International Journal of Biology, Pharmacy
and Allied Sciences (IJBPAS)
'A Bridge Between Laboratory and Reader'

www.ijbpas.com

DIFFERENTIAL GENE EXPRESSION IN SCRUB TYPHUS COMPARED TO OTHER ACUTE FEBRILE INFECTIONS BY BIOINFORMATIC APPROACHES

MANDAL M[1] AND MANDAL S[2]

1: Department of Physiology, MGM Medical College, Kishanganj-855107, India

2: Laboratory of Microbiology and Experimental Medicine, Department of Zoology,

University of Gour Banga, Malda-732103, India

*Corresponding Author: Dr. Shyamapada Mandal; E Mail: samtropmed@gmail.com

ABSTRACT

This communication aims to discriminate scrub typhus (ST), caused with the infection of *Orientia tsutsugamushi*, from other acute febrile infections, such as malaria (ML), dengue (DG) and murine typhus (MT) by differential gene expression analysis. The microarray gene expression profiles of GSE16463 dataset, from *O.tsutsugamushi* infected monocytes, was comparatively analysed along with biological process and functional pathways enrichment in disease pathogenesis, using STRING App of the cytoscape by maximal clique centrality method, GEO2R tools with GEOquery, limma R, and Biobase packages from the Bioconductor project. A unique 31 ST-associated signature genes were obtained in this study and the functional pathways significantly enriched were translocation of ZAP-70 to immunological synapse, combined with phosphorylation of CD3 and TCR zeta chains, involving PTPN22 and CD3G genes. The identified transcriptional signatures specific to ST could help distinguish the disease from other acute febrile illnesses: DG, ML and MT, that might be useful in the investigation of the molecular pathophysiology of ST, and discovering novel drug targets and vaccine development as well.

INTRODUCTION

Concerning the acute febrile infection (AFI) development of improved diagnostics, particularly the molecular diagnostic intervention with therapeutic potential, is an urgent need for infection control and treatment management. Scrub typhus (ST), which is caused with the infection of a bacterial pathogen *Orientia tsutsugamushi* (vectored by larval trombiculid mites *Leptotrombidium deliense*), presents as a febrile illness with the severity ranging from mild (no organ dysfunction), to moderate (with one organ dysfunction) to severe/lethal (two or more organ dysfunction: renal dysfunction, pulmonary dysfunction, cardiovascular dysfunction, or central nervous system dysfunction) [1].

The ST along with murine typhus (MT) is as prevalent as malaria (ML) and dengue (DG), displaying symptoms that include fever and chill, headache, body aches, and sometimes eschar or rash (at initial stages), and causes substantial morbidity and mortality worldwide [2, 3]. The ST (in absence of distinctive eschar at the inoculation site) mimics the other febrile diseases MT, DG and ML, and thus due to the non-specific clinical presentation distinction of ST is very difficult from MT, ML and DG in geographical areas where these diseases prevail, and in addition the

lack of reliable point-of-care (sensitive and specific) diagnostic tests, diagnosing ST is difficult. Therefore, early diagnosis of ST using effective molecular-based tools is crucial in order to provide prompt antimicrobial treatment to prevent the disease complications and reduce the mortality rate [1].

Bioinformatics approaches have been found useful in identifying signature genes as well as predicting diseases outcome. Herein we investigated the molecular response that shows up distinct to these four categories of infection (ST, from MT, ML and DG) with overlapping signs and symptoms using GSE1646 dataset [4], to track the transcriptome, to gauge the differentially expressed genes (DEGs), investigate underlying molecular interactions between DEGs, functional enrichment and significant pathways associated with the DEGs, protein-protein interaction (PPI) network of common DEGs, to identify the genes as predictive biomarkers as a basis for a targeted therapy that can help guide more accurate therapeutic clinical studies and, in the future, to provide treatment decisions. Detecting ST-associated signature genes have been reported to contribute in revealing the disease pathogenesis enabling the diagnosis, discovery of novel drug targets for ST [4],

and vaccine development as well **[5]**. Therefore, analysis of the DEGs in ST, along with other infectious febrile illnesses (ML, DG and MT), compared to healthy individuals might be an expedient tool in determining the biomarkers of clinical outcome of infections to facilitate the assessment of the need for clinical interventions.

MATERIALS AND METHODS

The dataset GSE1646 containinggenome-wide expression in peripheral blood mononuclear cells from patients with ST (n=4; with fever duration of 3 – 20 days from cases of 19 – 56 years of age), DG (n=7; with fever duration of 1 – 4 days from cases of 18 – 70 years), MT (n=7; with fever duration of 6 – 21 days from cases of 23 – 65 years), and ML (n= 4; with fever duration of 3 – 7 days from cases of 24 – 34 years), compared to two healthy controls (HC), deposited by **[4]**, was downloaded from the GEO (gene expression omnibus) database (www.ncbi.nlm.nih.gov/geo/) and extracted with GEO2R tools (http://www.ncbi.nlm.nih.gov/geo/geo2r/)in order to perform comparisons with GEOquery, limma R, Biobase packages from the Bioconductor project **[6, 7]**. The raw microarray data obtained was processed for "NA" filtering and normalized applying "normalize Between

Arrays" in limma, using quantile method followed by fitting with a linear model and comparison among disease groups: ST, DG, MT and ML, using empirical Bayes moderated t test, in limma, to generate log 2 Fold Change (log2FC) differential expression of ST, DG, MT and ML, the Pvalues and FDR (False Discovery Rate) adjusted P value (adj.P.value) with cutoff criterion for DEGs set as Pvalue <0.05 and |log2FC| >1**[8]**.

The regulatory network of the DEGs was built using cytoscape (https://cytoscape.org/), version 3.8.2 **[9]**. The PPI network was created using STRING (Search Tool for the Retrieval of Interacting Genes) (http://string-db.org/cgi/input.pl; version 1.6.0) App of the cytoscape **[10]**, version 3.8.2 (https://cytoscape.org/) **[9]**. The important hub genes within the PPI network were identified using the cytoscape plugin CytoHubba, version 0.1, ranked by higher scores as per maximal clique centrality (MCC) method, amongst several topological algorithms **[11]**. The significant modules containing cluster of densely connected regions within the PPI network were explored using the MCODE (molecular complex detection) version 2.0.0 cytoscape plugin, with threshold values of degree 2, node score 0.2, k-core 2 and maximum depth 100 **[12]**. The DEGs

were subjected to GO (gene ontology) enrichment analysis to explore the functional roles in terms of biological processes, cellular components and molecular functions using STRING Enrichment App of the cytoscape software, with threshold Pvalue <0.05 **[10]**. The DEGs were also examined to study the statistical enrichment in Reactome (https://reactome.org) pathways, using the STRING Enrichment App.

RESULTS

The ST, DG, MT and MLgene expression microarray signal intensity downloaded from NCBI GEO dataset GSE16463 before and after normalization have been indicated in **Figure 1A** and **Figure 1B**. The empirical Bayes moderated t distribution for the log2FC of ST, DG, MT, ML, HC compared to each other, average Expression (AveExpr), F-statistics (F), the Pvalues, and adj.P.values **(Figure 1C)**. Analysis of normalized expression data using GEO2R yielded 96 DEGs indicated as transcript IDs for the 24 samples of ST, DG, MT, ML and HC with a heatmap displaying hierarchical structural analysis among the samples and transcripts as well **(Figure 2)**.

The upregulated and downregulated DEGs among ST, DG, ML, MT compared to HC groups have been represented with volcano plots in **Figure 3**. The scrub typhus

regulatory networks consisted of 150 genes, of which 65 were upregulated and 85 downregulated with log2FC ranging from 1.006819243 (for PTPN22) to 3.87069874 by (for ANKRD22) and from -1.0598985 (for CBX7) to -2.288928002 (for GAL3ST4), respectively **(Figure 3A)**. Biological regulation, cell surface receptor signalling pathway, immune response and its regulation, cytokine-mediated signalling pathway, lymphocyte activation, protein localization to membrane were significantly overrepresented among the differentially expressed genes in ST disease **(Table 1)**. The most significantly enriched reactome pathways in ST group, from our study, were immune system, cytokine signalling, and immunoregulatory interactions between a lymphoid and a non-lymphoid cell **(Table 1)**.

The dengue regulatory network **(Figure 3B)**, consisted of 297 genes with 138 upregulated that exhibited log2FC ranging from 1.011207 (for ARSB) to 7.201346 by (for IFI27), and 159 downregulated genes for which log2FC ranged from -1.00412 (for IL32) to -0.31158 by (for LRRN3). Immune system and its innate type along with gamma interferon signalling, signalling by interleukins, ISG15 and OAS antiviral mechanism, L13a-mediated translational silencing of ceruloplasmin expression, negative regulators of

DDX58/IFIH1 signalling and TRAF3-dependent IRF activation pathway were among the over stimulated reactome pathways in DG-associated genes **(Table 2)**.

Among a total of 53 genes within the ML regulatory network **(Figure 3C)**, 28 were upregulated and 25 downregulated displaying log2 FC ranging from 1.000882 (for ZBED2) to 2.737276 (for IFI27) and from -1.01912 (for RNF125) to -2.42052 (for LOC649841), respectively. The immune system involving cytokine signalling was the most significant reactome pathway in ML group associated with genes including IFI27, RNF125, MT2A, TRIM21, LMNB1, SERPING1, UBE2L6, CXCL10, KLHL3, C1QB, SLAMF7, CD40LG, FBXO6, TNFRSF25, LGMN, JUP, IFI35, CFB, WWP1, ATOX1, and IL11RA **(Table 3)**.

The MT regulatory gene network **(Figure 3D)** displayed the involvement of a total 35 comprising 19 upregulated and 16 downregulated genes with log2FC ranging from 1.02326 (for LAP3) to 2.46668 in (for IFI27), and from -1.02326 (for C9orf123) to -2.46668 (for LRRN3), respectively. The immune system and DNA damage/telomere stress induced senescence were the most significant reactome pathways in MT group associated with genes including PSMA4, MT2A, OASL, LMNB1, SERPING1,

UBE2L6, CXCL10, KLHL3, C1QB, SLAMF7, NPDC1, IFIT3, FBXO6, JUP, IL11RA, IFI27, HIST2H2AC, H2AFJ and HIST2H2AA3 **(Table 3)**.

The over and under stimulated genes among the ST, DG, ML and MT disease groups are displayed in the heat map **(Figure 4A)**, which indicated the presence of 67 DEGs with highly overexpressed and under stimulated genes in DG in the upper and lower respective panels, while 3 genes including IL32, LOC729530 and PLTP were highly induced in ST. The number of unique signature genes for ST, DG, ML and MT were 31, 267, 4, and 3 respectively, while the number of overlapping genes in the ST group with the DG, ML and MT were 114, 42, and 54, respectively; 32 genes belonged commonly to all the four categories, as represented by the Venn diagram **(Figure 4B)**. The ST category consisted of 530 PPIs **(Figure 4C)**, with an enrichment of 10^{-16}. The most interacting proteins within this network were related to immune response including T-lymphocyte differentiation antigen T8/Leu-2 (CD8A) and T-cell surface glycoprotein CD8 beta chain (CD8B), T-cell surface glycoprotein CD3 gamma chain (CD3G), small and large ribosomal subunits RPS23 and RPL22 with interaction score 0.961 – 0.999. Among the enriched reactome pathway, interferon

signaling involving 22 genes namely USP18, OAS3, EIF2AK2, TRIM21, OASL, HERC5, RPS27A, UBE2L6, STAT2, MX2, OAS2, STAT1, IFIT1, IFIT3, IFIT2, ISG15, DDX58, RSAD2, IRF7, MX1, OAS1 and IFI35 were the key players in the pathogenesis of currently described four disease types **(Figure 4D and Table 4)**. Defense response to virus was the most significantly expressed biological process common to these diseases, associated with 22 genes including OAS3, EIF2AK2, OASL, RTP4, IFIH1, HERC5, CXCL10, STAT2, MX2, OAS2, STAT1, IFI44L, IFIT1, IFIT3, IFIT2, ISG15, DDX58, TLR7, RSAD2, IRF7, MX1 and OAS1 **(Table 4)**.

The DG disease category comprised of 1766 PPIs **(Figure 5A)**, with an enrichment of 10^{-16}, the most interacting proteins within this network included small and large ribosomal subunits involved in protein synthesis (RPS23, RPS27A, RPL13A, RPL22 and RPL37) with interaction score 0.999. The MT disease category demonstrated 47 PPIs with an enrichment of 5.89×10^{-9}(Figure 5B); the most interacting proteins within this network included core components of nucleosome H2A histone family member J (H2AFJ), histone cluster 2 (HIST2H2AC and HIST2H2AA3), interferon-induced protein with tetratricopeptide repeats 3 as an inhibitor of cellular as well as viral processes (IFIT3), interferon alpha-inducible protein 27 (IFI27), 59 kDa 2'-5'-oligoadenylate synthase-like protein (OASL), serpin peptidase inhibitor (SERPING1) and complement component 1 (C1QB), with interaction score 0.961 – 0.995. The ML disease expressed 50 PPIs with an enrichment of 2.36×10^{-16} (Figure 5C); the interaction score ranged from 0.96 to 0.979 for the mostly ML-associated proteins including mitochondrial glycine degrading aminomethyltransferase (AMT), bifunctional methylenetetrahydrofolate dehydrogenase/cyclohydrolase (MTHFD2), SERPING1, C1QB, HIST2H2AC, HIST2H2AA3, interferon-induced 35 kDa protein (IFI35) and IFI27.
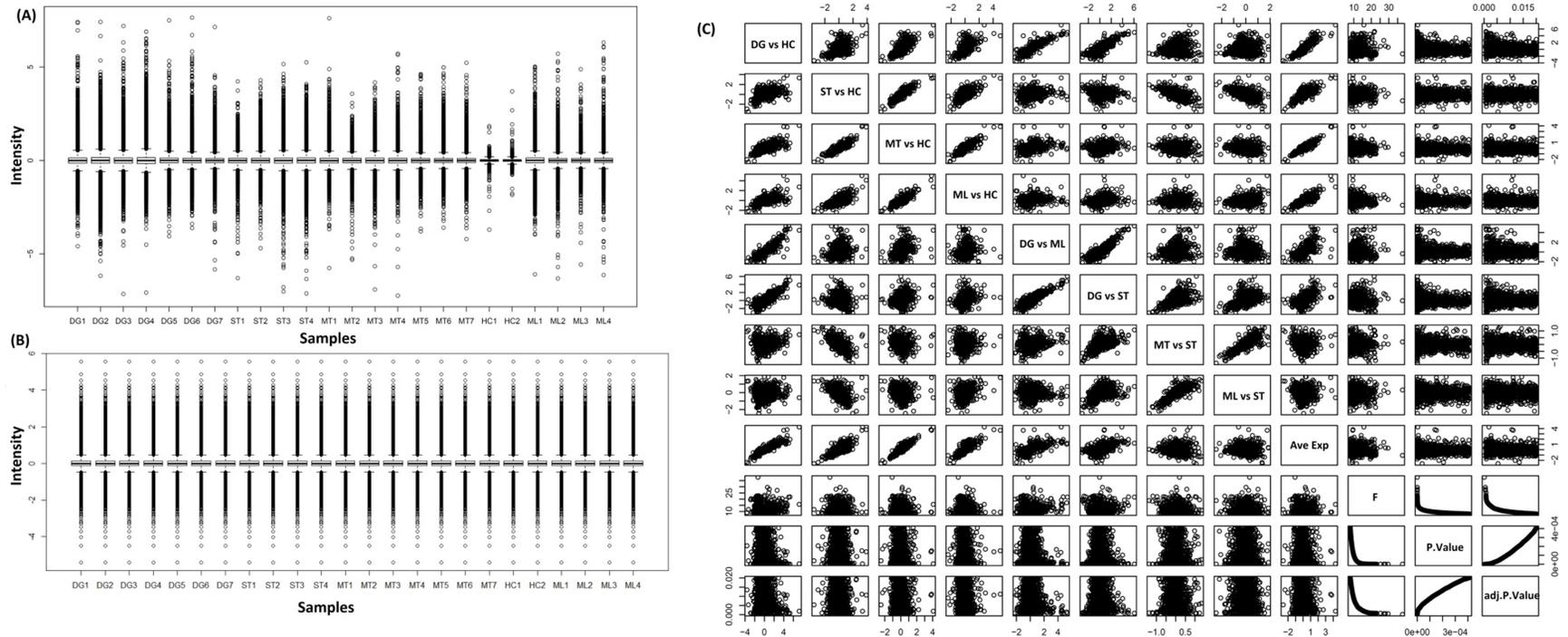
Figure 1: Microarray expression of RNA signal intensity from peripheral blood mononuclear cells in scrub typhus (ST; n=4), dengue (DG; n=7), murine typhus (MT; n=7), malaria (ML; n= 4) and healthy controls (HC; n=2): (A) before and (B) after normalization; (C) empirical Bayes moderated t distribution
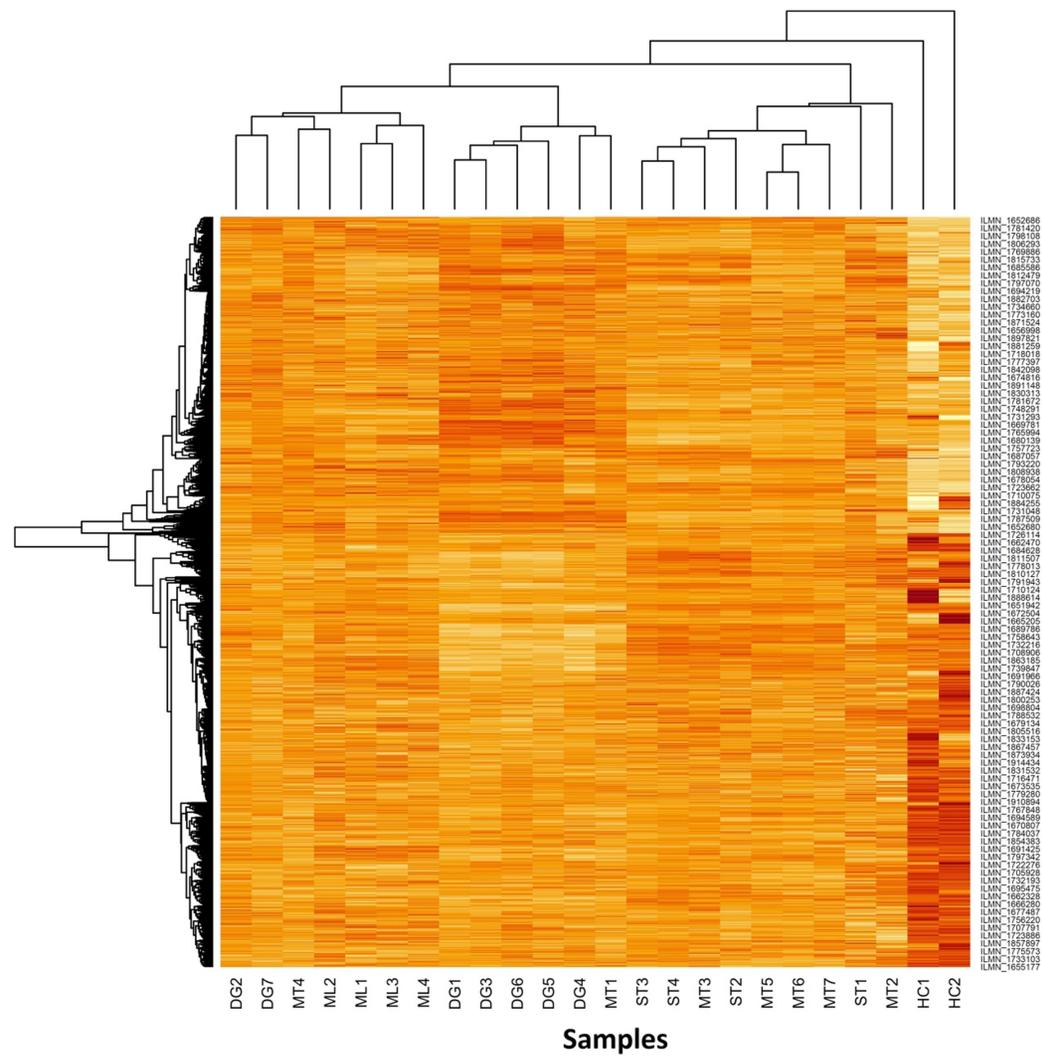
Figure 2: Heatmap of normalized microarrays in ST (scrub typhus), DG (dengue fever), MT (murine typhus), ML (malaria) and HC (healthy control)
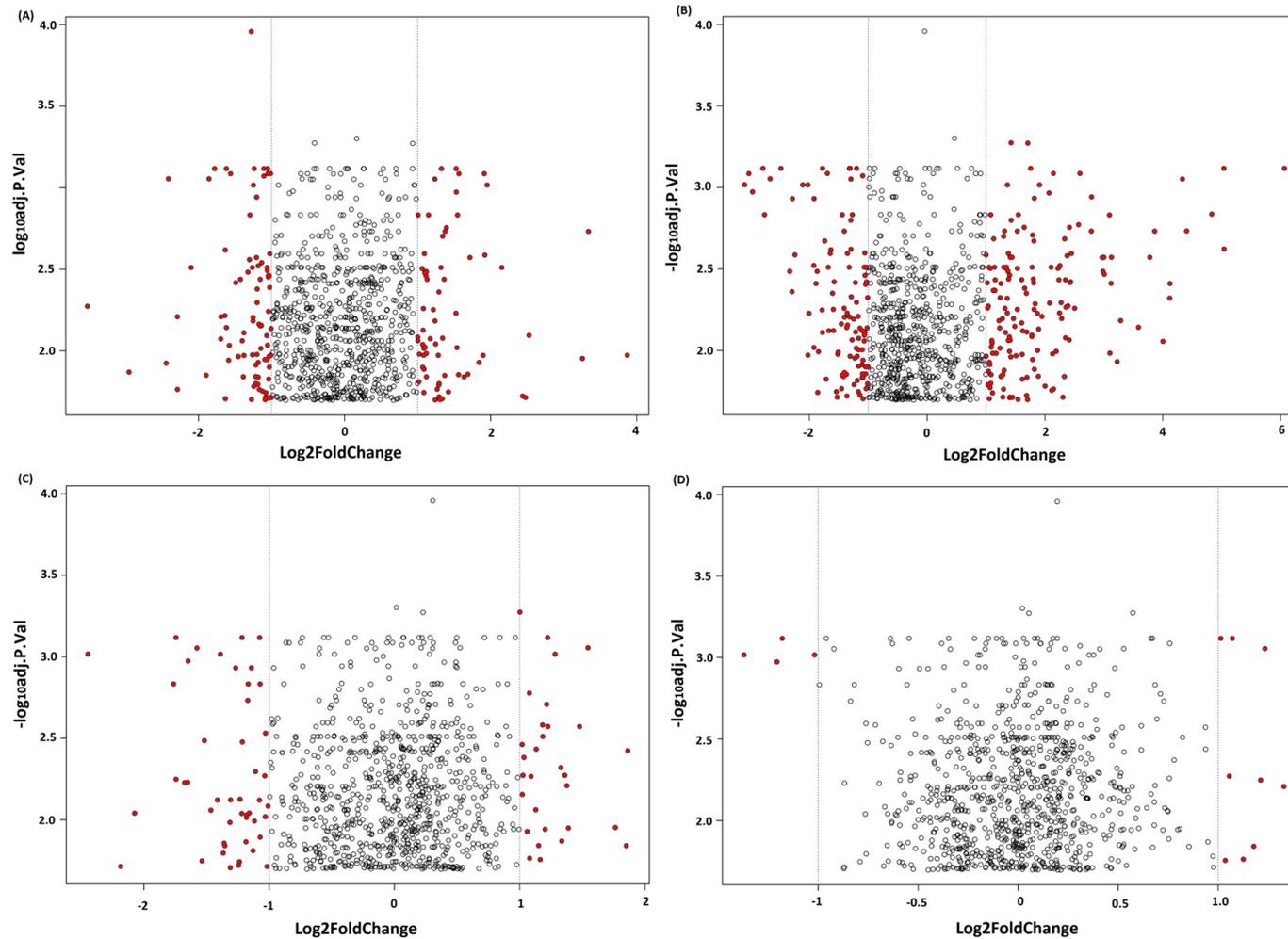
Figure 3: Volcano plots representing upregulated and downregulated expression of genes in (A) scrub typhus, (B) dengue, (C) malaria, (D) murine typhus, compared to healthy controls
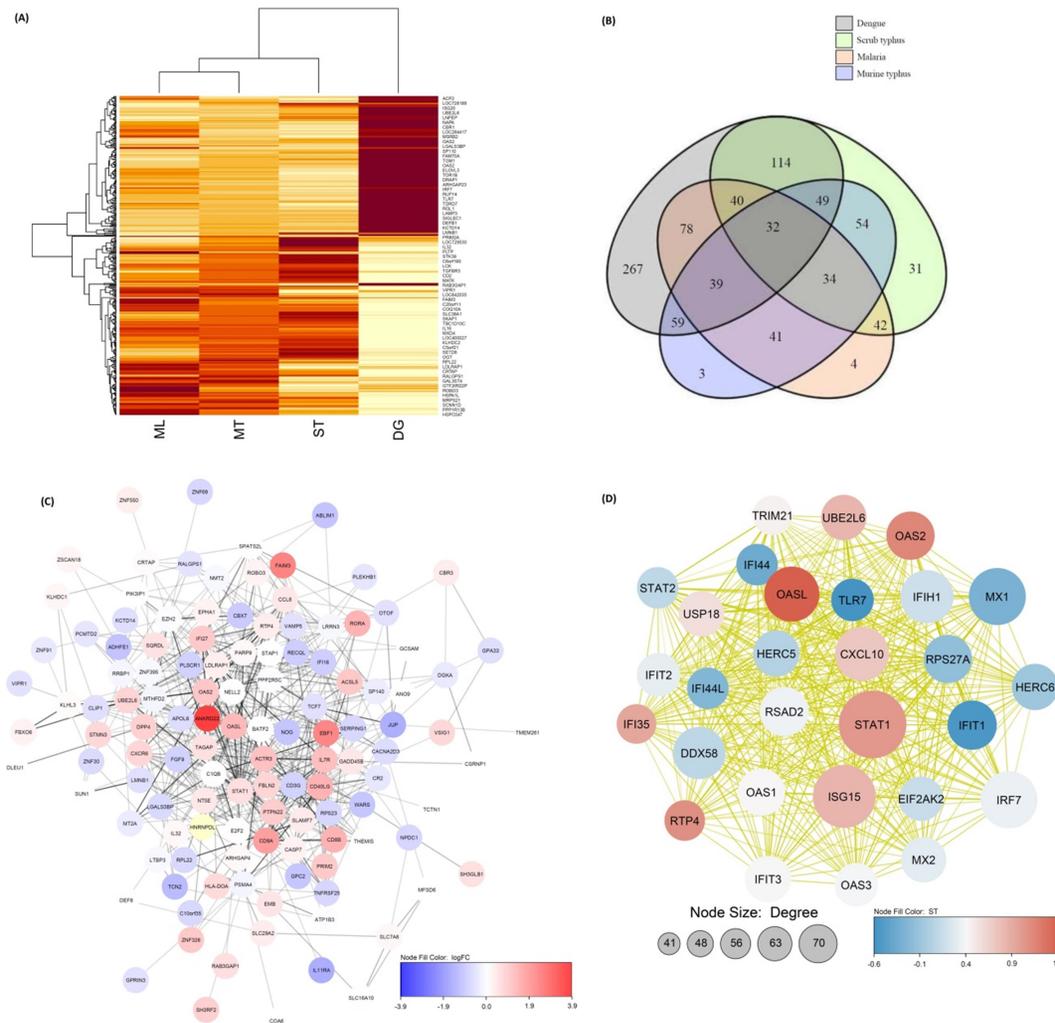
Figure 4: (A) Heatmap of differentially expressed genes among the ST, DG, ML and MT disease groups compared to health controls; (B) Venn diagram showing number of genes belonging to ST, DG, MT and ML disease groups; (C) protein-protein interaction (PPI) in ST; (D) top ranked 29 hub genes in ST, DG, MT and ML disease groups by maximal clique centrality
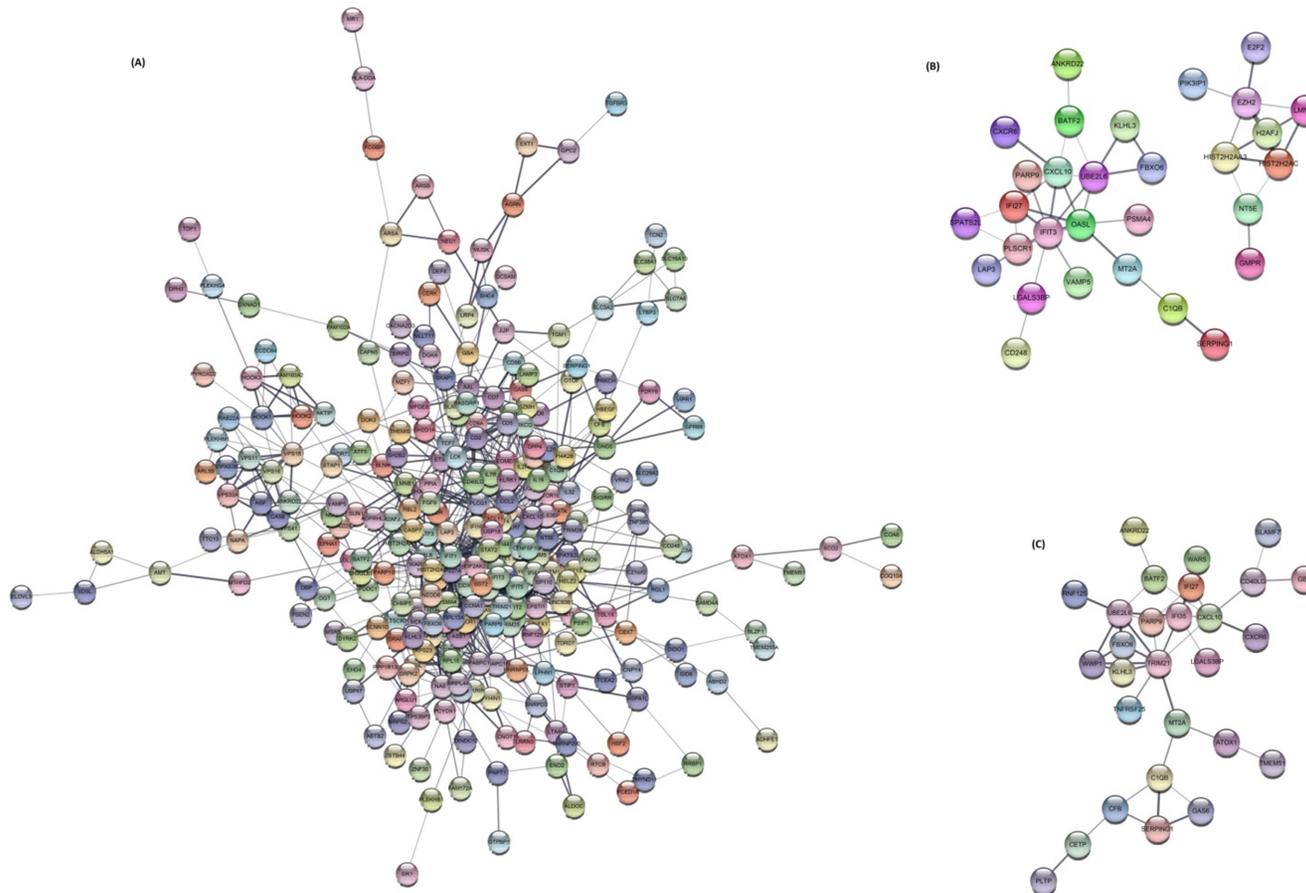
Figure 5: Protein-protein interaction (PPI) in (A) dengue, (B) murine typhus, (C) malaria

Table 1: Enriched GO processes and reactome pathways for DEGs in scrub typhus

Description	FDR	#Genes	Genes
I. GO process			
Biological regulation	0.0244	97	GADD45B, PIK3IP1, CBX7, HLA-DOA, MT2A, OASL, NT5E, RALGPS1, LMNB1, RORA, LGALS3BP, ACTR3, TMEM175, STAP1, DEF8, OTOF, CSRNP1, EPHA1, SERPING1, ATP1B3, UBE2L6, GPC2, RPS23, LTBP3, BATF2, IL7R, KLHL3, C1QB, EZH2, EBF1, CRTAP, VIPR1, SMDT1, NOG, DGKA, CD8B, ANO9, TCF7, ZNF326, OAS2, PLSCR1, RPL22, PLEKHB1, WARS, ACSL5, PTPN22, PARP9, DPP4, STAT1, E2F2, CR2, TAGAP, FAIM3, GPA33, SLAMF7, IFI16, THEMIS, SLC16A10, ARHGAP4, STMN3, CD40LG, NPDC1, LDLRAP1, FBXO6, TNFRSF25, RRBP1, TMEM8B, NMT2, FGF9, SP140, JUP, CCL8, TCTN1, ROBO3, C21orf2, WDR19, SUN1, FBLN2, CD8A, CXCR6, LRRN3, VSIG1, ZNF30, RAB3GAP1, PPP2R5C, NELL2, CACNA2D3, GCSAM, SH3RF2, CD3G, IL32, IL11RA, CLIP1, SH3GLB1, HNRNPDL, IFI27, PRIM2
Cell surface receptor signalling pathway	0.0024	34	MT2A, OASL, LMNB1, RORA, ACTR3, STAP1, CSRNP1, EPHA1, GPC2, IL7R, VIPR1, NOG, CD8B, TCF7, OAS2, PTPN22, STAT1, CR2, THEMIS, CD40LG, TNFRSF25, FGF9, CCL8, ROBO3, C21orf2, WDR19, CD8A, CXCR6, RAB3GAP1, CD3G, IL32, IL11RA, HNRNPDL, IFI27
Immune response	0.0071	25	HLA-DOA, MT2A, OASL, RORA, ACTR3, SERPING1, IL7R, C1QB, CD8B, TCF7, OAS2, PARP9, STAT1, CR2, SLAMF7, IFI16, THEMIS, CD40LG, TNFRSF25, JUP, CCL8, CD8A, CD3G, IL32, IFI27
Defense response	0.0024	24	MT2A, OASL, RTP4, LGALS3BP, ACTR3, SERPING1, BATF2, C1QB, OAS2, PLSCR1, PARP9, DPP4, STAT1, CR2, FAIM3, SLAMF7, IFI16, CD40LG, TNFRSF25, SP140, CCL8, CXCR6, IL32, IFI27
Regulation of immune response	0.0026	19	ACTR3, STAP1, SERPING1, IL7R, C1QB, CD8B, PLSCR1, PTPN22, PARP9, STAT1, CR2, SLAMF7, IFI16, THEMIS, CD40LG, NPDC1, CD8A, GCSAM, CD3G
Cytokine-mediated signalling pathway	0.0072	15	MT2A, OASL, LMNB1, RORA, IL7R, OAS2, STAT1, CD40LG, TNFRSF25, CCL8, CXCR6, IL32, IL11RA, HNRNPDL, IFI27
Lymphocyte activation	0.0024	13	RORA, IL7R, CD8B, TCF7, RPL22, PTPN22, DPP4, CR2, SLAMF7, THEMIS, CD40LG, CD8A, CD3G
Protein localization to membrane	0.0384	10	RTP4, ATP1B3, RPS23, VAMP5, RPL22, JUP, WDR19, RAB3GAP1, CLIP1, SH3GLB1
II. Reactome Pathways			
Immune system	0.0034	30	PSMA4, HLA-DOA, MT2A, OASL, LMNB1, RORA, ACTR3, SERPING1, UBE2L6, IL7R, KLHL3, C1QB, CD8B, OAS2, PTPN22, STAT1, CR2, SLAMF7, IFI16, CD40LG, NPDC1, FBXO6, TNFRSF25, JUP, CD8A, PPP2R5C, CD3G, IL11RA, HNRNPDL, IFI27
Cytokine signalling in immune system	0.0313	14	PSMA4, MT2A, OASL, LMNB1, RORA, UBE2L6, IL7R, OAS2, STAT1, CD40LG, TNFRSF25, IL11RA, HNRNPDL, IFI27
Immunoregulatory interactions between a lymphoid and a non-lymphoid cell	0.0393	6	CD8B, SLAMF7, CD40LG, NPDC1, CD8A, CD3G

Table 2: Enriched reactome pathways for DEGs in dengue fever

Reactome Pathways	FDR	#Genes	Genes
Immune System	1.35E-12	94	PSMA4, USP18, ARSA, IL2RB, RNF125, BLNK, CCL2, ALDOC, UNC93B1, LTA4H, OAS3, HLA-DOA, LNPEP, EIF2AK2, PLCG1, MT2A, DHX58, BST2, TRIM21, VHL, OASL, LMNB1, PRKCQ, IFIH1, CBLB, HERC6, HERC5, ASB1, ARSB, PML, RPS27A, SERPING1, UBE2L6, SIGLEC12, DEFB1, IL16, CXCL10, IL7R, ISG20, RASGRP1, KLHL3, C1QB, STAT2, TRIM25, CD8B, MX2, LCK, SIGLEC1, OAS2, IFI6, IL27, DOK3, STAT1, SLAMF7, IFI16, MB21D1, CD40LG, SH2D1A, ZBP1, IFIT1, IFIT3, IFIT2, NEU1, FBXO6, TNFRSF25, ISG15, DDX58, TRIM5, TLR7, RSAD2, STAT4, JUP, LGALS9, IRF7, SH3KBP1, MX1, CDKN1A, IFITM1, CD8A, OAS1, IFI35, SIGIRR, TRIM26, USP41, CFB, PPIA, WWP1, ATOX1, TBC1D10C, CYFIP2, IL11RA, CCL5, HNRNPDL, IFI27
Innate Immune System	0.034	38	PSMA4, ARSA, RNF125, ALDOC, UNC93B1, LTA4H, PLCG1, DHX58, BST2, TRIM21, PRKCQ, IFIH1, HERC5, ARSB, RPS27A, SERPING1, UBE2L6, DEFB1, RASGRP1, C1QB, TRIM25, LCK, DOK3, IFI16, MB21D1, ZBP1, NEU1, ISG15, DDX58, TLR7, JUP, IRF7, SIGIRR, CFB, PPIA, ATOX1, TBC1D10C, CYFIP2
Interferon signalling	3.05E-17	34	USP18, OAS3, EIF2AK2, PLCG1, MT2A, BST2, TRIM21, OASL, HERC5, PML, RPS27A, UBE2L6, ISG20, STAT2, TRIM25, MX2, OAS2, IFI6, STAT1, IFIT1, IFIT3, IFIT2, ISG15, DDX58, TRIM5, RSAD2, IRF7, MX1, IFITM1, OAS1, IFI35, TRIM26, USP41, IFI27
Signalling by Interleukins	0.0361	21	PSMA4, IL2RB, BLNK, CCL2, LMNB1, RPS27A, IL16, CXCL10, IL7R, STAT2, LCK, IL27, STAT1, STAT4, LGALS9, CDKN1A, SIGIRR, PPIA, IL11RA, CCL5, HNRNPDL
ISG15 antiviral mechanism	1.85E-7	14	USP18, EIF2AK2, PLCG1, HERC5, RPS27A, UBE2L6, TRIM25, MX2, STAT1, IFIT1, ISG15, DDX58, MX1, USP41
Interferon gamma signalling	5.2E-4	11	OAS3, MT2A, TRIM21, OASL, PML, TRIM25, OAS2, TRIM5, IRF7, OAS1, TRIM26
L13a-mediated translational silencing of ceruloplasmin expression	0.0413	9	RPS27A, RPL37, RPS23, RPL15, PABPC1, RPL22, RPL13A, EIF4B, EIF3L
Negative regulators of DDX58/IFIH1 signalling	2.0E-4	8	RNF125, IFIH1, HERC5, RPS27A, UBE2L6, TRIM25, ISG15, DDX58
OAS antiviral response	6.2E-4	5	OAS3, OASL, OAS2, DDX58, OAS1
TRAF3-dependent IRF activation pathway	0.034	4	IFIH1, TRIM25, DDX58, IRF7

Table 3: Enriched reactome pathways for DEGs in malaria and murine typhus

Reactome Pathways	FDR	#Genes	Genes
I. Malaria			
Immune System	0.0022	21	RNF125, MT2A, TRIM21, LMNB1, SERPING1, UBE2L6, CXCL10, KLHL3, C1QB, SLAMF7, CD40LG, FBXO6, TNFRSF25, LGMN, JUP, IFI35, CFB, WWP1, ATOX1, IL11RA, IFI27
Cytokine Signaling in Immune system	0.0328	10	MT2A, TRIM21, LMNB1, UBE2L6, CXCL10, CD40LG, TNFRSF25, IFI35, IL11RA, IFI27
II. Murine typhus			
DNA Damage/Telomere Stress Induced Senescence	0.0027	4	LMNB1, HIST2H2AC, H2AFJ, HIST2H2AA3
Immune System	0.003	16	PSMA4, MT2A, OASL, LMNB1, SERPING1, UBE2L6, CXCL10, KLHL3, C1QB, SLAMF7, NPDC1, IFIT3, FBXO6, JUP, IL11RA, IFI27
Interferon Signaling	0.0045	5	MT2A, OASL, UBE2L6, IFIT3, IFI27
Metabolism of vitamins and cofactors	0.0045	5	TCN2, NT5E, GPC2, PARP9, MTHFD2
Amyloid fiber formation	0.0134	3	UBE2L6, HIST2H2AC, HIST2H2AA3
Nicotinate metabolism	0.0263	2	NT5E, PARP9
Developmental Biology	0.0464	8	PSMA4, ABLIM1, EZH2, HIST2H2AC, JUP, NELL2, H2AFJ, HIST2H2AA3
Regulation of Complement cascade	0.0498	2	SERPING1, C1QB

Table 4: Hub genes of ST, DG, ML and MT disease groups

Description	FDR	#Genes	Genes
I. GO process			
Defense response to virus	1.00E-35	22	OAS3, EIF2AK2, OASL, RTP4, IFIH1, HERC5, CXCL10, STAT2, MX2, OAS2, STAT1, IFI44L, IFIT1, IFIT3, IFIT2, ISG15, DDX58, TLR7, RSAD2, IRF7, MX1, OAS1
Negative regulation of viral process	8.72E-15	10	OAS3, EIF2AK2, TRIM21, OASL, STAT1, IFIT1, ISG15, RSAD2, MX1, OAS1
Regulation of response to biotic stimulus	1.01E-08	10	USP18, IFIH1, HERC5, RPS27A, STAT1, IFIT1, DDX58, TLR7, RSAD2, IRF7
Regulation of type I interferon production	3.26E-12	9	TRIM21, IFIH1, HERC5, UBE2L6, STAT1, ISG15, DDX58, TLR7, IRF7
Protein modification by small protein conjugation or removal	5.80E-04	8	USP18, TRIM21, IFIH1, HERC6, HERC5, RPS27A, UBE2L6, DDX58
Regulation of protein modification process	0.0214	8	EIF2AK2, TRIM21, HERC5, RPS27A, CXCL10, STAT2, ISG15, TLR7
Interferon-gamma-mediated signalling pathway	4.31E-10	7	OAS3, TRIM21, OASL, OAS2, STAT1, IRF7, OAS1
Response to interferon-alpha	1.04E-08	5	EIF2AK2, MX2, IFIT3, IFIT2, OAS1
Response to bacterium	0.0094	5	HERC6, CXCL10, OAS2, IFI44, ISG15
II. Reactome pathways			
Interferon signalling	9.80E-36	22	USP18, OAS3, EIF2AK2, TRIM21, OASL, HERC5, RPS27A, UBE2L6, STAT2, MX2, OAS2, STAT1, IFIT1, IFIT3, IFIT2, ISG15, DDX58, RSAD2, IRF7, MX1, OAS1, IFI35
OAS antiviral response	5.65E-10	5	OAS3, OASL, OAS2, DDX58, OAS1
Negative regulators of DDX58/IFIH1 signalling	9.64E-10	6	IFIH1, HERC5, RPS27A, UBE2L6, ISG15, DDX58
TRAF6 mediated IRF7 activation in TLR7/8 or 9 signalling	7.18E-05	3	RPS27A, TLR7, IRF7
TRAF3-dependent IRF activation pathway	7.18E-05	3	IFIH1, DDX58, IRF7
Antigen processing: ubiquitination & proteasome degradation	0.0017	5	TRIM21, HERC6, HERC5, RPS27A, UBE2L6
Ub-specific processing proteases	0.0039	4	USP18, IFIH1, RPS27A, DDX58
Regulation of innate immune responses to cytosolic DNA	0.0046	2	TRIM21, RPS27A
Regulation of IFNA signalling	0.0084	2	USP18, STAT2
Interleukin-20 family signalling	0.0092	2	STAT2, STAT1
Influenza Infection	0.0168	3	EIF2AK2, RPS27A, ISG15
Signalling by NOTCH3	0.0264	2	RPS27A, STAT1

DISCUSSION

We performed a comparative analysis of microarray gene expression profiles from GSE16463 dataset using bioinformatics approaches, in connection with *O. tsutsugamushi* infection in human monocytes to identify the transcriptional signatures from gene interaction networks in ST discriminated from other acute febrile illnesses including DG, ML, MT accompanied by gene ontology and functional pathway enrichment analysis in disease pathogenesis.

It has been reported that in a network-based approach, 27 top ranked candidate genes were enriched in malaria patients through pathway analysis, with the expression of pyruvate metabolism, TCA cycle, integrin and vascular wall cell surface interactions, along with platelet adhesion to exposed collagen [13]. The complement system as a link between the adaptive and innate immune system, was implicated in viral antagonism and immune evasion; human proteins linked to the complement and coagulation cascade, the centrosome, and the cytoskeleton were enriched among the dengue viral interaction partners [14]. In ST, coagulation activation was prominent and linked to a strong proinflammatory response with high level of inflammatory cytokines expression (except IL12) compared to MT cases that displayed expression of high level of endothelium-

derived markers, as demonstrated in the early phase of infection with *R. typhi* and *O. tsutsugamushi* contributing to disease differentiation [15].

Earlier Park *et al*. [16] reported 70 upregulated proteins and 94 downregulated proteins ensuing with the *O. tsutsugamushi* infection compared to normal subjects, with the expression of proteins involved in immune responses, particularly the acute phase response signaling, complement system, LXR/RXR activation, FXR/RXR activation and coagulation system upregulation. The pathways enriched in differentially expressed genes in scrub typhusincluded surface proteins and adhesins, DNA replication, secreted effector protein, RAGE (rickettsial-amplified genetic element), and metabolic pathway, with upregulation of IFNB1 (interferon beta 1) gene, genes involved in regulating the type-I interferon response: IRF9 and STAT1/2, and interferon-stimulated genes (IFIT, OAS1) as well as proinflammatory chemokine (CXCL10, CXCL11) and cytokine receptor (IL13RA2, IL7R, IL15RA, IL3RA) genes [17]. Bacterial organisms stimulated the expression of a large panel of genes including type I interferon, interferon-stimulated, inflammatory, apoptosis-related genes and induced an M1-type gene response in macrophages, accounting for the local and systemic inflammation

observed in ST, and that interferon-gamma may be useful as an adjuvant treatment of patients with scrub typhus **[18]**.

In the current study, comparison of 635 gene sets screened from RNA microarrays of ST, DG, ML, MT cases, and healthy controls suggested that a subset of 31 genes were found unique to the ST group, which could be used as specific signature of scrub typhus significantly overrepresenting translocation of ZAP-70 to immunological synapse, and phosphorylation of CD3 and TCR zeta chains, involving PTPN22 and CD3G genes. The representative biomarkers of *O. tsutsugamushi* infection found in the current study will help in differential diagnosis of ST mimicking other acute febrile illnesses. These genes could be useful for investigating the molecular pathophysiology of ST and for discovering novel drug targets and vaccine developments.

Author's contribution: Both MM and SM designed the study, analysed and interpreted data, discussed and wrote the manuscript.

Declaration of competing interests: There is no conflict of interest by the authors.

REFERENCES

[1] Astrup E, Janardhanan J, Otterdal K, Ueland T, Prakash JAJ, Lekva T, *et al*. Cytokine network in scrub typhus: high levels of interleukin-8 are associated with disease severity and mortality. *PLoS Negl Trop Dis* 2014; **8**: e2648. DOI: 10.1371/journal.pntd.0002648

[2] Centers for Disease Control and Prevention. Scrub typhus. CDC 2020. Available from: https://www.cdc.gov/typhus/scrub/index.html (Accessed on December 14, 2020).

[3] Diaz FE, Abarca K, Kalergis AM. An update on host-pathogen interplay and modulation of immune responses during *Orientia tsutsugamushi* infection. *Clin Microbiol Rev* 2018. **31**: e00076-17. DOI: 10.1128/CMR.00076-17

[4] Tantibhedhyangkul W, Prachason T, Waywa D, El Filali A, Ghigo E, Thongnoppakhun W, *et al*. *Orientia tsutsugamushi* Stimulates an Original Gene Expression Program in Monocytes: Relationship with Gene Expression in Patients with Scrub Typhus. *PLoS Negl Trop Dis* 2011; **5**: e1028. DOI: 10.1371/journal.pntd.0001028

[5] Valbuena G, Walker DH. Approaches to vaccines against

Orientia tsutsugamushi. Frontiers Cell Infect Microbiol 2013; **2**. DOI: 10.3389/fcimb.2012.00170

[6] Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M**,** *et al*. NCBI GEO: Archive for functional genomics data sets-update. *Nucleic Acids Res* 2013; **41**: D991-5. DOI: 10.1093/nar/gks1193

[7] Huber W, Carey VJ, Gentleman R, Anders S, Carlson M, Carvalho B, *et al*. Orchestrating high-throughput genomic analysis with Bioconductor. *Nat Methods* 2015; **12**: 115-121.

[8] Mandal M, Mandal S. Computational approaches towards micro-RNA expression analysis for identification of biomarkers in inflammatory bowel disease subtypes. *Int J Clin Exp Physiol* 2020; **7**: 100-106.

[9] Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, *et al*. Cytoscape: A software environment for integrated models of biomolecular interaction net-works. *Genome Res*. 2003; **13**: 2498-2504.

[10] Doncheva NT, Morris JH, Gorodkin J, Jensen LJ. Cytoscape STRING App: Network analysis and visualization of proteomics data. *J Proteome Res* 2019; **18**: 623- 632.

[11] Chin CH, Chen SH, Wu HH, Ho CW, Ko MT, Lin CY, *et al*. CytoHubba: Identifying hub objects and sub-networks from complex interactome. *BMC Syst Biol* 2014; **8**: S11. DOI: 10.1186/1752-0509-8-S4-S11

[12] Bader GD, Hogue CW. An automated method for finding molecular complexes in large protein interaction networks. *BMC Bioinformatics* 2003; **4**: 1-27.

[13] Chen Y, Xu R. Network-based gene prediction for *Plasmodium falciparum* malaria towards genetics - based drug discovery. *BMC Genomics* 2015; **16:** S9. DOI: 10.1186/1471-2164-16-S7-S9

[14] Khadka S, Vangeloff AD, Zhang C, Siddavatam P, Heaton NS, Wang L, *et al*. A physical interaction network of dengue virus and human proteins. *Mol Cell Proteomics* 2011; **10**: M111.012187. DOI: 10.1074/mcp.M111.012187.

[15] Paris DH, Chansamouth V, Nawtaisong P, Lowenberg EC, Phetsouvanh R, Blacksell SD, *et al*. Coagulation and inflammation in scrub typhus and murine

typhus—a prospective comparative study from Laos. *Clin Microbiol Infect* 2012; **18**: 1221-1228.

[16] Park EC, Lee SY, Yun SH, Choi CW, Lee H, Song HS, *et al*. Clinical proteomic analysis of scrub typhus infection. *Clin Proteom* 2018**;** *15*: 6. DOI: 10.1186/s12014-018-9181-5

[17] Mika-Gospodorz B, Giengkam S, Westermann AJ, Wongsantichon J, Kion-Crosby W, Chuenklin S, *et al*. Dual RNA-seq of *Orientia*

tsutsugamushi informs on host-pathogen interactions for this neglected intracellular human pathogen. *Nat Comm* 2020; **11**: 3363. DOI: 10.1038/s41467-020-17094-8

[18] Tantibhedhyangkul W, Ben Amara A, Textoris J, Gorvel L, Ghigo E, Capo C, *et al*. *Orientia tsutsugamushi*, the causative agent of scrub typhus, induces an inflammatory program in human macrophages. *Microb Pathog* 2013; **55**: 55-63.