Inspection of MERS in the Clinical Point

Jaewoong Lee

2018.06.15.

Contents

List of Tables

List of Figures

1 Introduction

1.1 Coronavirus

Coronavirus, also know as CoV, is discovered in 1960s from patients with common cold, and these viruses were named human CoV 229E and OC43.[1] Coronaviruses cause a significant percentage of common colds in human. Coronavirus OC43 or 229E infection was detected in 17 of the 200 patients.[2] Common human coronaviruses usually cause mild to moderate upper-respiratory tract illnesses, however, some mutations of coronaviruses, for example SARS-CoV and MERS-CoV, have been known to frequently cause sever symptoms.[3]

1.2 MERS Coronavirus

MERS, also known as Middle East respiratory syndrome, is a viral respiratory infection caused by the MERS-CoV, but it may be an overestimate of the true mortality rate.[4] MERS-CoV causes many symptoms such as muscle pain, shortness or breath, and fever. Approximately 35% of reported patients with MERS have died.[4] MERS-Cov is derived from common CoV, but the fatality is significantly higher than common CoV, so we will find out the reason.

2 Object

2.1 How MERS-CoV can be detected?

As I mentioned later, ORF 1ab is used for screening MERS-CoV. The set of coronavirus are too big, so there are no appropriate strategy to divide with common cold and MERS-CoV. Hence, in this inspection, we divide coronavirus set with scoring about ORF.

2.2 How MERS-CoV avoid the antibody?

As I mentioned later, DPP4 affects to immune regulation. We assume that MERS-CoV has DPP4 somehow, so they can avoid immune response, and they have high fatality.

3 Material

3.1 ORF 1ab

ORF, also known as open reading frame, is the part of a reading frame that has ability to be translated. So an ORF is a continuous stretch of codons that contain a start codon and a stop codon. Also, ORF 1ab is used for detection MERS CoV with real-time reverse-transcription polymerase chain reaction.[5]

3.2 DPP4

DPP4, also known as dipeptidyl peptidase-4 or CD26, is a protein that is encoded by the DPP4 gene. The protein encoded by the DPP4 gene is an antigenic enzyme expressed on the surface of most cell types and is associated with immune regulation, signal transduction and apoptosis.

4 Method

4.1 Sequence Alignment

A sequence alignment is a way of arranging the sequences of DNA, RNA, or protein to identify regions of similarity that may be a consequence of functional, structural, or evolutionary relationships between the sequences. Aligned sequences of nucleotide or amino acid residues are typically represented as row within a matrix. Gaps inserted between the residues so that identical or similar characters are aligned in successive columns.

4.1.1 Global Alignments & Local Alignments

Global alignments, which attempt to align every residue in every sequence, are most useful when the sequence in the query set are roughly equal size. However, local alignments are more useful for dissimilar sequences that are suspected to contain regions of similar sequence motifs within their larger sequence context. In this inspection, I need to check that one sequence contains the other, so I used a local alignments only.

4.1.2 BLOSUM

BLOSUM, which stands for blocks substitution matrix, is a substitution matrix used for sequence alignment of proteins. Also, scoring matrix is a matrix which describes the probability of a biological meaningful amino acid residue pair occurring in an alignment. BLOSUM matrices with high number are designed for comparing closely related sequences, while those with low numbers are designed for comparing distant related sequences. In this inspection, I used the BLOSUM50 and BLOSUM62.
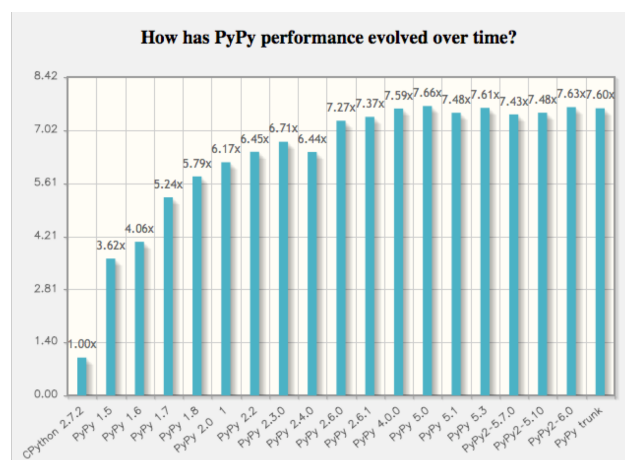
4.2 PyPy

Figure 1: Speed Improvement of PyPy

PyPy is an alternative implementation of the Python which often runs faster than the standard implementation of Python. PyPy is a just-in-time compiler while CPython, the standard, is an interpreter, so PyPy is faster than CPython. The speed improvement of PyPy is as shown as the figure 1.

5 Result

5.1 ORF with BLOSUM50

The local alignment score of ORF with BLOSUM50 is shown as the table 1. Also, the graph which from the table 1 is the graph 2.

Table 1: Local Alignment Score of ORF with BLOSUM62

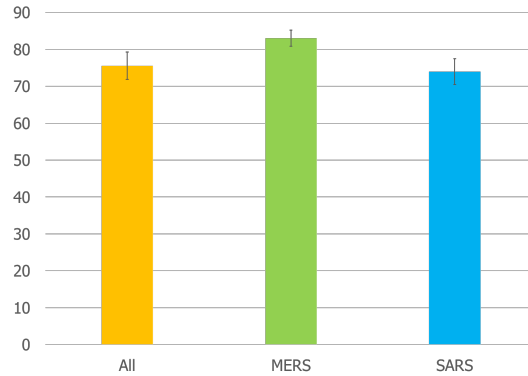|	Average	Standard Deviation
All	75.56527415	7.41806982
MERS	83	4.359748549
SARS	73.95859873	6.960462913

Figure 2: Local Alignment Score of ORF with BLOSUM50

5.2 ORF with BLOSUM62

The local alignment score of ORF with BLOSUM62 is shown as the table 2. Also, the graph which from the table 2 is the graph 3.

Table 2: Local Alignment Score of ORF with BLOSUM62

	Average	Standard Deviation
All	48.47650131	3.999522361
MERS	50.97794118	1.057368584
SARS	47.92993631	4.196702054

5.3 DPP4 with BLOSUM50

The local alignment score of DPP4 with BLOSUM50 is shown as the table 3. Also, the graph which from the table 3 is the graph 4.

Table 3: Local Alignment Score of DPP4 with BLOSUM50

	Average	Standard Deviation
All	66.9843342	6.744236175
MERS	63.77941176	3.882899216
SARS	67.65605096	7.014874491

5.4 DPP4 with BLOSUM62

The local alignment score of DPP4 with BLOSUM62 is shown as the table 4. Also, the graph which from the table 4 is the graph 5.
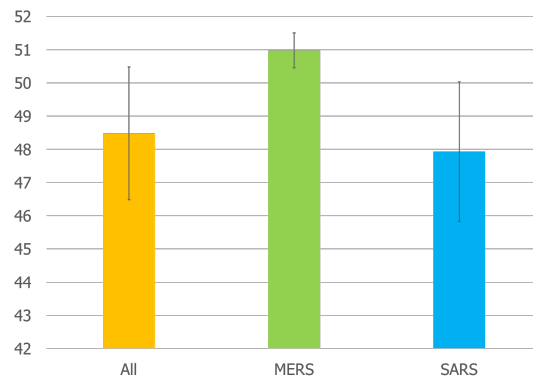
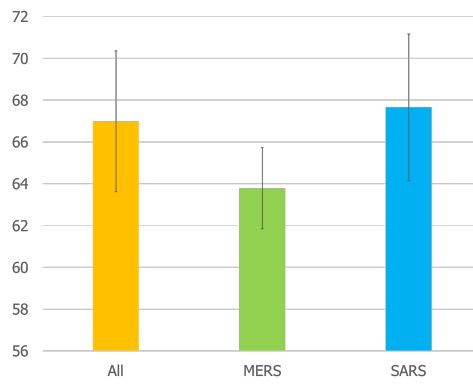Figure 3: Local Alignment Score of ORF with BLOSUM62

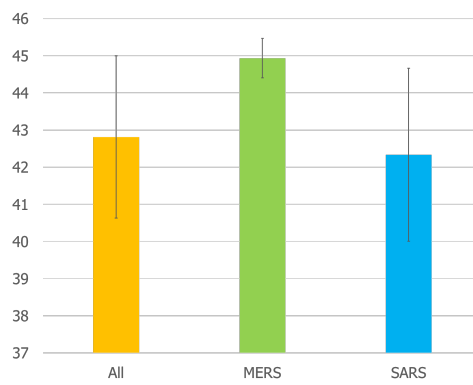Figure 4: Local Alignment Score of DPP4 with BLOSUM50

Figure 5: Local Alignment Score of DPP4 with BLOSUM62

Table 4: Local Alignment Score of DPP4 with BLOSUM62

	Average	Standard Deviation
All	42.81201044	4.363683984
MERS	44.93382353	1.055512555
SARS	67.65605096	7.014874491

6 Discussion

6.1 ORF

As comparing the figure 2 and the figure 3, they have similar appearance. As I mentioned hereinabove, ORF is used for screening MERS-CoV, so there should be sub-sequence of MERS-CoV which is similar to ORF. Hence, MERS-CoV has higher score than other group.

6.2 DPP4

In the figure 4, MERS-CoV has lower score than the other; however, in the figure 5, MERS-CoV has higher score than the other. The DPP4-similar region of MERS-CoV is came from host, in other word, MERS-CoV copied from host. Hence, there should be sub-sequence of MERS-CoV which is almost identical, over similar, to DPP4, so the score with BLOSUM50 is lower than with BLOSUM62.

7 Suggestion

7.1 Gene Set

In this inspection, we used some kind of coronavirus; common cold CoV, MERS-CoV, and SARS-CoV. However, the size of common cold CoV is only two, so we cannot calculate the statistics value.

7.2 Wobble

The given gene represented with DNA sequence, not amino acid sequence. So, in this inspection, I translate DNA sequence into amino acid sequence as cell does. Also, in DNA sequence, there are some wobbles, but I randomly choose what they translate into. For example, W can be A and T, but I assume that W only can be A. In the next research, I want to make the program to choose really random, and repeats that many times, then it can be more closer to real nature.

7.3 BLOSUM

In this inspection, I calculated sequence alignment with BLOSUM50 and BLOSUM62. If I do this inspection with more BLOSUM matrix, then I can confirm that my analysis is more exactly or not.

References

[1] Chloé Geller, Mihayl Varbanov, and Raphaël E Duval. Human coronaviruses: insights into environmental resistance and its influence on the development of new antiseptic strategies. *Viruses*, 4(11):3044–3068, 2012.

[2] Mika J Mäkelä, Tuomo Puhakka, Olli Ruuskanen, Maija Leinonen, Pekka Saikku, Marko Kimpimäki, Soile Blomqvist, Timo Hyypiä, and Pertti Arstila. Viruses and bacteria in the etiology of the common cold. *Journal of clinical microbiology*, 36(2):539–542, 1998.

[3] Centers for Disease Control and Prevention. Coronaviruses. `https://www.cdc.gov/coronavirus/about/symptoms.html`, 2017. Accessed: 2018-06-15.

[4] World Health Organization. Middle east respiratory syndrome coronavirus (mers-cov). `http://www.who.int/en/news-room/fact-sheets/detail/middle-east-respiratory-syndrome-coronavirus-(mers-cov)`, 2018. Accessed: 2018-06-15.

[5] World Health Organization. Laboratory testing for middle east respiratory syndrome coronavirus. `http://www.who.int/csr/disease/coronavirus_infections/mers-laboratory-testing/en/`, 2018. Accessed: 2018-06-15.