Untranslated region-5’ and viral protein 1-based genetic stability analysis of bulk polio in Indonesia 2010-2019

Andi Yasmon¹, Normasari²*⁴, Fithriyah¹, Fadilah³

¹Microbiology Department, Faculty of Medicine, Universitas Indonesia, Jakarta, ²Master Program in Biomedical Science, Faculty of Medicine, Universitas Indonesia, Jakarta, ³Chemistry Department, Faculty of Medicine, Universitas Indonesia, Jakarta, ⁴National Quality Control Laboratory of Drug and Food, National Agency of Drug and Food Control

ABSTRACT

Cases of vaccine-associated paralytic poliomyelitis (VAPP) continued increasingly from 2010-2019 in the world. Oral polio vaccine (OPV) is the live attenuated virus-based vaccine that could genetically revert to neurovirulent during the vaccine production process or when the virus replicates in the human body. The poliovirus neurovirulence is determined by the UTR-5’ region and VP1 coding region. UTR-5’ played a role in protein translation and VP1 was responsible for the immunogenicity of the virus. Some reported mutations in UTR-5’ and VP1 could affect the neurovirulence of poliovirus. In this study, we analyzed the genetic stability of the UTR-5’ and VP1 in the bulk of OPV types -1 and -3 produced in 2010 - 2019. The results of the analysis of UTR-5’ sequences in Sabin strain types-1 and -3 produced in 2010 - 2019. The results of the analysis of UTR-5’ sequences in Sabin strain types-1 and -3 and VP1 sequences on Sabin virus type 1 did not show any mutations. Meanwhile, the VP1 sequences in Sabin strain type 3 showed nucleotide mutation C2493U that caused the substitution amino acid Thr6Ile amino acid in all samples of the type 3 bulk polio test. Based on the results of in silico analysis, this mutation in VP1 did not contribute significantly to the neurovirulence of the virus.

ABSTRAK


Keywords:
polio virus
UTR-5’
VP1
Sabin
genetic characterization

*corresponding author: norma.hermansyah@gmail.com
INTRODUCTION

The oral polio vaccine (OPV) is the major vaccine used in polio eradication programs. The advantages of OPV chosen by many countries for polio eradication, such as easy administration, able to increase intestinal immunity to inhibit the replication of poliovirus in the gastrointestinal tract, and low cost. However, the weakness of using this live-attenuated vaccine is the change in the virulence of the Sabin strain that causes vaccine-associated paralysis poliomyelitis (VAPP) cases. Although it is very rare with an estimated 2.9-4.7 cases per million births.¹ The epidemiological study of VAPP by Platt et al.¹ reported that the trend of VAPP increases among countries. Approximately 90% of VAPP cases globally occur in low-middle income countries in South-East Asia, Africa, Western Pacific, and Eastern Mediterranean. However, data and publications in these area are low. World Health Organization (WHO) reported that VAPP cases in 2018-2019 occurred in 5 countries, namely Argentina, Iran, Egypt, Philippines, and Tunisia. It was also reported that VAPP cases associated with immunocompromised individuals vaccinated against OPV continued to increase by 66% from 2010-2019.² Although OPV plays a major role in polio eradication, on the other hand, OPV remains become problem as long as the use of OPV continues. Neurovirulence changes of the Sabin strain are caused by mutations that occur during virus passages in vaccine production and virus replication in the human gut.³⁴

Factors causing VAPP cases were the viral factor and the external factor. The viral factor associated with VAPP cases is the unstable RNA genome of the Sabin strain, causing nucleotide changes that can increase the neurovirulence of the virus. In VAPP cases, it turned out that mutations did not only occur in attenuating nucleotides but could occur at other nucleotide positions in UTR-5’ or in other genes or regions of the viral genome. Taherkhani et al.³ reported that mutations at nt 264 and 559 in UTR-5’ led to changes in neurovirulence of vaccine viruses. In addition, the study conducted by Zhang et al.⁶ reported that amino acid substitutions in VP1 protein of VDPV type 1, such as Ile90Met, Lys99Thr/Glu, and Thr106Ala, increased virus virulence. Meanwhile, the external factor related to VAPP is the host factor. Vaccine-associated paralysis poliomyelitis generally occurred in people with low immune systems, such as immunocompromised people. The low host immune system caused poliovirus to survive for a long time and replicate continuously in the gastrointestinal tract.⁷ It caused the accumulation of nucleotide changes that can affect the neurovirulence of Sabin strain. Because Sabin strain in OPV is genetically unstable, neurovirulence and vaccine safety monitoring should be carried out during vaccine production. In Indonesia, vaccine safety monitoring correlated to neurovirulence in vaccine production was carried out through the neurovirulence test in monkey (MNVT) and Mutant Analysis by PCR and Restriction Enzyme Cleavage (MAPREC) (only in bulk OPV type 3).⁸ The NVT in monkey was in vivo method that used to monitor viral neurovirulence in monkey. The MAPREC method was PCR-RFLP method that only monitored mutations at attenuating nucleotides in UTR-5’, nt 480 (type 1), 481 (type 2), and 472 (type 3), while mutations at other loci of the genome related to neurovirulence were not monitored. Therefore, in this study, DNA sequencing methods were performed on UTR-5’ and VP1 to analyze nucleotide changes that were not detected by the MAPREC assay and associated with the neurovirulence phenotype of the NVT assay in monkeys.
**MATERIALS AND METHODS**

**Ethics statement**

This study was approved by the Ethics Committee of the Faculty of Medicine, University of Indonesia/Dr. Cipto Mangunkusumo General Hospital, Jakarta with the certificate of approval no. KET-147/UN2.F1/ETIK/PPM.00.02/2021 for protocol no. 21-03-0203.

**Samples**

Twenty samples used in the study, 10 samples of OPV bulk type 1 and 10 samples of OPV bulk type 3, were retained from 2010 to 2019 that stored at -70 °C in BPPB, PPPOMN, BPOM. These bulk samples have passed MAPREC test and neurovirulence tests in monkeys.

**PCR primers**

The primer set used for UTR-5 amplification were U1F (5’-TGCGTGTGGTTGAAAACGACG-3’) and U1R (5’-ATAACTGTGAAATGGACTTT-3’) for OPV bulk type 1; and U3F (5’-CATGTACTTCGAGAACACTAGACGCTC-3’) and U3R (5’-AACAAACCAAGGGAGTAATTAATCTGATTCA-3’) for OPV bulk type 3. The primer pairs used for VP1 gene amplification were Y7F 5’-GGGTTTGTGTGCAGCCTGGAATGA-3’ and Q8R 5’-AAGAGGTCTCTTCACTTACAT3’.

**Extraction of viral RNA**

Two hundred microliters of thawed frozen-bulk OPV were extracted for viral RNA using the PureLink Viral RNA/DNA Mini Kit (Invitrogen). The RNA pellet was eluted in 50 µL of RNase free water. Details of methods used are given in the Supplementary Material.

**Reverse transcription**

Five microliters of extracted RNA were reverse transcribed using 200 U SuperScript III reverse transcriptase (Invitrogen) and 3 µg random primers (Invitrogen) in a 20-µL reaction volume according to the manufacturer's instructions. Details of methods used are given in the Supplementary Material.

**PCR and DNA sequencing**

Amplification of the UTR-5’ sequence and the VP1 gene in this study used the i-MAX™ II (INtRON) kit in a 35-µL reaction volume. Details of methods used are given in the Supplementary Material. The PCR product purification and sequencing procedures were carried out by PT Genetics Science. The DNA sequencing process is carried out in two reactions using forward and reverse primers.

**Mutation analysis and amino acid change**

Bidirectional sequences that obtained from DNA sequencing were analyzed using BioEdit. Overlap sequences were analyzed and edited using SeqScape v2.7 (Applied Biosystems) to generate consensus sequences. The consensus sequences were compared to the reference sequences Sabin type 1 (accession number: AY184219) and Sabin type 3 (accession number: AY184221). The nucleotide sequences of VP1 were translated into amino acid sequences using BioEdit. Alignment of the sample VP1 amino acid sequence with the Sabin reference strain VP1 amino acid sequence was performed with multiple alignment ClustalW.

**Protein conformation**

The three-dimensional conformation
of the VP1 protein from samples undergoing amino acid sequence changes were analyzed using the Robetta web server (https://robetta.bakerlab.org/).\textsuperscript{14} The protein conformational model was validated using the Ramachandran Plot (https://zlab.umassmed.edu).\textsuperscript{15} The VP1 protein structure of OPV bulk sample was aligned with VP1 protein structure of reference strain using PyMol software.

**Phylogenetic analysis**

Phylogenetic trees were constructed from VP1 coding region sequences using Neighbor Joining method in Mega X version 7.0.\textsuperscript{16} The evolutionary distances were estimated using Tamura-Nei method for UTR-5' and Jones-Taylor-Thornton (JTT) model for VP1. To construct phylogenetic tree of OPV bulk, this study used reference Sabin, wild-type, attenuated-type and VAPP/VDPV strains that were obtained from GenBank (Supplementary Material).

**RESULTS**

**Phylogenetic Analysis**

Phylogenetic investigations based on the UTR-5' sequence (FIGURE 1.A.) showed that all bulks of OPV type 1 had closest relationship to the reference strain Sabin 1 and VAPP isolates from Nigeria, India, US, Taiwan, Greece, and China. Meanwhile, phylogenetic investigations based on the VP1 coding region (FIGURE 1.B.) showed that all bulks of OPV type 1 had closest relationship with the strain reference Sabin 1 and VAPP/VDPV isolates from Nigeria, India, US, Taiwan, Greece, China, Russia, and UK.

![FIGURE 1. The phylogenetic tree of OPV bulk type 1 based on UTR-5' sequence (A) and amino acid sequences of VP1 (B). The samples of OVP bulk type 1 with black parallelogram, attenuated strains with black circle, wild-type poliovirus strains with black rectangles and VDVPV/VAPP strains with black triangles.](image-url)
Phylogenetic investigations of bulks of OPV type 3 based on UTR-5 sequences (FIGURE 2.A.) showed that all samples had closest relationship to reference strain Sabin 3 and other attenuated strains, ZhongIII2 and Leon 12 a1b as seed viruses used in OPV type 3 production. In addition, they were closely related to Sabin-like strain from Nigeria and VAPP strains from China as well. The VP1 gene phylogenetic tree (FIGURE 2.B.) showed that bulks of OPV type 3 were closely related to Sabin 3, attenuated strain P3/Leon12 a1b, and VAPP strains in US, Nigeria, and UK. In the VP1 phylogenetic tree, the bulks of OPV type 3 were in a different clade from ZhongIII2.

FIGURE 2. The phylogenetic tree of OPV bulk type 3 based on UTR-5' sequence (A) and amino acid sequences of VP1 (B). The samples of OVP bulk type 3 with black parallelogram, attenuated strains with a black circle, wild-type poliovirus strains with black rectangles and VDVP/VAPP strains with black triangles.

**Mutation analysis and amino acid changes**

To identify mutations in UTR-5' sequences, the obtained UTR-5' sequences from ten bulk samples of type 1 (231-690 bp) and ten bulk samples of OPV type 3 (267-643 bp) were aligned with UTR-5' sequences from reference strain Sabin 1 and 3, respectively. Analysis of UTR-5' sequences of OPV bulk polio types 1 and 3 showed those sequences similar to reference strain Sabin 1 and 3, respectively. Analysis of VP1 amino acid sequences (300 amino acids) of twenty bulk samples of OPV bulk type 1 and 3 showed those of OPV bulk type 1 similar to reference strain Sabin 1 (TABLE 1). On the other hand, VP1 OPV bulk type 3 differed from reference strain Sabin 3 at amino acid position six of VP1, Thr6Ile, as a missense mutation (TABLE 1). Sequence divergence in VP1 of 20 samples of bulk polio types 1 and 3 < 1% so these are still categorized as Sabin strain.
TABLE 1. Alignment of VP1 amino acids from 20 bulk samples of OPV type 1 and 3 produced in 2010-2019 at the hotspot of VP1 amino acid changes related to poliovirus virulence.

<table>
<thead>
<tr>
<th>Strain</th>
<th>Amino acid position in VP1 protein</th>
<th>90\textsuperscript{6,17,18}</th>
<th>99\textsuperscript{6,17,18}</th>
<th>106\textsuperscript{6,17,18}</th>
<th>134\textsuperscript{6}</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sabin 1*</td>
<td>CVAITVDNSASTKNDKLFTVWKI</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>Mahoney**</td>
<td>.T.M....P..T.....A....</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>S1-1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S1-2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S1-3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S1-4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S1-5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S1-6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S1-7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S1-8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S1-9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S1-10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>67\textsuperscript{17}</td>
<td>54\textsuperscript{19}</td>
<td>10\textsuperscript{319}</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sabin 3*</td>
<td>DLTSEVAQ</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P/Leon**</td>
<td>.I....</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td></td>
</tr>
<tr>
<td>S3-1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S3-2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S3-3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S3-4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S3-5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S3-6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S3-7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S3-8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S3-9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S3-10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: *reference strain of OPV, **wild-type poliovirus strain. The dot (.) represents the consensus of the amino acids of bulk polio that are similar to the reference strain Sabin.

**Protein conformation**

The amino acid substitution that occurred from this missense mutation was analyzed for the three-dimensional structure of the VP1 protein with the Robetta web server (https://robetta.bakerlab.org/). The 3D protein model VP1 of samples of OPV bulk type 3 and the reference strain Sabin 3 were validated by Ramachandran plots. The Ramachandran plot of the VP1 protein model of OPV bulk type 3 (FIGURE 3.A.) showed that there were 255 amino acids (97.32%) in the favored region, six amino acids (2.29%) in the allowed region, and one amino acid (0.38%) in the outlier’s region. Meanwhile, the Ramachandran plot of the VP1 protein model of reference strain Sabin 3 (FIGURE 3.B.) showed that there were 253 amino acids (96.57%) in the favored region, eight amino acids (3.05%) in the allowed region, and one amino acid (0.38%) in the outlier’s region. These results indicate the good quality of the protein model.
The result of overlapping VP1 protein structures from the bulk polio type 3 and VP1 Sabin 3 (FIGURE 4.) showed that Thr6Ile substitution caused a conformational change in the VP1 protein. The conformational change seen in the folding of the VP1 protein of OPV bulk type 3 tended to close compared to Sabin 3. These conformational change in the N-terminal region was suspected not affect the function of VP1 related to the interaction with the CD155 receptor or the function of immunogenicity.

FIGURE 4. Overlapping of VP1 protein models of OPV bulk type 3 and Sabin 3 using PyMol (https://pymol.org/2/). The VP1 protein model of OPV bulk type 3 shown as blue line. The VP1 protein model of the reference strain Sabin 3 shown as green line. Thr6Ile substitution is marked with a purple asterisk.
DISCUSSION

The phylogenetic tree of bulk OPV types 1 and 3 based on UTR-5' and VP1 showed that the bulk of OPV types 1 and 3 were close related to the reference strain (Sabin 1 and Sabin 3). These also showed that VAPP strains from various countries have close relationship with OPV bulk types 1 and 3. The VAPP strain, which was in the same clade as the bulk polio type 1 and 3 samples, did not have back mutations at attenuation points of 480 (type 1) or 472 (type 3). Therefore, the occurrence of VAPP might be due to mutations at other positions in UTR-5' or mutations in other genes that affect poliovirus neurovirulence.

Mutations in UTR-5' region could affect the regulation of protein translation and RNA replication of poliovirus. Several studies reported changes in UTR-5' region causing changes in the secondary structure and affinity for cellular factors that affect protein translation and viral replication. Protein translation and viral replication are important factors that influence viral virulence properties. The translational activity of the virus is closely related to the level of viral replication. In this study, the UTR-5' sequences from all bulks of OPV types 1 and 3 had no nucleotide changes. It also states that bulks of OPV types 1 and 3 had no reverse mutation in attenuating nucleotides, A480G (type 1) and U472C (type 3), while the secondary data from the MAPREC test on bulk polio type 3 used in this study reported the accumulation of the 472C revertant mutant with a proportion of <1%. It was the limitation of Sanger DNA sequencing method which only determine the nucleotide base sequence of nucleic acid/gene/genome sequences but cannot detect a mixture of sequence variants with a proportion of <1% due to background noise generated by the polymerase enzyme during the PCR process. However, there is a high-throughput and accurate sequencing method, Massively Parallel Sequencing or deep sequencing can analyze quasispecies of viruses that can identify the very small number of viral mutants (<1%), so it could be developed to quantify revertent mutations at the attenuation point.

Mutations in the VP1 gene could affect poliovirus neurovirulence as well because VP1 contributes greatly to antigenic neutralization sites and as a constituent of the outer capsid that interacts with the external environment and host cells. This study showed that all bulks of OPV type 1 had no change in the nucleotide sequences of the VP1 gene. Meanwhile, all bulks of OPV type 3 samples had C2493U mutation in VP1 gene. The results of alignment of amino acid sequence of VP1 in TABLE 1, showed C2493U mutation is a missense mutation, caused change in amino acid position six from threonine to isoleucine in VP1 protein of bulk OPV type 3. The results of this study were in line with the study conducted by Neverov and Chumakov which found an accumulation of the C2493U mutation in OPV type 3 analyzed by the Massively Parallel Sequencing method. In addition to the research conducted by Deng et al. also reported an accumulation of C2493U mutations in Pfizer and Zhong-3 strains of 13% and 99%, respectively, using the deep sequencing method.

The in vitro study showed that the C2493U mutation in VP1 correlated with a partial loss of attenuation properties of the Sabin 3. In addition, it was also reported that the Sabin 3 strain containing the 2493U mutation usually failed the neurovirulence test of transgenic mice with the CD155 receptor but passed the NVT test in monkeys. This is in line with the secondary data from the NVT test in monkeys for the ten bulk polio type 3 samples used in this study, which were not neurovirulent in monkeys. Based on
that, the C2493U mutation caused a non-significant increase in neurovirulence of Sabin 3 in vivo. It did not cause paralytic polio in monkeys.

The three-dimensional conformational analysis of the VP1 protein model in silico in figure 4. showed that there is a change in the conformation of the protein VP1 bulk polio type 3 with Thr6Ile substitution which tends closer than the conformational protein VP1 of the reference strain Sabin 3. This change could be due to differences in the properties of the acid. polar amino threonine with a hydroxyl side group and isoleucine amino acid which is nonpolar with a hydrocarbon side group. The conformational stability of the protein can be affected by a single amino acid substitution. However, in this study, analysis of the protein binding of VP1 of bulk polio type 3 with Thr6Ile substitution for the CD155 receptor was not carried out so that the effect of changes in confirmation of VP1 protein on the affinity of the receptor is unknown. The study conducted by Wien et al. reported that the amino acid position 6 is in the N-terminal region of VP1 and is located within the virion which is associated with the terminal regions of other capsid proteins to form a network that plays a role in the stability of virion assembly. Tatem explained that upon attachment of the virus to the host cell, the N-terminal region of VP1 is externalized from within the virion, which then exposes the N-terminal into the cell required for attachment to liposomes. Studies conducted by Kirkegaard also found that deletions in the N-terminal VP1 caused the failure of the release of viral RNA from the capsid into the host cell. Based on the results of these studies, it can be stated that the N-terminal in VP1 plays a role in the entry of the virus into the cell, does not play a role in the interaction of the virus with the host cell receptor.

CONCLUSION

In this study we found that UTR-5' sequence of Sabin strains of type -1 and -3 in OPV bulk was conserved. The VP1 gene from Sabin type 1 strain in bulk OPV was more conserve than Sabin type 3 in bulk OPV. Based on 3D analysis, mutation of Ile6Thr in VP1 of OPV bulk Sabin type 3 is predicted to not give any significant effect to viral neurovirulence. However, it is necessary to consider to do genetic stability monitoring of Sabin during OPV production process.

ACKNOWLEDGEMENTS

This study was funded by National Agency of Drug and Food Control (NADFC), Republic of Indonesia.

REFERENCES


