Heliyon 9 (2023) e18175

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Gene co-expression network identifies critical genes, pathways and regulatory motifs mediating the progression of rift valley fever in *Bos taurus*

John K. Gitau [a,*], Rosaline W. Macharia [a], Kennedy W. Mwangi [c], Nehemiah Ongeso [a], Edwin Murungi [b]

[a] *University of Nairobi, Biochemistry Department, P.O Box 30197, 00100, Nairobi, Kenya*
[b] *Kisii University, Department of Medical Biochemistry, P.O Box 408, 40200, Kisii, Kenya*
[c] *Jomo Kenyatta University of Agriculture and Technology, P.O Box 62000, 00200, Nairobi, Kenya*

ARTICLE INFO

Keywords:
Rift valley fever
Bos taurus
weighted gene co-expression network
Modules
Hub genes
Regulatory motifs
Innate immune response

ABSTRACT

Rift Valley Fever (RVF) is a mosquito-borne viral disease caused by the Rift Valley Fever Virus. The disease is a zoonosis that largely affects domestic animals, including sheep, goats, and cattle, resulting in severe morbidity and mortality marked by massive storm abortions. To halt human and livestock deaths due to RVF, the development of efficacious vaccines and therapeutics is a compelling and urgent priority. We sought to identify potential key modules (gene clusters), hub genes, and regulatory motifs involved in the pathogenesis of RVF in *Bos taurus* that are amenable to inhibition. We analyzed 39 *Bos taurus* RNA-Seq samples using the weighted gene co-expression network analysis (WGCNA) R package and uncovered significantly enriched modules containing genes with potential pivotal roles in RVF progression. Moreover, regulatory motif analysis conducted using the Multiple Expectation Maximization for Motif Elicitation (MEME) suite identified motifs that probably modulate vital biological processes. Gene ontology terms associated with identified motifs were inferred using the GoMo human database. The gene co-expression network constructed in WGCNA using 5000 genes contained seven (7) modules, out of which four were significantly enriched for terms associated with response to viruses, response to interferon-alpha, innate immune response, and viral defense. Additionally, several biological pathways implicated in developmental processes, anatomical structure development, and multicellular organism development were identified. Regulatory motifs analysis identified short, repeated motifs whose function(s) may be amenable to disruption by novel therapeutics. Predicted functions of identified motifs include tissue development, embryonic organ development, and organ morphogenesis. We have identified several hub genes in enriched co-expressed gene modules and regulatory motifs potentially involved in the pathogenesis of RVF in *B. taurus* that are likely viable targets for disruption by novel therapeutics.

1. Introduction

Rift Valley Fever (RVF) is a zoonotic disease caused by the Rift Valley Fever virus (RVFV), a triple segmented negative strand RNA

* Corresponding author.
E-mail address: johngitau51@gmail.com (J.K. Gitau).

https://doi.org/10.1016/j.heliyon.2023.e18175
Received 1 August 2022; Received in revised form 4 July 2023; Accepted 10 July 2023
Available online 14 July 2023

virus endemic throughout Africa [1]. Although RVF primarily infects domestic animals including sheep, goats, and cattle, resulting in wide ranging clinical symptoms including massive storm abortions and neonatal fatalities, it also causes acute febrile, largely self-limiting infection in humans [2]. However, 1–2% of infected individuals develop severe illnesses characterised by retinitis, hepatitis, delayed onset encephalitis, haemorrhagic diatheses, and miscarriage in pregnant women. Mosquitoes act as RVFV vectors and reservoirs [3], with the *Aedes* species enabling transovarial transmission while *Culex* and *Anopheles* species promote transmission to new hosts during routine blood-feeding. RVF outbreaks are closely linked to torrential rainfall and flooding, which provide an ideal environment for mosquito breeding and multiplication [1].

Although RVFV is known to infect an array of host cells, including hepatocytes [4], endothelial cells [5], and macrophages [6], the particular mechanism through which the virus enters and replicates within these cells is unknown. The host's immunity mounts robust responses following infection by deploying innate and adaptive components [5,6]. Decoding the molecular mechanism of host-RVFV interaction is critical for the design of effective interventions encompassing vaccines and antiviral drugs. A systems biology approach may be used to decipher distinct cellular and immune signaling pathways potentially involved in RVFV pathogenesis that could be targeted for disruption. Specifically, the popular Weighted Gene Co-expression Network Analysis (WGCNA) algorithm may be utilized to analyze RNA-Seq data obtained from hosts exposed to RVFV to uncover gene interactions and molecular relationships correlated with RVFV infection and pathogenesis. While differential gene expression analysis is focused on identifying individual genes that are upregulated or downregulated between different tissue states [7], such as diseased and normal, the WGCNA provides a broader perspective by analyzing the complex interactions among gene clusters [8,9]. By identifying groups of genes that are functionally related and work together in similar biological processes, WGCNA offers a more comprehensive understanding of the underlying molecular mechanisms that underlie bovine response to RVFV invasion.

The depiction of gene interactions and molecular relationships using biological networks is important for the understanding of biological systems. Informative biological systems include protein-protein interactions (PPIs) networks that illustrate protein interactions in cellular systems, metabolic networks that depict metabolic fluxes, and gene regulatory networks that represent the interactions of different cellular modulators of gene expression and protein synthesis [10,11]. In gene co-expression networks, genes are represented by nodes while the existence of co-expression relationships between genes is depicted as edges [9,12], with genes with similar expression profiles across samples are connected [9]. Co-expression networks aid in the understanding of system-level molecular interactions driving cellular processes. The modules (clusters) contained in these networks generally contain genes with similar expression patterns that plausibly modulate analogous biological processes [12,13]. Moreover, modular genes tend to contain similar regulatory motifs, short nucleotide sequences that regulate gene expression by binding transcription factors [12]. Transcription factors have been shown to be tractable drug and vaccine targets for cancer and parasitic diseases [13].

We have analyzed 39 RNA-Seq samples obtained from RVFV infected bovine cells using WGCNA and delineated clusters of functionally related genes co-expressed across the various samples. The identified gene clusters have allowed the unravelling of previously undetermined biological pathways and processes with potentially important roles in RVFV establishment in hosts. Notably, WGCNA enabled the determination of key (hub) regulatory genes that putatively modulate the expression of other genes within the network. The identified regulatory motifs may provide insights into gene expression modulation under different conditions and thus offer an avenue for interfering with cellular establishment of RVFV. WGCNA provides robust module preservation statistics that may be used to quantify similarity between conditions using the guilt-by-association approach [8,14], and investigate differences in the modular structure of networks. By leveraging guilt-by-association, previously unknown functions of RVFV genes may be inferred thus deepening the understanding of the biology of RVF infection. Collectively, we have deciphered genes, biological pathways, and regulatory motifs that potentially enable establishment and progression of RVF infection in *Bos taurus*. Our results provide a platform for further exploration and validation of vital genes involved in the establishment of RVF infection for the purposes of developing novel intervention strategies.

2. Materials and methods

2.1. Retrieval of RVF gene expression profiles

RVF gene expression profiles were obtained from the Gene Expression Omnibus database [15] (https://www.ncbi.nlm.nih.gov/geo/). The original study (GSE71417) used healthy 4-6-month-old *Bos taurus* heifer and steer calves housed in an ABSL2 Ag biocontainment facility. The animals were randomised into test groups and acclimated to the facility for 14 days prior to inoculation with the live, attenuated RVFV MP-12 vaccine. Specifically, inoculation involved administration of 1×105 PFU of MP-12 subcutaneously or intramuscularly to 3–5 animals in each group. On Days 0–7, 10, 14, and 21, whole blood was obtained from the animals for serum neutralisation studies (PRNT) or total RNA purification for RNAseq analysis. To minimize background noise in the gene co-expression network constructed in this study, non-coding and lowly expressed genes were removed following quality control, read mapping, and read count normalisation procedures. R version 4.0 (https://www.r-project.org/) was used to construct the gene network.

2.2. Construction and visualisation of a Bos taurus weighted gene co-expression network

The gene co-expression network was analyzed using the WGCNA version 1.69 R package [16], allowing the identification of modules containing genes with probable vital roles in the progression of RVF infection. Sample clustering was performed to identify and exclude outlier samples (Additional File 1). Scale independence and mean connectivity analysis (Additional File 2) were conducted on modules with different power values to determine the optimal soft threshold for module analysis, with power values ranging from 1

to 20. A power value was selected at a scale independence value of 0.9. Similar modules were merged using a cut height value of 0.3 (Additional File 3) and assigned specific colours for visualisation (Fig. 1). Subsequently, the *Bos taurus* gene co-expression network (Fig. 2) was computed using the determined power value, with a minimum module size set at 30. The network was visualized using Cytoscape version 3.9 [17].

2.3. Functional enrichment and pathway analysis

Enriched modules were identified through over-representation analysis (ORA) using the Goseq R package version 1.36.0 [18]. Subsequently, functional enrichment and pathway analysis of significantly enriched modules were conducted using g:Profiler [19] to determine gene ontology (GO) terms, including biological process (BP), molecular function (MF), and cellular compartment (CC), as well as biological pathways. Data sources for these pathways include Kyoto Encyclopedia of Genes and Genomes (KEGG), Reactome (REAC), WikiPathways (WP), Transfac (TF), miRTarBase (MIRNA), Human Protein Atlas (HPA), CORUM, and Human Phenotype Ontology (HPO). The top 10 most connected genes based on the degree of node/gene connectivity in each enriched module were identified using Cytohubba [20], a Cytoscape plugin (Additional File 4).

2.4. Hub genes and gene interaction analysis

Hub genes (the most connected genes) in each of the enriched modules, were identified and visualized using the STRING database [21]. The biological functions of both the hub genes and their interacting partners were inferred by conducting literature searches to ascertain the significance of gene interactions in the context of RVFV in bovine cells.

2.5. Identification and functional prediction of regulatory motifs

The Multiple Expectation Maximization for Motif Elicitation (MEME) suite [22] was utilized for the *de novo* identification of motifs with plausible gene regulatory roles at the 3′ end of genes in all enriched modules. DNA sequences of all genes from enriched modules were obtained from the BioMart-Ensembl database version 2.46.3 [23]. All duplicated gene sequences, genes lacking 3′ sequences, and sequences containing fewer than eight nucleotides were excluded from the input data for MEME. The motif with the lowest E-value in each enriched module was selected as being biologically relevant. To identify the probable biological roles of the uncovered regulatory motifs in humans, ontological analysis was performed using Gene Ontology for Motifs (GOMo) version 5.4.1 [24].

3. Results

3.1. Data pre-processing

A total of 45 raw RNA-Seq samples (Additional File 5) were downloaded for this study, out of which six (6) samples (five controls and one outlier) were excluded to give 39 samples for downstream analyses (Additional File 6). Moreover, to minimize noise in the network due to spurious gene correlation, non-coding and lowly expressed genes were removed. As recommended by the WGCNA

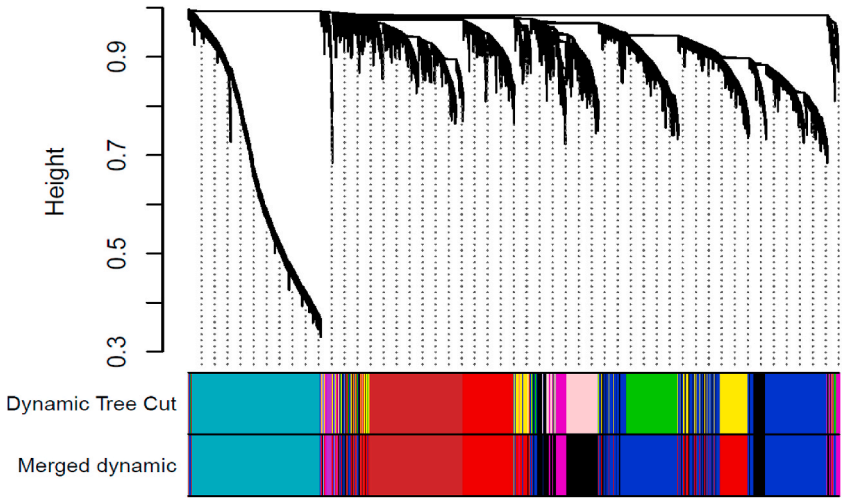


Fig. 1. Gene clusters (modules) of MP-12 inoculated *Bos taurus* gene co-expression network. Each colour on the x-axis represents a gene module. WGCNA's dynamic tree cut function trims the gene dendrogram at different heights based on expression similarities (y-axis), resulting in 11 module colours (x-axis). The number of genes contained in each module is represented by the width of the different coloured modules at the bottom of the figure. The wider the colour bar, the more genes contained in the module.

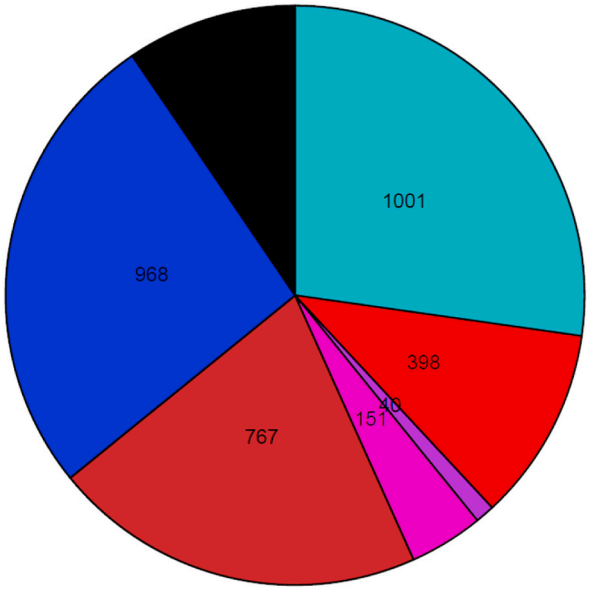


Fig. 2. Number of genes contained in the gene modules. Each module is represented by the gene module colour. As shown in the pie chart, seven modules were identified. The black module contains 349 genes.

methodology, only the top 5000 genes from the 39 samples were used in the construction of the *Bos taurus* gene co-expression network.

3.2. Weighted gene co-expression network construction

Prior to gene co-expression network construction, soft thresholding power analysis (Additional file 2) was performed to determine the appropriate power for the transformation of the gene similarity matrix into an adjacency matrix. A power of five (5), the power for which the scale-free topology fitting index (R2) was $\geq$0.8, was chosen. Using the dynamic tree cut function of WGCNA (Fig. 1), a total of 11 modules were generated. The grey module, consisting of nine (9) genes not assigned to any of the other modules, was excluded from subsequent analysis. The constructed gene co-expression network rendered in Cytoscape is depicted in Fig. 3.

To ensure that only distinct modules were identified and annotated, modules with a MEDissThreshold of $<$0.3 were merged (Additional file 3). Using the merged dynamic function in WGCNA, comparable gene clusters based on expression profiles were merged to give a total of seven modules (Fig. 2), out of which four (blue, brown, purple, and turquoise) were significantly enriched for GO terms associated with RVF disease (Additional File 7).

3.3. Functional enrichment and pathway analysis

Functional enrichment and pathway analysis identified GO terms and significant pathways in each of the four key enriched modules. In the purple module, 22 GO terms (MF:1, BP:20, CC:1) (Additional File 8) and 20 pathways (KEGG:6, REAC:3, WP:1, HP:10) (Additional File 9) were identified. In the blue module, 101 GO terms (MF:7, BB:62, CC:32) (Additional File 10), and 25 pathways (KEGG:10, REAC:11, WP:2, MIRNA:1, HP:1) (Additional File 11) were identified. In the turquoise module, 199 GO terms (MF:34, BP:102, CC:63) (Additional File 12) and 9 pathways (KEGG:6, REAC:2, CORUM:1) (Additional File 13) were identified. In the brown module, 85 GO terms (MF:21, BP:37, CC:27) (Additional File 14) and 5 pathways (KEGG:2, REAC:1, WP:2) (Additional File 15) were identified. Key GO terms and pathways within the four enriched modules (purple, blue, turquoise, and brown) were identified (Tables 1 and 2).

3.4. Hub genes and gene interaction analysis

Hub genes for each of the four significantly enriched modules were identified (Table 3) and the proteins they code for were uncovered. The interactions of the hub genes with other significant genes in the modules are depicted in Figs. 4–7.

3.5. Identification and functional prediction of regulatory motifs

The most abundant motif in each enriched module was identified (Table 4). Notably, the motifs were repeated at the 3′ end of the majority of the modules' genes, pointing to their potential importance in the functioning of the module. The top five (5) specific BP GO terms predicted in each of the four enriched modules are shown in Table 5. The most significant terms associated with bovine fetal development include gene expression, regulation of organ growth, tissue development, embryonic organ development, and organ

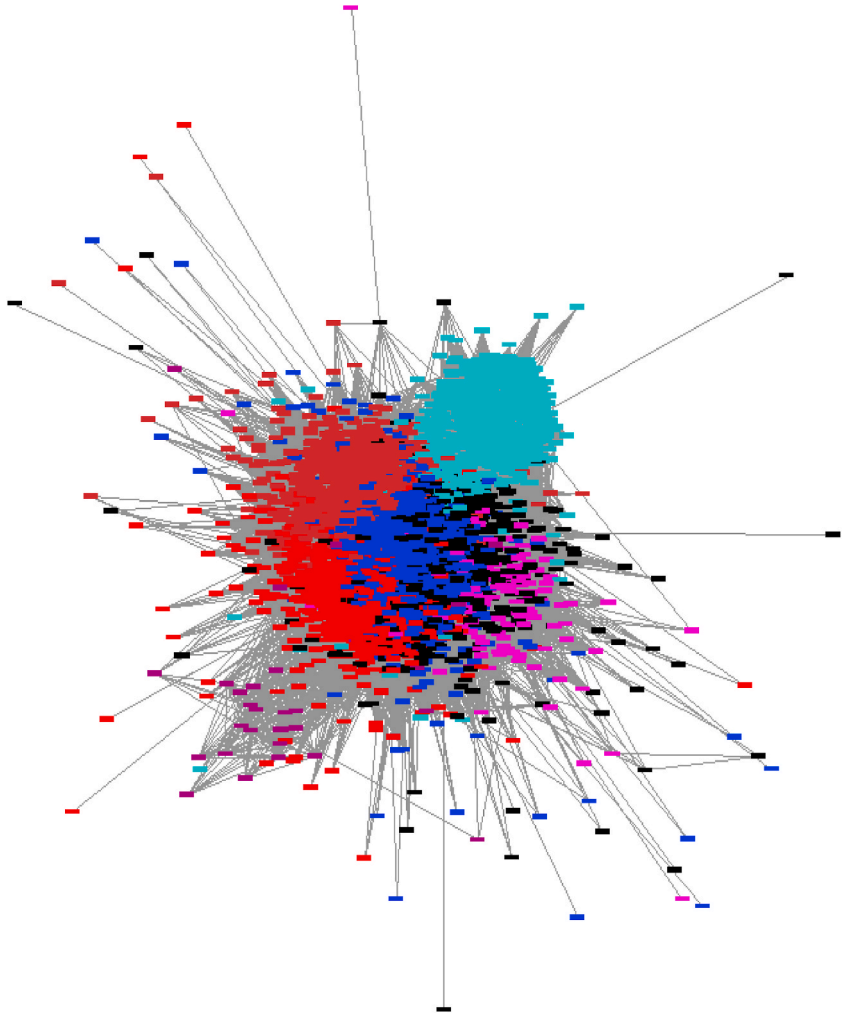


Fig. 3. An illustration of the interactions among components of the seven modules in the MP-12 inoculated *Bos taurus* gene co-expression network. Genes and interactions are depicted as nodes and edges respectively. Nodes are colour-coded to indicate their module membership. The gene co-expression network shown includes only genes with an interaction weight exceeding 0.05. This visualisation demonstrates the complexity of gene interactions in the MP-12 inoculated *Bos taurus* co-expression network. Only genes with significant co-expression relationships are shown.

morphogenesis.

4. Discussion

In this study, we utilized the WGCNA method to analyze RNA-Seq data from infected RVF cells and elucidate gene connections. Gene clusters, or modules, were identified from the constructed gene co-expression network. Enriched Gene Ontology (GO) terms and implicated biological pathways associated with RVF zoonosis were also identified from these modules. Furthermore, we identified the biological functions of intramodular hub genes, which play a prominent role within their respective modules and may serve as potential vaccine targets. Our findings demonstrate the significance of gene co-expression networks in understanding the genomic aspects of RVF progression, with a specific focus on enriched functions related to immunity and anatomical structure development.

The MSI2 gene, also known as Musashi RNA-binding protein 2, serves as the hub gene within the brown module (Fig. 4). It is a member of the Musashi family of RNA-binding proteins, which are known for their crucial role in the post-transcriptional regulation of gene expression [25,26], modulating the processing, localization, and translation of RNA molecules [27]. Proper regulation of MSI2 may be crucial for the normal development and functioning of bovine cells and tissues. RVFV has been shown to proliferate in placental tissues, causing harm and dysfunction [28,29]. This can lead to fetal distress, reduced placental blood flow, and ultimately fetal mortality and spontaneous abortion [30]. RVFV-induced inflammation and immune response may also interfere with prenatal development by activating the host's immune system, resulting in the production of pro-inflammatory cytokines and chemokines [31]. These inflammatory chemicals can harm placental tissues and disrupt embryonic development. Moreover, RVFV can disrupt hormonal regulation in pregnant cows [32], disturbing the delicate hormonal balance necessary for maintaining pregnancy, which can lead to

Table 1
An illustration of biological processes (BP) and gene ontology (GO) terms in the enriched modules (purple, turquoise, blue, and brown).

Module	source	term_name	term_id	adjusted_p_value
Purple	GO:BP	response to virus	GO:0009615	2.45431E-06
Purple	GO:BP	defense response to virus	GO:0051,607	6.99313E-06
Purple	GO:BP	defense response to symbiont	GO:0140,546	6.99313E-06
Purple	GO:BP	defense response to other organism	GO:0098,542	9.15728E-06
Purple	GO:BP	Innate immune response	GO:0045,087	1.87662E-05
Turquoise	GO:BP	localization	GO:0051,179	9.84E-17
Turquoise	GO:BP	Anatomical structure development	GO:0048,856	1.60E-16
Turquoise	GO:BP	developmental process	GO:0032,502	1.82E-15
Turquoise	GO:BP	Multicellular organism development	GO:0007275	8.86E-13
Turquoise	GO:BP	system development	GO:0048,731	1.88E-12
Blue	GO:BP	Cellular macromolecule metabolic process.	GO:0044,260	5.43E-18
Blue	GO:BP	cellular protein metabolic process	GO:0044,267	3.80E-17
Blue	GO:BP	organonitrogen compound metabolic process	GO:1,901,564	1.17E-16
Blue	GO:BP	protein metabolic process	GO:0019,538	1.58E-13
Blue	GO:BP	organic substance biosynthetic process	GO:1,901,576	1.17E-10
Brown	GO:BP	Cellular macromolecule metabolic process.	GO:0044,260	8.50E-11
Brown	GO:BP	organelle organization	GO:0006996	7.83E-08
Brown	GO:BP	macromolecule modification	GO:0043,412	6.57E-07
Brown	GO:BP	processes	GO:0006464	1.44736E-06
Brown	GO:BP	Protein modification process.	GO:0036,211	1.44736E-06

Table 2
Molecular functions (MF) and gene ontology (GO) terms in the enriched modules (turquoise, blue, and brown).

Module	Source	term_name	term_id	adjusted_p_value
Turquoise	GO:MF	Protein binding	GO:0005515	2.14E-12
Turquoise	GO:MF	Ion channel activity	GO:0005216	8.37E-08
Turquoise	GO:MF	Gated channel activity	GO:0022,836	1.58E-07
Turquoise	GO:MF	Passive transmembrane transporter activity	GO:0022,803	4.17E-07
Turquoise	GO:MF	channel activity	GO:0015,267	4.17E-07
Blue	GO:MF	catalytic activity	GO:0003824	7.44366E-06
Blue	GO:MF	Enzyme binding	GO:0019,899	3.80165E-05
Blue	GO:MF	Protein binding	GO:0005515	0.00017881
Blue	GO:MF	protein kinase binding	GO:0019,901	0.004823887
Blue	GO:MF	kinase binding	GO:0019,900	0.017100824
Brown	GO:MF	Protein binding	GO:0005515	1.35E-11
Brown	GO:MF	binding	GO:0005488	1.3665E-06
Brown	GO:MF	catalytic activity	GO:0003824	2.14812E-05
Brown	GO:MF	transferase activity	GO:0016,740	0.001408953
Brown	GO:MF	Adenyl ribonucleotide binding	GO:0032,559	0.002757682

Table 3
A representation of the enriched modules, number of genes in each module, hub genes, and proteins encoded by the hub genes in theMP-12 inoculated *Bos taurus* gene co-expression network.

Module colour	Number of genes	Most connected hub gene	Protein encoded by the hub gene
Brown	767	MS12	musashi RNA binding protein 2
Purple	40	EIF2AK2	eukaryotic translation initiation factor 2 alpha kinase 2
Turquoise	1001	ERC2	ELKS/RAB6-interacting/CAST family member 2
Blue	968	TMSB10	Thymosin beta-10

fetal death and spontaneous abortion.

The exact mechanisms by which RVFV affects bovine development are complex and may involve the interaction of multiple genes, including CSTF (cleavage stimulation factor) genes, TNPO (transportin) genes, and CPSF (cleavage and polyadenylation specificity factor) genes (Fig. 4). Like in other organisms, CSTF genes encode proteins that are involved in mRNA processing and polyadenylation in bovine cells, playing a role in the formation of the CSTF complex [33–35]. This complex is responsible for the cleavage and polyadenylation of pre-mRNAs, regulating gene expression in bovine cells. CPSF gene encodes proteins that are involved in the cleavage and polyadenylation of pre-mRNAs [35,36], an important step in mRNA maturation and regulation of gene expression in bovine species. Similarly, TNPO gene encodes proteins that play a role in nucleocytoplasmic transport, regulating various cellular processes, including gene expression, cell division, and response to stimuli [37].

The interaction of these genes (Fig. 4) within the brown module may suggest a coordinated working of their gene products, regulating multiple biological processes that may ultimately promote the invasion and progression of RVFV in bovine cells. Disruption

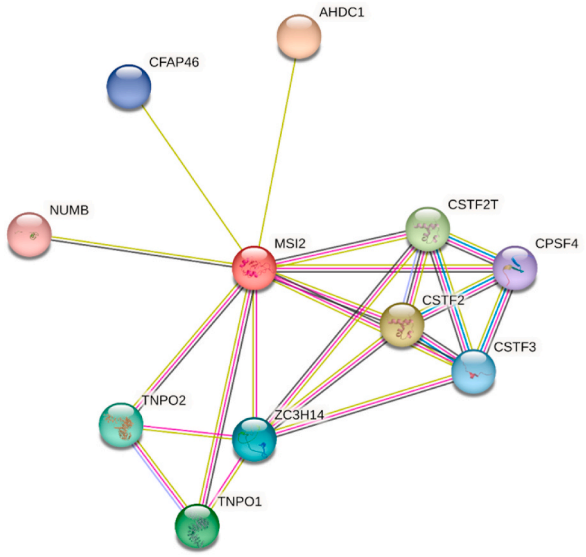


Fig. 4. Interactions of the MS12 hub gene within the brown module with other genes.

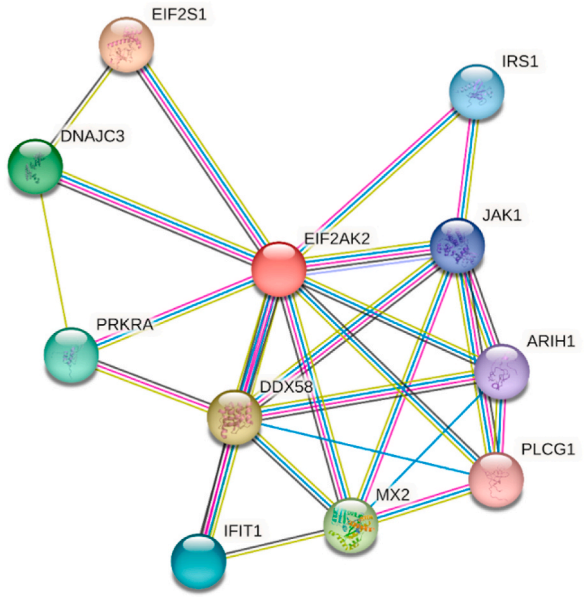


Fig. 5. Interactions of the EIF2AK2 hub gene within the purple module.

of these genes or their functions by RVFV may lead to alterations in gene expression, cellular processes, and hormonal balance, ultimately affecting bovine development and potentially leading to storm abortions in cattle.

EIF2AK2 is the hub gene in the purple module (Fig. 5). EIF2AK2 gene encodes the eukaryotic translation initiation factor 2 alpha kinase 2, which is involved in antiviral defense and cellular stress response. EIF2AK2 regulates protein synthesis and modulates the host's immunological response to viral infections. EIF2AK2 stimulates the integrated stress response pathway, leading to the phosphorylation of eukaryotic translation initiation factor 2 alpha (eIF2α) and inhibition of translation, thereby limiting viral replication. EIF2AK2 gene is connected to other genes within the purple module such as DDX58, JAK1, AR1H, PLCG1, MX2 and IFIT1. In humans, DDX58, also called retinoic-acid-inducible gene I (RIG-I) plays a critical part in the innate immune response to infections [38]. RIG-I detects viral RNA molecules and activates antiviral signaling pathways [39], resulting in the generation of interferons and other immune response molecules that aid in the fight against viral infections.

JAK1 (Janus kinase 1) is a protein that is involved in the signaling pathways triggered by type I and type III interferons [40], which are important components of the immunological response to viral infections in animals. JAK1 is involved in the phosphorylation and activation of downstream signaling molecules [41], which results in the stimulation of antiviral genes, the suppression of viral

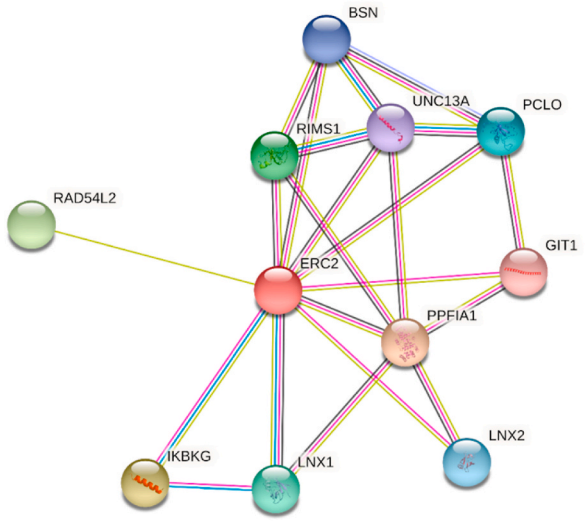


Fig. 6. Interactions of the ERC2 hub gene within the turquoise module.

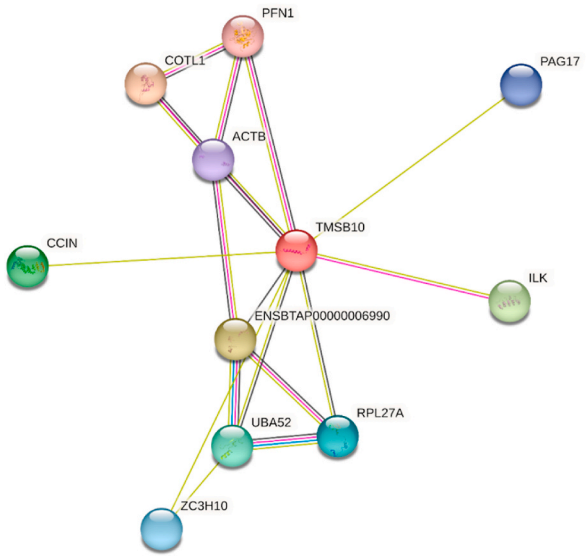


Fig. 7. Interactions of the TMSB10 hub gene within the blue module.

replication, and an improved immune response to viral infections.

The IFIT1 (Interferon-induced protein with tetratricopeptide repeats 1) gene may also play a crucial role in animal immunological response to viral infections. IFIT1 is a gene that encodes a protein that is activated by interferons, which are signaling proteins generated in response to viral infections [42,43]. IFIT1 inhibits viral replication by binding to viral RNA and blocking translation, decreasing viral protein production and replication [43]. Further investigation of these viral response genes and their collaborative interactions (Fig. 5) could provide insights into how they can be effectively targeted to elicit a robust immune response against Rift Valley fever virus (RVFV) invasion. This knowledge could potentially lead to improved strategies for RVF disease control, allowing for more accurate and effective management of RVF outbreaks.

The ERC2 gene is the most connected gene within the purple module (Fig. 6). ERC2 gene has no immune response related functions in bovine cells. In humans, the gene forms part of the cytomatrix at the active zone (CAZ) complex that regulates neurotransmitter release. Other genes within the purple module connected to ERC2 gene include PCLO, RIMS1, BSN and UNC13A, all of which as associated with neurotransmitter release along the synapse [44–48]. A rare characteristic of RVF disease in humans is encephalitis [49]. The association between RVFV and the nervous system is not well understood [49], and there is a significant knowledge gap in how RVFV may lead to brain pathology. Therefore, further research, especially in bovine species, is needed to better understand the role of the neurological system in RVF pathogenesis.

Table 4
Consensus regulatory motifs in each of the enriched modules.

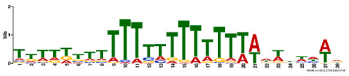
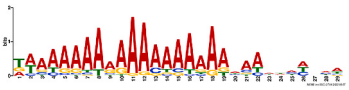
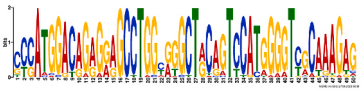
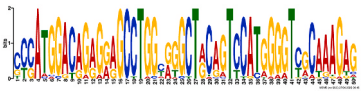
Module	Consensus regulatory motifs
Purple	
Turquoise	
Blue	
Brown	

Table 5
Predicted biological processes (GO terms) of the consensus regulatory motifs in the blue and brown modules.

GO term	Score	*P*-value	q-value	Specificity	GO category	GO name
GO:0009888	5.440e-03	1.187e-06	2.732e-03	~1%	BP	Tissue development
GO:0048568	8.244e-03	4.748e-06	6.829e-03	~8%	BP	Embryonic organ development.
GO:0009887	1.422e-02	2.603e-05	2.237e-02	~1%	BP	Organ morphogenesis
GO:0050767	1.437e-02	2.662e-05	2.237e-02	~14%	BP	Regulation of neurogenesis
GO:0045596	1.451e-02	2.722e-05	2.237e-02	~5%	BP	Negative regulation of cell differentiation

Additionally, the turquoise module also contains genes responsible for anatomical structure development, localization, and system development. Fetal malformation and storm abortions are the primary characteristics of RVFV infection. The PAFAH2 gene within the turquoise module has been implicated in various stages of reproduction, including implantation, fetal development, and parturition [50]. PAFAH2 promotes the formation of IP3 and DAG and increases intracellular calcium [51,52]. Calcium contributes to bone formation and fetal mineralization [29]. RVFV potentially disrupts PAFAH2 in order to interfere with normal fetal development and ultimately cause abortion. Our research findings highlight the crucial role of genes within the turquoise module in promoting bovine fetal development and parturition. This suggests that RVFV invasion and disruption of these genes may result in improper fetal structure development, leading to expulsion by the pregnant animal host.

TMSB10, the hub gene within the blue module (Fig. 7) has been linked to immune response and inflammatory modulation. TMSB10 has also been found to be upregulated in bovine cells in response to viral infections, implying a role in antiviral defense mechanisms. Additionally, TMSB10 has also been linked to the control of cell motility and migration, which could affect immune cells' ability to travel to sites of viral infection or tissue injury. TMSB10 is also connected to other genes such as ILK, PFN1, ACTB and UBA52. ILK (Integrin-linked kinase) is a serine/threonine protein kinase known to regulate immune cell signaling, cell adhesion, and cell survival [53]. PFN1 (Profilin 1), is involved in actin dynamics and cytoskeletal remodeling, which are critical for viral replication [54]. ACTB (Actin Beta) gene is involved in the cytoskeletal reorganization [55] while UBA52 (Ubiquitin A-52 Residue Ribosomal Protein Fusion Product 1) gene encode proteins involved in the ubiquitin-proteasome system, which plays a crucial role in degradation, cell immunity and cellular regulation [56].

Similarly, CLIC4 gene, connected to the TMSB10 hub gene, is a significant gene within the blue module. CLIC4 has been implicated in mid-gestational brain differentiation and neurogenesis [57] and several other cellular processes, including cell differentiation, cell-cycle control, and apoptosis. Cellular differentiation is critical for facilitating the replication of RVFV in mammalian hosts. Studies in human hosts have reported that monocytic cell differentiation enhances RVFV replication [57]. Thus, CLIC4 triggered cellular differentiation in *Bos taurus* probably aids the establishment of the virus and promotes infection progression. Consequently, disrupting CLI4 may be a viable way of halting RVFV establishment in the host. The functions of these genes are associated with cellular immune response, and understanding the mechanisms behind their working co-ordinately may provide insights on better and more efficacious RVF control at a genetic level.

The Gene Ontology (GO) analysis uncovered significant biological pathways linked to RVF zoonosis, as demonstrated in Tables 1 and 2 Further analysis using Reactome and WikiPathways identified several pathways involved in the host defense against RVF, including the interferon-signaling pathways that activate JAK-STAT signaling pathways. Notably, the JAK-STAT pathway was predominantly associated with genes in the purple module, with the JAK1 gene being connected to EIF2AK2, which is the hub gene of the purple module. This suggests that the upregulation of interferon-stimulated genes in this pathway activates the antiviral properties [58] of the bovine host, leading to immune responses against RVF infection. Additionally, the RIG-1-like receptor (RLR) signaling pathway was also associated with the purple module, as the DDX58 gene was significantly associated with the EIF2AK2 gene. This suggests that the activity of the RIG-1-like receptor (RLR) pathway within this module was also high, allowing for the detection of RVF

RNA and subsequent activation of downstream signaling pathways that stimulate the production of type I interferons and other pro-inflammatory cytokines.

The 3′ end regulatory motifs of genes identified in significantly enriched modules (Table 4) may represent potential targets for the development of drugs and vaccines against RVF. These regulatory motifs are DNA regions that play a vital role in controlling bovine gene expression and might be targeted to modulate the expression of RVF-associated genes. By targeting these regulatory motifs, it may be possible to modulate the expression of specific genes and influence the biological activity mediated by these motifs (Table 5), thereby impacting cellular processes or immune responses related to RVF zoonosis.

Manipulating the activity of these regulatory motifs could involve inhibiting or enhancing their function to influence the production of proteins or other cellular components that are involved in RVF pathogenesis. Modulating gene expression through these regulatory motifs may enable the development of drugs that selectively inhibit or enhance the expression of specific genes or biological pathways, leading to desired therapeutic effects against RVF progression, in efforts to control RVF in both bovine and human hosts.

The regulatory motif sequences identified in the four enriched modules (Table 5) are unique and merit further investigation. Notably, the blue and brown modules exhibit similarity in their regulatory motifs, suggesting potential congruency in gene function and a significant role for the identified motifs. The shared regulatory motif found in these modules may be implicated in bovine fetal tissue development and embryonic organ development.

5. Conclusion

The construction of the MP-12 inoculated *Bos taurus* gene co-expression network has facilitated the identification of critical genes and pathways that are potentially involved in the establishment and progression of RVF. Our findings highlight various host targets that could be inhibited or overexpressed to impede RVFV establishment and pathogenesis in *Bos taurus*. The identified genes, along with their associated regulatory motifs, represent potential novel targets for the development of new therapeutics against RVF.

Study limitation

Currently, there is a lack of a comprehensive database for bovine regulatory motifs, which can be used to assess the biological importance of identified regulatory motifs in bovine species.

Author contribution statement

John Kaniaru Gitau: Conceived and designed the experiments; Analyzed and interpreted the data, Contributed analysis tools and data; Wrote the paper.

Rosaline Wanjiru Macharia: Conceived and designed the experiments; Performed the experiments; Contributed analysis tools; Wrote the paper.

Kennedy Wanjau Mwangi: Conceived and designed the experiments, Analyzed and interpreted the data; contributed analysis tools and data; Wrote the paper.

Nehemiah Ongeso: Analyzed and interpreted the data; Contributed analysis tools and data; Wrote the paper.

Edwin Murungi: Performed the experiments; Analyzed and interpreted the data; Contributed analysis tools and data; Wrote the paper.

Data availability statement

Data associated with this study has been deposited at Gene Expression Omnibus repository - GSE71417.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2023.e18175.

References

[1] R.C. Laughlin, K.L. Drake, J.C. Morrill, L.G. Adams, Correlative gene expression to protective seroconversion in rift valley fever vaccinates, PLoS One 11 (1) (2016) 1–27, https://doi.org/10.1371/journal.pone.0147027.

[2] S. Lumley, et al., Competence of mosquitoes native to the United Kingdom to support replication and transmission of Rift Valley fever virus, Parasites Vectors 11 (1) (2018) 1–11, https://doi.org/10.1186/s13071-018-2884-7.

[3] S. Lumley, D.L. Horton, L.L.M. Hernandez-Triana, N. Johnson, A.R. Fooks, R. Hewson, Rift valley fever virus: strategies for maintenance, survival and vertical transmission in mosquitoes, J. Gen. Virol. 98 (5) (2017) 875–887, https://doi.org/10.1099/jgv.0.000765. Microbiology Society.
[4] A. Narayanan, et al., Reactive oxygen species activate NFκB (p65) and p53 and induce apoptosis in RVFV infected liver cells, Virology 449 (2014) 270–286, https://doi.org/10.1016/j.virol.2013.11.023.
[5] L. Odendaal, A.S. Davis, G.T. Fosgate, S.J. Clift, Lesions and cellular tropism of natural Rift Valley fever virus infection in young lambs, Vet. Pathol. 57 (1) (2020) 66–81, https://doi.org/10.1177/0300985819882633.
[6] L.M. Michaely, et al., NSG-mice reveal the importance of a functional innate and adaptive immune response to overcome RVFV infection, Viruses 14 (2) (2022) 1–17, https://doi.org/10.3390/v14020350.
[7] J. Plaza-Diaz, C. Gomez-Llorente, L. Fontana, A. Gil, Modulation of immunity and inflammatory gene expression in the gut, in inflammatory diseases of the gut and in the liver by probiotics, World J. Gastroenterol. 20 (42) (2014) 15632–15649, https://doi.org/10.3748/wjg.v20.i42.15632.
[8] W. Liu, et al., Construction and analysis of gene co-expression networks in Escherichia coli, Cells 7 (3) (2018), https://doi.org/10.3390/cells7030019.
[9] K.W. Mwangi, R.W. Macharia, J.L. Bargul, Gene co-expression network analysis of Trypanosoma brucei in tsetse fly vector, Parasites Vectors 14 (1) (2021) 1–11, https://doi.org/10.1186/s13071-021-04597-6.
[10] E. Alm, A.P. Arkin, Biological networks, Curr. Opin. Struct. Biol. 13 (2) (2003) 193–202, https://doi.org/10.1016/S0959-440X(03)00031-9.
[11] B. Zhang, S. Horvath, A general framework for weighted gene co-expression network analysis, Stat. Appl. Genet. Mol. Biol. 4 (1) (2005), https://doi.org/10.2202/1544-6115.1128.
[12] O.B. Morenikeji, E. Strutton, M. Wallace, K. Bernard, E. Yip, B.N. Thomas, Dissecting transcription factor-target interaction in bovine coronavirus infection, Microorganisms 8 (9) (2020) 1–20, https://doi.org/10.3390/microorganisms8091323.
[13] J.H. Bushweller, Targeting transcription factors in cancer — from undruggable to reality, Nat. Rev. Cancer 19 (11) (2019) 611–624, https://doi.org/10.1038/s41568-019-0196-7.
[14] C. Clarke, et al., Correlating transcriptional networks to breast cancer survival: a large-scale coexpression analysis, Carcinogenesis 34 (10) (2013) 2300–2308, https://doi.org/10.1093/carcin/bgt208.
[15] R. Edgar, M. Domrachev, A.E. Lash, Gene Expression Omnibus: NCBI gene expression and hybridization array data repository, Nucleic Acids Res. 30 (1) (2002) 207–210, https://doi.org/10.1093/nar/30.1.207.
[16] P. Langfelder, S. Horvath, WGCNA: an R package for weighted correlation network analysis, BMC Bioinf. 9 (2008), https://doi.org/10.1186/1471-2105-9-559.
[17] P. Shannon, et al., Cytoscape: a software Environment for integrated models of biomolecular interaction networks, Genome Res. 13 (11) (2003) 2498–2504, https://doi.org/10.1101/gr.1239303.
[18] M.D. Young, M.J. Wakefield, G.K. Smyth, A. Oshlack, Gene ontology analysis for RNA-seq: accounting for selection bias, Genome Biol. 11 (2) (2010), https://doi.org/10.1186/gb-2010-11-2-r14.
[19] J. Reimand, M. Kull, H. Peterson, J. Hansen, J. Vilo, G:Profiler-a web-based toolset for functional profiling of gene lists from large-scale experiments, Nucleic Acids Res. 35 (2) (2007) 193–200, https://doi.org/10.1093/nar/gkm226.
[20] C.H. Chin, S.H. Chen, H.H. Wu, C.W. Ho, M.T. Ko, C.Y. Lin, cytoHubba: identifying hub objects and sub-networks from complex interactome, BMC Syst. Biol. 8 (4) (2014) 1–7, https://doi.org/10.1186/1752-0509-8-S4-S11.
[21] D. Szklarczyk, et al., The STRING database in 2021: customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets, Nucleic Acids Res. 49 (D1) (2021) D605–D612, https://doi.org/10.1093/nar/gkaa1074.
[22] T.L. Bailey, J. Johnson, C.E. Grant, W.S. Noble, The MEME suite, Nucleic Acids Res. 43 (W1) (2015) W39–W49, https://doi.org/10.1093/nar/gkv416.
[23] D. Smedley, et al., BioMart - biological queries made easy, BMC Genom. 10 (2009) 1–12, https://doi.org/10.1186/1471-2164-10-22.
[24] X. Lu, C. Zhai, V. Gopalakrishnan, B.G. Buchanan, Automatic annotation of protein motif function with gene ontology terms, BMC Bioinf. 5 (2004) 1–11, https://doi.org/10.1186/1471-2105-5-122.
[25] A. Ratti, et al., A role for the ELAV RNA-binding proteins in neural stem cells: stabilization of Msi1 mRNA, J. Cell Sci. 119 (7) (2006) 1442–1452, https://doi.org/10.1242/jcs.02852.
[26] M. Christou-Kent, M. Dhellemmes, E. Lambert, P.F. Ray, C. Arnoult, Diversity of RNA-binding proteins modulating post-transcriptional regulation of protein expression in the maturing mammalian oocyte, Cells 9 (3) (2020) 1–20, https://doi.org/10.3390/cells9030662.
[27] F. Matalkah, B. Jeong, M. Sheridan, E. Horstick, V. Ramamurthy, P. Stoilov, The Musashi proteins direct post-transcriptional control of protein expression and alternate exon splicing in vertebrate photoreceptors, Commun. Biol. 5 (1) (2022) 1–15, https://doi.org/10.1038/s42003-022-03990-w.
[28] C. Pinkham, et al., Alterations in the host transcriptome in vitro following Rift Valley fever virus infection, Sci. Rep. 7 (1) (2017) 1–16, https://doi.org/10.1038/s41598-017-14800-3.
[29] Z. Heydarifard, et al., Potential role of viral infections in miscarriage and insights into the underlying molecular mechanisms, Congenital. Anom. 62 (2) (2022) 54–67, https://doi.org/10.1111/cga.12458.
[30] W. Yu, X. Hu, B. Cao, Viral infections during pregnancy: the big challenge threatening maternal and fetal health, Matern. Med. 4 (1) (2022) 72–86, https://doi.org/10.1097/FM9.0000000000000133.
[31] Y.D. Gwon, S.A.N. Mahani, I. Nagaev, L. Mincheva-Nilsson, M. Evander, Rift valley fever virus propagates in human villous trophoblast cell lines and induces cytokine mrna responses known to provoke miscarriage, Viruses 13 (11) (2021), https://doi.org/10.3390/v13112265.
[32] M. Grandadam, Rift valley fever virus, Mol. Detect. Hum. Viral Pathog. (2016) 679–690, https://doi.org/10.2174/156652405774962263.
[33] D.C. Di Giammartino, et al., RBBP6 isoforms regulate the human polyadenylation machinery and modulate expression of mRNAs with AU-rich 39 UTRs, Genes Dev. 28 (20) (2014) 2248–2260, https://doi.org/10.1101/gad.245787.114.
[34] P.F. Cstf, et al., Crossm Remodeling during the DNA Damage Response, 2018, pp. 1–17.
[35] J. Neve, R. Patel, Z. Wang, A. Louey, A.M. Furger, Cleavage and polyadenylation: ending the message expands gene regulation, RNA Biol. 14 (7) (2017) 865–890, https://doi.org/10.1080/15476286.2017.1306171.
[36] P. Dharmalingam, et al., Emerging roles of alternative cleavage and polyadenylation (APA) in human disease, J. Cell. Physiol. 237 (1) (2022) 149–160, https://doi.org/10.1002/jcp.30549.
[37] S. Moore, B.E. Rabichow, R. Sattler, The hitchhiker's guide to nucleocytoplasmic trafficking in neurodegeneration, Neurochem. Res. 45 (6) (2020) 1306–1327, https://doi.org/10.1007/s11064-020-02989-1.
[38] P. Yu, P. Liang, S. Pang, W. Yuan, Y. Zhao, Q. Huang, The function, role and process of DDX58 in heart failure and human cancers, Front. Oncol. 12 (2022) 1–12, https://doi.org/10.3389/fonc.2022.911309.
[39] J. Rehwinkel, M.U. Gack, RIG-I-like receptors: their regulation and roles in RNA sensing, Nat. Rev. Immunol. 20 (9) (2020) 537–551, https://doi.org/10.1038/s41577-020-0288-3.
[40] S.B. Fleming, Viral inhibition of the IFN-induced JAK/STAT signalling pathway: development of live attenuated vaccines by mutation of viral-encoded IFN-antagonists, Vaccines 4 (3) (2016) 1–26, https://doi.org/10.3390/vaccines4030023.
[41] J.J. Babon, I.S. Lucet, J.M. Murphy, N.A. Nicola, L.N. Varghese, The molecular regulation of Janus kinase (JAK) activation, Biochem. J. 462 (1) (2014) 1–13, https://doi.org/10.1042/BJ20140712.
[42] M.S. Diamond, M. Farzan, The broad-spectrum antiviral functions of IFIT and IFITM proteins, Nat. Rev. Immunol. 13 (1) (2013) 46–57, https://doi.org/10.1038/nri3344.
[43] V. Fensterl, G.C. Sen, Interferon-induced ifit proteins: their role in viral pathogenesis, J. Virol. 89 (5) (2015) 2462–2468, https://doi.org/10.1128/jvi.02744-14.
[44] C.H. Chen, Y.S. Huang, D.L. Liao, C.Y. Huang, C.H. Lin, T.H. Fang, Identification of rare mutations of two presynaptic cytomatrix genes bsn and pclo in schizophrenia and bipolar disorder, J. Personalized Med. 11 (11) (2021), https://doi.org/10.3390/jpm11111057.
[45] Ö. Genç, et al., Homeostatic plasticity fails at the intersection of autism-gene mutations and a novel class of common genetic modifiers, Elife 9 (2020) 1–32, https://doi.org/10.7554/eLife.55775.

[46] S.D. Fenster, C.C. Garner, Gene structure and genetic localization of the PCLO gene encoding the presynaptic active zone protein Piccolo, Int. J. Dev. Neurosci. 20 (3) (2002) 161–171, https://doi.org/10.1016/S0736-5748(02)00046-1.
[47] J. Ko, et al., Organization of the presynaptic active zone by ERC2/CAST1-dependent clustering of the tandem PDZ protein syntenin-1, J. Neurosci. 26 (3) (2006) 963–970, https://doi.org/10.1523/JNEUROSCI.4475-05.2006.
[48] D. Arancibia, et al., Serine–arginine protein kinase SRPK2 modulates the assembly of the active zone scaffolding protein CAST1/ERC2, Cells 8 (11) (2019), https://doi.org/10.3390/cells8111333.
[49] J.R. Albe, et al., Neutrophil and macrophage influx into the central nervous system are inflammatory components of lethal rift valley fever encephalitis in rats, PLoS Pathog. 15 (6) (2019) 1–26, https://doi.org/10.1371/journal.ppat.1007833.
[50] A. Sigdel, R.S. Bisinotto, F. Peñagaricano, Genes and pathways associated with pregnancy loss in dairy cattle, Sci. Rep. 11 (1) (2021), 13329, https://doi.org/10.1038/s41598-021-92525-0.
[51] A.R. Hughes, J.W. Putney, Inositol phosphate formation and its relationship to calcium signaling, Environ. Health Perspect. 84 (1990) 141–147, https://doi.org/10.1289/ehp.9084141.
[52] J.W. Putney, T. Tomita, Phospholipase C signaling and calcium influx, Adv. Biol. Regul. 52 (1) (2012) 152–164, https://doi.org/10.1016/j.advenzreg.2011.09.005.
[53] A.U. Ahmed, S.T. Sarvestani, M.P. Gantier, B.R.G. Williams, G.E. Hannigan, Integrin-linked kinase modulates lipopolysaccharide- and Helicobacter pylori-induced nuclear factor κB-activated tumor necrosis factor-α production via regulation of p65 serine 536 phosphorylation, J. Biol. Chem. 289 (40) (2014) 27776–27793, https://doi.org/10.1074/jbc.M114.574541.
[54] M.D. Figley, G. Bieri, R.M. Kolaitis, J. Paul Taylor, A.D. Gitler, Profilin 1 associates with stress granules and ALS-Linked mutations alter stress granule dynamics, J. Neurosci. 34 (24) (2014) 8083–8097, https://doi.org/10.1523/JNEUROSCI.0543-14.2014.
[55] T.M. Bunnell, B.J. Burbach, Y. Shimizu, J.M. Ervasti, β-Actin specifically controls cell growth, migration, and the G-actin pool, Mol. Biol. Cell 22 (21) (2011) 4047–4058, https://doi.org/10.1091/mbc.E11-06-0582.
[56] J. Mao, C. O'Gorman, M. Sutovsky, M. Zigo, K.D. Wells, P. Sutovsky, Ubiquitin A-52 residue ribosomal protein fusion product 1 (Uba52) is essential for preimplantation embryo development, Biol. Open 7 (10) (2018), https://doi.org/10.1242/bio.035717.
[57] R.M. Lewis, T.M. Cosgriff, C.J. Peters, J.C. Morrill, Differentiation of a human monocytic cell line associated with increased production of Rift Valley fever virus by infected cells, J. Med. Virol. 23 (3) (1987) 207–215, https://doi.org/10.1002/jmv.1890230302.
[58] K.M. Crosse, E.A. Monson, M.R. Beard, K.J. Helbig, Interferon-stimulated genes as enhancers of antiviral innate immune signaling, J. Innate Immun. 10 (2) (2018) 85–93, https://doi.org/10.1159/000484258.