Identifying novel biomarkers of the pediatric influenza infection by weighted co-expression network analysis

Mohadesheh Zarei Ghobadi 1,2†, Sayed-Hamidreza Mozhgani 3,4†, Mahdieh Farzanehpour 1 and Farida Behzadian 5*

Abstract

Background: Despite the high yearly prevalence of Influenza, the pathogenesis mechanism and involved genes have not been fully known. Finding the patterns and mapping the complex interactions between different genes help us to find the possible biomarkers and treatment targets.

Methods: Herein, weighted gene co-expression network analysis (WGCNA) was employed to construct a co-expression network among genes identified by microarray analysis of the pediatric influenza-infected samples.

Results: Three of the 38 modules were found as the most related modules to influenza infection. At a functional level, we found that the genes in these modules regulate the immune responses, protein targeting, and defense to virus. Moreover, the analysis of differentially expressed genes disclosed 719 DEGs between the normal and infected subjects. The comprehensive investigation of genes in the module involved in immune system and viral defense (yellow module) revealed that SP110, HERC5, SAMD9L, RTP4, C19orf66, HELZ2, EPSTI1, and PHF11 which were also identified as DEGs (except C19orf66) have the potential to be as the biomarkers and also drug targeting for the treatment of pediatric influenza.

Conclusions: The WGCN analysis revealed co-expressed genes which were involved in the innate immune system and defense to virus. The differentially expressed genes in the identified modules can be considered for designing drug targets. Moreover, modules can help to find pathogenesis routes in the future.

Keywords: Influenza, Co-expression network, Biomarker, Pathogenesis

Background

Influenza virus is one of the most incident infectious agent, which is mainly classified into three types of A, B, and C. The prevalence amount of type A is more than other influenza types in the world. Also, the burden of seasonal influenza virus caused the infection of 3–5 million cases with severe illness symptoms [1]. Influenza viruses affect the human life more than other respiratory illnesses. The pathogenesis of Influenza has not been yet well understood since it depends on the immune system and viral determinants. Moreover, the previous infection or vaccination causes the cellular immunity which affected the efficacy of infection with various seasonal, zoonotic, and pandemic influenza viruses [2]. Therefore, different hosts can have distinct effects on the incidence level of the disease.

Microarray is a high-throughput technique has the ability of simultaneous measuring of thousands of gene expressions and so generating tremendous data. In order to general and then detailed evaluation of the biological phenomena in each study, the special and sometimes complicated statistical analysis is required [3, 4].

Many genes are involved in the pathogenesis routes of influenza infection, which constitute the complicated networks. Among various genes, the change in expression levels of some genes regulate the expression of others, so they can put in a group called module. These modules have different biological responsibility. Also, each group may activate various pathways and control specific functions [5]. Finding the modules can help researchers to design the proper treatment route by
targeting key proteins with specific functions. Moreover, finding novel biomarkers including important functions in a pathogenic-related disease is of importance for prognosis, risk assessment, and progression monitoring of disease [6, 7]. Biomarkers can be detected by calculation of differentially expressed genes between the normal and infected subjects, however, discovery the proteins that each biomarker is in connection with them, help to find the key pathways functioning in a disease.

Exploration of the involved genes which have close correlation patterns could be achieved by weighted gene co-expression network analysis (WGCNA) [8]. Through this algorithm, the highly co-expressed genes are placed in one module and different modules may contain key genes implicated in the pathogenesis process [9].

In this study, we aimed to find modules containing highly co-expressed genes which involve in the pediatric influenza pathogenesis. Also, the key genes with the prognosis potential were identified. Moreover, the genes in each module were analyzed using Gene Ontology (GO) function and pathway enrichment methods.

**Methods**

**Gene expression data and preprocessing**
The microarray gene expression dataset GSE29366 was downloaded from the NCBI Gene Expression Omnibus (GEO). It contains 12 healthy subjects and 19 Pediatric influenza patients which their whole blood samples were analyzed. All of the involved children in this study were in the age range from 1.5 months to 17 months. All individuals developed primary infection and showed signs of the disease. The Illumina HumanWG-6 v3.0 expression beadchip (GPL6884) was used to produce the sequencing data. The goodSamplesGenes function in Weighted Gene Co-expression Network Analysis (WGCNA) was used to filter the samples and genes with too many missing entries and zero-variance genes [10].

**Weighted gene co-expression network**
The weighted gene co-expression network based on Pediatric influenza samples was constructed by R package “WGCNA” [11]. To this purpose, the adjacency matrix containing Pearson’s correlations between each gene pair was firstly generated with the optimized soft power and then was transformed into a Topological Overlap Matrix (TOM). The highly co-expressed genes were then grouped by hierarchical clustering. In the next step, the dynamic tree cut algorithm was utilized to cut clustering dendrogram branches and generation of the modules. The gene expression profiles of each module were summarized by the first principal component named as module eigengene (ME). Moreover, each module denotes as discrete colors. Finally, the similar modules (module eigengenes) were merged into a single module and was then used to further interpretation [12].

**Finding differentially expressed genes**
To detect the differentially expressed genes (DEGs) among normal and infected subjects, GEO2R tool in the GEO database was employed. Benjamini-Hochberg FDR-adjusted \( p \)-values < 0.05 was selected as a criterion for finding DEGs.

**Finding candidate genes and their functional annotations**
The top highly interconnected genes in each module were identified as hub genes. Afterward, they were submitted in the Search Tool for the Retrieval of Interacting Genes (STRING) database to build the protein-protein interaction network (PPIN) [13]. Then, the nodes with zero betweenness and degree one were set aside from the constituted network. The PPIN was visualized using Gephi 0.9.2 [14]. The gene ontology and pathway enrichment analysis were carried out in gprofiler website (https://biit.cs.ut.ee/gprofiler/) [15].

**Results**

**Data preprocessing and sample selection**
Gene expression dataset GSE29366 was firstly quantile normalized and log-transformed. Then, the probe sets with unknown gene name were removed. The samples were evaluated in terms of missing entries and zero-variance genes. The number of 15,436 probes remained from 28,742 probes. As shown in the cluster tree (Fig. 1), two samples (GSM725944 and GSM725948) are clustered in different branches and excluded for further analysis.

**Weighted gene co-expression network construction and modules identification**
To exploration the required criterion for WGCNA, the scale-free topology fit index was calculated for various soft-thresholding power. As Fig. 2 shows, it reaches 0.84 for a power of 7 while a relatively high mean connectivity remained. This value was then applied to measure the adjacency and topological overlap matrices. To specify the groups of co-regulated genes identified as modules, the dynamic cut-tree algorithm was utilized (Fig. 3). After merging the similar modules by applying a threshold of 0.25, 38 modules were identified to further analysis (Fig. 4a). Each row and column of the heatmap plot is in accordance with one module, in which red denotes positive correlation and blue reveals a negative correlation. Figure 4b shows the dendrogram and dynamic cut tree before and after merging modules. To determine the biologically meaningful modules, all modules were submitted into the STRING and ones which their proteins were highly connected, were selected. Figure 5 shows the protein-protein interaction
Fig. 1 Hierarchical cluster tree of samples

Fig. 2 Analysis of the scale-free fit index (left panel) and mean connectivity (right panel) for various soft-thresholding powers
networks (PPINs) of each module, in which mediumpurple2 has 18 nodes and 46 edges, skyblue has 41 nodes and 107 edges, and yellow has 60 nodes and 689 edges. The size and color of each node illustrate the degree value in which red and blue colors indicate the higher and lower degree, respectively. The chosen modules were under scrutiny in terms of gene ontology and pathway enrichment. The most related modules to the infection with influenza were identified as mediumpurple2, skyblue, and yellow. Table 1 demonstrates the genes in each module.

Differentially expressed genes
To evaluate which genes in each module probably affected the pathogenesis routes after infection by influenza virus, 719 differentially expressed genes were found between the normal and infected subjects. In the next step, the chosen modules were explored to find which DEGs were involved in them. The results disclosed the presence of RPS27A, CEP152, TPT1, UBA52 in the skyblue module, and EIF2AK2, C19orf66, PML, PARP12, TRIM22, CMPK2, SP100, SP110, TRIM21, CHMP5, IFI44L, IFIT5, FAM46A, ANKY1, IFIT1, LAMP3, TNFSF10, TNFSF13B, OAS2, PLSCR1, LGALS9, LBE2L6, ADAR, RTP4, IFIT2, IFI35, IFI16, HERC5, STAT2, OAS3, RSAD2, OAS1, MX1, IRF7, SAMD9L, DDX60, DDX58, HELZ2, IFIH1, TDRD7, USP18, SAMD9, EPSTI1, ZNFX1, FBXO6, DHX58, TRAFD1, PARP9, TRIM25, ZBP1, OASL, PHF11, TRIMS, IFI44, ISG15, MX2, and IFIT3 in the yellow module.

Gene ontology enrichment
To explore the biological relationship of the genes in each module, the gene ontology (GO) enrichment analysis was performed. The genes of mediumpurple2 module were enriched in different immune systems, neutrophil and leukocyte degranulation, antimicrobial humoral response, and defense response. The most related terms to the viral infection and immune system were specified from genes of yellow module which contains defense response to virus, innate immune response, defense response to other organism, response to type I interferon, response to cytokine, regulation of viral life cycle, immune response, viral process, negative regulation of viral life cycle, regulation of viral process, negative regulation of viral process, regulation of innate immune response, cellular response to virus, regulation of defense response to virus, negative regulation of viral release from host cell, and positive regulation of immune system process. As mentioned above, the genes in the mediumpurple2 and yellow modules are activated due to viral infection and the immune system provides defense against pathogenic agent. In the skyblue module, various protein targeting and localization such as co-translational protein targeting to membrane, protein targeting to ER, protein localization to endoplasmic reticulum were highlighted. Viruses usually use the complex membrane network existing in the host cell such as endoplasmic reticulum (ER) membrane to entry, replication, and assembly. Therefore, the advent of this module was expected.
Fig. 4 (a) Heatmap plot of the identified modules. Each row and column is in accordance with one module, in which red and blue shows the positive correlation and negative correlations, respectively. (b) Dendrogram of genes clustered based on a dissimilarity measure (1-TOM) with assigned module colors. The colored rows show the module membership obtained by the dynamic tree cut method and after merging modules.
Pathway enrichment analysis

To evaluate the pathway enrichment analysis, the Kyoto Encyclopedia of Genes and Genomes (KEGG) and Reactome Pathway Database were explored. The immune pathways, Neutrophil degranulation, Antimicrobial peptides, and Defensins were enriched by genes in the mediumpurple4. This module contains genes which involved in immune response system and have antimicrobial activity.

The genes pertaining the skyblue module were enriched in Axon guidance, signaling by NOTCH2, Infectious disease, Influenza Infection, Influenza Life Cycle, Influenza Viral RNA Transcription and Replication, Viral mRNA Translation, and rRNA processing. This module contains genes which their roles have been specified in the influenza infection pathways. Notch signaling involves several functions of cellular differentiation, cell fate, and cell survival. Viruses use several mechanisms to escape innate immune antiviral responses and cause cell survival. The downregulation of CREB1, RPS27A, and UBA52 in this pathway show the immune system domination.

Eventually, the genes grouped in yellow module were enriched in the viral and immune pathways including Influenza A, RIG-I-like receptor signaling pathway, NOD-like receptor signaling pathway, Herpes simplex infection, Cytosolic sensors of pathogen-associated DNA, Immune System, Cytokine Signaling in Immune system, Interferon Signaling, and Antiviral mechanism by IFN-stimulated genes. The RIG-I-like receptors activate innate immunity and inflammation after discernment viral RNA ligands [16]. RIG-I is required for type I IFN production in response to influenza. One of the cellular defense mechanism is performed by receptors that check uncommon RNA and DNA resulted from viral infection in the cytosol. It causes the limitation in the virus replication and activation of antiviral immunity process [17].

Discussion

In the present study, three PPI network were identified based on the WGCNA which are significantly related to the pediatric influenza infection. Functional analysis revealed that the genes in the mediumpurple2 module implicate in the immune responses through different routes, skyblue module in the protein targeting, and the yellow module in the defense and immune responses to virus. Moreover, the analysis of differentially expressed genes revealed 719 DEGs between the normal and infected subjects. Further analysis showed that most members of the yellow and some genes in the skyblue modules were among identified DEGs, so these modules were more discussed.

Children who did not experience previous exposures to influenza viruses are more vulnerable against infection and shed larger number of virus particles for a long time [18]. Moreover, the function of immune system may be different from adult-infected subjects. The initial defense and adaptive immune responses against viral infection are performed by type I IFNs (IFN-α/β) such as RTP4, GBP1, OAS1, IFI27, and IF44L [19]. They induce the IFN-stimulated genes (ISGs) through activation of the Janus kinase–signal transducer and activator of transcription pathway [20]. The influenza A viruses hinder type I IFN signaling via induction of the suppressor of cytokine signaling-3 (SOCS-3) protein [21]. As a result, the expression level of SPI10, OAS1, and IRF-1, which are IFN-induced genes, increases.

ISG15 has been recognized as an activator of NK cells and a driver of IFN-γ secretion [22]. It has common properties with other ubiquitin-like proteins (UBLs); however, its function is influenced by the innate immunity signaling pathways. The increase in the expression of ISG15 is found after type I interferon stimulation and viral infection. Similar to type I interferons, ISG15 involves in the disease tolerance against the influenza A
### Table 1  The genes involved in each module and identified DEGs.

| Modules   | Genes                                                                                                                                 |
|-----------|---------------------------------------------------------------------------------------------------------------------------------------|
| mediumpurple2 | ELANE, DEFA3, PRRT3, DEFA1, DEFA1B, MPO, DEFA4, BPI, CAMP, RNASE3, LTF, LCN2, C1, GLE1, CTSG, PGLYRP1, MA34A, OLR1 |
| skyblue    | RPS25, HBB, RPS27A, RPS27, RPLP2, RPS29, UBA52, RPS16, RPL41, UBA6, EEF1A1, TPT1, USP1, TRIM41, HT2A2, NARAS, CDKN1B, B2M, ACTB, TMA7, MBNL1, CREB1, H3F3A, DNM1L, CEP152, C14orf2, CDC5L, OAZ1, HBA1, RPS12, F2R, VAMP4, ORC6, SEC24B, LILRB3, EED, CLASRP, TUBD1, SUPT3H, KIAA0907, ROCK2 |
| yellow     | ISG15, OAS1, MX1, DDX58, IFIT1, RSAD2, IFIT3, IFIT7, IFIT2, OAS2, IFI35, UBE2L6, MX2, SP100, HERCS, TRIM22, TRIM25, EIF2AK2, OASL, ADAR, STAT2, TRIM21, TRIM5, PML, IFIH1, CMPK2, IFI44, FBXO6, ZBP1, IFI44L, DDX60, OAS3, IFITS, SAMD9, SAMD9L, DDX58, RTFP, PARP12, P1, LSLCR1, HEL22, AZ22, IFI16, PARP9, TRNSF10, TRD7, PHF11, TRNSF13B, EPST1, ANKY1, ZNFX1, LGALS9, FAM46A, USP18, NTS3C, NMI, TRAFD1, C19orf66, CHMPS, LAMP3 |

**DEGs (between normal and infected subjects)**

| Genes                                                                 |
|-----------------------------------------------------------------------|
| OTOF, IFI44L, XAF1, IFI6, EPST11, IFI44, SAMD9L, EIF2AK2, IFI35, IFI35, OAS3, HESX1, OAS1, STAT1, ISG15, MX1, DDX60, PARP9, RNFR13, IFITM3, LY6E, HEL22, FBXO18, IFI7, ADAR, RSAD2, MOV10, IFIH1, IFIT3, PML, STAT2, FBXO6, SAMD9, SC02, CMPK2, NEXN, H1F0, TNK2, C2, OAS2, MYO5, SP100, DHX58, CMT1R, RT4P, PARP10, SHIG5, SLC26A8, IFIT1, TRAP1, SPATS2L, OAS1, RNFI3, GPR84, RP4A, LOC100419583, TAP1, PML, SLC24A4, IFIT1M, LGALS3BP, UBE2L6, FCER1A, TOMM7, MASTL, M54A44, RPL135P, SSH1, REC8, MEF2A, BST2, PHF11, ABCD1, MX2, JAK2, IL15RA, SPI10, PARP12, DEFB1, SLC1A3, CCR3, BMX, TAPBP, LENG8, PLSCLR1, HERCS, HIST2H2AA3, TRIM5, IFIP9, PARP14, SERPING1, BATF2, TOR1B, STAB1, RTN2, PKNO, LAP3, IL37, STRADA, NAPBP1, NRRC2, EEF1B2, CD1C, RN5K, CLECSA, GBA, LG1, HP, GRN, CCDB2, FCGR1B, CARD16, NCOA7, ZBP1, EPB41L3, IFI27, TRIM22, ARHGEF11, KIAA1958, MARK3, C1QC, TRIM25, MT1G, TMRE1, TMXPE, MT2A, HIST2H2AC, ADRA1D, TDRD7, ZNFX10, LTBR4, ARGL1, LDLR, CAPN3, SLC30A1, BRK51, GPD2, CLU1, TMYP, AGRN, MAFB, RPL14, ZNF839, RPL23, IFIT5, SUSD1, FXYD6, ECE1, TRAFD1, EIF3L, RIMBP3, RPLS, AXL, DNASE2, NTSC3A, CNHI, LGL1, SUCX, LGALS9, YEATS2, MICB, MTHFD1, IFI3F, HLA-DR, NELL1, DTX3L, TRIM14, ANKRD22, HOOK3, IL1RN, PIWIL4, TMEM51, CAPN2, PRA1M, TRIM16, TCTN1, SPP522, MTF1, RPL37A, FCCRG1, CCR1A, NTNG2, AIM2, C2L2, GARSAP1, CCR1, C3AR1, ATP8B, EEF1G, ZCCHC2, REX6, SLC9A1, SOSC3, RPS27A, HTR3A, RAB27A, UBA7, PAQR6, MMP9, ZNF341, TAOK2, CEP152, EXOG, FLVCR2, CTGF, HIST1H2BD, TRIM38, MAP_ 5K7CL, CDC25C, SLC16A6, PAG1, HGD, VDR, TMEM62, DDX58 RPL4 TIAF1 ANKFK1 ADAMA9 RPL23AP32 MT1A, ADAP2 FBL, FOXDL1, IFIT2, AT3, RPS9 TNC2, MAFG, MDM2, OLG3 CHMPS, LMBN1, DPH5, AGFG1, RIPK3, LGALS9C IFI16 MB21D1, MERTK SIGLEC16, CKAP4, DFRXP4 BAZ1A DISC1 SCAR2B, CDKN1A, TMEM110, WARS, TXN, APOL2, SNX20, GPR155, CTRTAP TXNDC12, RPL31 TRIM21, EIF4G3 TET2 CTSD, CXCRRP1, PAF3, INPP4A, RHT1, C12orf57, LGALS9B, MED15, SRGAP2, RPL3, AT3, SLC9A8, TMOD3, GYGI, HVCN1, ASGR2, FAM209B, GCH1, OGF, NELL2, SLC27A3, STBD1, NEFAL2, NDST1, TRIM69, RNF19A, CD36, VNN1, CEPT1, LHX6, SLC22A18, TBCD12, HIST2H2AB, ITIH4, GLS, PTP4A1, MYD88, RPL13A, RPS14, PLEC, UGOT2, RBM43, GBP1, SYT17, FAM46A, SORT1, ADAMTS14, RPL23AP6, USP18, ATR11C, C12orf8, ARSA, EXT1, MR1, OPLAH, RPS3, BCL2L11, HK3 LOXL3, MCM8, WWP2, CD177, MCEMP1, TRPM4, SLC6A12, ABC10, RPL21, SLITRK4, CALM4L, STX6, B4GALT5, ARID3A, VAMPS, DNM2, JYF1, DIAP1 |
virus infection. Indeed, ISG15 conjugation of cellular target proteins such as Mx1 and Mx2 is needed for the antiviral activities. Moreover, ISG15 conjugation of a viral protein blocks some indispensable functions [23].

**HERC5** is another IFN-induced gene that mediates ISGylation of protein targets. It is induced by Influenza infection and elevates ISGylation as an ISG15 E3 ligase [20, 24, 25]. It also functions as a critical antiviral protein against Influenza A virus by restricting IAV-NS1.

**Samd9l** is a Sam domain containing protein which has important roles during virus infection and innate immunity [26]. The expression of **Samd9l** is increased by type I interferon and its role in the control of influenza virus infection and also pathogenesis has been proposed [27].

In particular, we also identified the C19orf66 gene which previously reported as an IFN-induced transcript that suppresses dengue virus [28]. C19orf66 named as RyDEN can constitute a complex with the indispensable cellular mRNA binding proteins, PABPC1 and LARP1, for the efficient replication of DENV. Moreover, it was proposed that C19orf66 is a new antiviral effector which has a prominent role in the inhibition of virus replication. The increase in the expression of C19orf66 was accompanied with up-regulation of **IFIT** genes [29, 30] which can inhibit the translation initiation, binding and sequester uncapped viral RNA, and viral protein in the cytoplasm.

Another substantial gene found in this study is **HELZ2** (Helicase With Zinc Finger 2) which is an IFN stimulated gene and immediate early gene (IEG). It was reported that the HELZ2 knockout in mice causes the enhancement of the dengue virus infectivity. Also, the mediatory of HELZ2 in the IFN antiviral response was disclosed, so that the up-regulation of the HELZ2 transcriptional and protein in the nucleus, and activation of a transcriptional program were involved. Moreover, IEGs were introduced as the biomarkers due to their influence on the host response to viral infections [31].

**EPSTI1** is an IL-28A-induced ISGs which was identified as an anti-HCV. Also, the knockdown of **EPSTI1** caused the enhancement of virus. It was stated that **EPSTI1** may actuate PKR promoter and induce IFN-β, IFIT1, OAS1, and RNase L, which are PKR-dependent genes and responsible for the **EPSTI1**-mediated antiviral activity. Therefore, **EPSTI1** was presented as a proper therapeutic target to treat HCV infection [32]. In this study, the expression level of **EPSTI1** has increased in the influenza-infected subjects versus normal cases. The up-regulation of the IFIT proteins especially IFIT1 and also OAS1 may induce the overexpression of **EPSTI1**, so it can be considered as the biomarker or a target for designing a drug.

**PHF11** is a transcriptional co-activator of IL2 and IFNG which its knock-down causes up-regulation of the pro-inflammatory chemokine IL-8, therefore the contribution of **PHF11** in epidermal recovery was proposed [33]. The overexpression of **PHF11** due to the infection with Japanese encephalitis virus [34], avian influenza A virus [35], and Epstein-Barr Virus [36] were reported previously. Similarly, the expression of **PHF11** was increased after influenza infection which can confirm its role in the diminution of pro-inflammatory chemokine and increase the IFNG due to the infection.

**IRF7** which is overexpressed due to the influenza infection regulates the antiviral response. The viral infection causes the phosphorylation and translocation of **IRF7** to the nucleus and as a result the induction of expression of type-I interferons which in turn activates **IRF7** transcription through STAT2 [37, 38].

**EIF2AK2** is the interferon-induced dsRNA-dependent protein kinase which has a prominent role in the innate immune response to viral infection, apoptosis, cell proliferation, and differentiation. MX1 as a member MxA protein belongs to the GTPase family and suppresses the influenza virus replication via targeting the viral nucleoprotein [39]. OAS proteins family are IFN-stimulated proteins which have prominent roles in the innate immune responses. They also catalyze the synthesis of 2′-5′-linked oligoadenylates, which in turn cause the activation of RNase L and degradation of viral and cellular RNAs [40]. ZBP1 has been recently identified as an innate immune sensor of influenza virus. It regulates NLRP3 inflammasome activation and induces the apoptosis, necroptosis, and pyroptosis in the influenza-infected cells [41].

The genes involved in the mediumpurple4 were enriched in the immune pathways and neutrophil degranulation. Neutrophils are granulocytes that comprise innate phagocytic cells packed with granules containing proteins with antibacterial function [42, 43]. In this study, we identified proteins belonging to three types of neutrophil granules: primary (azurophilic: ELANE, MPO, PRTN3, BPI), secondary (specific: LTF, DEFA4, DEFA1, DEFA1B), and tertiary (gela-tinase: PGLYRP1) granules. These genes were upregulated in the influenza-infected subjects. The neutrophil granule proteins have antibacterial function and have key roles in the innate immune defense against bacteria. The up-regulation of neutrophil granule proteins has been reported in viral infections and RSV-stimulated neutrophils [44–46].

In the skyblue module, **UBA52** and some genes belonging to RP family were identified in consistent with the required interaction between **UBA52** and RP family for virus replication. The previous study revealed that the knockdown of **UBA52** in the chicken cells causes the diminution of the progeny viral titer denoting the substantial function of **UBA52** in the H5N1 influenza A virus infection [47]. However, **UBA52** and most of the **RP** genes were down-regulated in this study after the influenza infection with respect to the normal cases. It can
be due to the overexpression of genes involved in the defense response to virus and innate immune response including EIF2AK2, C19orf66, PML, TRIM22, IFIH44L, IFITS, IFIT1, OAS2, PLSCR1, ADAR, RTP4, IFIT2, IFI16, HERC5, STAT2, OAS3, RSAD2, OAS1, MX1, IRF7, DDX60, DDX58, IFIH1, DHX58, PARP9, TRIM25, OASL, TRIM5, ISG15, MX2, and IFIT3. The innate immune system responds to the entry Influenza virus by members of Toll-like receptors, and RIG-I and the NOD-like receptor family. Finally, the inter-individual variation in expression of the identified differentially expressed genes was low. It is found that inter-individual variation in gene expression profiles is related to sex, age, and the time of sample collection [48, 49]. Moreover, differences in genotype, epigenetic or environmental factors can be cause of the intrinsic differences in expression patterns.

Conclusions

Overall, the systems virology approach based on the application of WGCNA helped us to find three modules of co-expressed genes related to pediatric influenza. Moreover, SP110, HERC5, SAMD9L1, RTP4, C19orf66, HELZ2, ENSTI1, and PHEF11 were remarkably co-expressed with other known genes which were involved in innate immune system and defense to virus. These genes and subsequently related proteins are proposed as the candidate biomarkers and also drug targeting. However, further studies are required to evaluate and confirm them.

Abbreviations

DEGs: Differentially expressed genes; GEO: Gene Expression Omnibus; GO: Gene Ontology; KEGG: Kyoto Encyclopedia of Genes and Genomes; ME: Module eigengene; PPI: Protein-protein interaction network; STRING: Search Tool for the Retrieval of Interacting Genes; TOM: Topological Overlap Matrix; WGCNA: Weighted gene co-expression network analysis.

Acknowledgements

Not applicable.

Authors’ contributions

MZG designed and analyzed results and also wrote the manuscript. SHM provided analysis tools or data and supervised the manuscript.

Funding

There is no funding provided for this project.

Availability of data and materials

All data generated or analysed during this study are included in this published article.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

1. Department of Virology, School of Public Health Tehran University of Medical Sciences, Tehran, Iran. 2. Institute of Biochemistry and Biophysics, University of Tehran, Tehran, Iran. 3. Department of Microbiology, School of Medicine, Alborz University of Medical Sciences, Karaj, Iran. 4. Non-communicable Diseases Research Center, Alborz University of Medical Sciences, Karaj, Iran. 5. Department of Bioscience and Biotechnology, Malek Ashtar University of Technology, Tehran, Iran.

Received: 18 May 2019 Accepted: 9 October 2019
Published online: 29 October 2019

References

1. Dunning J, Blankley S, Hoang LT, Cox M, Graham CM, James PL, Bloom CL, Chaussabel D, Banchereau J, Brett SJ. Progression of whole-blood transcriptional signatures from interferon-induced to neutrophil-associated patterns in severe influenza. Nat Immunol. 2018;19:625–35.

2. Cox R, Brockstad K, Ogra P. Influenza virus: immunity and vaccination strategies. Comparison of the immune response to inactivated and live, attenuated influenza vaccines. Scand J Immunol. 1994;59:1–15.

3. Raza K. Clustering analysis of cancerous microarray data. J Chem Pharm Res. 2014;6:488–93.

4. Vollmer E, Goldmann T. Pathology on the edge of interdisciplinarity. A historical epiphenomenon. Romanian J Morphol Embryol. 2011;52:223–30.

5. Mete M, Tang F, Xu X, Yuysk I. A structural approach for finding functional modules from large biological networks. In: BMC Bioinformatics BioMed Central. 2008;8:19.

6. Munshi SU, Pand D, Holli P, Rewari BB, Jameel S. MicroRNA-150 is a potential biomarker of HIV/AIDS disease progression and therapy. PLoS One. 2014;9:e95920.

7. Ngo Y, Munteanu M, Moulos D, Charlotte F, Imbert-Bismut F, Thabut D, Lebray P, Thibault V, Benhamou Y, Moussali J. A prospective analysis of the prognostic value of biomarkers (FibroTest) in patients with chronic hepatitis C. Clin Chem. 2006;52:1887–96.

8. Yang Y, Han L, Yuan Y, Li J, Hei N, Liang H. Gene co-expression network analysis reveals common system-level properties of prognostic genes across cancer types. Nat Commun. 2014;5:3231.

9. Yuan L, Chen L, Qian K, Qian G, Wu C-L, Wang X, Xiao Y. Co-expression network analysis identifies six hub genes in association with progression and prognosis in human clear cell renal cell carcinoma (cCcrCC). Genomics data. 2017;14:132–40.

10. You L, Wang J, Zhang F, Zhang J, Tao H, Zheng X, Hu Y. Potential four-miRNA signature associated with T stage and prognosis of patients with pancreatic ductal adenocarcinoma identified by co-expression analysis. Mol Med Rep. 2019;6:941–51.

11. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. BMC bioinformatics. 2008;9:559.

12. Guilletti M, Occhipinti G, Principato G, Piva F. Weighted gene co-expression network analysis reveals key genes involved in pancreatic ductal adenocarcinoma development. Cell Oncol. 2016;39:379–88.

13. Säkkärzyk J, Franceschini A, Wyder S, Forland K, Heller D, Huerta-Cepas J, Simonovic M, Roth A, Santos A, Tsafou KP. STRING v10: protein–protein interaction networks, integrated over the tree of life. Nucleic Acids Res. 2014;43:D447–52.

14. Bastian M, Heymann S, Jacomy M. GeNi: an open source software for exploring and manipulating networks. In: Third international AAAI conference on weblogs and social media. 2009.

15. Reimand J, Arak T, Adler P, Kolberg L, Reisberg S, Peterson H, Vilo J. g:Profiler—a web server for functional interpretation of gene lists (2016 update). Nucleic Acids Res. 2016;44:W883–9.

16. Pichlmair A, Schulz O, Tan CP, Näsland T, Ljeström P, Weber F, Sousa C, RE. RIGI-mediated antiviral responses to single-stranded RNA bearing 5′-phosphates. Science. 2006;314:997–1001.

17. Goubaud D, Dedduche S, Sousa C, RE. Cytosolic sensing of viruses. Immunity. 2015;38:553–69.

18. Krammer F, Smith GJD, Fouchier RAM, Peiris M, Kedzierska K, Doherty PC, Palese P, Shaw ML, Treanor J, Webster RG, García-Sastre A. Influenza. Nature Rev. Disease Primers. 2018;4:3.

19. Palese P, Shaw ML, Treanor J, Webster RG, García-Sastre A. Influenza. Nature Rev. Disease Primers. 2013;3:855–69.

20. Krammer F, Smith GJD, Fouchier RAM, Peiris M, Kedzierska K, Doherty PC, Palese P, Shaw ML, Treanor J, Webster RG, García-Sastre A. Influenza. Nature Rev. Disease Primers. 2018;4:3.

21. Lebray P, Thibault V, Benhamou Y, Moussali J. A prospective analysis of the prognostic value of biomarkers (FibroTest) in patients with chronic hepatitis C. Clin Chem. 2006;52:1887–96.

22. Yang Y, Han L, Yuan Y, Li J, Hei N, Liang H. Gene co-expression network analysis reveals common system-level properties of prognostic genes across cancer types. Nat Commun. 2014;5:3231.
20. Wong JY, Pung YF, Sze NS-K, Chin K-C. HERCS is an IFN-induced HECT-type E3 protein ligase that mediates type I IFN-induced ISGylation of protein targets. Proc Natl Acad Sci. 2006;103:10735–40.

21. Pauli E-K, Schmolke M, Wolff T, Viemmann D, Roth J, Bode JG, Ludvig S. Influenza virus inhibits type I IFN signaling via NF-κB-dependent induction of SOCS-3 expression. PLoS Pathog. 2008;4:e1000196.

22. Care MA, Stephenson SJ, Barnes NA, Fan J, Zougnia A, B-Sherby NY, Vital EM, Westhead DR, Tooke RM, Doody GM. Network Analysis Identifies Proinflammatory Plasma Cell Polarization for Secretion of ISG15 in Human Autoimmunity. Journal of Immunology (Baltimore, Md : 1950). 2015;197:1447–59.

23. Hsiang T-Y, Zhao C, Krug RM. Interferon-induced ISG15 conjugation inhibits influenza virus gene expression and replication in human cells. J Virol. 2009;83:5971–7.

24. Dastur A, Beaudenon S, Kelley M, Krug RW, Hulbrecht JM. Herc5, an interferon-induced HECT E3 enzyme, is required for conjugation of ISG15 in human cells. J Biol Chem. 2006;281:43348–48.

25. Tang Y, Zhong G, Liu X, Shan Y, Feng H, Bu Z, Chen H, Wang C. Interferon-stimulated cellular inhibitor against dengue virus replication. PLoS Pathog. 2016;12:e1005357–7.

26. Liu J, Wennier S, Zhang L, McFadden G. M0E2 is a host range factor essential for myxoma virus pathogenesis in mice and functions as an antagonist of host SAMD9 in human cells. J Virol. 2011;85:3270–82.

27. Boon AC, Williams RW, Sinasac DS, Webby RJ. A novel genetic locus linked to pro-inflammatory cytokines after virulent H5N1 virus infection in mice. BMC Genomics. 2014;15:1017.

28. Suzuki Y, Chin W-X, Han QE, Ichiyama K, Lee CH, Eyo ZW, Ebina H, Takahashi C, Tan BH, et al. Characterization of RyDEN (C19orf66) as an interferon-stimulated cellular inhibitor against dengue virus replication. PLoS Pathog. 2015;12:e1005357–7.

29. Lietzén N, Ohman T, Rintahaka J, Julkunen I, Attockkallio T, Matikainen S, Nyman TA. Quantitative subcellular proteome and secretome profiling of influenza virus-infected human primary macrophages. PLoS Pathog. 2011;7:e1001540–10.

30. Zhou X, Michal JI, Zhang L, Ding B, Lunney J, Liu B, Jiang Z. Interferon-induced IFIT family genes in host antiviral defense. Int J Biol Sci. 2013;9:200–8.

31. Fusco DN, Pratt H, Kandilas S, Cheon S, Lin W, Cronkite DA, Basavappa M, Jefferies KL, Anselmo A, Sadreyev R, et al. HELZ2 is an IFN effector mediating suppression of dengue virus. Front Microbiol. 2017;8.

32. Meng X, Yang D, Yu R, Zhu H. EPSTI1 is involved in IL-28A-mediated inhibition of HCV infection. Mediat Inflamm. 2015;2015:716315

33. Yuan J, Cahir-McFarland E, Zhao B, Kieff E. Virus and cell RNAs expressed during Epstein-Barr virus infection. J Virol. 2008;82:3624–37.

34. Dastur A, Beaudenon S, Kelley M, Krug RW, Hulbrecht JM. Herc5, an interferon-induced HECT E3 enzyme, is required for conjugation of ISG15 in human cells. J Biol Chem. 2006;281:43348–48.

35. Reemers SS, van Leenen D, Koerkamp MJG, van Haarlem D, van de Haar P, van Eden W, Vervelde L. Early host responses to avian influenza a virus are prolonged and enhanced at transcriptional level depending on maturation of the immune system. Mol Immunol. 2010;47:1675–85.

36. Honda K, Yanai H, Negishi H, Asagiri M, Sato M, Mizutani T, Shimada N, Hasegawa T, Takahashi C, Tan BH, et al. Characterization of RyDEN (C19orf66) as an interferon-stimulated cellular inhibitor against dengue virus replication. PLoS Pathog. 2016;12:e1005357–7.

37. Tingley P, van Eden W, Vervelde L. Early host responses to avian influenza a virus are prolonged and enhanced at transcriptional level depending on maturation of the immune system. Mol Immunol. 2010;47:1675–85.

38. Brown PO. Individuality and variation in gene expression patterns in human blood. Proc Natl Acad Sci. 2003;100:1896–901.

Publisher’s Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.