScan

Variation within the repeat tracts of meningococcal surface proteins is crucial for phase variation. The length of each repeat tract was estimated by the GeneScan technique.

Specific primers labelled with fluorescent dyes were used to amplify particular SSRs for five phase variable genes: - *opc*, *hpuA*, *nalP*, *hmbr* and *nadA*. The amplification of the SSRs with four genes was successful while the *nadA* was not found in the samples (N419, N436 and N428) (Fig. 1).

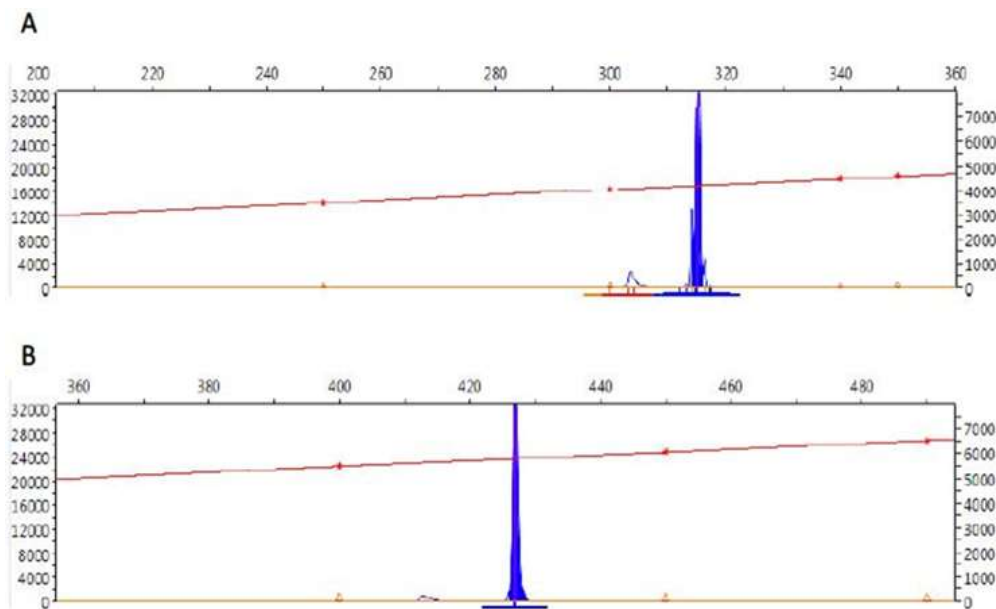


Fig. 2: GeneScan analysis for two phase variable genes form

Panel A: GeneScan analysis of *nalP* with peak size 315 in blue colour, which correlates to nine repeats, panel B: GeneScan analysis of *hnbr* with peak size 426 in blue colour, which correlates to eight repeats.

To be more precise, sequencing of a sub-set of repeat tracts was used to confirm that changes in the PCR product size was due to changes in the length of repeat tract among phase variable genes (Fig. 3).

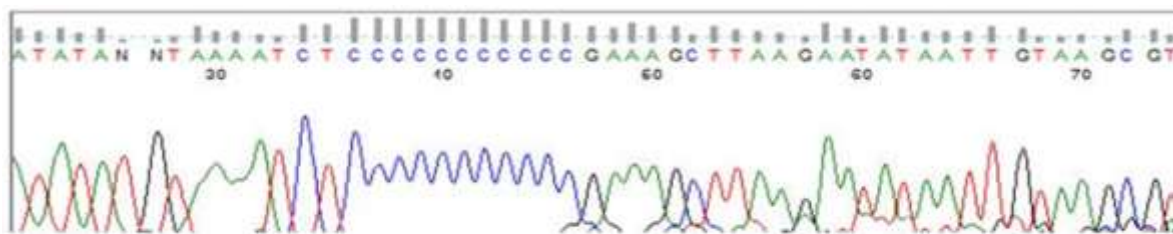


Fig. 3: Example of sequence of the repeat tract of *hpuA* gene from isolate N419.3. The repeat sizes obtained from genescan were confirmed by sequencing.

Distribution of tract lengths During Carriage

In general, Alamro et al. (2014) determined the repeat tract length for four genes under the study in up to 6 colonies of carriers V64 (for strains of type N64 in first time point, N257 in second time point, N348 in third point), V124 (for strains of type N124 in first time point, N290 in second time point, N336 in third point) and V54 (for strains of type N54 in first time point, N237 in second time point, N343 in third point). In the current study, the repeat tracts length were determined for four genes in up to 6 colonies from the fourth time point of same carriers V64 (for strain of type N436) and V124 (for strain of type N419) and up to 3 colonies from the fourth time point of volunteer V54 (for strain of type N428). The strains in the fourth time point were from different CCs than those in the earlier time points.

The results showed that high frequencies of phase variation occurred during the fourth time point in all four genes. The repeat tract of *opc* is located within the core

promoter and changes in the repeat lead to changes in the transcription level of the gene. Sarkari et al. (1994) showed that the *opc* gene varied in its expression depending on the length of repeat tract with high expression of *opc* is correlated with 12Cs, 13Cs, 11Cs, and 14Cs while the low or intermediate expression of *opc* is correlated with less or equal to 10Cs or larger and equal to 15Cs. In carrier V54, in the first time point, the *opc* repeat numbers were 11Cs with high level of expression then in the second time point; *opc* repeat numbers were 10Cs with four colonies with low or intermediate expression and two colonies with 11Cs with high expression. In the third time point, *opc* repeat numbers were 11Cs with three colonies (high expression), 10Cs with two colonies (low or intermediate expression) and 12Cs with one colony (high expression) (Alamro et al., 2014). In this study, in carrier V54, in the fourth time point, the *opc* repeat numbers were 10Cs with two colonies (low or intermediate expression) and 14Cs with one colony (high expression). In carrier V64, in the first time point, the *opc* repeat numbers were 13Cs, 14Cs with high expression. In

the second time point, the *opc* repeat numbers were 1 colony with 15Cs and correlated with low or intermediate expression and 5 colonies with 14Cs and correlated with high expression. In the third time point, the *opc* repeat numbers were 2 colonies with 15Cs and correlated with low or intermediate expression and 4 colonies with 14Cs and 12Cs correlated with high expression (Alamro et al., 2014). In this study, in the fourth time point, in carrier V64, the *opc* repeat numbers were 9Cs and 10Cs with two colonies with low or intermediate expression and 4 colonies with 13Cs and 14Cs with high expression. In carrier V124, in the first, the *opc* repeat numbers were 3 colonies correlated with low or intermediate expression with 15Cs and 3 colonies correlated with high expression with 13Cs, 14Cs. In the second, the *opc* repeat numbers were 13Cs, 14Cs with five colonies (high expression) and 15Cs with one colony (low or intermediate expression). In the third time point, the *opc* repeat numbers were also 15Cs with one colony (low or intermediate expression) and 14Cs with five colonies (high expression) (Alamro et al., 2014). In this study, in the fourth time point, in carrier V124, the *opc* repeat numbers were less than 10Cs therefore all with intermediate or low expression.

It has been reported that the poly G in the *hpuA* gene is located within the gene therefore the change in the repeat tract leads into frameshift mutations. Tauseef et al. (2011) showed that the repeat numbers 7Gs, 10Gs, 13Gs, 16Gs, and 19Gs were associated with ON expression state while other repeat numbers were OFF expression state. In carrier V54, in the first time point, most colonies had a repeat number of 12Gs therefore; *hpuA* was in the OFF state. In the second time point, the *hpuA* repeat numbers were 5 colonies with 13Gs (ON state) and one colony with 11Gs with (OFF state) while in the third time point, the *hpuA* repeat numbers were 13Gs with one colony (ON state) and

11Gs or 12Gs with (OFF state) (Alamro et al., 2014). In this study, in the fourth time point, in carrier V54, the *hpuA* repeat numbers were 10Gs with one colony (ON state) and two colonies with less than 10Gs (5Gs, 6Gs) (OFF state). In carrier V64 and V124, in the first, second and third time point, the *hpuA* repeat numbers were 10Gs therefore all the phase variable genes with ON state (Alamro et al., 2014). In this study, in the fourth time point, in carrier V64 and V124, the *hpuA* repeat numbers were 5Gs, 7Gs, 11Gs, 12Gs and 15Gs so mostly in the OFF state. The *nalP* repeat tract is also located within the open reading frame. In carrier V54, in the first and second time point, the *nalP* repeat numbers were 10Cs therefore all the phase variable genes were in the ON state. In the third time point, the *nalP* repeat numbers were between 11Cs and 12Cs and in an OFF state (Alamro et al., 2014). In this study, in the fourth time point, in carrier V54, the *nalP* gene was 9Cs and also in the OFF state. In carrier V64 and V124, in the first, second and third time points, and the *nalP* repeat numbers were 10Cs therefore, all the phase variable genes had an ON state (Alamro et al., 2014). In this study, in the fourth time point, in carrier V64 and V124, the *nalP* repeat numbers were within a range from 8Cs to 10Cs. There were five colonies with 10Cs with ON state while there was one colony with 8Cs and an OFF state. For strain of types N419 the *nalP* repeat numbers were 8Cs and 9Cs hence in the OFF state.

The *hmbr* gene has a repeat tract in the coding sequence. In this study, in carrier V54, in the fourth time point, the *hmbr* gene had 9Gs or 12Gs with an ON state. In carrier V64, the *hmbr* gene was also between 9Gs and 12Gs with ON state. While for V124, there were two colonies with 8Gs with OFF state and three colonies with 9Gs and one colony with 12Gs with ON state and hence 4 colonies within ON state (Fig. 4).

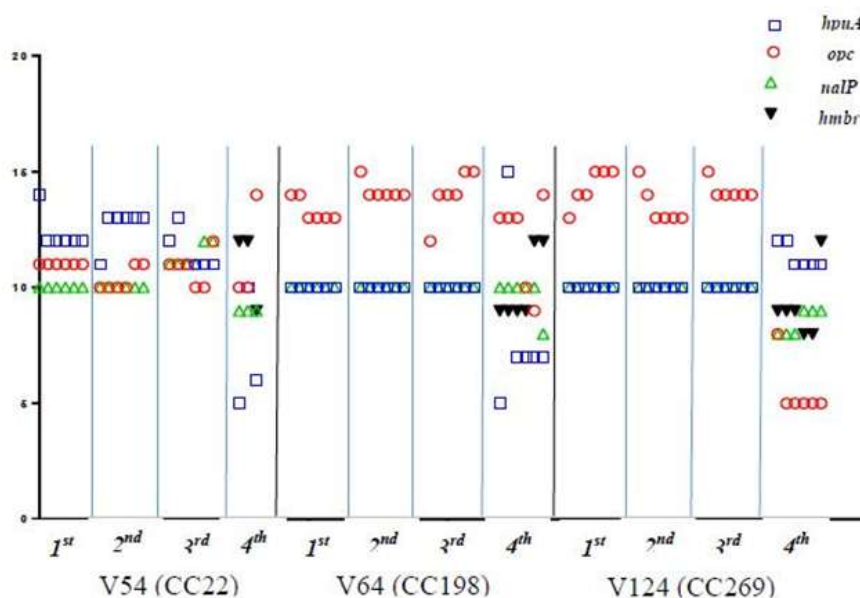


Fig. 4: SSRs of four phase variable genes in fourth time of three carriers.

The *hpuA* gene, blue square, *opc*, red circle, *nalP*, green triangle and *hmbR*, black triangle. 1st, 2nd and 3rd indicate first, second and third time point, these data were taken from (Alamro *et al.*, 2014). The data in this study represented in the fourth time points (4th).

Conflict of Interest: None.

References

1. Alamro, M., Bidmos, F.A., Chan, H., Oldfield, N.J., Newton, E., Bai, X. Phase variation mediates reductions in expression of surface proteins during persistent meningococcal carriage. *Infect Immun* 2014;82:2472-84.
2. Bayliss, C.D., Hoe, J.C., Makepeace, K., Martin, P., Hood, D.W., Moxon, E.R.,. *Neisseria meningitidis* escape from the bactericidal activity of a monoclonal antibody is mediated by phase variation of IgtG and enhanced by a mutator phenotype. *Infect Immun* 2008;76:5038-48.
3. Metruccio, M.M., Pigozzi, E., Roncarati, D., Scorza, F.B., Norais, N., Hill, S.A.,. A novel phase variation mechanism in the meningococcus driven by a ligand-responsive repressor and differential spacing of distal promoter elements. *PLoS Pathog.* 2009;5, e1000710.
4. Moxon, R., Bayliss, C., Hood, D. Bacterial contingency loci: the role of simple sequence DNA repeats in bacterial adaptation. *Annu.Rev.Genet.* 2006;40:307-33.
5. Bayliss, C.D.,. Determinants of phase variation rate and the fitness implications of differing rates for bacterial pathogens and commensals. *FEMS Microbiol Rev* 2009;33:504-20.
6. Bayliss, C.D., Field, D., Moxon, E.R.,. The simple sequence contingency loci of *Haemophilus influenzae* and *Neisseria meningitidis*. *J Clin Investig* 2001;107:657-66.
7. Tauseef, I., Harrison, O.B., Wooldridge, K.G., Feavers, I.M., Neal, K.R., Gray, S.J. Influence of the combination and phase variation status of the haemoglobin receptors *HmbR* and *HpuAB* on meningococcal virulence. *Microbiology.* 2011;157:1446-56.
8. Sarkari, J., Pandit, N., Moxon, E.R., Achtman, M.,. Variable expression of the *Opc* outer membrane protein in *Neisseria meningitidis* is caused by size variation of a promoter containing poly-cytidine. *Mol Microbiol* 1994;13:207-17.

How to cite this article: Al-Maeni MAR, Analysis of genome variation in Simple sequence repeat (SSR) of meningococcal isolates. *J Pharm Biolog Sci* 2019;7(1):