Investigation the possibility of simple sequence repeats for generating a phase variation with the entire *N. meningitidis* genome

Mohammad Abdul Rahmman Al-Maeni, Shaymaa Fouad Rasheed AlKhazraji
Biology department-college of Science-University of Baghdad-Baghdad-Iraq
Corresponding author: A.S.T. Figueiredo
E-mail: alex.figueiredo@sakata.com.br / alexstfigueiredo@gmail.com

Genet. Mol. Res. 16 (2): gmr16029643
Received February 5, 2021
Accepted February 16, 2021
Published February 26, 2021
DOI http://dx.doi.org/10.4238/gmr16029643

ABSTRACT

Phase variation mediating simple sequence repeat (SSR) in sipped strand mispairing is considered the most important mechanisms of genetic variation in *N. meningitidis* and in defense against host attack strategies despite elicitation of the immune system against commensal and pathogenic bacteria. We aimed to achieve in a silico study for determining the possibility of SSR within different positions of *N. meningitidis* genome to enroll within a phase variation mechanism.

We applied different criteria to judge on SSR as it generates a phase variable mechanism relying on different works of literature. These criteria represented by the length, polymorphic, instability of SSR and the value of Z score using a Markov model and synonymous shuffling model of SSR and the position of SSR within the gene or promoter. We detected 67 out of 327 putative phase variable genes spread on N. meningitidis genome that fit with all criteria and we are highly recommended to provide an experimentally evidence on those putative phase variable genes.
Key words: simple sequence repeats; phase variation; *N. meningitidis*

INTRODUCTION

Phase variation is one of the crucial mechanisms of pathogenic and commensal of N. meningitidis. This process can be described as hypermutation mediating repeat tracts or hypervariable methylation mediating a specific target of DNA sequences in a reversible manner. These mechanisms comprise different processes which are an epigenetic modification, homologous recombination, site-specific recombination and slipped-strand mispairing (Bayliss et al., 2008). However, the major player in the processing of repeat changing in the phase variable gene of N. meningitidis is Slipped-strand mispairing mediating SSR. The Slipped-strand mispairing occurs through DNA replication thereby the repeat tract will be changed within the open reading frame or promoter regions. Changing the repeat tract in the position of open reading frame leads to an abnormal or missing product while changing the repeat within the promoter triggers changing within the distance between elements of promoter hence leads to change in a level of expression (Metruccio et al., 2009). Different factors can influence on an amount of changing in a repeat pattern of phase variable genes which are trans-acting factors such as DNA replication and repair factors and cis-acting factors such as repeat length (Moxon et al., 2006; Bayliss, 2009; Bayliss et al., 2001).

Firstly the shorter repeat that identified experimentally as phase variable, they have long repeat in other intrastrains, therefore, it is necessary to include a big chunk of data to see whether there is a really a repeat tract with length lead to phase variability. There is possibilities of genes which are in their ancestral were not phase variable but they transfer to phase variable in due to selection (Saunders et al., 2000). Secondly, the strains have been selected from a carrier as well invasive therefore we suspect to find all variable repeat tracts with the whole genome which works as a phase variation.

MATERIAL AND METHODS

Criteria used to predict SSR generating a phase variation

Target DNA sequence holding SSR can be predicted as a phase variable gene relying on two approaches are comparative and probabilistic analysis methods (Saunders et al., 2000). The criteria have been taken from the principle of both previous methods which they are as follows;

Determination of polymorphism and stability of SSR within intrastrains

Unique simple sequence repeats were identified using MICAS program (Microsatellite Extraction Web server) (Sreenu et al., 2003) from the whole genome sequence of id-20026 invasive isolate which was collected from MRF Meningococcus Genome Library.

Determination of Z score using a Markov chain of order N-2

Markov chain used to calculate Z score for each SSR. The method used to count the expected value as a null model.
The method used to count the expected value could be explained in the following example, the formula used to count expected value for the word "ABCDE"

Determination of Z score using synonymous shuffling model

The global codon shuffling was done by scanning all codons in the genome from which random synonymous CDS were constructed, from which 5 globally shuffled copies of synonymous sequences were generated.

RESULTS

The number of repeat patterns extracted from id-20026 using MICAS program was 45 (table-1 excel sheet appendix). Initially, the criteria used to indicate if SSR involved in phase variable repeat was the same as previously described in Snyder and Nigal's work in which the length and the presence of a repeat is the factor that associated with interpretation of the sequence context (Saunders et al., 2000; Snyder et al., 2001). The SSR identified relying on only the length of the repeat within our cut off were 217, 214, 200, 208, 213, 216 and 200 putative phase variable genes detected in CC1157-N73, CC167-N64, CC174-N59, CC23-ST1655-N258, CC23-ST23-N188, CC60-N114 and Serob-N119 strains respectively.

The alignment and location of SSR were checked thereby there were some SSR that located at the end of counting therefore they were excluded (table-2 excel sheet appendix). On the other hand, there were 57 a new putative phase variable genes which have been identified in invasive isolates and have been not detected in the seven carrier genomes.

DISCUSSIONS

Polymorphism of each SSR was identified depending on an idea that has been taken from the fact that evidence of genome plasticity through slipped-strand mispairing, called phase-variable loci, leads to different polymorphisms in length of repeat tract. Moreover, sequence-specific mutational biases favour the instability of sequences, therefore, it becomes necessary to analyse the distribution of homopolymeric tract lengths and identify the unstable repeat which leads to phase variability. Meanwhile, Polymorphism refers to the number of SSR that have different length for a particular repeat pattern above the cut off within different compared strains, stability refers to the possibility of SSR to be a presence with a stable or variable length within different compared strains. In addition, the variability of repeat is considered as a strong evidence on phase variability because that the instability of repeat probably leads to change in frame shift of genes if it is located within the content of the genes or changes in their length would be expected to alter level of expression of genes if it is located within promoter, therefore, we looked for the stability of repeat tracts.

REFERENCES

Snyder, L. A., Butcher, S. A. & Saunders, N. J. (2001). Comparative whole-genome analyses reveal over 100 putative phase-variable genes in the pathogenic Neisseria spp. Microbiology 147 , 2321–2332.

Van der Ende, A., Hopman, C. T. & Dankert, J. (2000). Multiple mechanisms of phase variation of PorA in Neisseria meningitidis Infect Immun 68 , 6685–6690.

Van Passel MW, Ochman H (2007) Selection on the genic location of disruptive elements. Trends Genet 23: 601–604.

Orsi RH, Bowen BM & Wiedmann M (2010) Homopolymeric tracts represent a general regulatory mechanism in prokaryotes. Genomics 11 : 102.

Martin, P., T. van de Ven, N. Mouchel, A. C. Jeffries, D. W. Hood, and E. R. Moxon. 2003. Experimentally revised repertoire of putative contingency loci in Neisseria meningitidis strain MC58: evidence for a novel mechanism of phase variation. Mol. Microbiol. 50:245–257.

Saunders, N. J., A. C. Jeffries, J. F. Peden, D. W. Hood, H. Tettelin, R. Rappuoli, and E. R. Moxon. 2000. Repeat-associated phase variable genes in the complete genome sequence of Neisseria meningitidis strain MC58. Mol. Microbiol. 37:207–215.

Janulczyk R, Masignani V, Maione D, Tettelin H, Grandi G & Telford JL (2010) Simple sequence repeats and genome plasticity in

Streptococcus agalactiae . J Bacteriol 192: 3990 – 4000.

Lin WH & Kussell E (2012) Evolutionary pressures on simple sequence repeats in prokaryotic coding regions. Nucleic Acids Res 40 : 2399 – 2413.

YC Li, AB Korol, T Fahima, E Nevo (2004) Microsatellites Within Genes: Structure, Function, and Evolution. Mol. Biol. Evol. 21(6):991–1007.

Conflicts of interest

The authors declare no conflict of interest.

ACKNOWLEDGMENTS

The authors are grateful to CNPq, Capes and Finep by all financial support.

REFERENCES

Argillier O, Barrière Y, Panel A, Aziac B, et al. (1998). Variability of digestibility criteria in maize elite hybrids submitted for registration in the French official catalogue. *Agronomie* 18: 639-648. http://dx.doi.org/10.1051/agro:19981003

Argillier O, Méchin V and Barriére Y (2000). Inbred line evaluation and breeding for digestibility-related traits in forage maize. *Crop Sci.* 40: 1596-1600. http://dx.doi.org/10.2135/cropsci2000.4061596x

Barrière Y and Argillier O (1998). *In vivo* silage feeding value of early maize hybrids registered in France between 1958 and 1994. *Euphytica* 99: 175-182. http://dx.doi.org/10.1023/A:1018379931874

Bello OB, Ige SA, Azeez MA, Afolabi MS, et al. (2012). Heritability and genetic advance for grain yield and its components characters in maize (*Zea mays* L.). *Int. J. Plant Res.* 02: 138-145. http://dx.doi.org/10.5923/j.plant.20120205.01

Cruz CD (2013). Genes - a software package for analysis in experimental statistics and quantitative genetics. *Acta Sci. Agron.* 35: 271-276. http://dx.doi.org/10.4025/actasciagron.v35i3.21251

Cruz CD, Regazzi AJ and Carneiro PCS (2012). Modelos biométricos aplicados ao melhoramento genético. UFV, Viçosa.

de Oliveira FCL, Jobim CC, da Silva MS, Calixto Jr M, et al. (2011). Productivity and nutritional value of silage of corn hybrids with different heights of harvest. *Rev. Bras. Zootec.* 40: 720-727.

de Souza Neto IL, Pinto RJB, Scapim CA, Jobim CC, et al. (2015). Diallel analysis and inbreeding depression of hybrid forge corn for agronomic traits and chemical quality. *Bragantia* 74: 42-49.

Ertiro BT, Zeleke H, Friesen D, Blummel M, et al. (2013). Relationship between the performance of parental inbred lines and hybrids for food-feed traits in maize (*Zea mays* L.) in Ethiopia. *Field Crops Res.* 153: 86-93. http://dx.doi.org/10.1016/j.fcr.2013.02.008

Fernandes EH, Schuster I, Scapim CA, Vieira ES, et al. (2015). Genetic diversity in elite inbred lines of maize and its association with heterosis. *Genet. Mol. Res.* 14: 6509-6517. http://dx.doi.org/10.4238/2015.June.12.3

Gomes MS, Von Pinho RG, Ramalho MAP, Ferreira DV, et al. (2004). Variabilidade genética em linhagens de milho nas características relacionadas com a produtividade de silagem. *Pesqui. Agropecu. Bras.* 39: 879-885. http://dx.doi.org/10.1590/S0100-204X2004000900007

Gralak E, Faria MV, Possato O, Junior., Rossi ES, et al. (2014). Capacidade combinatória de híbridos de milho para caracteres agronômicos e bromatológicos da silagem. *Rev. Bras. Milho e Sorgo* 13: 187-200. http://dx.doi.org/10.18512/1980-6477/rbms.v13n2p187-200

Guerrero CG, Robles MAG, Ortega JGL, Castillo IO, et al. (2014). Combining ability and heterosis in corn breeding lines to forage and grain. *Am. J. Plant Sci.* 5: 845-856. http://dx.doi.org/10.4236/ajps.2014.56098

Hallauer AR, Carena MJ and Miranda Filho JB (2010). Quantitative genetics in maize breeding. Springer-Verlag, New York.

Idris AE and Abuali AI (2011). Genetic variability for vegetative and yield traits in maize (*Zea mays* L.) genotypes. *Int. Res. J. Agric. Sci. Soil Sci.* 01: 408-411.