

Complexity in correlated cytokine networks associated with COVID-19

Hiroshi Morimoto

Abstract—The COVID-19 pandemic has led to considerable global morbidity and mortality. The cytokine storm is thought to be a major feature among patients with COVID-19. Therefore, profound understanding of cytokine network is necessary. For the study of an inter-correlated cytokine network, identifying the complexity of cytokine network will be an important issue.

Most of the research studies on cytokine networks have attempted to find significant cytokines by observing the change of DNA expression levels in cases and comparing them to controls. However, fewer researches have delved into the features of the network itself. In this paper, we proposed a practical measure to identify the complexity of the cytokine network by applying mathematical graph theory.

Herein, we constructed a correlation network of cytokines from DNA expression data derived from NCBI. By applying the mathematical graph theory to these networks, we looked for prominent sub-networks and extracted several interesting types of sub-networks of cytokines and described their characteristics as a network. In particular, we found a complex type of network in the correlation network associated with COVID-19 patients.

These findings suggest a mechanism by which the correlated cytokine network is induced by COVID-19. Our findings also suggest potential novel prediction of cytokine storm for COVID-19 disease.

Index Terms—COVID-19, graph theory, network, cytokine, cytokine storm.

I. INTRODUCTION

The on-going coronavirus disease (COVID-19) pandemic is still spreading worldwide and resulting in a large number of deaths. It is postulated that a hyper-inflammatory response, such as the cytokine storm, is a major cause of the acute respiratory distress syndrome (ARDS) in patients with COVID-19. Therefore, further information on the complexity of the cytokine correlation network is required to understand the cytokine storm induced by COVID-19.

A cytokine correlation network plays an important role in the pathogenesis of other diseases. For example, a positive feedback network including IL-17 and IL-6 is known to result in the induction of severe rheumatism [1],[2],[3]. When this IL-17 and IL-6 cascade network is driven, it may result in pain, swelling and progressive destruction of the small joints of the hands and feet, accompanied by loss of function. There are cases of rheumatic diseases that are closely related with COVID-19 patients [4], therefore the features of cytokine correlation network become a very important issue even for COVID-19.

Thus, research is required to understand the complexity associated with the cytokine correlation network. However, there are no established measures to evaluate the complexity of the cytokine network. Most research studies on the

cytokine network tried to find significant cytokines which increased in COVID-19 patients and compared them to those who were not infected to COVID-19 from a calculation of \log_2FC .

In this paper, we developed a practical measure to describe the complexity of the cytokine network by applying the mathematical graph theory to a cytokine network in a mathematical graph format. This idea enabled us to apply the fruitful results of the graph theory to correlated cytokine network. A reference for graph theory was given in method section. The necessary results of graph theory were also described in the same section.

We explored a subset of networks (sub-networks) among the cytokine correlation network derived from DNA expression data of COVID-19 patients. The expression data were derived from NCBI with GEO accession number GSE156701. We extracted several types of sub-networks of cytokines and described their characteristics as a network. Based on the graph theory, there was a prominent complex sub-network.

Our findings will contribute to develop our knowledge of features of networks and will suggest a practical method to find therapeutic targets of cytokine storms induced in COVID-19 disease.

II. METHODS

We extracted the gene expression data, which were originally contributed to NCBI by [5], using the GEO accession number GSE156701. The organism used was *macaca mulatta*, the samples were those in SARS-CoV-2-Infected Rhesus Macaques. The samples, from nine adult rhesus macaques (6 to 12 years of age), showed robust induction of interferon and proinflammatory markers. Bronchoalveolar lavage (BAL) data were collected 1, 2, 4, 7, 10 and 14 days post-infection. The data also included information on 6 additional uninfected rhesus macaques. Thus, the total number of samples was 60, i.e., (9 infected rhesus macaques)*(6 assessment time-points) + (6 control samples). Since the total number of genes was 18292, the whole dataset was expressed as a matrix of 18292 rows and 60 columns.

Using all the DNA expression data (excluding data of controls), we constructed a cytokine correlation network using the software “R”, especially “bioConductor” packages (such as “GeneNet”).

We used the mathematical method, called “graph theory”, as an analytical method referring to [6] as a reference for this field. Since the correlation network obtained by GeneNet could be considered as a “graph” from a mathematical point of view, we could apply fruitful theorems of graph theory to this cytokine correlation network as follows.

To define the “complexity” of a network, the graph theory supplied the notion of “planarity” [6, Chapter 5]. A graph is

called a planar graph if it can be drawn in a plane without crossings, i.e., so that no two edges intersect geometrically except at a vertex to which both are incident.

The graph theory supplied the following criterion for planarity:

Theorem 1 ([6], p.67 Corollary 13.4 (i))

If G is a simple connected planar graph with n (greater than or equal to 3) vertices and m edges, then m less than or equal to $3n - 6$

There is another notion that expresses complexity, i.e., the “genus”. A surface is of genus g if it is topologically homeomorphic (mostly equal as a graph) to a sphere with g handles. A graph that can be drawn without crossings on a surface of genus g , but not on one of genus $g - 1$, is called a graph of genus g .

Thus a theorem that estimates genus is stated as:

Theorem 2 ([6], p.71 Corollary 14.3)

The genus $g(G)$ of a simple graph G with n vertices and m edges satisfies the inequality

$g(G)$ greater than or equal to $\lceil \frac{(m-3n)}{6} + 1 \rceil$,

here the symbol $\lceil x \rceil$ represents the smallest integer not less than x .

We applied these two theorems to clarify the “complexity” of network that we found from the DNA expression data associated to COVID-19 in the results’ section.

We used also a method of \log_2FC in order to find DNAs, the expression of which showed stochastically significant increase compared with controls, and calculated \log_2FC by using package “DESeq2” written for the software R.

III. RESULTS

In the methods’ section we described our DNA expression data, which were extracted from NCBI which were originally contributed by [5]. Recall that the dataset was expressed as a matrix of 18292 rows and 60 (=9 infected)*(6 times) + (6 controls) columns.

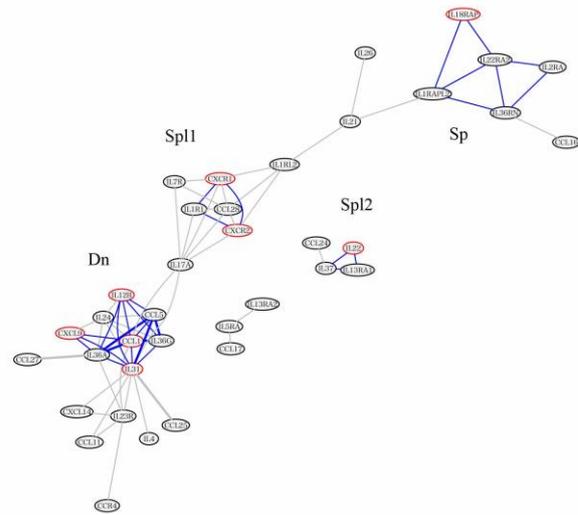
We selected only genes of interleukins and chemokines from the 18292 genes in the whole matrix. From this sub-matrix, we constructed a cytokine correlation network using “GeneNet” library.

We then obtained a cytokine correlation cytokine network including the following 37 cytokines.

IL23R, IL24, IL1R1, IL1RL2, IL18RAP, IL37, IL36G, IL36A, IL36RN, IL5RA, IL21, IL7R, IL4, IL12B, IL17A, IL22RA2, IL2RA, IL26, IL22, IL31, IL1RAPL2, IL13RA2, IL13RA1, CXCR2, CXCR1, CCR4, CXCL9, CCL28, CXCL14, CCL24, CCL27, CCL17, CCL11, CCL1, CCL5, CCL16, CCL25.

The result is shown in Fig.1.

Figure 1. Correlation network of cytokine (interleukins and chemokines)



Note: Four clusters of sub-networks Dn, Sps, Spl1 and Spl2, are named Dense, Sparse, Simple-1, and Simple-2, respectively. They are connected by blue lines (edges) and are positively correlated. The cytokines in red circles show significant cytokines from \log_2FC calculation. The cytokines with grey edges are negatively correlated.

In Fig.1, we identified the following four clusters in this correlation network (neglecting nodes with one edge):

- (1) Spl1: IL1R1, CXCR2, CXCR1 (Simple-1)
- (2) Sp: IL18RAP, IL1RAPL2, IL22RA2, IL36RN, IL2RA (Sparse)
- (3) Spl2: IL37, IL22, IL13RA1 (Simple-2)
- (4) Dn: IL36G, CCL5, IL12B, IL31, CCL1, IL36A, CXCL9 (Dense)

Evidently, the cluster (4) Dn, appears dense and complex, whereas the other clusters ((1)~(3)) appear rather simple and sparse. We then clarified the “complex” feature of this network using the notions of graph theory by applying the theorems of graph theory listed in methods’ section.

We first applied Theorem 1 (in Method section) to our dense network Dn. Since m and n are the number of edges and nodes respectively, and for case of Dn, we have $m=17$, $n=7$.

Therefore,

$$m > 3*n - 6.$$

From Theorem 1 (in the methods’ section), we have the following assertion for planarity of Dn:

- (1) The network Dn can not be described in the plane without crossing edges.

Secondly, we applied theorem 2 to Dn. Since

$$(m-3*n)/6 + 1 = 1/3,$$

the smallest integer that cannot be less than 0.333.. is equal to 1. Since

$g(G)$ is greater than or equal to $\lceil \frac{(m-3*n)}{6} + 1 \rceil$,

we have the following assertion for genus of Dn:

- (2) The genus of Dn, $g(Dn)$, is greater than or equal to 1.

In Summary, COVID-19 patients have a dense network with has genus >1 that cannot be described in the plane without crossing edges.

The dataset GSE156701 contained also six controls, i.e., six additional rhesus macaques who were not infected. Comparing cases with these controls, we calculated log₂FC for these clusters of sub-networks.

We found the following stochastically significant genes in each cluster:

Dn IL12B, IL31, CCL1, CXCL9
Sp IL18RAP
Spl1 CXCR2, CXCR1
Spl1 IL22.

Notably, four significant cytokines were found in the dense sub-network Dn which were originally composed of seven cytokines.

IV. DISCUSSION

Using the dataset GSE156701 in NCBI database, we constructed a correlation network for cytokines (including interleukin and chemokine). In this network, we found four clusters. In particular, a complex and dense cluster, “Dn (Dense)”, was notable from a mathematical point of view. From a graph theory, “Dn” showed a complexity, which was clarified from the notions “planarity” and “genus”. We also calculated stochastically significant up-regulated cytokines in the four clusters. For Dn, there were four significantly up-regulated cytokines among seven cytokines in Dn.

We compared our results of networks with those of other research works. For example, Y. Liu [7] found consistently elevated levels of IL-6, IL-8 and IL-5 in COVID-19 patients who were deceased after mild/moderate or severe disease. In our case, calculating log₂FC showed very high levels of increase of IL6 as shown below:

Symbol	log ₂ fc	pvalues	order
IL6	6.7679276	1.046467e-06	21
IL8	3.0814075	8.715799e-07	473

In [8], investigators performed extensive cellular (leukocytes and T-lymphocyte subsets) and serological immune profiling (cytokines, soluble cell surface molecules, and SARS-CoV-2 antibodies) at hospital admission and every 3-4 days during hospitalization. They found that IL-6, CCL2 and CXCL10 levels were higher at baseline in the infected group. We observed similar results with IL-6 and CCL2 but not CXCL10.

Because Psoriasis is common immune-mediated disease in European populations, its cytokine networks have been evaluated [9]. These results showed that IL-17A, and IL-23 are upregulated in the mouse skin; similar to our results regarding IL17A.

Most of the research works on cytokine networks including works by [7], [8], [9] document significantly increased cytokines. However, they have not studied the features of networks. If we can clearly define the complexity of networks, we will be able to accurately identify them.

In this paper, we defined the complexity of the correlation network and described its’ complexity using a mathematical method. We defined the planarity and genus of networks and

found that the Dn (dense sub-network of cytokines) was really “complex”.

To clarify the role of this complex network Dn (Dense network), we further investigated the roles of cytokines included in Dn, which included IL36G, CCL5, IL12B IL31, CCL1, IL36A, CXCL9.

Cytokines such as CXCL9, IL-12 and CCL5 play a crucial role in basic inflammatory processes. CCL1 is involved in inflammatory processes through leukocyte recruitment and plays an important role in angiogenesis and other viral and tumoral processes. CXCL9 leads to Th1 polarization through CXCR3 for immune cell differentiation. IL-12 is involved in the differentiation of naive T cells into Th1 cells. CCL5 activates recruiting leukocytes into inflammatory sites, and induces the proliferation and activation of NK (natural-killer) cells.

Cytokines, IL31 and IL-36(A, G) are associated autoimmune diseases. IL31 induces cytokine and chemokine production from human bronchial epithelial cells, and plays an important role in allergic skin inflammation and atopic dermatitis [10]. It was reported that IL-36 expression and function might correlate with the pathogenesis of Systemic Lupus Erythematosus (SLE) [11], where SLE is an autoimmune disease characterized by damage to multiple systems and organs such as skin.

Therefore Dense sub-network “Dn” is a collection of cytokines playing crucial roles in both basic inflammatory systems and autoimmune diseases.

Our notion of networks’ complexity is a useful indication of the complexity of networks and can be used as guideline to compare correlation networks associated with cytokine storms.

V. CONCLUSION

We defined the notion “complexity” of a correlation network, and found a “complex” sub-network in the cytokine correlation network in those infected to COVID-19. The “complexity” was clarified using the notion of “planarity” and “genus” from a mathematical point of view. This “complex” network contains cytokines which play crucial roles in both basic inflammatory systems and autoimmune diseases.

REFERENCES

- [1] Ogura H, Murakami M, Okuyama Y, Tsuruoka M, Kitabayashi C, et al. (2008) Interleukin-17 promotes autoimmunity by triggering a positive- feedback loop via Interleukin-6 induction. *Immunity* 29: 628–636.
- [2] Morimoto.H. (2017) A probabilistic approach for links between rheumatic diseases and weather. *Internal Medecine and Care*, Vol.1, p1-4. <http://dx.doi.org/10.15761/IMC.1000104>
- [3] Morimoto H (2018) Interleukins as a Common Indicator for the Effects of Both Weather and Obesity on Human Health. *Int J Biomed Data Mining*, vol.7, 1, p1-6. <http://dx.doi.org/10.4172/2090-4924.1000131>
- [4] Espinosa,G. Araujo,O. Amaro,S. Bodro,M. Moreno, P.J. Moreno,R. Ugarte,A. Cervera,R. (2021) COVID-19 and Behçet’s disease: clinical case series. *Ann Rheum Dis*, Vol 80, No 3. <http://dx.doi.org/10.1136/annrheumdis-2020-217778>
- [5] Aid, M, Busman-Sahay, K, Vidal ,SJ, Maliga, Z et al. (2020) Vascular Disease and Thrombosis in SARS-CoV-2-Infected Rhesus Macaques. *Cell* 2020 Oct 9. PMID: 33065030. <https://doi.org/10.1016/j.cell.2020.10.005>

- [6] Wilson, R.J. (1996) introduction to Graph Theory, Pearson Education Limited, England 1996. ISBN 0-582-24993-7
- [7] Liu, Y. Chen, D. Hou, J. et al. (2021) An inter-correlated cytokine network identified at the center of cytokine storm predicted COVID-19 prognosis. *Cytokine* Volume 138, February, 155365. <https://doi.org/10.1016/j.cyto.2020.155365>
- [8] Schrijver, B. Assmann, L.C. van Gammeren, A.J. et al. (2020) Extensive longitudinal immune profiling reveals sustained innate immune activation in COVID-19 patients with unfavorable outcome. *European Cytokine Network*, Volume 31, numéro 4. <https://doi.org/10.1684/ecn.2020.0456>
- [9] Blumberg, B. Dinh, H. Dean, C. Jr. Trueblood, E.S. Bailey, K. Shows, D. Bhagavathula, N. Aslam, M.N. Varani, J. Towne, J.E. and Sims, J.E. (2010) IL-1RL2 and Its Ligands Contribute to the Cytokine Network in Psoriasis, *J Immunol*, 185 (7) 4354-4362. <https://doi.org/10.4049/jimmunol.1000313>
- [10] Wai, K. Wong, C.K. Li, M.L.Y. Li, P.K. Cheung, P. F. Y. Lam, C. W. K. (2007) Interleukin-31 induces cytokine and chemokine production from human bronchial epithelial cells through activation of mitogen-activated protein kinase signalling pathways: implications for the allergic response. *Immunology*, 122(4): 532–541. <https://doi.org/10.1111/j.1365-2567.2007.02668.x>
- [11] [Zhi] Yuan, Z.C. Xu, W.D. Liu, X.Y. Liu, X.Y. Huang, A.F. and Su, L.C. (2019)

Hiroshi Morimoto Professor emeritus of Nagoya University. He graduated from Mathematical Institute of Nagoya University. He was first employed by Nagoya University as a pure Mathematician. Then he explored many fields including biology, global climate change, human health and DNA, using methods of data mining.