Journal of Communicable Diseases
Volume 53, Issue 4 - 2021, Pg. No. 84-96
Peer Reviewed & Open Access Journal

Research Article

Prediction of Putative Protein Interactions between Zika Virus and Its Hosts Using Computational Techniques

Surendra Kumar Sagar[1], Manoj Kumar[2], Prithvi Singh[3], Shweta Sankhwar[4], Ravins Dohare[5]

[1]Department of Zoology, Swami Shraddhanand College (University of Delhi), Delhi, India.
[2]Centre for Economic Studies and Planning, Jawaharlal Nehru University, New Delhi, India.
[3,5]Centre for Interdisciplinary Research in Basic Sciences, Jamia Millia Islamia, New Delhi, India.
[4]Department of Computer Science, Maitreyi College (University of Delhi), New Delhi, India.
DOI: https://doi.org/10.24321/0019.5138.202178

I N F O

Corresponding Author:
Ravins Dohare, Centre for Interdisciplinary Research in Basic Sciences, Jamia Millia Islamia, New Delhi, India.
E-mail Id:
ravinsdohare@gmail.com
Orcid Id:
https://orcid.org/0000-0002-0602-3719
How to cite this article:
Sagar SK, Kumar M, Singh P, Sankhwar S, Dohare R. Prediction of Putative Protein Interactions between Zika Virus and Its Hosts Using Computational Techniques. J Commun Dis. 2021;53(4):84-96.

Date of Submission: 2021-09-22
Date of Acceptance: 2021-12-25

A B S T R A C T

Generally, protein-interaction prediction between the proteins of any host and the virus's proteins is quite crucial for the infection and the pathogenesis of the virus, which makes it striking target for the development of the therapeutics. The major aim of the present study was to utilize the structure-based approach to predict proteins responsible for the propagation of the ZIKV infection in the host machinery. A computational structure-based approach has been applied for the prediction of interacting proteins. From this methodology, we come up with the interactions which are very crucial for the virus infection propagation into the host's cellular system. As there is a notable relationship between the Zika virus and the neurodevelopment abnormalities, still there is no specific system underlying which impaired neurological development has not been determined. We encounter some of the interactions which are predicted from the methodology adopted in our work, through which we can say that these are some interactions which cause neuron disorders as the major problem associated with this viral infection.

Keywords: ZIKV, PPI; hZIKV-similar, Protein Structure, Protein-Interaction Prediction

Introduction

Zika virus (ZIKV) is a vector borne disease which is of the family *Flaviviridae*, genus *Flavivirus*. It was first reported in 2007 on Yap Island.[1] The epidemic of ZIKV was reported in October 2013 in French Polynesia[2] where a large population estimated around 28,000 (11% of whole population) suffered illness and sought medical care.[3] *Aedes* mosquito is a vector, which plays a crucial role for the transmission of ZIKV. It has been reported that the virus infection also transmits through sexual contacts with the infected person, as well as from mother to her babies. *Aedes* mosquitoes are also having a major role during the transmission of dengue fever and yellow fever. Mild headaches, joint pains, fever, malaise, and conjunctivitis and maculopapular rash are common symptoms of ZIKV infection. So far, it is a mild disease and only 20% of patients may develop symptoms. However, for pregnant women, especially those that become infected in the first trimester of pregnancy, Zika virus infection

Copyright (c) 2021: Author(s). Published by Advanced Research Publications

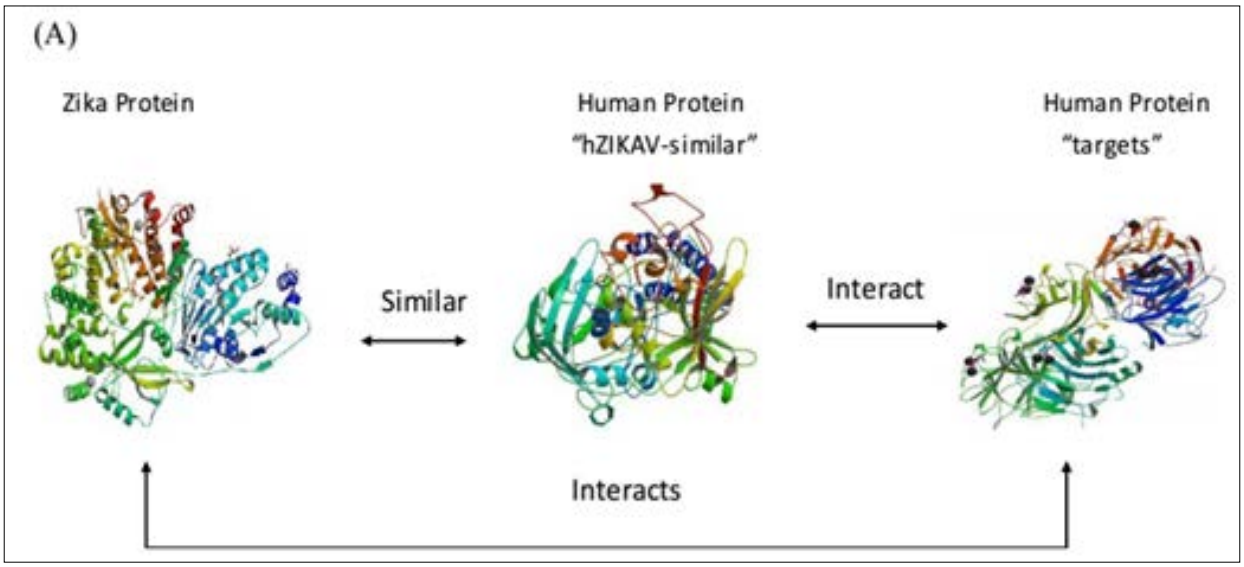
can damage brain and can cause microcephaly. Zika virus accomplishes its invasion by a gene which encodes only 10 proteins. For the successful attack to its each hosts, ZIKV is capable to alter its host at its very molecular level. The alteration can be done by specific interactions between particular proteins which provide them a way to bow with the cellular pathway of the existing system to make them alive in the lifecycle of the virus. Owing to the complexities of the virus-host dynamics and also because of the deficiency of of some other model, it is pretty difficult to understand the interaction of pathogen with its host's cellular system.

Thus, various computational methods give us an efficient tool to study the interactions involved in the pathogen-host system, to be specific, for the predicted interactions of proteins between the viruses and their hosts using various computational methods. This could give us a hint for the identification of the atypical interaction targets to enhance our knowledge of various experimental tasks and also finding particular therapeutics to curb their viral infections.[4] Finding Protein-Protein Interactions (PPI) within a species is common, while predicting PPI among different species is rare; therefore, it is also important to predict pathogen and their host's PPI. Recently, several computational method have been used for host-pathogen interactions, such as, between human and *P. falciparum*, which is mainly based on one of the computational techniques, i.e., orthologous-based approach.[5,6] Tastan *et al.* used data mining technique for the prediction of virus and its host's interactions on the basis of knowledge of known interactions. Also, Evans *et al.* considered the basis for prediction of protein interactions, conserved sequence motifs in the host, and different pathogen *i.e.* human and HIV.[7]

An approach, which is based on structures, has been widely used for the prediction of PPI.[8,9] In this approach, interactions can be predicted between the set of proteins when other pair of protein structures and their interaction are known. This approach has been applied to many non-viral pathogens, as well as HIV-human interactions and dengue-human interactions. But, this approach is not always applied for many problems, which involve pathogen and its host's interactions. We are not aware of any study which deals with the prediction of interactions of the protein between ZIKA and human and its other vector *Aedes*.

Some of the attempts have been made to predict host-pathogen PPI using different computational methods. Information gained from the structure of the proteins is also used in the predictions of PPI.[8,9] If we think of a pair of proteins, which have structures that are similar to a known interacting pair of proteins, it can be believed that the former are likely to interact in a way that is structurally similar to that of the latter. Structure-based approach, which is based on the prediction by comparative modelling, was also widely used for host-pathogen interactions.

In this study, we developed network of the predicted interactions between ZIKA proteins and its both host proteins (human proteins and insect host's protein) on the basis of structural similarity. First, we identify structural similarity between host and pathogen proteins using DaliLite web server, which compared 3D structures of the proteins. We called host proteins with high-similarity to ZIKV proteins as "hZIKV-similar" which can be understood using Figure 1(A) and as "dZIKV-similar" which can be understood using Figure 1B. Thereafter, we determined known protein interactions within the species for these hZIKV-similar and dZIKV-similar proteins, which we refer to as "target" proteins. This could be understood using Figure 1(A-B). In this way, target proteins were predicted in both human host and *Aedes aegypti* (insect host). These are further validated experimentally and targeted protein must be useful for pharmaceutical intervention against ZIKV disease.



ISSN: 0019-5138
DOI: https://doi.org/10.24321/0019.5138.202178

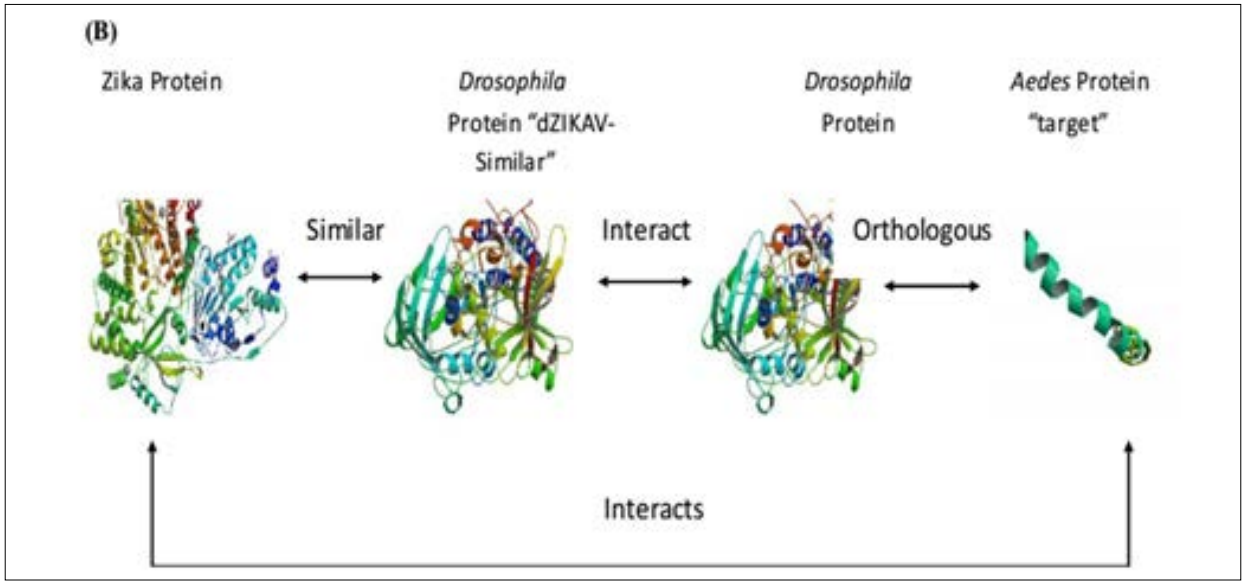


Figure 1.Diagrammatical representation of prediction modelling: (A) ZIKA virus proteins showing structural similarity to one or more human proteins. Interactions for these hZIKV-similar proteins with other human proteins are then identified. Following appropriate filtering, this methodology predicts the existence of a physical interaction between the proteins of ZIKA virus and the human proteins, which are termed as "targets". (B) Prediction of Interaction for the ZIKA virus and one of its important host is also determined in the same way as (A), except we have done some add-on task of determining orthologous of the *D. melanogaster* target proteins in the actual host i.e. *A. aegypti*.

Materials and Methods

Data Sources

Structures of ZIKV proteins had been derived from PDB, and those proteins whose structures were not present in the PDB were modelled using I-TASSAR.[10,11] The I-TASSAR web server made use of the sequence of the proteins whose structure had to be determined i.e. NS2B, NS4A, NS4B, GLUCA, and CAPSID with the default setting, with neither constraints nor exclusion/selection of any templates. When the structures of all ZIKV protein's structure had been determined, each ZIKV protein's structure had to be run over the DaliLite v.3 web-server, with the purpose of finding the protein structure which were structurally similar to the ZIKV proteins.[12,13] Known Human PPI were taken from HPRD Release 7, and PPI of *D. melanogaster* were obtained from DroIDv5.0.[14,15] The *D. melanogaster* protein's orthologs of *A. aegypti* were taken from Fly Base v. FB2009 10 database. Every database and the different sources of literature have their own scheme of identifiers. The codes from PDB were taken from DaliLite v.3 web-server which can be used to map in their analogous taxonomy and the accessions number from the Uniprot database. It becomes easier to integrate data from diverse databases. Mapping of other identifiers was done with the use of Uniprot ID or Gene ID conversion using DAVID. The network of PPI were drawn using iGraph package of R and various images of protein structures were made using MacPyMol.

Finding Proteins with Similar Structure between ZIKV and Its Host

Using DaliLite v.3 web-server, we determined the proteins which were structurally similar. The DaliLite v.3 web-server did the comparison of 3D structures of two PDB entries through the alpha carbon alignment process, which made use of algorithm based on the distance matrix and, on the structural similarity approach, it also gave similarity score. For this study, each protein structure whether known and predicted, of ZIKV, ran through the DaliLite v.3 web-server, that finds structurally similar proteins against every entry from PDB, with a *z-score* above . Default settings of a score cut off 40 bits and sequence overlap cut off 50% were used. From the result, we only took structures which were from the *H. sapiens* and *D. melanogaster* and refer these proteins of Human as "hZIKV-similar" proteins and that of fly proteins as "dZIKV-similar" proteins.

Interaction Prediction

For the prediction of different proteins among humans, which might involve in the interaction with different proteins of ZIKV, we particularly relied on those human target proteins during cellular processes which take part in the interaction with human proteins that are similar to ZIKA virus proteins, say hZIKV-similar. For the above purpose, we need to identify interaction between hZIKV-similar proteins with human target proteins, with the help of data, which can be downloaded from the database, namely, Human

ISSN: 0019-5138
DOI: https://doi.org/10.24321/0019.5138.202178

Protein Reference Database (HPRD); it provides literature that curates interaction between pair of different proteins among human.[12] At this stage, we can assume that all the targeted proteins that are recognized to take part in the interaction with hZIKV-similar protein might also be take part in the interaction with that of corresponding ZIKV proteins. We can assume this theme for each and every hZIKV-similar protein.

Like the above basic idea, a similar protocol was also implemented for the prediction of protein interactions involving ZIKV proteins and different proteins of *A. aegypti*, but with an extra effort of recognizing the orthologous among *D. melanogaster* and proteins of *A. aegypti* host. Determined interactions between the dZIKV-similar proteins and other *D. melanogaster* proteins can be taken from DroID.[14,15] After which, using Fly Base, the orthologs proteins of *D. melanogaster* were found for *A. aegypti*.[15] Then, we were able to make prediction that target proteins of *A. Aegypti* interact with the various ZIKV proteins.

GO Term Enrichment

The Gene Ontology (GO) gives us an arrangement of terms to express and annotate genes and gene products in any organism. GO term enrichment was performed using the DAVID Functional Annotation Chart tool.[16] It is arranged as tree-structured, as the distance from the root increases the terms becoming more exact. That's why, to keep away from very general and the terms which do not give any useful information, we have used the level 4 for GO terms. The Bonferroni procedure was used for the correction of p-value for numerous testing and it was transformed into .

Validation of Prediction

Predicted interactions must be validated as there might be some redundancy since numerous PDB structures are there in DaliLite v.3 to embody the same protein. In many cases, for the same viral protein, several PDB structures were found that are analogous to numerous PDB structures for a ZIKV-similar protein, which leads us to the identical protein interaction prediction. Hence, the interaction predictions have to be counted as distinctive pairs. Supports for the interaction which are predicted were taken from the various literatures. As very less number of interactions is known between ZIKA virus and human proteins, we have to make a check whether any of them was predicted with the help of our methodology, but we came up with no such literature describing this work by taking ZIKA virus.

Results and Discussion

Identification of ZIKA-Similar Host Proteins

Initially, 3D structures of the ZIKV proteins were obtained with the help of two sources. Structures, which are determined experimentally, were extracted from the Protein Data Bank (PDB); and those, whose structures are not experimentally determined in the PDB, were modelled using the web-server I-TASSAR, to predict structure of NS1, NS2A, NS4A, and NS4B proteins.[10,11] Then, we investigated required interactions for each ZIKV protein. After this, using the DaliLite v.3 web-server for the determination of structurally similar proteins of the hosts, comparisons with the other protein structures in PDB was done against ZIKV proteins using DaliLite v.3 web-server.[11] But, we will focus on those, which are significantly structural matches with the ZIKV host's proteins. 45 human proteins were similar to ZIKV protein, which we called hZIKV-similar proteins. Yet, we encountered no similarity between ZIKV proteins and *A. aegypti* proteins. As of now, structures of 10 proteins of *A. aegypti* are there in PDB. Hence, we determined similarity between ZIKV proteins and the fly, *Drosophila melanogaster*, which we come up with 64 proteins which are similar to the ZIKV proteins, that we called dZIKV-similar proteins. Number count of the similar human protein corresponding to the Zika virus proteins and that of similar *D. melanogaster* proteins corresponding to the Zika virus proteins are shown in Table 1.

Table 1.Number of hZIKV-similar and dZIKV-similar proteins

S. No.	Name of ZIKA Virus	Structure Type	Number of hZIKV-Similar Proteins	Number of dZIKV-Similar Proteins
1.	5tfr	Known (available in PDB)	1	-
2.	5jhm	Known (available in PDB)	7	1
3.	5k6k	Known (available in PDB)	5	4
4.	5jmt	Known (available in PDB)	5	2
5.	Ns2a	Known (available in PDB)	6	8
6.	Ns2b	Predicted from I-Tassar	-	-
7.	Ns4a	Predicted from I-Tassar	5	8
8.	Ns4b	Predicted from I-Tassar	9	36
9.	Gluca	Predicted from I-Tassar	3	1
10.	Capsid	Predicted from I-Tassar	4	4

ISSN: 0019-5138
DOI: https://doi.org/10.24321/0019.5138.202178

Protein-Interactions Prediction

Mainly, the central problem in the study of host-pathogen protein interactions is the deficiency of appropriate data. Presently, we know only of a few protein interactions among the ZIKV and its hosts. After the determination of the potential host protein interactions, which have structure similar to the ZIKV proteins, we investigated for the already determined interactions for every ZIKV-similar proteins participation in its host i.e. protein-protein interaction within the host. For the protein set of hZIKV-similar, we got known interactions between human proteins through the web server, HPRD; it contains around 37,000 interactions which are established from the literature. After this, we considered ZIKA virus proteins which might interact with interaction partners of their corresponding hZIKA-similar proteins, under the assumption that proteins with highly-similar structures are likely to be involved in similar protein interactions.[12] The count of 978 predicted potential interactions between pathogen-host are found, involving 802 distinct proteins of human and also (list can be seen in Supplementary Table 1), 110 host-pathogen interactions in the case of the insect host. The summary of these predicted proteins can be seen in Table 2. Following (Figures 2 and 3) are the predicted PPI networks, which are showing interactions between the different Zika virus proteins and the target human proteins (Figure 2) and interactions between Zika virus proteins and its insect host's, i.e. A. aegypti, target proteins (Figure 3). From the network between Zika and human target proteins, we can see that envelope protein of the Zika virus interacts with the highest number of target proteins; for any pathogen virus invasion into the host machinery, this protein plays the major role as this is the starting point from where virus starts to incubate into the host cell pathways. After this, we can view the highest number of interaction involving NS4B and NS4A of Zika virus' protein. These can be the proteins which lead to the huge impairment of the process of neurogenesis of human neuronal cells and the up-regulation of autophagy for viral replication. Autophagy comes into the sphere of those catabolic processes that help in the development of the immune response during evolution. However, hosts evolved autophagy for the maintenance of the cellular homeostasis and also it limits the infection created by the pathogen.

Table 2.Interaction prediction summary corresponding to every ZIKV proteins

S. No.	ZIKV Proteins (PDB IDs)	hZIKA-Similar Proteins	Number of Human Targeted Proteins	dZIKV-Similar Proteins	Number of *D. melanogaster* Targeted Proteins
1.	5tfr	1	1	-	-
2.	5jhm	7	272	1	-
3.	5k6k	5	56	4	2
4.	5jmt	5	86	2	-
5.	Ns2a	6	135	8	4
6.	Ns2b	-	-	-	-
7.	Ns4a	5	130	8	20
8.	Ns4b	9	249	36	37
9.	Gluca	3	19	1	-
10.	Capsid	4	30	4	47
Total		45	978	64	110

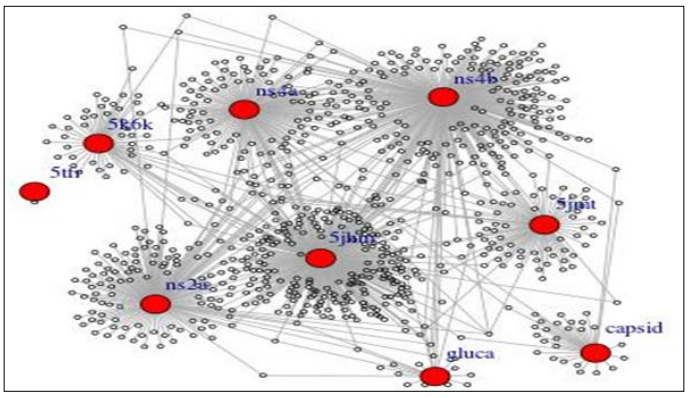


Figure 2.Protein-Protein Interaction network between human targets and Zika virus proteins

ISSN: 0019-5138
DOI: https://doi.org/10.24321/0019.5138.202178

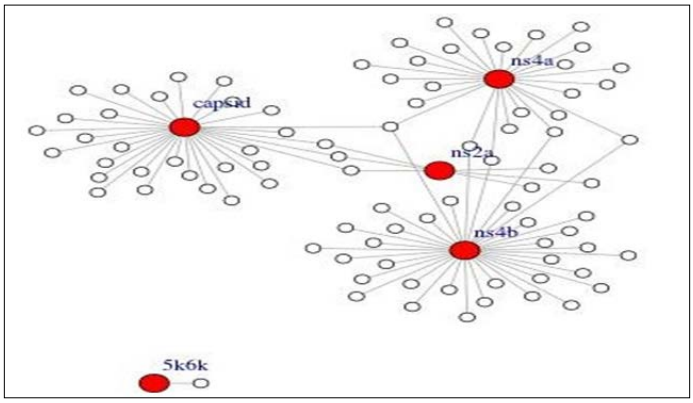


Figure 3.Protein-Protein Interaction network between A. aegypti targets and Zika virus proteins

Assessment of Predictions

To evaluate whether the predicted interactions are accurately predicted or whether there is any similarity between the known predicted interactions and our interactions, to this point, we have to assess our predicted interactions, so that we can determine which proteins are necessary for the viral invasion into the human and its insect hosts of the Zika virus. Assessment can be done into two folds, first, on the basis of the GO term enrichment and secondly, sub cellular co-localization and also additionally, it can also done by literature filtering. Hence, we validated through two ways discussed in following sections.

GO Term Enrichment

Because of lack of the known pathogen-host protein interactions from which we can match up to and to assess our predicted interactions, the count of different GO terms enriched in the similar proteins of ZIKA virus as well as its targeted proteins were determined. From this assessment in Figure 4, we can see that many most significant enriched terms are from the processes or functions, which are notorious for important for ZIKA virus infection (the details of GO term (BP) can be seen in Supplementary Table 2).

As till now we do not know which protein interactions are responsible for the Zika virus infection propagation, to this end, we only assume that our result is consistent with the study of changed expression of proteins during the infection of ZIKA virus, among which many of identified proteins, which have functions belong to Gene Ontology terms, positive regulation of RNA process, positive regulation of macromolecular metabolic process, reg. of gene expression which are enriched in our predictions.

Blue blocks in Figure 4, which represents human target proteins, terms involving such as cell death involved in cell development, positive regulation and negative regulation of the cellular process are more frequent, which is consistent with the study that Zika virus increases the cell death and deregulated the cell-cycle progression, which results into an attenuated hNPC (Human Neural Progenitor Cells) growth.[19]

Human Zika-similar proteins enriched to the terms like cell cycle process and positive and negative regulation are more frequent used which is the process involved in the virus's course of infection. As we get no significant terms in the enriched terms for biological process, neither for ZIKV similar in *A. aegypti,* nor forzikv targets in *A. aegypti*.

Most of the enriched term for Molecular functions for ZIKV-*A.aegypti* among our predicted interactions generally involve in the GTPase activator and regulator activity. This can be noted that in the terms which are enriched are also included in DNA binding (Figure 5 and the details of GO term (MF) can be seen in Supplementary Table 3). Terms such as P53 binding are also common within our prediction as we can also observe that gene from predicted interactions network.

TP53 is involved in the interaction with envelope protein of the Zika virus as this result is consistent with the work of Teng *et al.*, who came up with the result that P53 is the hub of the genetic regulatory network for ZIKV-related and proteins, which is found to be associated with microcephaly and, also, P53 cell death pathways play an important role in the infection associated with ZIKV and microcephaly.[20] And also, in a study, Zang *et al.*, established that Zika virus infected strain from Asia triggered much greater inborn response from immune system than the strain from African sample, which includes a greater number of the gene *TP53 expression*. Additionally, it inhibited to a greater extent, the strain from Asia's pathogenicity–harm causing ability–compared to the strain taken from that of Africa, when treated with p53 inhibitors.

Literature Filtering

As from the prediction summary of Zika virus and human protein, we can say that the highest number of interactions is of 5jhm, which is the Envelop (E) Glycoprotein of the Zika virus, which is the same protein as is responsible for the initial interactions between any virus and their host's proteins. And, also the second and third highest number of interactions is of Ns4b and Ns4a, as per the study of Liang

ISSN: 0019-5138
DOI: https://doi.org/10.24321/0019.5138.202178

et al. In their work they came up with the results that these Zika virus proteins deregulate Akt-mTOR signalling in Human Fetal Neural stem cells to inhibit neurogenesis and induce autophagy. In fact, DENV replication requires autophagy to control processing of lipid droplets and triglycerides. Thus, these findings suggest that ZIKV may also require host autophagy pathways to create membrane structures to serve as viral replication sites.[21] By screening the 10 ZIKV-encoding potential proteins, we found that NS4A and NS4B cooperate to induce efficient autophagy by suppressing the Akt-mTOR signalling pathway that is essential for controlling simulation-induced autophagy. Similar to DENV NS4A and NS4B, ZIKV NS4A and NS4B are small hydrophobic proteins with potential transmembrane spanning regions.

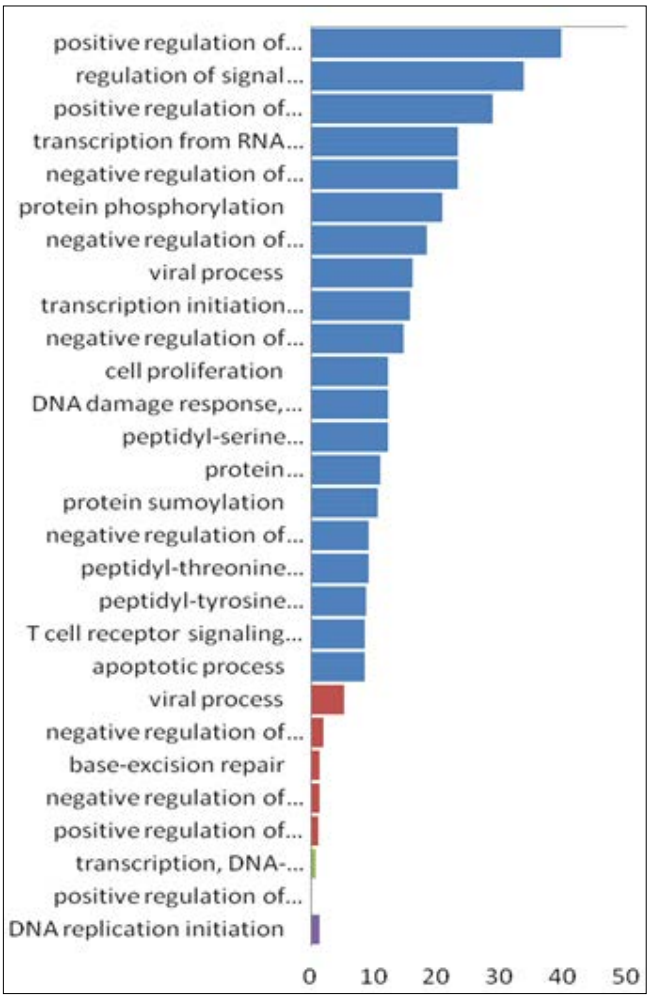


Figure 4.GO enriched terms for Biological process. Blue blocks denote terms for protein targets among humans whereas blocks in red are terms for proteins among hZIKV-similar

Zika virus (ZIKV) interferes with the cellular machinery controlling cell division and alters the expression of hundreds of genes responsible for guiding the information and development of brain cells, according to findings released by Scientific Reports.[22] Also, as we can see in the GO terms involved in the molecular functions, the P53 binding to much larger extent; from the literature, we can also see that many of the terms are also involved in the apoptosis process; so, we can also conclude that p53 binding turns to the apoptosis process, which appears to be one of the active metabolic pathways in the Zika virus infection.[23]

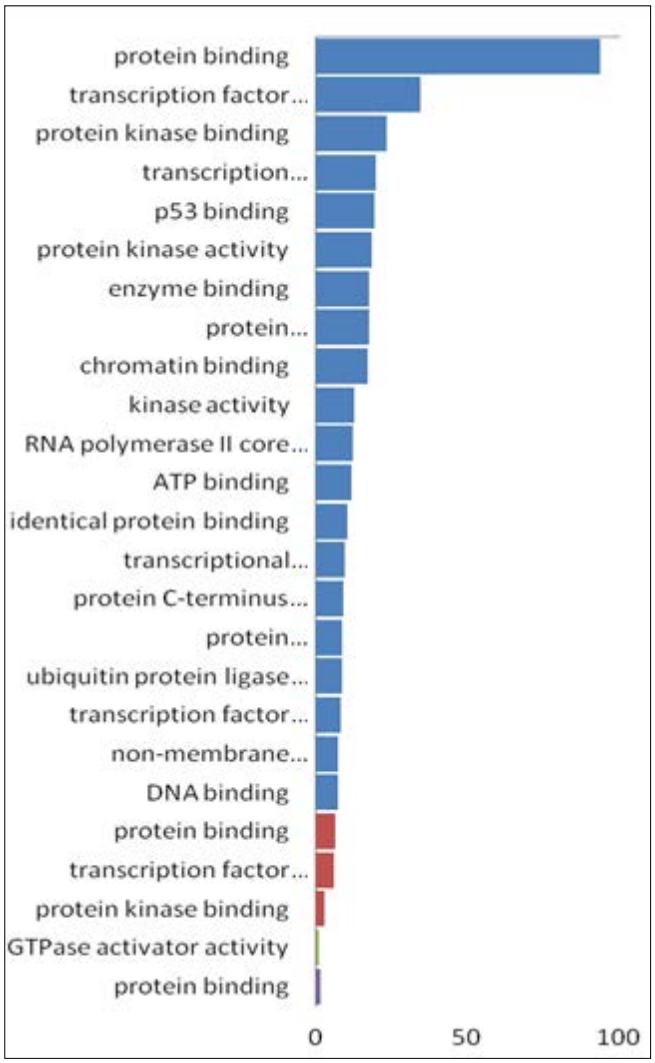


Figure 5.GO enriched terms for Molecular Function. Blue blocks represent terms among human target proteins, red is for terms among hZIKV-similar proteins, green for *A. aegypti* targets, and purple for dZIKV-similar proteins

Conclusions

With the help of above-mentioned protocol, we were able to predict interactions about which proteins of the host organism may impact the most viral infection propagation by interacting with some specific virus proteins, in our case, ZIKA virus. We can also note down that the methodology based on the criteria of similar structures here gave us

ISSN: 0019-5138
DOI: https://doi.org/10.24321/0019.5138.202178

a bigger depiction of the network of protein-protein interaction. The various networks of protein interactions presented here might help us give a set of assumption for future clinical investigations, intervention of likely therapeutic, as well as provide us a better understanding on the life cycle of ZIKA virus and other similar viruses.

Source of Funding

We would like to thank Science and Engineering Research Board (SERB), Govt. of India (with grant no. EEQ/2016/000509). This fund is provided for other projects but it is useful in the development of basic infrastructure of lab where we did this work.

Conflict of Interest: None

References

1. Duffy MR, Chen TH, Hancock WT, Powers AM, Kool JL, Lanciotti RS, Pretrick M, Marfel M, Holzbauer S, Dubray C, Guillaumot L, Griggs A, Bel M, Lambert AJ, Laven J, Kosoy O, Panella A, Biggerstaff BJ, Fischer M, Hayes EB. Zika virus outbreak on Yap Island, Federated States of Micronesia. N Engl J Med. 2009 Jun;360(24):2536-43. [PubMed] [Google Scholar]
2. Hancock WT, Marfel M, Bel M. Zika Virus, French Polynesia, South Pacific, 2013. Emerg Infect Dis. 2014 Nov;20(11):1960. [PubMed] [Google Scholar]
3. Musso D, Roche C, Robin E, Nhan T, Teissier A, Cao-Lormeau VM. Potential sexual transmission of Zika virus. Emerg Infect Dis. 2015 Feb;21(2):359-61. [PubMed] [Google Scholar]
4. Doolittle JM, Gomez SM. Mapping protein interactions between dengue virus and its human and insect hosts. PLoS Negl Trop Dis. 2011 Feb;5(2):e954. [PubMed] [Google Scholar]
5. Lee SA, Chan CH, Tsai CH, Lai JM, Wang FS, Kao CY, Huang CY. Ortholog-based protein-protein interaction prediction and its application to inter-species interactions. BMC Bioinformatics. 2008 Dec;9 Suppl 12(Suppl 12):S11. [PubMed] [Google Scholar]
6. Dyer MD, Murali TM, Sobral BW. Computational prediction of host-pathogen protein protein interactions. Bioinformatics. 2007 Jul;23(13):i159-66. [PubMed] [Google Scholar]
7. Evans P, Dampier W, Ungar L, Tozeren A. Prediction of HIV-1 virus-host protein interactions using virus and host sequence motifs. BMC Med Genomics. 2009 May 18;2:27. [PubMed] [Google Scholar]
8. Lu L, Lu H, Skolnick J. MULTIPROSPECTOR: An algorithm for the prediction of protein-protein interactions by multimeric threading. Proteins. 2002 Nov;49(3):350-64. [PubMed] [Google Scholar]
9. Aloy P, Russell RB. InterPreTS: protein interaction prediction through tertiary structure. Bioinformatics. 2003 Jan;19(1):161-2. [PubMed] [Google Scholar]
10. Holm L, Sander C. Protein structure comparison by alignment of distance matrices. J Mol Biol. 1993 Sep;233(1):123-38. [PubMed] [Google Scholar]
11. Holm L, Kääriäinen S, Rosenström P, Schenkel A. Searching protein structure databases with DaliLite v.3. Bioinformatics. 2008 Dec;24(23):2780-1. [PubMed] [Google Scholar]
12. Keshava Prasad TS, Goel R, Kandasamy K, Keerthikumar S, Kumar S, Mathivanan S, Telikicherla D, Raju R, Shafreen B, Venugopal A, Balakrishnan L, Marimuthu A, Banerjee S, Somanathan DS, Sebastian A, Rani S, Ray S, Harrys Kishore CJ, Kanth S, Ahmed M, Kashyap MK, Mohmood R, Ramachandra YL, Krishna V, Rahiman BA, Mohan S, Ranganathan P, Ramabadran S, Chaerkady R, Pandey A. Human Protein Reference Database--2009 update. Nucleic Acids Res. 2009 Jan;37(Database issue):D767-72. [PubMed] [Google Scholar]
13. Yu J, Pacifico S, Liu G, Finley RL Jr. DroID: the Drosophila Interactions Database, a comprehensive resource for annotated gene and protein interactions. BMC Genomics. 2008 Oct 7;9:461. [PubMed] [Google Scholar]
14. Hermjakob H, Montecchi-Palazzi L, Lewington C, Mudali S, Kerrien S, Orchard S, Vingron M, Roechert B, Roepstorff P, Valencia A, Margalit H, Armstrong J, Bairoch A, Cesareni G, Sherman D, Apweiler R. IntAct: an open source molecular interaction database. Nucleic Acids Res. 2004 Jan;32(Database issue):D452-5. [PubMed] [Google Scholar]
15. Crosby MA, Goodman JL, Strelets VB, Zhang P, Gelbart WM; FlyBase Consortium. FlyBase: genomes by the dozen. Nucleic Acids Res. 2007 Jan;35(Database issue):D486-91. [PubMed] [Google Scholar]
16. Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc. 2009;4(1):44-57. [PubMed] [Google Scholar]
17. Berman HM. The protein data bank. Nucleic Acids Res. 2000 Jan;28(1):235-42. [PubMed] [Google Scholar]
18. Zhang Y. I-TASSER server for protein 3D structure prediction. BMC Bioinformatics. 2008 Jan;9:40. [PubMed] [Google Scholar]
19. Garcez PP, Nascimento JM, de Vasconcelos JM, Madeiro da Costa R, Delvecchio R, Trindade P, Loiola EC, Higa LM, Cassoli JS, Vitória G, Sequeira PC, Sochacki J, Aguiar RS, Fuzii HT, de Filippis AM, da Silva Gonçalves Vianez Júnior JL, Tanuri A, Martins-de-Souza D, Rehen SK. Zika virus disrupts molecular fingerprinting of human neurospheres. Sci Rep. 2017 Jan;7:40780. [PubMed] [Google Scholar]

20. Teng Y, Liu S, Guo X, Liu S, Jin Y, He T, Bi D, Zhang P, Lin B, An X, Feng D, Mi Z, Tong Y. An integrative analysis reveals a central role of P53 activation via MDM2 in Zika virus infection induced cell death. Front Cell Infect Microbiol. 2017 Jul;7:327. [PubMed] [Google Scholar]
21. Garcez PP, Nascimento JM, de Vasconcelos JM, da Costa RM, Delvecchio R, Trindade P, Loiola EC, Higa LM, Cassoli JS, Vitória G, SequeiraPC, Sochacki L, Aguiar RS, Fuzii HS, de Filippis AMB, Vianez Júnior JLSG, Tanuri A, Martins-de-Souza D, Rehen SK. Zika virus disrupts molecular fingerprinting of human neurospheres. Sci Rep 7: 40780. DOI: 10.1038/srep40780. [PubMed] [Google Scholar]
22. Heaton NS, Randall G. Dengue virus-induced autophagy regulates lipid metabolism. Cell Host Microbe. 2010 Nov;8(5):422-32. [PubMed] [Google Scholar]
23. Lee J, Shin O. Advances in Zika virus–host cell interaction: current knowledge and future perspectives. Int J Mol Sci. 2019 Mar;20(5):1101. [PubMed] [Google Scholar]

ISSN: 0019-5138
DOI: https://doi.org/10.24321/0019.5138.202178

Supplementary Tables

Table 1.List of predicted pathogen-host interactions

S. No.	ZIKV Proteins	Human Proteins
1.	5jhm	GRB2; PCSK6; INPP1; PCNA; MRE11A; MYD88; IL1RAP; IRAK2; SIGIRR; TICAM2; IL1F10; FCGR3A; IL2RA; LGALS1; PTPRC; CD53; CD5; LCK; CD2; DPP4; CXCL12; CCR5; LAT; IL16; UNC119; SELL; PIP; PI4KA; CD82; HLA-DQA2; SPG21; 01-03-2004; CD4; KCNAB2; COPS3; CD82; CD226; ITGAD; CYTH2; ARIH2; FXYD6; MAD2L1BP; -; WDR33; BTBD2; MPHOSPH6; DLEU1; ACTA1; CR2; CD40LG; SMAD2; SMAD3; SCAMP1; CCNA2; SP3; SP1; UBB; MDM2; CCDC106; FAM173A; THAP8; ZCCHC10; TOP1; TOP2A; TOP2B; KAT5; PLK1; HDAC1; HDAC2; HDAC3; EGR1; CHEK2; NFYA; NFYB; NR4A1; PPP2R2B; FBXO11; EEF2; EPHA3; TEC; BTK; SMARCD1; ESR1; ERCC3; ERCC6; NR3C1; HSPA1A; HSP90AA1; HTT; IFI16; BCL2; YBX1; CDC25C; PPP1CA; HMGB1; HMGB2; NFKBIA; EIF2AK2; MAPK1; PRKCA; NEDD8; TP53; POLA1; UBE2A; BCL2L1; TBP; CREBBP; BRCA2; PTGS2; HNF4A; ELL; SMN1; CSNK1A1; HSPA9; CHUK; KPNA2; CSNK1D; PRKDC; PIN1; MAPK8; YWHAZ; TSG101; ING1; CCNG1; SMARCB1; UBE3A; TEP1; PTEN; MAPK3; ZNF148; TP73; NDN; VRK1; STK11; KLF4; USP7; MNAT1; AURKA; MDM4; KPNB1; UBE2K; MAPK9; MAPK10; PLK3; TADA3; PRMT1; E4F1; CHEK1; ATF3; DAXX; SMARCA4; TP63; HIF1A; PRKRA; CDC14A; CDC14B; MTA1; MED17; PTTG1; MED1; COPS2; BLM; WRN; EFEMP2; RRM2B; BRF1; TAF1A; TAF1B; TAF1C; STK4; GSK3B; WWOX; TP53BP1; YWHAG; HIPK2; WT1; SHISA5; CUL9; ATM; HIPK1; GNL3; TAF9B; HSP90AB1; KAT2B; COPS5; PIAS4; COPS3; THRB; RCHY1; ATR; SIRT1; PRIM1; UBE2I; CCNH; CHD3; PPP1R13L; TP53INP1; SIN3A; ING4; CEBPZ; COPS4; ARID3A; ZHX1; TOPORS; TP53BP2; GPS2; TP53RK; ANKRD2; CABLES2; GPS1; PLAGL1; PIAS1; SETD7; MDC1; ING5; CABLES1; BANP; COPS6; COPS7A; COPS8; CAPN1; CSNK2A1; CSNK2B; BARD1; DDX5; DHCR24; CDK2; RFWD2; PPP2CA; MTA2; TFAP2C; TAF9; MSX1; S100A2; KLF6; ANXA3; CDC42; GSTM4; BCR; PNP; PPA1; SNRPN; TK1; NQO1; HABP4; MNDA; CDK5; CDK9; EP300; PTK2; S100A8; RAB4A; TRIAP1; MAPKAPK5; NMT1; NMT2; SYVN1; NPM1; NAP1L1; STRA13; SUMO1; PPP1CC; SMYD2; ZNHIT1; HUWE1; PBK; BAK1; MIF; PSME3; DHFR; HNRNPUL1; PPM1D
2.	5jmt	UPF3B; TUBB3; BAT1; YWHAG; BAT1; DNM2; TDG; RPS15A; HIPK2; PIAS2; EXOSC9; THOC4; RBM39; ZHX1; DDX39; HNRPLL; SARNP; SFRS12; POU3F2; THRA; EP300; NR2F1; HSF1; PSMC2; GTF2B; FOXF2; TBPL1; TAF9; TCEA2; TRAM2; ESR1; PSMC5; ESRRA; DHX9; CREBBP; YWHAE; VDR; CTCF; GTF2B; XRCC5; CCNK; TCERG1; CCNL2; CDK12; PPIG; PCIF1; TAF10; TRAM2; XAB2; CDK9; POLR2H; MED21; TRAK1; CSH2; MCM3; ERCC5; ITCH; KLK2; SPIB; BRF2; PAX6; KDM5A; GTF2B; FOXF2; EDF1; TAF1L; CREG1; TAF13; TAF5; GTF3C3; TAF10; TAF11; TRAM2; ABT1; PAX5; POU3F2; RB1; MYC; HNRNPK; TAF12; SNAPC1; SPI1; CAND2; SP1; TEAD1; YWHAE
3.	5k6k	RPA2; NUP62; ITGB2; JAM3; TLN1; KNG1; ITGB2; ICAM4; ITGAM; RDX; HP; ICAM2; ICAM3; FUT4; ICAM1; FCER2; ITGAX; PRKCA; PRKCB; PRKCD; ITGB2; SYK; PTK2; PTK2B; ESM1; ICAM5; ILK; FHL2; DOK1; DAB1; VNN2; NUMB; RANBP9; PRKCH; COPS3; CD82; CD226; ITGAD; CYTH2; PRDM1; SIX3; HES6; HHEX; GSTM4; PEX2; RPA2; TK1; SAT1; HSPE1; TLE1; MSX1; PFN1; POLB; SIX2; SIX6; SIX1
4.	5tfr	HMGA1
5.	capsid	CNOT2; CTNNB1; MAPRE1; PRLR; PPIB; ADRA1D; ADRA1B; ADRB1; ADRB2; HSP90AA1; PFKM; PTPN6; PRKCA; ARG1; SNTA1; CAV3; DYNLL1; PRKACA; PTPRN; CTBP1; DLG4; DLG2; CAMK1; CAMK2A; CAMK4; RASD1; DLGAP2; NOS1AP; ZDHHC23; NOSIP
6.	gluca	HBZ; CPM; HBE1; HBA2; HBB; AHSP; TP53; CASP3; SFRS1; CASP6; SUMO1; SFRS11; NCOA6; TDP1; SFPQ; UBE2I; TOPORS; BTBD1; BTBD2

ISSN: 0019-5138
DOI: https://doi.org/10.24321/0019.5138.202178

7.	Ns2a	MAPK1; XRN1; SNRNP70; CD8A; NFYA; U2AF1; ZRSR2; C20orf158; C1QBP; SFRS11; SIP1; NXF1; SRPK2; PSIP1; SFRS2IP; PRPF4; CLK1; PRPF4B; TCERG1; CDK12; YTHDC1; SRPK1; LUC7L3; SFRS12; TNPO3; U2AF2; SFRS17A; NUP153; SSRP1; NUP50; GADD45G; XPO5; RANGRF; GIT1; IPO7; IPO11; NEK9; NUTF2; RANBP10; NXT1; XPO4; BMPR1B; EP300; ZNF646; PHC1; CHMP4B; ATF2; UBA2; SUMO1; ARNT; TOP1; UBE2I; SP100; ARNTL; FAS; NR3C1; SLC2A1; SLC2A4; HSF2; TCF3; MITF; PRPF40A; CHD3; TOPORS; RAD52; HIST4H4; IPO13; PPM1J; ETV6; PIAS1; DACH1; GMEB1; GMEB2; STRA13; CEBPE; KCNK1; TTRAP; LMNA; ATXN1; HNRNPC; ZBED1; HNRNPK; SATB1; IKZF1; RIPK2; UNC119; RAD54B; HGS; IKZF3; ESR1; HABP4; SOX4; SOX10; NIN; MKL1; HNRNPD; ELK1; UBE2I; HNF4A; NR1D2; NR1H2; RORB; TRAF4; CHMP1A; SKIL; CDH7; -; BOC; CDH24; CDH9; AJAP1; GRIN1; MAGI2; NEURL2; CDON; SOX17; PTPN13; PTPRM; BCL3; SPN; BCL9; FLT1; KDR; TFAP2A; PITX2; FSCN1; RAPGEF2; GLIS2; CDC73; PICK1; NDRG1; MITF; TCF4; IGF2BP1; ESR1
8.	Ns4a	E2F1; CALD1; EGFR; ESR1; TCF3; INSR; MYOD1; MYOG; MYF5; MYF6; RAB3B; PPEF1; ESR2; NEUROD1; ASCL2; KCNQ2; PPEF2; TCF4; GRM7; GRM5; KCNQ3; EDF1; KCNQ5; AKAP9; MINK1; PLCB3; -; KCNQ5; -; GSK3B; CSNK2A1; CDH1; PSEN1; CTNNA1; NDRG1; PPP1CA; ITGAE; ANAPC7; HDAC1; HDAC2; MAD2L2; IL1R1; CD5; NFKBIA; PIK3CA; PIK3R1; PDGFRB; PECAM1; PFN1; PTPN11; TYK2; CD28; CD7; CD4; GTF2H1; HRAS; FES; SRC; ERBB3; TGFBR1; TUBG1; SH3KBP1; WAS; ERAS; CSF2RA; ARAF; SYN1; RRAS2; MST1R; TEK; TIE1; SHB; TYRO3; SLC9A2; SHC1; PTK2; IRS2; INPP4A; GHR; MAPK8; PTK2B; CRKL; RAC1; TUBA1B; IL1RAP; ADAM12; ARHGAP1; PIK3CD; PIK3CB; BCAR1; IRS4; SOCS1; CD2AP; HGS; GAB1; TOM1L1; TRAT1; YWHAG; AGAP2; VAV3; GAB2; PIK3AP1; GAB3; CD3E; -; PPM1A; WBP11; PTPN6; NTRK1; TUB; CBLB; ABL1; AXL; KHDRBS1; ANK3; WASF3; IL7R; JAK3; TSHR; CD40; FGFR1; AKT1; CRK; MET; CBL; SSTR2; PSMB5; ARHGAP17; INSR; LCK
9.	Ns4b	FCN2; ASF1A; CHAF1A; CHAF1B; TCF3; INSR; MYOD1; MYOG; MYF5; MYF6; RAB3B; PPEF1; ESR2; NEUROD1; ASCL2; KCNQ2; PPEF2; TCF4; GRM7; GRM5; KCNQ3; EDF1; KCNQ5; AKAP9; MINK1; PLCB3; -; KCNQ5; RIT1; SSX2IP; NRXN2; SORBS2; RIT2; PVRL4; SMAD2; ESR1; FOXO1; CASP1; PAK1; CDK5R1; HGS; NCK2; PDPK1; ARHGEF7; ABI3; OXSR1; PAK1IP1; CDK11B; RHOJ; PPM1F; MAPK1; ACVR1; BMPR1B; EGFR; GRB2; HIST1H4A; DSCAM; MBP; AKT1; NCF1; YWHAG; CDK5; CRIPAK; DYNLL1; SORBS2; SHC1; SNAPC3; CTBP1; MNAT1; MDM4; RBBP9; RBBP7; RBBP4; E4F1; SNW1; CBX4; SMARCA4; PRKRA; USP4; CCNT2; CCNA1; RBBP8; TRIP11; ENC1; PIK3R3; TRAP1; FRK; SPIB; NDC80; EID1; HSPA8; RNF40; KDM5A; THOC1; MORF4L1; ARID3B; CREG1; PELP1; AATF; HBP1; LIN9; RBAK; CDK9; CDK1; HDAC3; PAX5; CEBPE; CDK2; KDM5B; PIK3R1; NCOA6; BNC2; PURA; BDP1; GTF3C2; BRF1; PPIA; PRKCB; CASP6; CASP8; CASP7; CASP2; CASP10; CASP9; MAPK9; MNDA; CDK4; CCND2; SKP2; INS; PSMD10; DGKZ; L3MBTL; PRMT2; PPP1CB; PPP1CC; RAB34; RAB38; RAN; RASD2; RASL12; RHEBL1; RHOD; RHOJ; RIT1; RPS27A; SMAD2; LEMD3; CREBBP; NFYC; TSC2; FOXG1; NOTCH4; GTF2I; CDK4; TP53; MAPK8; C16orf28; ERBB2IP; -; PIAS4; STUB1; AKT1; PPM1A; USP9X; BTBD2; OS9; PSG9; IRS1; KRT18; MLLT4; NEFL; ATP5A1; BCR; YWHAG; PAK4; ARHGEF7; KIF1B; RASAL2; GSK3A; ERC1; TJP2; SRRM2; RASSF8; KIF5B; THRAP3; PRPF38B; FAM82A2; EML3; TMEM102; DYRK1A; CLK1; PRPF4B; HDAC7; CDK17; SH3BP4; CLASP1; DCAF7; RAI14; YAP1; OSBPL3; PARD3B; MAP3K2; MYCBP2; SAMD4A; ZNF295; SRRM1; ARHGEF2; RAB11FIP2; ANKS1A; CGN; SAMD4B; PRPF40A; LARP1; MICALL1; EEF1G; PI4KB; SF3B3; PPIG; CRTC2; SMCR7L; RALGPS2; SRPK1; BCLAF1; LUC7L3; DOCK7; FARP2; LSR; LUC7L2; MPRIP; TBC1D1; SRGAP2; UCP3; UCP2; CDK11B; YWHAE; MAP3K3; BAD; CFL1; CDKN1B; ING1; CENPJ; SH3BP5L; HSPB6; KCNK3; KCNK9; KCNK15; ABL1; MDM4
10.	Ns2b	-

ISSN: 0019-5138
DOI: https://doi.org/10.24321/0019.5138.202178

Table 2.List of GO Terms (Biological Process) at 4 level vs. -log 10 (P- values)

GO Terms (Biological Process) at 4 level	"-log 10 (p-values)"
Positive regulation of transcription from RNA polymerase II promoter	39.71296221
Regulation of signal transduction by p53 class mediator	33.89936657
Positive regulation of transcription, DNA-templated	28.93274197
Transcription from RNA polymerase II promoter	23.44211252
Negative regulation of transcription from RNA polymerase II promoter	23.28803334
Protein phosphorylation	20.94969658
Negative regulation of apoptotic process	18.3503342
Viral process	16.21992159
Transcription initiation from RNA polymerase II promoter	15.67563073
Negative regulation of transcription, DNA-templated	14.72247016
Cell proliferation	12.35405248
DNA damage response, signal transduction by p53 class mediator resulting in cell cycle arrest	12.35405248
Peptidyl-serine phosphorylation	12.35405248
Protein autophosphorylation	11.12360355
Protein sumoylation	10.62165872
Negative regulation of cell proliferation	9.223075784
Peptidyl-threonine phosphorylation	9.183497417
Peptidyl-tyrosine phosphorylation	8.857122828
T cell receptor signaling pathway	8.505186264
Apoptotic process	8.497867549
Viral process	5.205076366
Negative regulation of transcription from RNA polymerase II promoter	2.011336161
Base-excision repair	1.454283832
Negative regulation of transcription, DNA-templated	1.319907479
Positive regulation of transcription, DNA-templated	1.236345674
Transcription, DNA-templated	0.733004633
Positive regulation of transcription from RNA polymerase II promoter	0.182491774
DNA replication initiation	1.459885542

ISSN: 0019-5138
DOI: https://doi.org/10.24321/0019.5138.202178

Table 3.List of GO Terms (Molecular Function) at 4 level vs. -log 10 (P- values)

GO Term (Molecular Function) at level 4	"-log 10 (p-values)"
Protein binding	93.73734298
Transcription factor binding	34.25158695
Protein kinase binding	23.15163882
Transcription coactivator activity	19.66972884
P53 binding	19.33789876
Protein kinase activity	18.58299495
Enzyme binding	17.54458753
Protein serine/threonine kinase activity	17.46710826
Chromatin binding	17.20112675
Kinase activity	12.53003479
RNA polymerase II core promoter proximal region sequence-specific DNA binding	12.10406606
ATP binding	11.68493675
Identical protein binding	10.19625177
Transcriptional activator activity, RNA polymerase II core promoter proximal region sequence-specific binding	9.430354152
Protein C-terminus binding	8.892845371
Protein heterodimerization activity	8.726441129
Ubiquitin protein ligase binding	8.43535408
Transcription factor activity, sequence-specific DNA binding	8.352835087
Non-membrane spanning protein tyrosine kinase activity	7.185448029
DNA binding	7.066875221
Protein binding	6.41702886
Transcription factor binding	5.946513191
Protein kinase binding	2.636509916
GTPase activator activity	1.232341905
Protein binding	1.558652794

ISSN: 0019-5138
DOI: https://doi.org/10.24321/0019.5138.202178