Bocavirus

**Abstract**Human bocavirus is a parvovirus discovered in 2005 that causes infection in respiratory tract, most commonly in children. The virus was first detected by PCR method using nasopharyngeal swab samples, even though that it can be detected in blood and fecal samples. HBoV strains, such as HBoV2, HBoV3 and HBoV4, have been detected worldwide, without any regional or geographic restrictions. Bocavirus is a small non-enveloped virus that contains single-stranded DNA and three open reading frames that encodes four proteins: NS1, NP1, VP1 and VP2. Patients reported various symptoms after infection with this virus, most commonly pneumonia, fever, and vomiting.

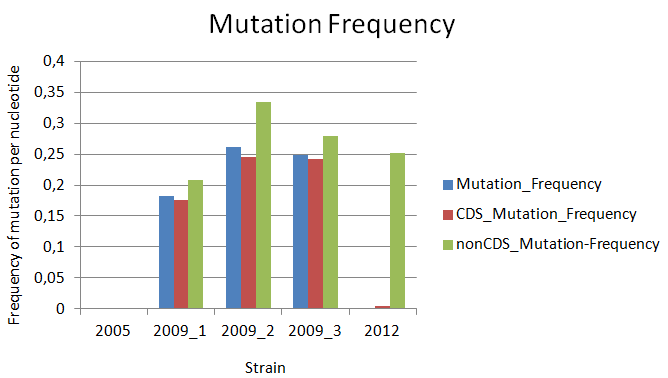
**Introduction**

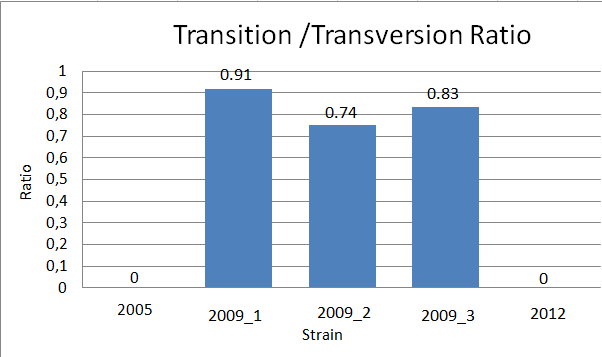
Human bocavirus (HBoV) is a parvovirus with a worldwide distribution that primarily infects children aged between six and twenty-four months, causing a respiratory infection. The HBoV genotypes belong to the family Parvoviridae, subfamily Parvovirinae, genus Bocavirus, which causes infection only in vertebrates. The family Parvoviridae also includes the subfamily Densovirinae, which infects arthropods. The first bocavirus (HBoV1) was discovered in 2005 by PCR from nasopharyngeal swabsamples, after which three more strains termed HBoV2, HBoV3 and HBoV4 were discovered, but their clinical significance in symptomatic infections has not been determined yet. The name Bocavirus derives from the combination of names of its two close relatives, Bovine Parvovirus (BPV) and Canine Minute Virus (CMV).

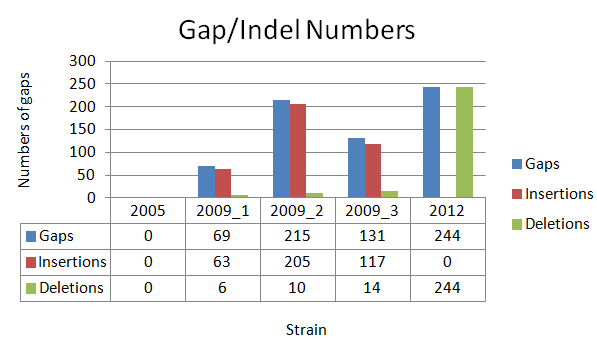
Bocavirus is a small, single stranded DNA virus of nearby 5300 nucleotides that doesn't contain an envelope. The genome of bocavirus is organized in three open reading frames. The first open reading frame encodes two forms of the non-structural protein NS1, the second open reading frame encodes a supplementary non-structural protein, the nuclear phosphoprotein NP1, while the third open reading frame encodes the two structural viral capsid proteins VP1 and VP2. The exact mode of transmission of bocavirus has not yet been determined. However, this virus is thought to be transmitted like other parvoviruses, by inhalation and by getting in contact with respiratory secretions. During the infection, capsid proteins of the virus bind to cellular receptors. Virus replication takes place in the nucleus of the cell, where viral DNA replication is mediated by DNA polymerase of the host cell. After congregation, viral progeny can be released by cell lysis. Although the bocavirus primarily causes respiratory infections, diarrhea, fever, cyanosis and vomiting are accompanying symptoms. Furthermore, this virus is characterized by the fact that it stays in the respiratory tract for a longer time than other viruses. (Guido *et al.*, 2016)

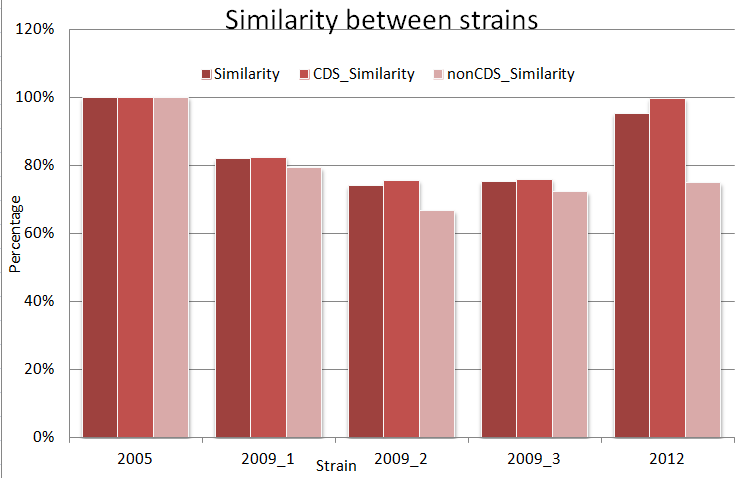
**Materials and Methods**  
  
The National Center for Biotechnology Information page provides us with all information required for sequence analysis. After searching for Genome Assembly and Annotation report of a chosen virus, the bocaparvovirus replicons were available, which lead us to the whole sequences of five bocaparvovirus strains: three human and two primate strains.  
We used entire sequences of those strains for Multiple Sequence Alignment, which was done in Clustal Omega program. After selecting the 'DNA' as a sequence type, we copied all our sequences, ordered them hierarchical and pressed submit. The Clustal Omega program provided us with all information needed for further analysis.

**Results**Length of a whole sequence is 5555 base pairs, from which coding sequence is 4582 base pairs long while the lengths of non-coding sequence is 973 base pairs.

  
The 2005 strain has overall accumulated 0 mutations, hence the mutation frequency is 0.  
Compared to 2005 strain, 2009\_1 strain has accumulated 1011 mutations in total, from which 809 mutations in coding and 202 mutations in non-coding region. Overall mutation frequency is 0,18 per nucleotide.  
When compared to strain from 2005, 2009\_2 strain overall accumulated 1450 mutations, 1125 in coding and 325 mutations in non-coding region. Overall mutation frequency is 0,26 per nucleotide.  
In comparison to 2005 strain, 2009\_3 strain accumulated 1380 mutations in total, wherefrom 1190 mutations in coding region and 271 mutations in non-coding region. Overall mutation frequency is 0,24 per nucleotide.  
The 2012 strain has accumulated 263 mutations in total, 18 in coding and 245 in non-coding region. Overall mutation rate is 0,04 per nucleotide.

  
Further analysis revealed that 2005 and 2012 strains have zero transition/ tansversion ratio.  
The transition/ transversion ratio of a 2009\_2 strain is 0,74.  
The transition/ transversion ratio of a 2009\_3 strain is 0,83 while the 2009\_1 strain has the highest transition/transvertion ratio of 0,91.

  
It is discovered that strain from 2005 has zero gaps in sequence along with zero deletions and insertions.   
When compared to 2005 strain, the 2009\_1 strain has in total 69 gaps accumulated, from which 15 in coding and 54 in non-coding region. This sequence accumulated 63 insertions, from which 15 in coding and 48 in the non-coding region. Overall, 6 deletions were discovered, and they are all found in the non-coding region.  
In comparison with 2005 strain, 2009\_2 strain has 215 gaps accumulated, from which 27 are found in coding and 188 in non-coding region. This sequence accumulated 205 insertions in total, from which 27 are found in coding and 178 in non-coding sequence. In total, 10 deletions were discovered, all of them in non-coding regions.   
Compared with the 2005 strain, 2009\_3 strain accumulated 131 gaps, 30 in coding and 101 in non-coding region. Collectively, 117 insertions were discovered, 30 in coding and 87 in the non-coding region. Fourteen deletions were found, all of them were present in the non-coding region.   
When compare to 2005 strain, the strain from 2012 has 244 gaps and 244 deletions, all of them were located in the non-coding region. This sequence accumulated zero insertions.

  
  
2005 strain is 100% similar to itself.  
2009\_1 strain is 81.80% similar to the 2005 strain.  
2009\_2 strain is 73.89% similar to 2005 strain.   
2009\_3 strain is 75.15% similar to 2005 strain.  
The strain from 2012 is 95.26% similar to 2005 strain.

**Discussion**  
  
  
Compared to 2005 strain, three strains; 2009\_1, 2009\_2 and 2009\_3 accumulated more mutations in coding regions, while 2012 strain accumulated more mutations in the non-coding region. However, the mutation frequency in non-coding regions of all four strains is bigger than a mutation frequency in coding regions.   
Most commonly, mutations occur in non-coding regions, which don’t have any noticeable effect since those genes are not expressed. Mutations in coding regions rarely occur because these mutations can often be harmful. Given that as many as three of our virus strains have more mutations in the coding region, we can assume that these are for example silent mutations, such as changes in a single base in the nucleotide sequence, that do not have a significant impact on the organism.

Throughout the analysis of the virus strains, it was discovered that the 2012 strain is most similar to the 2005 strain, with a similarity of 95.26%. It is understandable that these two strains are most similar because they are derived from humans, while the other three strains are derived from primates.  
However, previous studies discovered that primate strains of bocavirus differed by ≥8 % from the next closest human bocavirus in their nucleotide sequence, this high degree of similarity was an unexpected finding.

The strain from the year of 2012 has 244 gaps in a sequence and analysis revealed that they are all deletions, hence this strain accumulated 0 insertions.  
On the other side, all three strains from the year of 2009 have the higher number of insertions than deletions. It can be assumed that primate virus strains are more prone to certain types of mutations, in this particular case insertions, while human strains are more prone to deletions.

**Health effects**  
Considering that the bocavirus is relatively recently discovered, only a small proportion of its health effects have been determined up to this point.   
It is known that, in humans, this virus primarily causes respiratory tract infection, followed by diarrhea, fever, cyanosis and vomiting. Patients also reported symptoms of gastroenteritis and shortness of breath.  
In 2017, a group of scientists published a study related to the clinical features of a bocavirus. In total, 185 patients diagnosed with bocavirus infection participated in the study. Nasopharyngeal samples were taken when patients had symptoms of acute respiratory infection. Out of 185 patients with HBoV infection, 75 patients had upper respiratory tract infection, while 110 of them had pneumonia. One hundred and nineteen patients became infected by contact with other people. Also, seventy-three patients (39.5%) visited the emergency department because of a HBoV infection.(Lee *et al.*, 2019)  
In children, infection with this virus included following symptoms: cough 79%, fever 67%, rhinorrhea 66%, hypoxia 40%, tachypnea 35% and wheezing 27%. (Allander, 2008)  
Since this type of virus in humans is quite similar to the virus that infects primates, the clinical effect is also expected to be similar. Nevertheless, bocavirus in primates causes somewhat milder symptoms, followed by formation of lesions in the lungs. It should be taken into account that the infection can spread from respiratory to gastrointestinal tract. (Kumakamba *et al.*, 2018)  
It is known that this virus in humans, as well as in primates, stays in the respiratory tract for a longer time than other viruses. However, it is not yet known what the incubation period may be for bocavirus infections.

**Conclusion**  
Based on the current data, the pathogenic roles of the various bocaparvovirus genotypes are still undetermined. Since HBoV strains have the worldwide distribution, without any geographic restrictions, it is important to optimize methods for HBoV identificatication, which is so far done only trough PCR, rarely is it done with serological methods.   
In order to survive, viruses mutate constantly and very quickly, resulting in the frequent emergence of new strains. Bocavirus is still being researched and there are many things that have not yet been identified. It is expected that in the near future we will have more information about this virus, as well as about the differences between strains, that will contribute to faster development of an effective vaccine to control the spread of infection.

**References**

Allander, T. (2008) ‘Human bocavirus’, *Journal of Clinical Virology: The Official Publication of the Pan American Society for Clinical Virology*, 41(1), pp. 29–33. doi:10.1016/j.jcv.2007.10.026.

Guido, M. *et al.* (2016) ‘Human bocavirus: Current knowledge and future challenges’, *World Journal of Gastroenterology*, 22(39), pp. 8684–8697. doi:10.3748/wjg.v22.i39.8684.

Kumakamba, C. *et al.* (2018) ‘DNA indicative of human bocaviruses detected in non-human primates in the Democratic Republic of the Congo’, *The Journal of General Virology*, 99(5), pp. 676–681. doi:10.1099/jgv.0.001048.

Lee, H.N. *et al.* (2019) ‘Human Bocavirus Infection in Adults: Clinical Features and Radiological Findings’, *Korean Journal of Radiology*, 20(7), pp. 1226–1235. doi:10.3348/kjr.2018.0634.