

Comparison of episodes of mosquito borne disease: Dengue, Yellow Fever, West Nile, and Filariasis with Decision tree, Apriori Algorithm

Seung Hye Song*, Yijeong Choi*, and Taeseon Yoon**

*Natural Science, Hankuk Academy of Foreign Studies, Wangsanri Yongin, Republic of Korea

** Computer Science and Engineering, Korea University, Anamdong Seoul, Republic of Korea

sarasong98@hafs.hs.kr, ejeong100@hafs.hs.kr, tsyoon@scie.kr

Abstract—This study focuses on the mosquito borne diseases including dengue-1, dengue-4, yellow fever, west nile virus infection, and filariasis. These are the diseases that are typically shown in the African continent and Eastern Asia, which are the places that suffer from poverty the most. Vaccines for some of the diseases have already been made but the ones who inhabit in those areas do not have the ability to afford it. The research will use 3, 5, 7 windows of decision tree and apriori algorithm. By finding the similarities between the amino acid sequences between these viruses the study to find the curing for all diseases will be in progress. Emerging filariviruses: the spread and resurgence of Japanese encephalitis, West Nile and dengue viruses

Keywords—Dengue-1, Dengue-4, Yello Fever, West Nile Virus, Filariasis, Dicision Tree, Apriori Algorithm

I. INTRODUCTION

Allegedly, there are numerous kinds of endemic diseases across Africa. This study was held in order to compare the transmissions of the diseases that are infected by mosquitos. The infection method varies in many kinds however, this study focuses on five specific diseases: dengue fever, yellow fever, west nile virus infection, filariasis, and malaria. To find the relations between these infections, research includes an comparison of amino acid sequence through decision tree, neural network, and apriori algorithm. And by this research we tried to find the similarities between the amino acid sequences between these viruses the study to find the curing for all diseases will be in progress.

II. MATERIAL AND METHOD

An easy way to comply with the conference paper formatting requirements is to use this document as a template and simply type your text into it.

A. Dengue fever

Dengue fever is a mosquito-borne tropical disease caused by the dengue virus, whose symptoms include fever, headache, muscle and joint pains, and a characteristic skin rash that is similar to measles[9]. There are 4 distinct, but closely related,

serotypes of the virus that cause dengue (DEN-1, DEN-2, DEN-3 and DEN-4). Recovery from infection by one provides lifelong immunity against that particular serotype[10]. However, cross-immunity to the other serotypes after recovery is only partial and temporary. Dengue is transmitted by several species of mosquito within the genus *Aedes*, particularly *A. aegypti*. No specific treatment or vaccine to protect against dengue exists but can be prevented by reducing the habitat and the number of mosquitoes and limiting exposure to bites[11].

B. Yellow fever

Yellow fever virus is an RNA virus of the genus *Flavivirus*. Like other Arboviruses which are transmitted by mosquitoes, the yellow fever virus is taken up by a female mosquito *Aedes aegypti* when it ingests the blood of an infected human or other primate[12]. In most cases, this virus accompanies symptoms such as fever, chills, loss of appetite, nausea, muscle pains particularly in the back, and headache which typically improve within five days. Absolute diagnosis of the virus can be confirmed by blood sample testing which requires polymerase chain reaction. No specific treatment against yellow fever exists[13] but vaccination for prevention exists and it can also be prevented by mosquito control.

C. WNV(West Nile Virus)

WNV is a mosquito-borne zoonotic arbovirus whose structure is similar to the dengue fever; both belong to the genus *Flavivirus* within the family *Flaviviridae*. Mainly the transmission of MNV is progressed via various species of mosquitoes, which are the prime vector, with birds and reptiles being the most commonly infected animal and serving as the prime reservoir host [1][2]. Infection with WNV is either asymptomatic (no symptoms) in around 80% of infected people, or about 20% of people who become infected with WNV will develop West Nile fever. Symptoms include fever, headache, tiredness, and body aches, nausea, vomiting,

occasionally with a skin rash and swollen lymph glands[3]. Until now, no specific vaccine against MNV was made.

D. Filariasis

Filariasis is commonly infected by the filarial worms [4]. When bitten by the mosquito, the parasites move into the host and grow into an adult worm. The worm affects the lymphatic system and leads to symptoms such as elephantiasis[1], edema [2], abdominal pain, and thickening skin [5]. There are diverse kinds of Filariasis emerging in Africa and it tends to episode throughout all body parts. Found in a primary stage, it could be treated with anti-filarial drugs but in lethal situations, the symptoms must be taken care of through surgery for hydrocele and care of the skin and exercise [5].

E. Decision Tree

This method is a decision support tool that generates a tree-like graph of decisions, in order to apply it into a certain algorithm[7]. In this study, we use this method to extract the specific patterns of amino acid sequences in each virus. The procedure done is called the 'Data mining.' Data mining refers to extracting patterns from raw or warehoused data, in other words, efficiently extracting previously unknown but potentially useful patterns from data. The extracted patterns can be used to classify the data, understand the characteristics of each data set, and predict outcomes for future situations. In this process, decision tree is used as a predictive model which decides the value of given data sets.

F. Apriori Algorithm

Apriori[8] is an algorithm for frequent item set mining and association rule learning over transactional databases. It proceeds by identifying the frequent individual items in the database and extending them to larger and larger item sets as long as those item sets appear sufficiently often in the database. The frequent item sets determined by Apriori can be used to determine association rules which highlight general trends in the database: this has applications in domains such as market basket analysis.

III. EXPERIMENT

In this experiment class 1 stands for Bru, class 2 stands for Dengue, class 3 stands for Yellow fever, and class 4 stands for West Nile.

A. Decision tree

TABLE 1. RULE EXTRACTION OF DECISION TREE. (9WINDOW)

class	rule
Class1	Pos1=F Pos8=L
Class2	Pos6=H Pos9=S Pos6=K Pos9=S

Class3	Pos6=W Pos9=S
Class4	Pos8=A Pos9=K

TABLE 2. RULE EXTRACTION OF DECISION TREE.(13WINDOW)

class	rule
Class1	Pos4=F
Class2	Pos2=T Pos4=R
Class3	Pos2=C Pos5=E
Class4	Pos1=T Pos2=A

TABLE 3. RULE EXTRACTION OF DECISION TREE. (17WINDOW)

class	rule
Class1	Pos15=F
Class2	Pos5=V Pos15=K
Class3	Pos2=G Pos10=W
Class4	Pos3=I Pos5=R

After performing experiments, we selected the results that had the frequency higher than 0.8. According to the results under 9window, the feature of class1 is very clear. The most distinctive amino acid, amino acid F could be the one that demonstrates Class 1. However, the results under other windows are quite unclear. So, we made the conclusion that there are clear differences between classes. Therefore, we can distinguish four viruses very easily based on these decision tree results.

B. Apriori Algorithm

In 9 window result, few similar points are found through comparing Dengue virus and Yellow fever virus. It seems there are similarities between West Nile virus and Brugia malayi mitochondrion. But they are not associated viruses at all.

TABLE 4. RULE EXTRACTION OF DECISION TREE (9WINDOW).

9win	Dengue	Brugia malayi Mitochondrion	Yellow Fever	West Nile
A				45
F		113		
G	43		47	
I		61		
L		100		47
R	53		42	
S	49			

T				24
V		55		

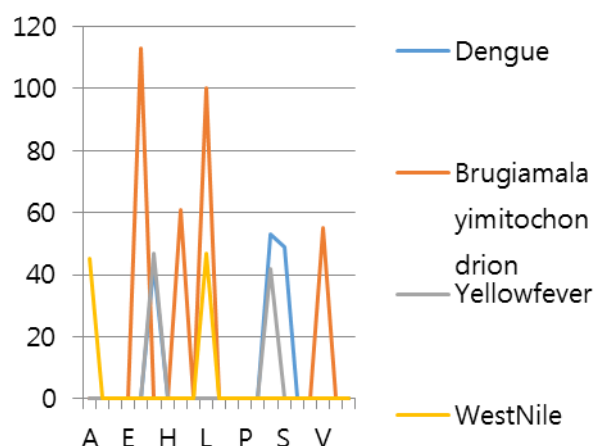


Figure 1. Amino acid an image with 9window

In 13 window result, similarities between Dengue virus and Yellow fever virus are detected. Like 9 window, there are some similarities between West Nile virus and Brugiayimichondrion. Superficially, they are different viruses, but, they are similar in the way that they both have amino acid Leucine. Through this we may estimate that Leucine might be the chronic factor that relates West Nile virus and Brugiayimichondrion.

TABLE 5. RULE EXTRACTION OF DECISION TREE (13WINDOW).

13win	Dengue	Brugiayimichondrion	Yellow Fever	West Nile
A				29
F		82		
G	33		35	32
I		41		
L		68		34
R	32			
S	34		30	
T				30
V		40		

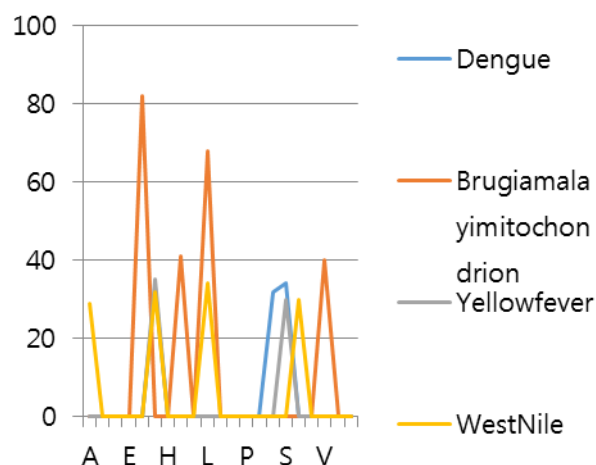


Figure 2. Amino acid an image with 13window

In 17 window result, the results are similar to 13 window result. We can find definite correlation between Dengue virus and Yellow fever virus. Thus, similarities between West Nile virus and Brugiayimichondrion are shown, too. They all have Leucine in common.

TABLE 6. RULE EXTRACTION OF DECISION TREE (17WINDOW).

17win	Dengue	Brugiayimichondrion	Yellow Fever	West Nile
A				26
F			26	
G		64		
I	24		32	25
L		37		
R	21			
S		53		28
T	27		25	
V	24		21	

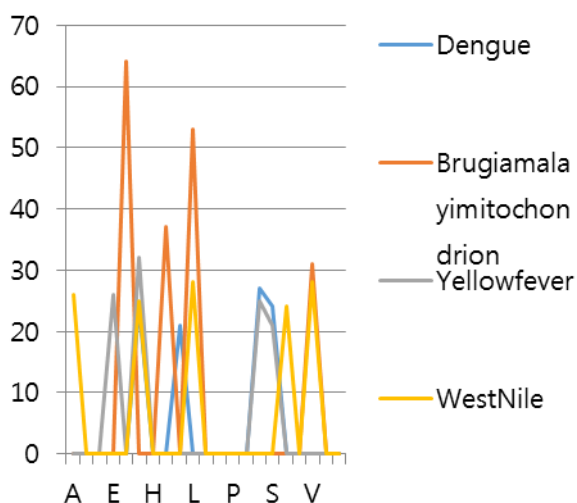


Figure 3. Amino acid an image with 17window

IV. Conclusions

Usually mosquito borne diseases such as dengue-1, dengue-4, yellow fever, west nile virus infection, and filariasis were primarily found in the African continent and Eastern Asia. Two areas share common social issues regarding poverty which leads to an assumption that even though cure for some diseases mentioned above exists, people can't afford to buy it. So by finding a concomitant feature between the amino acid sequences between these viruses we tried to make a progress in making a cure that can cure the disease all at one swoop. First of all, analysis using Decision Tree, trying 9, 13, and 17 windows for analysis, results of under 13,17 windows were quite unclear which we concluded as the four viruses are clearly distinguishable. However, results under 9windows, we could find the vivid feature of Brugia Malayimitchondrion demonstrated as amino acid F. In sequence, analysis using Apriori Algorithm, trying 9, 13, and 17 windows of analysis, commonly, in all three windows, we were able to notice that there are some similar traits between Dengue virus and Yellow fever virus. Futhermore, WestNile and Brugia Malayimitchondrion also share similar traits in this case amino acid Leucine.

By research, we were able to approach a conclusion that even though it might seem that there are similar traits between Dengue virus, Yellowfever virus and WestNile, Brugia Malayimitchondrion, actually overall results of the experiment seems to show more strongly that they are

distinctive from each other. In addition we could assume that controlling Leucine can attribute in developing a unique cure which is effective in both WestNile and Brugia Malayimitchondrion.

REFERENCES

- [1] Steinman, A.; Banet-Noach, C.; Tal, S.; Levi, O.; Simanov, L.; Perk, S.; Malkinson, M.; Shpigel, N. (2003). "West Nile Virus Infection in Crocodiles". *Emerging Infectious Diseases* **9** (7): 887–889..
- [2] Klenk, K.; Snow, J.; Morgan, K.; Bowen, R.; Stephens, M.; Foster, F.; Gordy, P.; Beckett, S.; Komar, N.; Gubler, D.; Bunning, M. (2004). "Alligators as West Nile Virus Amplifiers". *Emerging Infectious Diseases* **10** (12): 2150–2155
- [3] "West Nile Virus: What You Need to Know CDC Fact Sheet". www.CDC.gov. Retrieved 2012-04-09.
- [4] Center for Disease Control and Prevention. "Lymphatic Filariasis". Retrieved 18 July 2010.
- [5] "Lymphatic filariasis". *Health Topics A to Z*. Source: The World Health Organization. Retrieved 2013-03-24.
- [6] Gaur RL, Dixit S, Sahoo MK, Khanna M, Singh S, Murthy PK. Anti-filarial activity of novel formulations of albendazole against experimental brugian filariasis. *Parasitology*. 2007.
- [7] Quinlan, J. R. "Induction of Decision Trees." *Machine Learning* 1.1 (1986): 81-106..
- [8] Rakesh Agrawal and Ramakrishnan Srikant Fast algorithms for mining association rules in large databases. *Proceedings of the 20th International Conference on Very Large Data Bases, VLDB*, pages 487-499, Santiago, Chile, September 1994..
- [9] Guha-Sapir, Debarati, and Barbara Schimmer. "Dengue fever: new paradigms for a changing epidemiology." *Emerg Themes Epidemiol* 2.1 (2005): 1.
- [10] World Health Organization. "Dengue." (2012).
- [11] Gubler, Duane J., and Gary G. Clark. "Dengue/dengue hemorrhagic fever: the emergence of a global health problem." *Emerging infectious diseases* 1.2 (1995): 55.
- [12] Strode, George K. "Yellow Fever." *Yellow fever* (1951).
- [13] Staples, J. Erin, Mark Gershman, and Marc Fischer. "Yellow fever vaccine: recommendations of the Advisory Committee on Immunization Practices (ACIP)." *MMWR*.