Analysis of gene encoding haemolysin A of Vibrio cholerae isolated in Vietnam

Phân tích gene mà hóa haemolysin A của Vibrio cholerae phân lập ở Việt Nam

Research article

Ha Thi Quyen*

VNU University of Engineering and Technology, 144 Xuan Thuy Road, Cau Giay, Ha Noi, Vietnam

Vibrio cholerae is the cholera causing agent, divided into two biotypes, including the classical biotype and ElTor biotype. Both of these biotypes caused cholera epidemics in the world. The classical biotype caused 6th cholera pandemic (from 1921 to 1961), and ElTor biotype caused 7th cholera pandemic (from 1961 to the 70s). Haemolysin A, a hemolytic protein of V. cholerae ElTor biotype, is encoded by the hlyA gene. This gene is often used for analyzing genetic relationship between strains in the same species or between species in the same Vibrio genus. Results of analyzing nucleotide and amino acid sequences of hlyA gene of V. cholerae strain causing cholera in Vietnam (named hlyA.VN) showed that: the hlyA.VN gene sequence was similar to the hlyA gene sequences of V. cholerae strains of the 6th and 7th cholera epidemics. The hlyA gene of the 6th cholera epidemic strain was deficient in 11 nucleotides (this deficiency leading to the loss of 4 amino acids in the haemolysin A protein) comparing to hlyA.VN gene and hlyA gene of the 7th cholera epidemic strain. The results of genetic distance analysis as well as phylogenetic tree construction also confirmed V. cholerae causing cholera in Vietnam was closely related to the strains causing cholera pandemics in the world. It is great significance for the surveillance of molecular epidemiology to prevent cholera effectively.

Vibrio cholerae là tác nhân gây bệnh tả, được chia thành hai тип sinh học, đó là тип sinh học cổ điển và тип sinh học ElTor. Cả hai type này đã từng gây ra các dịch tả trên thế giới. Type sinh học cổ điển đã từng gây ra đại dịch tả lần thứ 6 (từ năm 1921 đến 1961), còn type sinh học ElTor đã từng gây ra đại dịch tả lần thứ 7 (từ 1961 đến những năm 70). Haemolysin A, một protein có chức năng làm tan máu của V. cholerae типа sinh học ElTor, được mà hóa bởi gen hlyA. Gene này thường được sử dụng cho các phân tích quan hệ di truyền giữa các chủng trong cùng một loại V. cholerae hay giữa các loại trong cùng một chi Vibrio. Kết quả phân tích trình tự nucleotide và axit amin gen hlyA của chủng V. cholerae gây bệnh ở Việt Nam (hlyA.VN) cho thấy: trình tự gen hlyA.VN có sự tương đồng lớn với trình tự gen hlyA của chủng gây đại dịch tả 6 và 7. Gen hlyA của chủng gây đại dịch tả 6 bị thiếu hụt 11 nucleotide (sự thiếu hụt này dẫn tới sự mất 4 axit amin trong phân tử haemolysin A) so với gen hlyA.VN và gene hlyA của chủng gây đại dịch tả 7. Kết quả phân tích khoảng cách di truyền cũng như xây dựng cây phát sinh chủng loại cũng đã khẳng định: chủng gây bệnh ở Việt Nam có quan hệ rất gần với các chủng gây đại dịch tả trên thế giới. Nhận định này có ý nghĩa rất lớn đối với công tác giám sát dịch tả hoắc phân tử để ngăn chặn bệnh tả hiệu quả.

Keywords: classical biotype, El Tor biotype, hlyA gene, Vibrio cholerae

1. Introduction

So far, there are 7 major cholera pandemics in the world caused by Vibrio cholerae. V. cholerae is divided into two biotypes including classical biotype and El Tor biotype. The El Tor biotype is found in all of countries, and the classical biotype is only detected in Bangladesh (Phung Dac Cam, 2003).

In 1905, V. cholerae ElTor biotype was isolated from cholera corpse in ElTor isolating station of Egypt by Gotschlich (Phung Dac Cam, 2003).
V. cholerae can produce haemolysin – a red blood cells dissolving protein. What does haemolysin play a role in the pathogenesis of V. cholerae? Some previous studies showed that haemolysin did not play a role in the disease mechanism of V. cholerae, it only played the role in survival of V. cholera in the natural environment (Byun R. et al, 1999). The ability of producing haemolysin is a new feature acquired during evolution process. This may be due to a mutation or insertion of the transposon gene, because of comparing the nucleotide sequences of the structural coding region for haemolysin of classical biotype and El Tor biotype showed a lack of 11 nucleotides in classical biotype (Alm R.A. et al, 1988; Byun R. Et al, 1999; Rader A.E. et al, 1988).

Haemolysin of V. cholerae El Tor biotype is a product of hlyA gene encoding pre-haemolysin. After removing signal peptides, pre-haemolysine becomes haemolysin. The hlyA gene is located on the HLYA-VIBHIC locus, on chromosome II. This gene was studied by several authors for evaluating the genetic relationship between V. cholerae strains (Brown M.H. et al, 1985; Kotetishvili M. et al, 2003).

In this study, hlyA gene of V. cholerae strain isolated in Vietnam were compared to hlyA genes of some V. cholerae strains that were isolated from different geography regions for evaluation genetic relationship between them.

The evaluation of the genetic relationship between V. cholerae strains is very important for the cholera epidemic surveillance. It provides data at the molecular level to propose accuratesolutions to epidemiological field.

2. Materials and methods

2.1. Materials

Nucleotide sequence of gene encoding haemolysin A (named hlyA.VN) of V. cholerae O1, that was isolated from cholera epidemics in Vietnam. Some homologous nucleotide sequences of hlyA on Genebank (Table 1).

| Name of strains         | Isolated sources       | Toxicity   | hlyA gene     | Name of genes | Accession No. |
|-------------------------|------------------------|------------|---------------|---------------|---------------|
| V. cholerae O1 I389     | Clinical isolate       | toxic      | hlyA          | VN            | AF117833      |
| V. cholerae O1 M793     | Clinical isolate, 7th pandemic | toxic     | hlyA.M793     | AF117835      |
| V. cholerae O1 M645     | Clinical isolate, before 7th pandemic | toxic     | hlyA.M645     | AF117835      |
| V. cholerae O1 569B     | Clinical isolate, 6th pandemic | toxic     | hlyA.M569B    | AF117834      |
| V. cholerae O1 M536     | Environmental isolate | non-toxic  | hlyA.M536     | AF117837      |
| V. cholerae non-O1/non-O139M554 | Environmental isolate | non-toxic  | hlyA.M554     | AF117843      |
| V. mimicus              | Clinical isolate       | toxic      | vmhA          |               | U68271        |

2.2. Methods

Clustal X and MEGA version 6.0softwares were used for molecular evolutionary and genetic analysis.

3. Results and discussion

3.1 The comparison of nucleotide and amino acid sequences of hlyA gene

Vibrio mimicus is responsible for gastroenteritis and is closely related phylogenetically to Vibrio cholerae. For the hlyA genes analysis, vmhA gene sequence of V. mimicus was used as an out of group model because vmhA was identified as the gene encoding heat-resistant haemolysin of V. mimicus (Byun R. et al, 1999). Most of the phenotypic characteristics of V. mimicus are similar to V. cholerae. The only difference between them is sucrose fermenting ability because V. mimicus has ability of fermenting more sugar sources than V. cholerae. In addition, they have common antigens as well as virulence-relating genes, so V. mimicus is also an assistive agent in cholera endemics.

Table 1. Information on sequences used in comparative analysis of hlyA gene

The nucleotide sequence of hlyA gene of V. cholerae in Vietnam was mentioned by Ha Thi Quyen et al. (2008). However, in this report, the nucleotide and amino acid sequences of hlyA genes containing 1047bp after alignment by Clustal X software were compared and analysed in detail. 1047bp of these hlyA genes were compared each other (data not be shown). With 1047bp of nucleotide sequence, 349 amino acids were inferred by MEGA 6.0 software (Figure1).

Nucleotide sequence comparison showed that there was a 9-nucleotide interruption in the haemolysin genes of V. cholerae strains compared to haemolysin gene of V. mimicus. Moreover, the hlyA gene of V. cholerae from the 6th cholera pandemic also lost 11 nucleotides in comparison to hlyA of V. cholerae of Vietnam and other strains (data not shown). That’s why V. cholerae strain of the 6th cholera pandemic lost blood dissolving function.

Because of the disruption in hlyA gene sequence of the V. cholerae strains, the translated proteins also lost four amino acids at 142, 147, 148 and 149 positions comparing to the one of mhA gene. For hlyA gene of 6th pandemic strain, beside the loss of above mentioned four amino acids, it also lost 4 amino acids at positions from 237 to 240, including isoleucine, histidine, leucine and aspara-
gene. This is a major indicator for distinction between El Tor biotype (having the ability of red blood cells dissolving) and classical biotype (not capable of red blood cells dissolving).

3.2 The statistical characteristics of nucleotide and amino acid sequences of hlyA genes

The position of a nucleotide or an amino acid in one sequence is called a characteristic and the change of the nucleotide or amino acid at these positions is called characteristic variation. The statistical characteristics for nucleotide transformation in the hlyA genes were presented in Table 2.

Table 2. The statistical characteristics for nucleotide transformation in the hlyA genes

| Statistical targets | Common between 7 sequences | hlyA.VN and hlyA.M793 | hlyA.VN and hlyA.M569B | hlyA.VN and hlyA.M536 | hlyA.VN and hlyA.M569A 645 | hlyA.VN and hlyA.M554 | hlyA.VN and vmhA |
|---------------------|-----------------------------|------------------------|------------------------|------------------------|----------------------------|------------------------|-----------------|
| Total of characteristics | 1.047 | 1.047 | 1.047 | 1.047 | 1.047 | 1.047 | 1.047 |
| Variable characteristics (%) | (27.1) | (0.29) | (0.29) | (2.3) | (4.2) | (6.0) | (23.3) |
| Characteristics with parsimony | (43) | (0) | (0) | (0) | (0) | (0) | (0) |

Figure 1. Comparison of amino acid sequences of the hlyA.VN gene with some homology sequences
The number of variable characteristics between seven sequences was 27.1%, while the number of characteristics with parsimony information was low (4.1%). This is due to the sequences used in statistical analysis are mostly strains of a common species, so the characteristics of the parsimony are much lower when comparing at the species or genus level. The difference in the sequence is expressed by the ratio of variable characteristics. The ratio of variability of hlyA.VN compared to vmhA is very high (23.3%) because vmhA is a gene of different species. The ratio of variability of hlyA.VN (0.29%) was very low in comparison to hlyA of 6th and 7th pandemic strains (M793 and M569B strains, respectively), but was higher in comparison to non-toxic strain M536 and before 7th pandemic strain M645. This indicates that the pathogenic strain in Vietnam has sequence of hlyA gene that is closer to the hlyA sequences of the pandemic strains.

The statistic result for changing characteristics of the amino acid sequences, that was translated from the hlyA genes, is lower than the changing characteristics of the nucleotide sequences (Table 3). It proved that many changes of nucleotides are synonymous so there is little change in the amino acid sequences. The variation of amino acids between hlyA.VN gene and hlyA genes of V. cholerae is still much lower than vmhA of V. mimicus. For the same species of V. cholerae, the number of amino acid changing characteristics between the strain of Vietnam and the two strains causing 6th and 7th pandemics (strains 569B and M793) have the same value (0.29%).

Table 3. The statistical characteristics for amino acid transformation in the hlyA genes

| Total of characteristics | 349 |
|--------------------------|-----|
| Common between 7 sequences | 73 (20.9) |
| hlyA.VN and hlyA.M793 | 1 (0.29) |
| hlyA.VN and hlyA.569B | 1 (0.29) |
| hlyA.VN and hlyA.M536 | 5 (1.4) |

Table 4. Genetic distance of hlyA genes

| [1] | [2] | [3] | [4] | [5] | [6] |
|-----|-----|-----|-----|-----|-----|
| hlyA.M793 | - | - | - | - | - |
| hlyA.569B | 0.000 | - | - | - | - |
| hlyA.VN | 0.003 | 0.003 | - | - | - |
| hlyA.M536 | 0.020 | 0.020 | 0.023 | - | - |
| hlyA.M645 | 0.040 | 0.040 | 0.043 | 0.040 | - |
| hlyA.M554 | 0.058 | 0.058 | 0.061 | 0.072 | 0.056 | - |
| vmhA | 0.234 | 0.234 | 0.236 | 0.234 | 0.226 | 0.239 |

3.4 The phylogenetic analysis

Genetic relationship between microbial groups are usually presented in geometric form called phylogenetic tree. The end of each tree branch represents the groups of survival organisms. The branching points of the tree express their near and far ancestors. The length of tree branches indicates evolutionary time of the organisms or different levels of DNA sequences. The methods for constructing phylogenetic tree from DNA sequences are based on different principles and evolutionary models described by statistical algorithms. These methods express relationships based on calculating the length of the tree branches. We used neighbor-joining (NJ) method combining maximum-parsimony (MP) method for constructing phylogenetic tree, which would give more accurate analysis results (Nei M. Et al, 2000).
The NJ tree is constructed according to p-distance model with total length of the branches (SBL) of 0.29765. It is a non-rooted tree (Figure 2).

The length of MP tree is 320, the CI (consistency index) is 0.9438 and the RI (retention index) is 0.7049. The CI and RI indicate the degree of homoplasy, meaning the rate of inverse change, convergence and parallelism of nucleotides in evolution process. The CI and RI values are between 0 and 1. The higher the value, the lower the rate of homoplasy. The high rate of homoplasy variation will falsify results of genetic relationship analysis. Conversely, the rate of homoplasy is low, meaning CI and RI values are high, the MP tree will be more reliable. According to the analysis result, CI and RI values were high (≥0.7), so the obtained MP tree is reliable for phylogeny (Figure 3).

4. Conclusion

Based on MEGA version 6.0 software, the nucleotide and amino acid sequences of hlyA gene of *V. cholerae* isolated in Vietnam were compared to some of homologous sequences in Genebank. The results of statistical analysis and phylogenetic tree construction showed that the hlyA.VN gene is closely related to the hlyA genes of the 6th and 7th cholera pandemic strains. This results provided data at molecular level for surveillance of cholera epidemiology in Vietnam.

5. References

[1] Alm R.A., Stroecher U.H., Manning P.A. (1988) Extracellular proteins of Vibrio cholerae: nucleotide sequence of the structural gene (hlyA) for the hemolysin of the hemolytic El Tor strain 017 and characterization of the hlyA mutation in the non-hemolytic classical strain 569B, Molecular Microbiology, 2(4), 481-488.

[2] Brown M.H., Maning P.A. (1985) Haemolysin genes of Vibrio cholerae: presence of homologous DNA in non-haemolytic O1 and haemolytic non-O1 strains. FMES Microbiology Letters, 30, 197-201.

[3] Byun R., Elbourne L. D.H., Lan R., Reeves P.R. (1999) Evolutionary relationships of pathogenic clones of Vibrio cholerae by sequence analysis of four housekeeping genes, Infection and Immunity, 67(3), 1116-1124.

[4] Ha Thi Quyen, Dinh Duy Khang (2008) Genetic relationship of pathogenic Vibrio cholerae clones by sequence analysis of two housekeeping genes, Journal of Preventive Medicine, 18(6) (98) 17-24.

[5] Hillis D.M., Bull J.J. (1993) An empirical test of bootstrapping as a method for assessing confidence in phylogenetic analysis, Systematic Biology, 42, 182-192.

[6] Kotetishvili M., Stine O., Chen Y., Morris Jr.J. (2003) Multilocus Sequence Typing Has Better Discriminatory Ability for Typing Vibrio cholerae than Does Pulsed-Field Gel Electrophoresis and Provides a Measure of Phylogenetic Relatedness, Journal of Clinical Microbiology, 41(5), 1125-1131.

[7] Nei M. and Kumar S. (2000) Molecular evolution and phylogenetic, Oxford University Press.

[8] Phung Dac Cam (2003) Vibrio cholerae and cholera, Medical Publishing House.

[9] Rader A.E., Murphy J.R. (1988) Nucleotide sequences and comparison of the hemolysin determinants of Vibrio cholerae El Tor RV 79 (Hly+) and RV 79 (Hly-) and classical 569B (Hly-), Infection and Immunity, 56(6), 1414-1419.