Protein network analysis to prioritize key genes in amyotrophic lateral sclerosis

Rupesh Kumar,1 Shazia Haider1,1

Department of Biotechnology, Jaypee Institute of Information Technology, Noida, Sec-62, Uttar Pradesh, India

ARTICLE INFO

Keywords:
ALS
ALS-PPIN
CDC5L
SNW1
TP53
SOD1
VCP
Bottleneck-hubs

ABSTRACT

Amyotrophic Lateral Sclerosis (ALS) is a fatal disease, progressive nature characterizes by loss of both upper and lower motor neuron functions. One of the major challenges is to understand the mechanism of ALS multifactorial nature. We aimed to explore some key genes related to ALS through bioinformatics methods for its therapeutic intervention. Here, we applied a systems biology approach involving experimentally validated 148 ALS-associated proteins and construct ALS protein-protein interaction network (ALS-PPIN). The network was further statistically analysed and identified bottleneck-hubs. The network is also subjected to identify modules which could have similar functions. The interaction between the modules and bottleneck-hubs provides the functional regulatory role of the ALS mechanism. The ALS-PPIN demonstrated a hierarchical scale-free nature. We identified 17 bottleneck-hubs, in which CDC5L, SNW1, TP53, SOD1, and VCP were the high degree nodes (hubs) in ALS-PPIN. CDC5L was found to control highly cluster modules and play a vital role in the stability of the overall network followed by SNW1, TP53, SOD1, and VCP. HSPA5 and HSPA8 acting as a common connector for CDC5L and TP53 bottleneck-hubs. The functional and disease association analysis showed ALS has a strong correlation with mRNA processing, protein deubiquitination, and neoplasms, nervous system, immune system disease classes. In the future, biochemical investigation of the observed bottleneck-hubs and their interacting partners could provide a further understanding of their role in the pathophysiology of ALS.

1. Introduction

Amyotrophic lateral sclerosis (ALS) is one of the fatal neurodegenerative disease mainly occurs due to the loss of motor neuron and neuronal death (Martin et al., 2017). It leads to death by progressive paralysis and respiratory failure within 2–4 years (Chio et al., 2009). About 90–95% of ALS cases are sporadic (sALS) and 5–10% are inherited through family (fALS) (Van Rheenen et al., 2016). Many genetic factors have been identified, involving aggregations for RNA metabolism (Butti and Patten, 2018), weakened protein homeostasis (Cykowski et al., 2019), damaged DNA repair (Brenner et al., 2016), disruption of nucleocytoplasmic transport (Kim and Taylor, 2017), excitotoxicity (Foerster et al., 2013), oxidative stress (Mitsumoto et al., 2008), and axonal transport disturbance (De Vos et al., 2007). Mutations observed in several genes related to ALS such as SOD1 (Rosen et al., 1993), FUS (Kwiatkowski et al., 2009), C9ORF72 (DeJesus-Hernandez et al., 2011; Renton et al., 2011), AXTN2 (Elden et al., 2010), OPTN (Maruyama et al., 2010), VCP (Johnson et al., 2010), PFN1 (Wu et al., 2012), MATR3 (Johnson et al., 2014), SETX (Hirano et al., 2011), UBQLN2 (Deng et al., 2011). Initial discovery of a risk associated C9ORF72 locus in ALS (GWAS), a pathology associated risk hexanucleotide-repeat expansion (G4C2) updated the field of ALS genetics and biology (Laaksovirta et al., 2010). Since 2015, others gene such as TBK1 (Freschmidt et al., 2015), C21ORF2 (Van Rheenen et al., 2016), Neki (Brenner et al., 2016), CCNF (Williams et al., 2016), MOBP, SCFD1 (Van Rheenen et al., 2016), KIF5A (Nicol et al., 2018), LGALS (Gelfman et al., 2019), GLT8D1 (Cooper-Knock et al., 2019), DNAJC7 (Farhan et al., 2019), TUBA4A (Smith et al., 2014), ANXA11 (Topp et al., 2017), UBQLN4 (Edens et al., 2017),

Abbreviations: ALS, Amyotrophic Lateral Sclerosis; ALSSoD, Amyotrophic Lateral Sclerosis online database; ALS-PPIN, Amyotrophic Lateral Sclerosis Protein-Protein Interaction Network; BC, Betweenness centrality; Bn-H, Bottleneck-hub; CDC5L, Cell division cycle5-like protein; FUS, Fused in sarcoma; MND, Motor neuron disease; MCODE, Molecular Complex Detection; SOD1, Superoxide dismutase; SNW1, SNW domain-containing protein 1; SMA, Spinal muscular atrophy; SMN, Survival of motor neuron; TP53, Tumor protein p53; VCP, Valosin containing protein.

* Corresponding author.

E-mail addresses: 20401005@mail.jiit.ac.in (R. Kumar), shazia.haider@jiit.ac.in (S. Haider).

ORCID ID: 0000-0002-1237-4400.

https://doi.org/10.1016/j.ibneur.2021.12.002

Received 29 September 2021; Received in revised form 18 November 2021; Accepted 5 December 2021

Available online 7 December 2021

2667-2421© 2021 The Author(s). Published by Elsevier Ltd on behalf of International Brain Research Organization. This is an open access article under the CC
or RNA interference, are frequently time-consuming and expensive. The viability of computational techniques to predict gene essentiality and required for survival and that their absence results in lethal phenotype. Thus, to gain an understanding of the function of specific molecules in a variety of cellular processes. Dysregulation of numerous biological processes, cellular component pathways, and molecular activity has a downstream effect such as motor neuron degeneration which has a significant risk associated with ALS (Saez-Artenzer et al., 2021). In our study, we identify, modules that controlled the network from dissortivity and also maintain the flow of information throughout the network of ALS. A set of ALS-associated genes/proteins were used to generate the Amyotrophic Lateral Sclerosis Protein-Protein Interaction Network (ALS-PPIN). We also identified the biological significance, the molecular activity of the modules, and their association with different disease classes. Our interest was to identify the Bn-H and their interaction which may provide an opportunity for a better understanding of ALS.

2. Methods

2.1. Construction of ALS protein-protein interaction network

The genes associated with ALS disease were retrieved from the Amyotrophic Lateral Sclerosis online database (ALSoD) (Abel et al., 2012) and literatures (Chia et al., 2018; Dervishi et al., 2018; Le Gall et al., 2020; Mejzini et al., 2019; Srinivasan and Rajasekaran, 2020). ALSoD database provides comprehensive information about genomic, proteomic, and bioinformatics in association with ALS (Wroe et al., 2008). It also includes patient information like age, clinical data, family history, survival data, sex (Abel et al., 2012). The ALSoD was constructed using genotype, phenotype, and geographical information with associated analysis tools and it transformed from a single gene storage facility recording mutations in the SOD1 gene to multigene ALS repository ALSoD database links with others databases like ALSgene provides evidence of association to complement the genotype-phenotype association given in ALSoD. It is also used to perform multiple alignment and mutations on SOD1 gene using ClustalW and Jalview to perform multiple sequence alignment in others species for selected genes (Waterhouse et al., 2009). Other links and databases integrated with ALSoD are GeneMANIA used to predict interaction of selected gene (Warde-Farley et al., 2010), a Google Earth API is used to visualise maps...
of mutations, risk and exposure distributions, HGNC, Enterz Gene, UCSC browser, Protein Structure, OMIM, Genecards, ProtScale, KEGG, UniProt, iHOP, GeneTest, AmiGO, Ensemble, NCB1, Life Science DB, GeneWiki, WolframAlpha, and WikiGenes. The redundancy in selected genes was removed (duplicate names and aliases names used for the gene). A set of 148 genes were used to generate the Amyotrophic Lateral Sclerosis Protein-Protein Interaction Network (ALS-PPIN) (Fig. 1). These ALS associated genes were used to fetch PPI and to analyse a large number of interactions, we first retrieved the PPIN from five different source databases, namely STRING (von Mering et al., 2003), the General Repository of Interaction Datasets (BioGRID) (Stark et al., 2006), Database of Interacting Proteins (DIP) (Xenarios et al., 2001), IntAct (Hermjakob et al., 2004), Molecular Interactions Database (MINT) (Chatr-aryamontri et al., 2007). We selected those interactions commonly present in at least two of the databases to generate the ALS-PPIN. The constructed network shows a graph denoted by $G (N, E)$, where, $N$ represents sets of nodes with $N = \{1, 2, \ldots, N\}$ and $E$ the sets of edges with $E = \{(i, j) | i, j = 1, 2, 3, \ldots, N\}$.

### 2.2. Statistical analysis of the network

To analyse the topological properties of ALS-PPIN we used various statistical parameters like degree distribution, clustering coefficient, neighborhood connectivity, and centrality parameters (betweenness, closeness, eigenvector) using Cytoscape plugins, Network Analyzer (Assenov et al., 2008), and CytoNCA (Tang et al., 2015).

#### 2.2.1. Degree($k$) and probability of degree distribution ($P(k)$)

Degree($k$) is a basic characteristic that has an impact on a node’s centrality and is represented by the number of connections a node has to others in a network. The probability of degree distribution($P(k)$) is represented by the given equation (Eq. 1).

$$P(k) = \frac{N_k}{N}$$

Where $N_k$ represents the total number of nodes with degree $k$ and $N$ represents node. The $P(k)$ of random and small-world networks follows poison distribution against degree, but for scale-free network, it obeys power-law distribution $P(k) \sim k^{-\gamma}$ where $4 \geq \gamma \leq 2$. The value of $\gamma$ represent many structural and organizing properties: (i) if $\gamma$ is $2 \geq \gamma \leq 3$, several smaller hubs are integrated with few large hub and a large number of separate nodes together, (ii) if $\gamma \leq 2$, consist modules/clusters structure showed that roles of modules are more important, known as hierarchical network, (iii) if $\gamma > 3$ the hubs are unrelated, losing many scale-free characteristics (Ravasz and Barabasi, 2003).

#### 2.2.2. Clustering coefficient $C(k)$

It indicates the overall organization of formation of clusters in networks and also characterizes the strength of internal connections between the nodes neighborhood in the network. For any a particular node, it is calculated by $C(k_i) = \frac{2m_i}{k_i(k_i-1)}$ where $m_i$ represent the total number of connections with its close neighbors. For scale-free and random networks, $C(k) \sim$ constant (Barabasi and Oltvai, 2004) and independent on $k$ and for hierarchical network, $C(k) \sim k^{-\alpha}$ where $\alpha \sim 1$ it follows power law-distribution (Borgatti and Everett, 2006).

#### 2.2.3. Neighborhood connectivity $C_N(k)$

Is the number of neighbors connected with a node and it defines the correlation pattern of connectivity for the interacting nodes of the network. $C_N(k)$ can be calculated by (Eq. 2).

$$C_N(k) = \sum_q qP(q|k)$$

![Fig. 1. Schematic workflow implemented to study Amyotrophic Lateral Sclerosis protein-protein interaction network.](image-url)
where \( P(q|k) \) is the conditional probability of creating a link from a node with a \( k \) degree to another node having a \( q \) degree (Traag et al., 2013). \( C_0(k) \sim \text{constant} \) for scale-free network while for hierarchical network structure \( C_0(k) \) follows power law scaling against degree \( k \), i.e., \( C_0(k) \sim k^\beta \), where, \( \beta \sim 0.5 \). If \( C_0(k) \) is an increasing function of \( k \), the positive and negative value of exponent \( \beta \) showed assortivity or dissortivity nature of the network topology respectively (Virkar and Clauset, 2014).

2.2.4. Betweeness centrality (\( C_{B} (v) \))

It measures a node occurring several times to bridge along the shortest path between nodes \( i \) to \( j \). It is calculated by (Eq. 3).

\[
C_{B} (v) = \sum_{i \neq j \neq k} \frac{d_{ij}(v)}{d_{ij}}
\]

where \( d_{ij}(v) \) denotes the number of geodesic paths connecting node \( i \) to node \( j \) passing through node \( v \). If a node in a network has a high betweenness centrality value, it indicates that the node lies on a path with many other nodes and have the significant ability to propagate information in the network (Borgatti and Everett, 2006).

2.2.5. Closeness centrality (\( C_{c} \))

It is recognised in terms of ‘shortest path lengths’ among the pair of nodes in a network. It can be calculated in terms of farness and is given by (Eq. 4).

\[
C_{c} (k) = \frac{N}{\sum_{j} d_{ij}}
\]

where \( d_{ij} \) is the geodesic distance between pair of nodes \( i \) and \( j \), and \( N \) is the nodes present in the network. The higher value of \( (C_{c}) \) for a node means more capability to propagate information in the entire network and the low value represents higher receiving capabilities of information.

Eigenvector centrality \( (C_{e}) \) is proportional to the total sum of the centrality of all neighborhood nodes. It describes the effect of a node on signal processing. It is calculated by (Eq. 5).

\[
C_{e} (i) = \frac{1}{d_{i}} \sum_{j \text{ non}(i)} v_{j}
\]

Where \( \text{non}(i) \) represents the nearest neighbor of \( i \) node in the network, with eigenvalue \( \lambda \) and eigenvector \( v_i \) of eigenvalue equations, \( A v_i = \lambda v_i \) where \( A \) is the network adjacency matrix. \( C_e \) the function of a node is a powerful indicator of information transmission and depends on the centralities of its neighbors, it varies across different networks (Bonacich, 1987).

2.3. Tracing of bottleneck-hubs in the network

In a ALS-PPIN, the nodes with high degree, considered as hubs, are important nodes, because they might be related to disease-causing genes (Barabasi et al., 2011; Lim et al., 2006; Stelzl et al., 2005). The nodes with the high Betweeness Centrality (BC) were defined as bottlenecks. It is believed that these nodes will play an important role in information flow and controlling capability. In the protein-protein interaction network of ALS, we mainly focused to identify the hubs, bottlenecks and bottleneck-hubs (Bn-H) that were the central to the ALS-PPIN. We selected the top 1% highest degree(k) and BC nodes from the overall nodes (1949) present in a network. Further, we identified the overlapping of nodes between hubs and bottleneck, which were considered as Bn-H. We also characterise non-Bn-H, non-hub bottlenecks in ALS-PPIN. In order to validate the importance of Bn-H, we also carried out computational knockout experiments. Removal of Bn-H may led to perturbation in the network the phenomenon referred as centrality-lethality rule (Jeong et al., 2001). In our study, to access the network’s organisation, change in the absence of the most influential bottlenecks-hubs one at a time, then computed the topological parameters of the modified/reorganized network to assess regulating capacities by evaluating the degree distribution and BC.

2.4. Module construction and their interaction with bottleneck-hubs

In the literature, several different types of clustering approaches have been suggested (Huang et al., 2013; Jeong et al., 2001; modeling, 1999; Vidal et al., 2011; Wu et al., 2008). Many frequently used methods are described in terms of comparable data assumptions (e.g., k-means and k-medoids). The k-means (MacQueen, 1967) method has been frequently employed by researchers (Wu et al., 2008) in partitional techniques. The number of groups (k) and a distance measure are required input parameters for this approach. Each data point is first assigned to one of the k clusters based on its distance from the centroids (cluster centres) of each cluster. We constructed the modules/clusters, using the “Molecular Complex Detection” (MCODE) v2.0.0 (Bader and Hogue, 2003), a cytoscape plugin that identify the nodes that are highly interconnected in the form of clusters representing relatively stable, multi-protein complexes that function as a single entity in the network. Clusters in a protein-protein interaction network are commonly protein complexes and pathways, whereas clusters in a protein similarity network are protein families. To separate the dense areas according to provided parameters, the approach uses vertex weighting by local neighbourhood density and outward traversal from a locally dense seed protein. The algorithm has the advantage of having a directed mode, which permits fine-tuning of clusters of interest without considering the rest of the network and analysis of cluster interconnectivity, which is important in protein networks. A known high rate of false positives in data from high-throughput interaction methods has no effect on the algorithm (Bader and Hogue, 2003). The algorithm uses a three-stage process: (i) Weighting: the nodes with the most linked neighbors receive a higher score. (ii) Molecular complex prediction: recursively add nodes to the complex that are over a specified threshold, starting with the highest-weighted node (seed). (iii) Post-processing: filters are applied to increase cluster quality (haircut and fluff). We used defaults values of MCODE, node score cutoff (0.2), haircut, node density cutoff (0.1), K-score (2), maximum depth (100). We selected the top modules based on M-score having values greater than 4. The interaction between the Bn-H and the modules was identified using Cytoscape v3.8.2. which provided the opportunity for a more precise understanding of the biological functions, providing valuable clues for biologists (Nafis et al., 2015).

2.5. Gene ontology enrichment and disease association of modules

The five modules were functionally annotated using the gProfiler package (Reimand et al., 2016) to correlate the biological significance. We showed the Gene Ontology classifications like molecular activity, biological process, cellular components of modules using filtering domain size set to “only annotated”, default g:SCS method for multiple testing correction for p-values, maximum p-value set to 0.05, numeric IDs as prefix ENTERZGENE_ACC. Further, we identifying disease-gene correlations for each module because it helps in the understanding of disease mechanisms, which have several applications including disease diagnosis, therapy, and prevention (Luo et al., 2019). The gene-disease class association is identified by the versatile platform “disgenet2r” an R package (Pinero et al., 2020). We used gene2disease function (dis-gene2r package) to retrieve the Gene–Disease class association for a given list of genes of each module with search options source database “ALL”, DisGeNET score (0–1). Higher the DisGeNET score shows a stronger correlation between the gene and disease and it showed the Gene-Disease class association in the terms of percentage. Diseases are grouped by their MeSH disease classes, and the Gene-disease association
is proportional to the percentage of diseases in each MeSH disease class. The Gene-Disease association score GDAs takes into account the number and type of sources (level of curation, organisms), and the number of publications supporting the association.

3. Results

3.1. Characterization and statistical analysis of network

We extracted the PPINs commonly present in at least two of interaction databases out of the five. The ALS-PPIN consists of 1949 nodes and 13087 edges. The results of the topological analysis of degree and BC of all 1949 nodes in Supplementary Table 1. We retrieved the nodes with top 1% highest degree and BC of total nodes (1949) (Table 1). Our network showed a disassortative nature due to the calculated negative value of $\rho_1$, the exponent of connectivity parameter and reflects that the Bn-H are still a significant part in regulating the stability of the network. To recognize the importance of the Bn-H nodes strength in signal processing in a network, we used three topological centrality parameters like Closeness centrality ($C_c$), Betweenness centrality ($C_B$) and Eigenvector centrality ($C_v$). In ALS-PPIN these parameters followed power law against degree(K) and showed positive exponents values indicate the strong regulating behaviour of the leading Bn-H. The calculated values of exponents and correlation coefficient(r) of $C_c(r = 1.67, r = 0.92)$, $C_B(r = 0.99, r = 0.87), C_v(r = 0.87, r = 0.93)$ are respectively. The correlation coefficient value for centrality measure are very high and good fitted with the data (Fig. 2E-G). The graph of betweenness against degree showed that high-connecting nodes have more controlling strength to outspread signal throughout the network (Fig. 2E). Additionally, the closeness centrality parameter showed that high degree nodes are quickly spread the information as compared to the lower degree nodes which considered as good receiver of propagated signal (Fig. 2F). Our network followed a hierarchical scale-free nature means network having modular structure and system level of organization.

3.2. Tracing of bottleneck-hubs in the network

In the protein-protein interaction network of ALS, we calculated the topological parameters of each node using Network analyzer (Assenov et al., 2008), the plug-in of Cytoscope v3.8.2 (Tang et al., 2015). The first two parameters, Degree distribution and BC were used to filter the hubs and bottleneck in a network. We selected top 19 proteins (1% of total nodes) in each parameter with highest degree as hubs and BC as bottlenecks. Here, we found 17 nodes are common in both the parameter as Bn-H, 2 non-Bn-H (DLST & EP300), and 2 non-hub bottlenecks (PARK7 & MAPT) (Fig. 3) (Table 1). Out of 17 bottleneck hubs, we identified the significant hubs, namely TP53, SOD1, CDC5L, SNW1, and VCP (Fig. 2), with the rationale of high degree nodes (Table 1) as the top five hubs in ALS-PPIN. The node TP53 had the highest degree (236) followed by SOD1 (211), CDC5L (186), SNW1 (173) and VCP (172). These Bn-H proteins represent as backbone of the network which has great influence on information flow and more control over the network.

In general, important genes may be identified by a series of independent gene knockout experiments. In order to validate the Bn-H, we also carried out computational knockout experiments. Removal of Bn-H may lead to perturbation in the network the phenomenon referred as centrality-lethality rule verified by various genomic investigations (Jeong et al., 2001). The removal of Bn-H does not cause network breakdown, the close interaction of these Bn-H can control network properties and regulatory mechanisms. The accumulation of mutations in any gene causes the dysfunction or removal of expression of a specific protein to become dysfunctional or absent in the cell. In general, this situation re-examining the network’s topological properties change may be shown by removing the corresponding hub/hubs (important genes/-genes that may have been altered) from the network and then (Peng et al., 2015). If the removal of hub/hubs results in a significant change in network properties, such as breakdown and change in information flow, the network is said to be controlled by the centrality-lethality rule (Jeong et al., 2001). In our study, to access the network’s organisation, change in the absence of the most influential Bn-H one at a time, then computed the topological parameter; Degree distribution and BC of the modified/reorganized network to assess the hubs’ regulating capacities. The hub removal analysis showed a significant magnitude of changes in network metrics in particular, such as degree and betweenness centrality. In betweenness centrality, all 17 bottleneck-hub were analysed by removing them from the network systematically. We found few Bn-H (AR, SOD1, TP53, CDC5L, VCP, EGFR and APP) that allowed significant changes in BC could enhance local and global signal propagation in regulating ALS (Supplementary Fig 1A). The removal of all Bn-H showed a significant control over the interactions of other nodes in the ALS-PPIN. The degree distribution showed mutual control over the

### Table 1

| S.NO | Name  | Degree (K) | Name  | Betweenness ($C_B$) |
|------|-------|------------|-------|---------------------|
| 1    | TP53  | 236        | VCP   | 0.06587026          |
| 2    | SOD1  | 211        | TP53  | 0.06480069          |
| 3    | CDC5L | 186        | SOD1  | 0.0560067           |
| 4    | SNW1  | 173        | AR    | 0.05264571          |
| 5    | VCP   | 172        | DISC1 | 0.05158762          |
| 6    | AR    | 171        | ATXN1 | 0.0441985           |
| 7    | DLD   | 155        | SQSTM1| 0.03839771          |
| 8    | DISC1 | 148        | APP   | 0.03540357          |
| 9    | NEK4  | 138        | CDC5L | 0.03374739          |
| 10   | HSPP90AA1 | 128     | PSEN1 | 0.0306112           |
| 11   | EGF   | 125        | EGFR  | 0.0313865           |
| 12   | DLST  | 124        | SNW1  | 0.02913541          |
| 13   | PDHA1 | 122        | HSPP90AA1 | 0.02680708 |
| 14   | ATXN1 | 119        | Htt   | 0.02414826          |
| 15   | Htt   | 117        | DLD   | 0.0236974           |
| 16   | SQSTM1| 117        | PDHA1 | 0.0223344           |
| 17   | APP   | 114        | NEK4  | 0.02047508          |
| 18   | PSEN1 | 110        | PARK7 | 0.02043984          |
| 19   | EP300 | 108        | MAPT  | 0.01992754          |
Fig. 2. The protein-protein interaction network of ALS and its topological properties. (A) The ALS-PPIN is represented in terms of nodes (proteins) and edges (physical interaction). The total number of nodes (1949) and edges (13087) are filled circles (orange) and lines (gray), respectively. The top five significant bottleneck-hubs such as TP53 (red), SOD1 (blue), CDC5L (green), SNW1 (yellow), VCP (light pink) in the order of size according to a degree, respectively. (B-G) The topological and centrality properties of network represented with correlation coefficient values ($r$) (B) probability of degree distribution $P(k)$, (C) average clustering coefficient $C(k)$, (D) average neighborhood connectivity ($C_N(k)$), (E) betweenness centrality ($C_B$), closeness centrality ($C_C$), eigenvector centrality ($C_E$). All these properties follow the power law scale and show the scale-free hierarchical nature of the network.
modules. Interestingly, we found SNW1, TP53, SOD1, VCP had the largest number of connections with five modules followed by (2, 1) (Figs. 5 and 6). They also showed an interaction with each other mediators of the modules (Table 3). Bn-H, which have the possibility of cross-talk among the modules. (Fig. 4) (Table 2). These modules have a system-level organization that respectively, with the corresponding edges of 103, 34, 176, and 140. (Fig. 4) (Table 2). These modules have a system-level organization that was maintained by connecting with other nodes and provide overall functional properties.

3.3. Modules of network and their cross-talk with bottleneck-hubs

We identified the 29 modules that are highly interconnected regions representing relatively stable, multi-protein complexes that function as a single entity in the network (Supplementary Table 2). We selected the top 5 modules based on the high M-score and the first module had 11 nodes and 44 edges with scoring (8.8). The second, third, fourth, and fifth modules had 41 (5.15), 15 (4.85), 76 (4.69), and 65 (4.37) nodes, respectively, with the corresponding edges of 103, 34, 176, and 140. (Fig. 4) (Table 2). These modules have a system-level organization that was maintained by connecting with other nodes and provide overall functionality to the network (Dong and Horvath, 2007). Out of the five modules, module-4 (CDC5L) and module-5 (SNW1) consists of one Bn-H, which not only controls the internal regulation of modules respectively; but also affected other modules by interacting with different nodes. Whereas other Bn-H, SOD1, TP53, and VCP were not observed in any of the five modules suggested that these Bn-H indirectly connected with modular function (Fig. 4). Cross-talk between the modules may be possible due to interaction with common Bn-H and removal of such regulators can affect the functionality of modules and leads to dis- sortivity of the network. These modules were found to be linked via Bn-H, which have the possibility of cross-talk among the modules. Further we also analysed the interaction of 17 Bn-H with the five modules. Interestingly, we found CDC5L, SNW1 and TP53 showed highest strength of interaction with all five functional modules. CDC5L, had the largest number of connections with five modules followed by SNW1, TP53, SOD1, VCP indicating that these proteins were the key mediators of the modules (Table 3). CDC5L and SNW1 were found to be highly connected with module-4 (28, 26) whereas least with module-3 (2, 1) (Figs. 5 and 6). They also showed an interaction with each other in the ALS-PPIN. TP53, SOD1, and VCP showed the interactions with modules-1, 2, 4, and 5 with different strength and control the functions of the modules (Figs. 7–9). SNW1, SOD1 and TP53 had the largest number of connections with a module-5 followed by CDC5L and VCP. From the analysis of Bn-H and module interaction, we identified that CDC5L and SNW1 highly regulates one cluster of nodes (module-5, 2 & 4) and TP53 regulates another cluster of nodes (module-5 & 4). SOD1 interacted with VCP in ALS-PPIN and also showed an interaction with modules-1, 2, 4, and 5. We also identified that VCP highly interacted with the module-4 and equally interacted with module-1, 2, and 5. Interestingly we also found among all 17 Bn-H, only CDC5L, SNW1, PDHA1 and PSEN1 showed an interaction with Module-3 (Table 3).

3.4. Functional enrichment analysis of modules

We performed the functional enrichment analysis for all five modules. Module-1 was enriched with proteins having proteasome-activating activity, endopeptidase activity and are associated with biological processes such as protein deubiquitination, protein modification, proteasome protein catabolic process. These proteins were found to be localized in proteasome complex, endopeptidase complex, peptide complex (Fig. 10A). The Module-2 proteins were found to be enriched in RNA binding function and regulate mRNA processing, mRNA splicing via spliceosome, mRNA metabolic process. These proteins were found mainly in the ribonucleoprotein complex (Fig. 10B). The proteins of module-3 were present in the spliceosome complex and are involved in Golgi vesicle transport process by binding to the SNARE complex. These proteins were also involved in SNAP receptor activity (Fig. 10C). Module-4 showed the involvement of the protein in binding to various macromolecules such as an enzyme, nucleic acid, RNA, DNA that perform RNA splicing via transesterification, mRNA splicing via spliceosome, and are mainly found in nucleoplasm, membrane-bound lumen, intracellular organelle lumen (Fig. 10D). Proteins associated with module-5 are mainly present cytosol, protein-containing complex, nucleus, membrane enclosed lumen performing molecular functions such as enzyme binding, protein binding, kinase binding thus regulating the viral process, cell cycle, protein modification process (Fig. 10E). We also analysed that the CDC5L present in module-4, and from the result of functional enrichment it was related to mRNA metabolic process, suggesting that the influence of the CDC5L in mRNA metabolism using spliceosome and transesterification activities. Furthermore, the SNW1 also present in module-5, which is functionally associated with enzyme binding as molecular function and also involved in viral process in the cytoplasm.

3.5. Module-disease class associations

The module-disease association is identified by DiSGenet tool of each module (module-1–5). We found that the proteins of all five modules were associated with different disease classes like neoplasm, nervous system disorder, digestive, respiratory tract, cardiovascular, mental disorder, immune system, hemic and lymphatic, congenital, hereditary, neonatal, and others. Disease-associated with these modules depicts the majority of genes of module-1 showed association with neoplasm disease (40%), immune system disease (30%), and nervous system disease (30%). PSMC4 gene present in module-1 showed high association with neoplasm (40%) as well as nervous system (30%) while it does not show association with immune system disease (Supplementary Fig. 2). In module-2 proteins are highly associated with neoplasm (50%), pathological conditions (40–50%), nervous system (30%), cardiovascular disease (40%), and digestive system disease (20%). SNRPD3 and U2AF2 showed (50%) the highest association; DDXX, DDX17, and SNRPF in between (40–50%) with neoplasm class. SNX32 the proteins showed 50% association with the pathological condition and cardiovascular disease. ATXN1, ATXN2, FMRI, FXR2, NUP93, PRPF6, PRPF8, PSEN1 were mostly associated with nervous system disorder (30–40%) (Supplementary Fig. 3). Module 3 proteins were highly associated with neoplasm disease (40–50%), nervous system disease (60%), congenital, hereditary, and neonatal disease and abnormalities (20–30%) while these proteins were least associated with chemically induced disorders (1%), otohinolaryngologic disease (<10%) and male urogenital disease.
Fig. 4. Structure of five modules in PPI network. All the top five modules are constructed and analysed using MCODE. (A) All the nodes in module 1 are the filled circle (green cyan) consist nodes(11) edges(44) with scoring value(8.8); (B) module 2 are the filled circle (pink) consist nodes(41) edges(103) with scoring value(5.15); (C) module 3 are the filled circle (green) consist nodes(15) edges(34) with scoring value(4.85); (D) module 4 are the filled circle (red) consist nodes(76) edges(176) with scoring value(4.69); (E) module 5 are the filled circle (voilet) consist nodes(65) edges(140) with scoring value(4.37) with the corresponding edges in lines (gray). One of the significant bottleneck-hub (CDC5L) present in module 4 as filled circle (dark green) and SNW1(yellow) in module 5.
Among genes of module-3, NAPA showed the strongest association (60%), SCFD1, and STX5 (30–40%) suggesting their possible role in the etiology of nervous system diseases (Supplementary Fig. 4). Among all the modules, module-4 showed the highest association with different disease classes. The majority of proteins are associated with neoplasm disease, pathological conditions, nervous system diseases, digestive system disease. These proteins have a negligible association with occupational disease, chemically-induced disorders. HSPD1 and PSMB4 were observed to show strongly interrelated with neoplasm disease (80%) while GEMIN6 is highly interleaved with nervous system diseases (80%). SNRPG showed 50% association with a pathological condition, skin, and connective tissue disease class (Supplementary Fig. 5). Furthermore, the major group of proteins in module-5 were strongly linked to neoplasms (60%), nervous system disease (40%), pathological conditions (30%). Several proteins like CPSF7, EEF1G, RNF11, RPA3, RPL30, SH3GL3, SKP2, SNW1 were highly correlated with the neoplasms’ disease (50–60%). The nervous system disease class consists of proteins APP, DYNC1H1, DYNC1I2, HINT1, LSM2, RNF11, SNCA in between 30% and 40% association. One protein DYNLT1 showed a major association with respiratory tract (60%), cardiovascular and male urogenital (30–40%) (Supplementary Fig. 6).

4. Discussion

In this study, we applied a computational systems biology approach to investigate the ALS-PPIN. To construct the ALS-PPIN, we used five different human databases STRING, BIOGRID, DIP, IntACT, and MINT. Finally, we constructed ALS-PPIN consist of 1949 nodes and 13087 edges. It followed a hierarchical scale-free behaviour. The nodes with number of links that greatly exceeds the average value were considered as hubs (Chen et al., 2019; Pang et al., 2016). In the ALS-PPIN, we considered high-degree nodes as hubs. In order to filter the potential hubs, we selected 19 highest degree nodes (1% of total nodes (1949)) as hubs (Chen et al., 2019; Pang et al., 2016). In the ALS-PPIN, we considered 1% of BC nodes which were considered as bottleneck hubs (Chen et al., 2019; Pang et al., 2016). In the ALS-PPIN, we considered 19 highest BC nodes which were considered as bottleneck hubs (Chen et al., 2019; Pang et al., 2016). Further, out of 17 significant bottleneck hubs, we identified the significant hubs, selected 1% of BC nodes which were considered as bottleneck hubs (Chen et al., 2019; Pang et al., 2016).

Among the 19 hubs, we found only 17 hubs were also present in selected 1% of BC nodes which were considered as Bn-H, could provides the stability and control the information flow in the network (Table 1). Further, out of 17 bottleneck hubs, we identified the significant hubs, namely TP53, SOD1, CDC5L, SNW1, and VCP with the rationale of high degree nodes in the ALS-PPIN (Fig. 2). CDC5L, TP53, SNW1, SOD1, VCP were top hubs as well as Bn-H which were the dominant preservers of the topological properties of the network. Interestingly, we identified SOD1 as a Bn-H which is already reported as a high-risk gene in ALS (Rosen et al., 1993). CDC5L, TP53, SNW1, SOD1, and VCP have significant roles because they interact with functional modules with high number of strengths as compared to other Bn-H. The correlation between Bn-H and modules gave an idea that Bn-H also controlled the regulation of these modules. Based on Bn-H and module interaction we found CDC5L, had the largest number of connections with five modules followed by SNW1 and TP53; indicating that these proteins were the key mediators of the modules (Table 3). Bn-H could preserve the stability of the network, information was fast and molecules were quickly accessible by directly interacting with the nodes in the module (Nafis et al., 2015).

Out of 17 significant Bn-H, only CDC5L was present in module 4 and SNW1 in module 5, which was the influencing node with strong cross-talking among the module that interacted with it. SOD1, TP53, and VCP were not present in any of the five modules, acted as mediators to cross-talk among the modules and also indirectly interfered with modular properties and activities. Because these Bn-H are strongly sensitive in preserving the topological properties of the network, their absence causes the breakdown of the network. Therefore, acting as the key mediator in the regulation of ALS-PPIN. Modules that perform significant functions in the network were constructed. In this study, the correlation between Bn-H and modules interaction gave an idea that Bn-H were also responsible in controlling the regulation of these modules.

In module-1, CDC5L interacts with three proteins PSMC4, PSMD2, and PSMC5 which mainly participate in protein homeostasis and cell processes like DNA damage, apoptosis, cell cycle progression. In module-2, five proteins (PSMD1, DDX5, PCBP1, ELAVL1, and HNRNPK) were interacting with both CDC5L and TP53 which are functionally enriched with mRNA splicing mechanism which is known to be highly affected in ALS (Butti and Patten, 2018). Interestingly in module-4, we found CDC5L connected with FUS protein, which is one of the high-risk proteins of both familial and sporadic type ALS (Volk et al., 2018). Dementia (Mompean and Laurents, 2017), and Parkinson’s disease (Yan et al., 2010). TAF15 and FUS affect the turnover of their RNA targets in ALS (Kapeli et al., 2016). The interaction of CDC5L with module-5 proteins which were functionally associated with enzyme activity, neoplasm, and nervous system diseases.

| S.NO | Name of bottleneck-hubs | Module 1 | Module 2 | Module 3 | Module 4 | Module 5 | Total |
|------|-------------------------|----------|----------|----------|----------|----------|-------|
| 1    | CDC5L                   | 3        | 14       | 2        | 28       | 13       | 60    |
| 2    | SNW1                    | 3        | 13       | 0        | 22       | 14       | 57    |
| 3    | TP53                    | 4        | 9        | 0        | 16       | 9        | 49    |
| 4    | NEK4                    | 0        | 11       | 0        | 10       | 15       | 29    |
| 5    | HSP90A1                 | 3        | 1        | 0        | 10       | 15       | 29    |
| 6    | SOD1                    | 5        | 4        | 0        | 5        | 14       | 28    |
| 7    | EGRF                    | 1        | 5        | 0        | 6        | 13       | 25    |
| 8    | DLD                     | 4        | 5        | 0        | 8        | 8        | 25    |
| 9    | VCP                     | 5        | 5        | 0        | 8        | 5        | 23    |
| 10   | AR                      | 0        | 4        | 0        | 6        | 12       | 22    |
| 11   | PDIHA1                  | 0        | 3        | 2        | 5        | 6        | 16    |
| 12   | ATXN1                   | 1        | 4        | 0        | 2        | 7        | 14    |
| 13   | FSEN1                   | 3        | 4        | 1        | 5        | 5        | 18    |
| 14   | DSCC1                   | 0        | 1        | 0        | 6        | 4        | 11    |
| 15   | APP                     | 2        | 3        | 0        | 5        | 12       | 20    |
| 16   | SQSTM1                  | 1        | 3        | 0        | 11       | 15       | 20    |
| 17   | HTT                     | 4        | 5        | 0        | 10       | 13       | 32    |
Fig. 5. Cross-talk between the significant bottleneck-hub \textit{CDC5L} and five modules in the network. The \textit{CDC5L} bottleneck-hub showed at the centre (green) and its interacting nodes(yellow) of five modules.
Fig. 6. Cross-talk between the significant bottleneck-hub SNWI and five modules in the network. The SNWI bottleneck-hub showed at the centre (orange) and its interacting nodes(yellow) of five modules.
Fig. 7. Cross-talk between the significant bottleneck-hub TP53 and five modules in the network. The TP53 bottleneck-hub showed at the centre (red) and its interacting nodes(yellow) of five modules.
Fig. 8. Cross-talk between the significant bottleneck-hub SOD1 and five modules in the network. The SOD1 bottleneck-hub showed at the centre (blue) and its interacting nodes(yellow) of five modules.
Fig. 9. Cross-talk between the significant bottleneck-hub *VCP* and five modules in the network. The *VCP* bottleneck-hub showed at the centre (light pink) and its interacting nodes(yellow) of five modules.
Fig. 10. Manhattan plot that represents the enrichment analysis of five modules. The x-axis represented functional term’s Molecular function (GO: MF) is red; Biological process (GO: BP) is orange and Cellular component (GO: CC) is green. The sizes of the filled circle according to the term size, means larger terms have larger circles. The y-axis shows the adjusted enrichment p-values in the negative log10 scale.
In module-1, SWN1 showed interaction with three proteins (PSMC5, PSMC4, PSMD2) involved in protein metabolism and other cellular function like DNA damage, cell cycle progression. Thirty proteins showed interaction with SWN1 in module-2 (Fig. 6) which are functionally enriched with mRNA processing process. SWN1 also showed the interaction with NSF in module-3, which is functionally associated with SNARE binding. NSF protein functionally important for the delivery of cargo proteins to all compartment of the Golgi stack and highly associated with neoplasm and nervous system diseases. In module-4, SWN1 interacting with more than 25 proteins which are functionally enriched with mRNA processing. SWN1, have been associated with cell migration in glioblastoma (Ramos et al., 2019) and also reported in axonal transport implication during tau organization related to Alzheimer’s disease (Ishashko-Pachima and Gozes, 2019). Interestingly, SWN1 showed direct interaction with CDC51 (157–536 residue), it functions as a coactivator that may couple vitamin D receptor-mediated transcription and RNA splicing (Zhang et al., 2000). In module-5, SWN1 showed interaction with 14 proteins involved functionally in enzyme activity mostly observed in ALS disease progression (Srinivasan and Rajasekaran, 2020). Two proteins of module-5, LSM2 (also known as snRNP) and HSPA8 both are reported as ALS causative proteins. Mutation in FUS/TLS protein reduced genes formation, altered snRNPs in patients’ fibroblasts and transgenic mice. Along these lines, it is also observed that FUS/TLS mutations reduce interaction with U1-snRNP (Sun et al., 2015; Yu et al., 2015). Post transcriptional inhibition of HSPA8 expression leads to synaptic vesicle cycling defects in multiple models of ALS (Coyne et al., 2017).

TP53 showed interaction with (PSMD4, PSMD2, PSMC5, RAD23) of module-1 involved in the protein deubiquitination process. In module-2, the interaction of TP53 with HDAC2, DDX5, PCBP1, PHSN, ELAVL1, DDX20, PSMD1, UBC, HNRNPK were enriched in mRNA splicing. The dysregulation or inhibition of protein deubiquitination (Cykowski et al., 2019) and mRNA splicing (Butti and Patten, 2018) processes favor ALS. Phosphorylation of HNRNPK by cyclin-dependent kinase 2 controls the cytosolic accumulation of TDP-43 in ALS (Moujalled et al., 2015). Interestingly, HNRNPK and DDX5 proteins are transcriptional activators of TP53 and help to regulate the intrinsic apoptotic pathway in response to the DNA damage process observed in ALS (Pelisch et al., 2012). DDX20 is associated with a neurogenerative disease SMA (Charroux et al., 1999). Functional loss of DDX20 was observed in ALS on disruption of high-risk gene FUS (Cacciottolo et al., 2019). In modules-4, TP53 interacting proteins are majorly enriched with mRNA splicing function and highly associated with neoplasms diseases. SMN1 interacted with TP53 in module 4 reported in SMA (Singh and Singh, 2019), and ALS (Chi et al., 2018). TP53 showed interaction with modules – 5 proteins which are functionally enriched with enzyme activity that altered process in ALS (Srinivasan and Rajasekaran, 2020).

Association of SOD1 with module-1 proteins (PSMD2, ADRM1, RAD23A) were involved in ALS-S4G with or without Frontotemporal dementia (Huttlin et al., 2017). The proteins of module-2 (PSME2, PSMA6, ELAVL1, YWHAQ) and module-4 (OPN, PSMC2, PSMA2, PSMC6, BANFI1) are enriched with RNA processing and protein homeostasis whose disruptions function observed in ALS (Butti and Patten, 2018; Mejzini et al., 2019). Interestingly, SOD1 showed interaction with OPN in involved in neuroinflammation, autophagy, and vesicular trafficking in ALS (Maruyama et al., 2010), FTD (Hortobagyi et al., 2011), Alzheimer’s disease, and Huntington’s disease (Schwab et al., 2012). SOD1 showed interaction with module-5 which is highly associated with enzymatic activity. The proteins (SNCA, HSPA8, HINT1) are associated with neurodegenerative diseases like Parkinson’s, Alzheimer’s, Huntington’s, Prion disorders, and Fronto-temporal dementia (Lin et al., 2020; Matsumura et al., 2013; Shchagina et al., 2020; Sung et al., 2005; Wytenbach and Arrigo, 2009). The nodes in module-5 (RUVBL1, RUVBL2) interacts with HINT1 and modulates TP53 levels and TP53-mediated apoptosis (Weiske and Huber, 2006).

Many studied reported that more than 50 missense mutations in gene coding VCP is causative in many neurodegenerative diseases characterised by ALS, FTD, IBM, CMT2Y, and PBD (Al-Tahan et al., 2018). Approximately 9% of patients with VCP mutations had ALS phenotype, 4% with Parkinson’s disease, and 2% has been diagnosed with Alzheimer’s (Al-Obedi et al., 2018). In module-1, VCP interacts with 5 proteins which are functionally associated with protein metabolism and mostly associated with neoplasms disorders. VCP showed interaction with (ELAVL1, HNRNPK, ATXN1, PSM7, USP14) of module-2 play important role in protein homeostasis process which is mostly affect in ALS. A functional deficiency of VCP observed contributes to impaired DNA repair in multiple polyglutamine diseases. Although normal and mutant polyglutamine proteins (ATXN1) interact with VCP, only mutant protein affect dynamism of VCP (Fujita et al., 2013). In module-3 VCP showed no interaction with any of the proteins. In module-4 VCP interacts with 8 proteins (HSPP0AA1, USP7, PSMA1, PSMA2, PSMA4, FUS, RANBP2, OPTN) functionally enrichment with mRNA processing. Both FUS and OPTN are well known studied protein to cause ALS and cause aberrant protein homeostasis due to various mutations. FUS protein reported as one of the top 4 high risk genes to cause both sALS and iALS (Yolk et al., 2018). Q290X mutation leads to FUS mRNA degraded by nonsense-mediated decay, which results in loss of FUS functions and cause ALS (Deng et al., 2014). OPTN is an autophagy receptor and mutations in the OPTN gene result in familial glioma (E50K) and ALS (E478G) reportedly abolishes its NF-κB suppressive activity (Nakazawa et al., 2016; Shen et al., 2015). In module-5, VCP shown interaction with 5 proteins (AKTI, CA1, HSPA8, DNM1L, TKT) functionally associated with enzyme binding processes and most of these genes are mostly related to neoplasms disorder followed by nervous system diseases. role of DNM1L seen in abnormal mitochondrial dynamics, mitochondrial fragmentation, autophagy/mitophagy, and neuronal damage in Alzheimer’s disease and other neurological diseases, including Parkinson’s, Huntington’s, ALS, multiple sclerosis, diabetes, and obesity (Olive and Reddy, 2019; Vantaggiato et al., 2019). Analysis reveals that LanCL1 is a positive regulator of AKT1 activity, and LanCL1 overexpression restores the impaired AKT1 activity in ALS model mice (Tan et al., 2020). Neuron-targeted CAV1 improves neuromuscular function and extends survival in SOD1G93A transgenic mice (Sawada et al., 2019).

We also found 29 node/protein, showed a common connection between the five Bn-H (CDC51, SNW1, TP53, SOD1, and VCP) and others also in ALS-PNP. HSPPA5 nodes showed interaction with 10 Bn-H (APP, AR, ATXN1, CDC51, EGRF, HSPP0AA1, PSEN1, SNW1, SQSTM1, and TP53) whereas HSPPA8 (with APP, ATXN1, CDC51, HSPP0AA1, HTT, PSEN1, SNW1, SOD1, TP53, and VCP). HSPPA5 majorly associated with digestive system disorders and neoplasm (Liver carcinoma) (Feng et al., 2019; Shu et al., 2020). HSPPA5 (GRP78) activates the Wnt/HOXB9 pathway to promote invasion and metastasis of hepatocellular carcinoma by chaperoning LRP6 (Xiong et al., 2019). Post transcriptional inhibition of HSPPA5 expression leads to synaptic vesicle cycling defects in multiple models of ALS (Coyne et al., 2017). HSPPA5 mostly reported in many neurodegenerative diseases like Parkinson’s, Alzheimer’s, Huntington’s, Prion disorders, and Fronto-temporal dementia (Lin et al., 2020; Matsumura et al., 2013; Sung et al., 2005). Whereas, GAPDH is another node showed interactions with 9 Bn-H (APP, AR, ATXN1, CDC51, HSPP0AA1, HTT, PSEN1, SOD1, TP53). The alteration in GAPDH function is associated with oxidative stress in ALS (Pierce et al., 2008). S-nitrosylated GAPDH mediates neuronal apoptosis induced by ALS-associated mutant SOD1G93A (Lee et al., 2016). GAPDH expression defects were also found in muscles from ALS patients (Deselle et al., 2017). The nodes (PAR7, HSPPA4, CCA2) were found as an intermediate for SOD1, TP53, AR and HSPP0AB1, RUVBL1 for AR, SOD1, CDC51. In future, biochemical investigation of the observed Bn-H and their interacting partners could provide further understanding to prioritize key genes and their role in the pathophysiology of ALS.
Ethics declarations
Not applicable.

CRediT authorship contribution statement
Rupesh Kumar: Conceptualization, Data curation, Investigation, Writing – original draft. Shazia Haider: Conceptualization, Writing – review & editing, Supervision.

Conflicts of Interest
The authors declare that they have no conflict of interest.

Acknowledgment
R.K gratefully acknowledge the Council for Scientific and Industrial Research (CSIR), New Delhi- 110005, India. for providing research fellowship. Sincerely thank the Jaypee Institute of Information Technology, Noida, Sec-62, Uttar Pradesh, India. We acknowledge Dr. Kaalairasan Ponnusamy for careful reading of the Manuscript and critical inputs.

Appendix A. Supporting information
Supplementary data associated with this article can be found in the online version at doi:10.1016/j.jneurb.2021.12.002.

References
Abel, O., Powell, J.F., Andersen, P.M., Al-Chalabi, A., 2012. ALSDo: a user-friendly online bioinformatics tool for amyotrophic lateral sclerosis genetics. Hum. Mutat. 33 (9), 1345–1351. https://doi.org/10.1002/humu.22157.
Al-Obeidi, E., Al-Tahan, S., Surampalli, A., Goyal, N., Wang, A.K., Hermann, A., Abel, O., Powell, J.F., Andersen, P.M., Al-Chalabi, A., 2012. ALSoD: a user-friendly online version at doi:10.1016/j.ibneur.2011.12.002.

Chiu, R., Chio, A., Traynor, B.J., 2018. Novel genes associated with amyotrophic lateral sclerosis: diagnostic and clinical implications. Lancet Neurol. 17 (1), 94–102. https://doi.org/10.1016/S1474-4422(17)30401-5.
Chio, A., Logroscino, G., Hardiman, O., Swingler, R., Mitchell, D., Beghi, E., Traynor, B. G., Eur Atlas, 2009. Prognostic factors in ALS: A critical review. Amyotroph Lateral Scler 10 (5–6), 310–323. https://doi.org/10.3109/17482908026560274.
Clauzet, A., Shalizi, C.R., Newman, M.E.J., 2009. Power-law distributions in empirical data. SIAM Rev. 51 (4), 661–703. https://doi.org/10.1137/070710111.
Cooper-Knock, J., Moll, T., Ramesh, T., Castell, L., Beare, A., Roberts, H., Fox, L., Nemerowner, I., Van Damme, P., Moisesc, M., Bobrova, L., Hardiman, O., Panades, M.P., Ansalouli, A., Mora, J.S., Basak, A.N., Morrison, K.E., Shaw, C.E., Al-Chalabi, A., Landers, J.E., Wyles, M., Heath, P.R., Higginbottom, A., Walsh, T., Kazakova, M., McDermott, C.J., Hauertgrube, G.M., Kirby, J., Shaw, P.J., 2019. Mutations in the glycyrretinaldehyde transfer domain of GUSB1 are associated with familial amyotrophic lateral sclerosis, 2298-2306.e2295 Cell Rep. 26 (9). https://doi.org/10.1016/j.celrep.2019.02.006.
Coyne, A.N., Lorenzo, I., Chou, C.C., Torvund, M., Rogers, R.S., Starr, A., Zebeleif, B.L., Levy, J., Johannesmees, J., Schwartz, J.C., Niithumine, H., Zisemke, K., Rossol, W., Sattler, R., Zarnescu, D.C., 2017. Post-transcriptional inhibition of Hs70-4/ISPA8 expression leads to synaptic vesicle cycling defects in multiple models of ALS. Cell Rep. 21 (1), 110–125. https://doi.org/10.1016/j.celrep.2017.09.026.
Cykowski, M.D., Dickson, D.W., Powell, S.Z., Arumanayagam, A., Rivera, A.L., Appel, S.H., 2019. Dipeptide repeat (DRP) pathology in the skeletal muscle of ALS patients with C9ORF72 repeat expansion. Acta Neuropathol. 138 (4), 667–670. https://doi.org/10.1007/s00401-019-01930-8.
De Vos, K.J., Chapman, A.L., Tennant, M.E., Manser, C., Tudor, E.L., Lau, K.F., Brownlee, J., Ackerley, S., Shaw, P.J., McLoughlin, D.M., Shaw, C.E., Leigh, P.N., Miller, C.C., Grisson, A.J., 2007. Familial amyotrophic lateral sclerosis-linked SOX5 mutants perturb mitochondrial transport to reduce axonal mitochondria content. Hum. Mol. Genet. 16 (22), 2720–2728. https://doi.org/10.1093/hmg/ddm226.
Djelasus-Hernandez, M., Mackenzie, I.R., Beove, B., Boxer, A.L., Baker, M., Rutherford, N.J., Nicholson, A.M., Finch, N.C.A., Flynn, H., Adamson, J., Kouri, N., Wojtas, A., Sengdy, P., Hsing, Y.G.R., Karydas, A., Seeley, W.W., Josephs, K., Coppola, G., Geschwind, D.H., Wiszek, Z.K., Feldman, H., Knapson, D.S., Petersen, R.C., Miller, B.L., Dickson, D.W., Boylan, K.B., Graft-Redford, N.R., Rademakers, R., 2011. Expanded GGGGCC hexanucleotide repeat in noncoding region of ORFX2 causes chromosome 9p-linked FTD and ALS. Nature Neurol 72 (2), 245–256. https://doi.org/10.1038/nn.2811.2011.09.031.
Deng, H., Gao, K., Jankovic, J., 2014. The role of FUS gene mutations in neurodegenerative diseases. Nature Rev 14 (9), 373–387. https://doi.org/10.1038/nrn3809.
Deng, H.X., Chen, W., Hong, S.T., Boycott, K.M., Gorrie, G.H., Siddique, N., Yang, Y., Fecto, F., Shi, Y., Zhai, H., Jiang, H., Hirano, M., Rampaesaud, E., Jansen, G.H., Dencker, B.R., Bigio, E., Lim, H., Okada, T., McAllister, K., Ackerley, S., Shaw, P.J., 2019. Mutations in the glycyrretinaldehyde transfer domain of GUSB1 are associated with dominant X-linked juvenile and adult-onset ALS and ALS/dementia. Nature 477 (7363), 211–215. https://doi.org/10.1038/nature10535.
Dervishi, I., Gouotok, O., Munra, K., Gautam, M., Heller, D., Bigio, E., Ondler, P.H., 2018. Protein-protein interactions reveal key canonical pathways, upstream regulators, interactome domains, and novel targets in ALS. Sci. Rep. 8 (1), 14732. https://doi.org/10.1038/s41598-018-32902-4.
Desselle, C., Deforges, S., Biondi, O., Hervezine, L., Amico, D., Lamaziere, A., Caradeuc, C., Bertho, G., Brunetseau, G., Weill, L., Bustin, J., Djoos, F., Salachas, F., Lopes, P., Chanoine, C., Massaad, C., Charbonnier, F., 2017. Specific physical exercise improves energetic metabolism in the skeletal muscle of amyotrophic-lateral-sclerosis mice. Front. Mol. Neurosci. 10, 332. https://doi.org/10.3389/fnins.2017.00332.
Dong, J., Horvath, S., 2007. Understanding network concepts in modules. BMC Syst. Biol. 1 (1), 24. https://doi.org/10.1186/1752-0509-1-24.
Eden, B.M., Yan, J., Miller, N., Deng, H.X., Siddique, T., Ma, Y.C., 2017. A novel ALS-associated variant in UBQLN4 regulates motor axon morphogenesis. elife 6. https://doi.org/10.7554/eLife.25453.
Eden, A.C., Kim, H.J., Hart, M.P., Chen-Plotkin, A.S., Johnson, B.S., Fang, X., Armaka, M., Gese, F., Gretz, R., La, M.M., Padmanabhan, A., Clay-Falcone, L., McCluskey, L., Eiman, L., Ju, D., Gruber, P.J., Rüb, U., Auburger, G., Trojanowsk, J.Q., Lee, V.M.Y., Van Deeen, V.M., Bonini, N.M., Gitler, A.D., 2010. Ataxin-2 intermediate-length polyglutamine expansions are associated with increased risk for ALS. Nature 466 (7310), 1069–1075. https://doi.org/10.1038/nature09320.
Farhan, S.M., Howigran, D.P., Abbott, L.E., Bynet, C.E., Churchrise, C., Phatnai, H., Phatnai, C.N., Topp, S., Remonstad, E., Wu, G., Wu, J., Gubin, A.H., Klim, J.H., Mordes, D.A., Ghosh, S., Eggan, K., Rademakers, R., Mccluskey, J.L., Schüle, R., Zücher, S., Benatar, M., Taylor, J.P., Nalls, M.A., Traynor, B., Shaw, C.E., Goldstein, A.,
2015. Phosphorylation of MrkNpK by cyclic-depend kinase 2 controls cytoxic accumulation of TDP-43. Hum. Mol. Genet. 24 (6), 1655–1669. https://doi.org/10.1093/hmg/ddv168.

Nafis, S. Kalaairasan, B., Projen Singh, R.K., Husain, M., Bamezai, R.N., 2015. Apoptosis regulatory protein-protein interaction demonstrates hierarchical scale-free fractal network. Brief Bioinform. 16 (4), 675–699. https://doi.org/10.1093/bib/bbv006.

Nakatani, N., Okabe, A., Higashihara, T., Yaku, T., Hasegawa, A., Nakano, B.N., Mann, L., Topp, S.D., Abravaya, K., Hanka, N.K., Suvanto, P., Kamei, K., Takeyoshi, I., Kawahama, H., Iwai, K., Hatada, I., Sawasaki, T., Ito, H., Nureki, O., Tokunaga, F., 2016. Linear ubiquitination is involved in the pathogenesis of optineurin-associated amyotrophic lateral sclerosis. Nat. Commun. 7, 12547. https://doi.org/10.1038/ncomms12547.

Nicolas, A., Kenna, A., Renton, A.E., 2014. How RNA structure dictates the usage of a critical exon of the spinal muscular atrophy gene. Biochim. Biophys. Acta Gene Regul. Mech. 1862 (6), 699. https://doi.org/10.1016/j.bib.2014.05.046.

Pelisch, F., Pozzi, B., Risso, G., Munoz, M.J., Seebrow, A., 2012. DNA damage-induced heterogeneous nuclear ribonucleoprotein K sumoylation regulates p53 transcriptional activity. Biochim. Biophys. Acta 1827 (36), 30789–30799. https://doi.org/10.1016/j.bbamem.2012.04.001.

Peng, E., Hao, Y., Sun, Y., Lin, K., 2016. Differential variation patterns between hubs and bottlenecks in human protein-protein interaction networks. BMC Evol. Biol. 16 (6), 210. https://doi.org/10.1186/s12862-016-0840-4.

Pastor-Satorras, R., Vazquez, A., Vespignani, A., 2001. Dynamical and correlation networks. Phys. Rev. Lett. 87 (25), 258701. https://doi.org/10.1103/PhysRevLett.87.258701.

Penny, J., Ramírez-Anguita, J.M., Sauch-Pitarch, J., Ronzano, F., Centeno, E., Sanz, F., Purlong, L.I., 2020. The DeGeNet knowledge platform for disease genomics: 2019 update. Nucleic Acids Res. 48 (D1), D845–D855. https://doi.org/10.1093/nar/gkz1021.

Pierce, A., Mirzaei, H., Muller, F., De Waal, E., Taylor, A.B., Leonard, S., Van Remmen, H., Regnier, F., Richardson, A., Chauvidia, A., 2019. GDPi is conformationally and functionally altered in association with oxidative stress in mouse models of amyotrophic lateral sclerosis. J. Mol. Biol. 432 (5), 1195–1210. https://doi.org/10.1016/j.jmb.2018.07.088.

Pinero, J., Renton, A.E., Vazquez, A., Vasilimon, V., Girdhar, A., Patel, B.K., 2019. Molecular mechanisms of TDP-43 misfolding and pathology in amyotrophic lateral sclerosis. In: Ramos, A.R., Ghosh, S., Dekkere, M., Robe, P.A., Rogister, B., Erneux, C., 2019. Lipid phosphates skip and SHIP2 regulate fibronectin-dependent cell migration in glioblastomas. FEBS J. 286 (6), 1120–1135. https://doi.org/10.1111/febs.14769.

Ravasz, E., Barabási, A.L., 2003. Hierarchical organization in complex networks. Phys. Rev. E Stat. Nonlin. Soft Matter Phys. 67 (2 Pt 2), 026112. https://doi.org/10.1103/PhysRevE.67.026112.

Reimand, J., Arak, T., Adler, P., Linde, E., Peterson, H., Vilo, J., 2016. g:Profiler—a web server for functional interpretation of gene lists (2016 update). Nucleic Acids Res. 44 (W1), W83–W89. https://doi.org/10.1093/nar/gkw199.

Renton, A.E., Majounie, E., Waite, A., Simón-Sánchez, J., Rollinson, S., Gibbs, J.R., Schymick, J.C., Lassakowska, E., van Swieten, J.C., Myllykangas, L., Kalimo, H., Poletto, M., Uramono, Y., Malaspina, A., Aisen, P., Poirier, M., Girulli, A., Doherty, J., Darvasi, F., Dupre, E., Lai, C., Parati, G., Wallin, E., Nalls, M.A., Tipping, E., Mora, J.S., Van Blitterswijk, M., Morrison, K., Basak, A.N., Wu, X., Wang, J., Wang, J., Wu, F.X., Pan, Y., 2015. Rechecking the centrality-lethality mechanism of TDP-43. Neurobiol. Aging 36 (2), 971–985. https://doi.org/10.1016/j.neurobiolaging.2014.10.016.
