Identification of viral-mediated pathogenic mechanisms in neurodegenerative diseases using network-based approaches

Anna Onisiforou and George M. Spyrou

Corresponding author. George M. Spyrou, Bioinformatics Department, Cyprus School of Molecular Medicine, Cyprus Institute of Neurology & Genetics, 6 Iroon Avenue, 2371 Nicosia, Cyprus. Tel.: +35722392852; E-mail: georges@cing.ac.cy

Abstract

During the course of a viral infection, virus–host protein–protein interactions (PPIs) play a critical role in allowing viruses to replicate and survive within the host. These interspecies molecular interactions can lead to viral-mediated perturbations of the human interactome causing the generation of various complex diseases. Evidences suggest that viral-mediated perturbations are a possible pathogenic etiology in several neurodegenerative diseases (NDs). These diseases are characterized by chronic progressive degeneration of neurons, and current therapeutic approaches provide only mild symptomatic relief; therefore, there is unmet need for the discovery of novel therapeutic interventions. In this paper, we initially review databases and tools that can be utilized to investigate viral-mediated perturbations in complex NDs using network-based analysis by examining the interaction between the ND-related PPI disease networks and the virus–host PPI network. Afterwards, we present our theoretical-driven integrative network-based bioinformatics approach that accounts for pathogen–genes–disease-related PPIs with the aim to identify viral-mediated pathogenic mechanisms focusing in multiple sclerosis (MS) disease. We identified seven high centrality nodes that can act as disease communicator nodes and exert systemic effects in the MS-enriched Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways network. In addition, we identified 12 KEGG pathways, 5 Reactome pathways and 52 Gene Ontology Immune System Processes by which 80 viral proteins from eight viral species might exert viral-mediated pathogenic mechanisms in MS. Finally, our analysis highlighted the Th17 differentiation pathway, a disease communicator node and part of the 12 underlined KEGG pathways, as a key viral-mediated pathogenic mechanism and a possible therapeutic target for MS disease.

Key words: virus–host–disease interactions; viral perturbations; neurodegenerative diseases; multiple sclerosis; protein–protein interactions

Introduction

Neurodegenerative diseases (NDs) are chronic degenerative diseases of the central nervous system (CNS), and currently, there are no effective pharmacotherapies for their treatment, thus highlighting the need for novel therapeutic interventions.

Several etiologies have been identified to contribute to their development and are considered to be multifactorial diseases involving the complex interaction of both genetic and environmental factors [1]. The 'multiple hit' hypothesis states that their development requires the combination of several factors;

Table 1. Viruses associated with NDs

Neurodegenerative disorder	Associated viruses	References
AD	Herpes simplex virus 1 (HSV-1), human immunodeficiency virus (HIV), herpes simplex virus 2 (HSV-2), human cytomegalovirus (HCMV/HHV-3), Epstein–Barr Virus (EBV/HHV-4), varicella-zoster virus (VZV/HHV-5), human herpesvirus 6 (HHV-6), human herpesvirus 7 (HHV-7), hepatitis C virus (HCV)	[5–10]
MS	EBV, HHV-6, HSV-1, measles *Morbillivirus* virus, rubella virus, VZV, human endogenous retrovirus W (HERV-W), human T-cell leukemia virus type 1 (HTLV-1), John Cunningham virus (JCV)	[11–20]
PD	West Nile virus, Japanese encephalitis B virus, Saint Louis encephalitis virus, HIV, HSV, EBV, HCMV, VZV, influenza virus type A, measles *Morbillivirus* virus, coxsackievirus, poliovirus	[5, 21–24]
ALS	HIV, HTLV-1, human endogenous retrovirus K (HERV-K), enteroviruses (poliovirus, coxsackievirus, echovirus, enterovirus-A71 and enterovirus-D68)	[5, 25–27]

however, how the interplay of these factors contributes to their development remains elusive [2]. A growing body of evidences has associated pathogenic organisms, such viral infections, as possible risk factors for the development of several NDs as described in Table 1 [2–4].

The viral infectious hypothesis of NDs is best supported in the case of MS as viral infections are the strongest environmental risk factor associated with its development, with Epstein–Barr virus (EBV) infection showing the strongest association [18, 28]. Several other viruses have been also implicated in the development and/or progression of MS, as described in Table 1. Neurotropic viruses that belong to the Herpesviridae family, such as herpes simplex virus (HSV) and EBV, have also been associated with the development of Parkinson's disease (PD) as they can cause acute and chronic parkinsonism [21]. In addition, infection with West Nile virus (WNV), a neurotropic virus of the Flaviviridae family that causes meningoencephalitis, can also result in the development of Parkinson-like symptoms that usually resolve over time [29, 30]; however, in some cases, the symptoms can persist, resulting in chronic PD [31]. Similar to WNV, parkinsonian-like symptoms can also be caused by other viruses of the Flaviviridae family [21, 32], including Japanese encephalitis virus [33] and Saint Louis encephalitis virus [34]. Amyloid lateral sclerosis (ALS) is a deadly ND involving the degeneration of motor neurons in the brain and spinal cord. Enteroviruses, including poliovirus and coxsackievirus, are suspected to be involved in the development of ALS as they can target motor neurons and can cause TAR DNA-binding protein-43 pathology *in vivo*, which is a major pathological hallmark in sporadic ALS [25, 35, 36]. Moreover, analysis of blood and brain samples of patients with ALS indicated increased expression of HERV-K, and experimental evidence revealed that infection of neurons with HERV-K can result in motor dysfunction in transgenic mice [37, 38]. Alzheimer's disease (AD) is a progressive ND and the most common form of dementia in the elderly, with its main pathophysiological hallmarks being the formation of amyloid plaques and neurofibrillary tangles in the brain. Individuals seropositive to HSV-1 show increased risk of developing AD, and increased presence of HSV-1 was found in AD post-mortem brains, including in Aβ plaques, especially in apolipoprotein E4 gene carriers, a major susceptibility gene in sporadic AD [7, 39, 40]. The presence of infectious agents in the brain of AD has led to the emergence of the 'antimicrobial protection hypothesis of AD', which suggests that increased microbial burden in the brain might lead to the disposition of β-amyloid due to the role of Aβ in

innate immune responses against pathogens, leading to sustain neuroinflammation that propagates neurodegeneration [41].

Viruses can induce neuronal degeneration via direct and/or indirect actions, for example they can cause neuronal dysfunction directly through their cytolytic effects and indirectly via various mechanisms such as by expressing viral genes that interfere with the host's immune system and cellular processes, via bystander inflammatory reactions, or by inducing apoptosis [42, 43]. Viruses are intracellular obligate parasites that lack their own replication machinery, thus to ensure their survival and reproduction they rely on the host, and via physical interactions they manipulate and exploit the host's cellular machinery [44, 45]. In addition, through the evolutionary process, viruses have developed an array of adaptive immune evasion strategies including interfering with antigen presentation and mimicking immune processes [43]. Several evidence indicates that neuroinflammation is involved in the progression of NDs, by directly or indirectly contributing to neuronal loss; however, whether the immune system participates in the initiation of these diseases remains undetermined [5, 46]. Immune system activation within the CNS can also be elicited by viral infections, either by neurotropic viruses that infect the CNS or by viruses that infect peripheral tissues and can cause a strong inflammatory response, resulting in the infiltration of peripheral leukocytes into the CNS and the subsequent activation of resident microglia cells [47]. Infection by certain neurotropic viruses can also lead to the formation of neuropathological and/or immunopathological lesions with the pattern of the resulting lesions having similar distribution as that of the common NDs [48, 49]. It is therefore possible that viruses may be directly or indirectly involved in the development of NDs, by either causing dysregulation of the immune system and/or by interfering with the host's cellular and immune system components.

However, despite the association of viral infections with NDs, it is highly unlikely that viruses are solely responsible for the development and/or progression of NDs, as the prevalence of viruses that are linked to NDs is very common in the general population, whereas the rate of individuals with NDs is significantly lower [42]. Therefore, it is unlikely that viral infections are the self-determining factor in the pathogenesis of NDs, and other factors particularly genetic susceptibility or other environmental factors could contribute in association with viral infections. Genetic susceptibility is not only associated with NDs, but is also an important factor that determines how a person's immune system responds to a pathogen and how efficiently
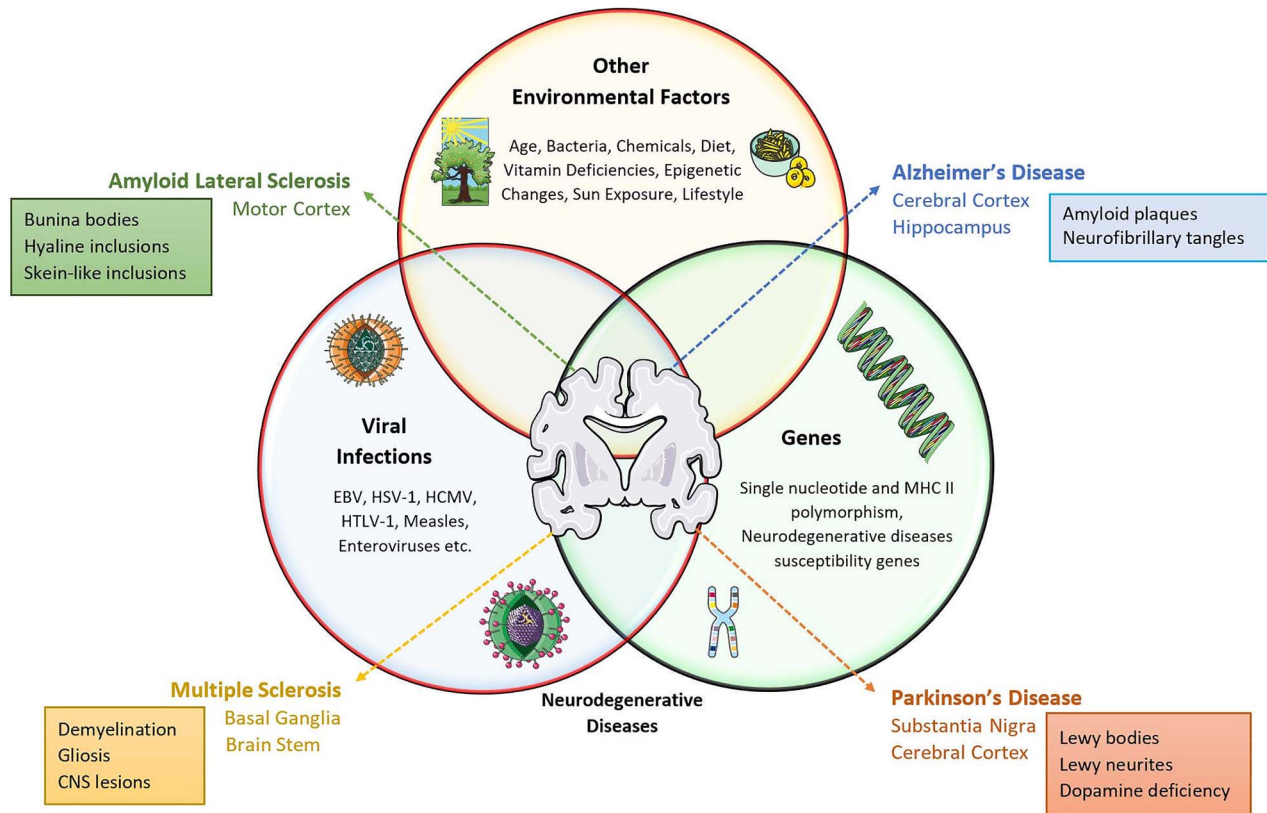
Figure 1. A schematic representation of the multifactorial origin of a group of pathologically distinct but related neurological diseases. NDs are complex disorders that are caused by the combination of environmental and genetic factors; however, how the interplay of these factors contributes to their development still remains unclear. One hypothesis is that NDs might be caused by the combinatorial effects of viral infections, other environmental factors and genetic susceptibility. The combination of these factors contributes to the clinical heterogenicity and histopathological diversity of the central nervous lesion that characterizes these diseases, as they differ in the subset of neurons and anatomical structures that are affected and have both common and distinct pathological abnormalities. Figure contains illustrations obtained from Servier medical art (https://smart.servier.com/), provided free and licensed under the Creative Commons Attribution 3.0 Unported License.

it can eliminate acute viral infections or control chronic viral infections [42, 46]. Therefore, the individual's genome determines the genetic sensitivity of the immune system towards different pathogens. However, genetic sensitivity to infections does not imply that there is a general weak immune surveillance towards pathogens [42], but rather that the individual's 'immunogenome' is more susceptible towards specific pathogens. The 'immunogenome' can change virus–host symbiosis and predispose an individual to certain viral-mediated immune-associated diseases by various mechanisms such as altering immune response to specific viruses, shifting viral cell tropism and changing neurovirulence [50–52]. In addition to genetic susceptibility, viruses can lead to the development of NDs in combination with other environmental factors [4]. For example, advanced age is considered as a major environmental risk factor in most NDs [53], and it can possibly act as a cofactor in the microbial-mediated pathogenesis of NDs. Aging causes alterations on the blood-brain barrier (BBB) and is also associated with increased oxidative stress in the brain, these age-mediated effects make the brain more permeable to viruses and neurons more vulnerable to viral insults [3]. Therefore, understanding the interaction between viruses, genes and environmental factors is not only essential for elucidating the role of viruses in the pathogenesis of NDs but can also aid to determine how the combinatorial effects of different risk factors can lead to the development of NDs (Figure 1).

Integration of virus–host PPIs and host PPIs in ND interactomes for the investigation of virus-mediated pathogenesis

During the course of a viral infection, virus–host PPIs are a key infection mechanism that allows viruses to evade host immune responses, replicate and hence survive within the host. These molecular interactions can lead to the dysregulation of normal biological processes within the host, causing numerous diseases from cancer to NDs [54]. These interspecies PPIs can be represented as a network, where the nodes represent the proteins and the edges their interactions [55]. This physical interaction network between viral proteins and host's cellular targets can be used to provide further insights in the disease etiology [56]. Understanding pathogen–host PPIs might enable the identification of mechanisms that directly or indirectly lead to the development of certain diseases.

Human PPI networks exhibit a scale-free behavior, where the majority of proteins in the network have few connections to other proteins whereas few ones, termed 'hubs', are connected with multiple proteins [57]. Viruses can be classified based on their genome. Comparative interactome analysis between DNA versus RNA viruses, revealed that the DNA viruses-human PPI network follows a scale free behavior, whereas the RNA viruses-human PPI network does not follow a scale free behavior [58]. In addition, unlike the cooperative nature of the evolutionary
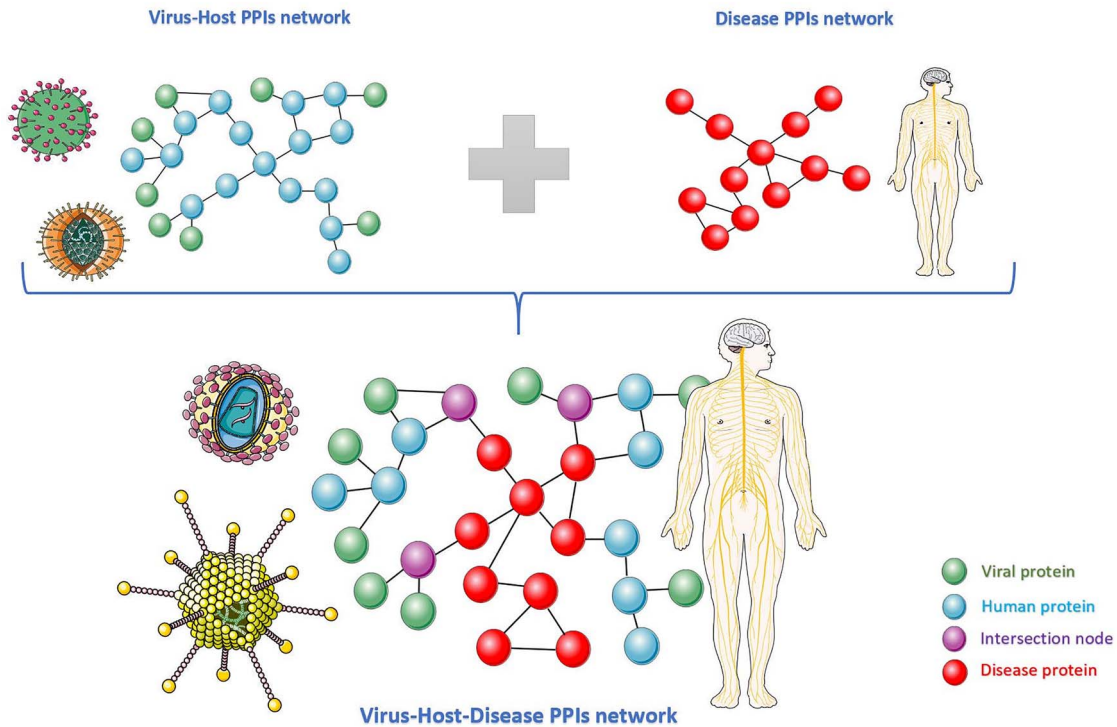
Figure 2. Schematic representation of the construction of the integrated virus–host–ND PPI network, where the virus–host PPI network is merged with a ND-related PPI network. Figure contains illustrations obtained from Servier medical art (https://smart.servier.com/), provided free and licensed under the Creative Commons Attribution 3.0 Unported License.

principles that govern PPIs of the host's biological processes, the evolutionary trajectories of PPIs in the virus–host interaction network are antagonistic, as there is a constant competition between virus and host interactions where a change in a protein of the virus might result in a reciprocal counterchange in a host's protein and vice versa [59, 60]. This constant arm race of evolutionary actions and counteractions between virus–host processes determines the evolutionary nature of the virus–host PPI network, which is not static and hence its topology changes over time.

Topological analysis studies of virus–human PPI networks revealed several important characteristics about the interaction of viral proteins with human targets. Viral proteins have a preference to target human proteins that are hubs, have high betweenness (bottlenecks), are articulation points and belong to rich clubs [61–64]. Hub proteins play an essential role in maintaining the connectivity of the network. The removal of hub nodes causes network failure, a property known as lethality–centrality, whereas the deletion of a random node with few connections does not influence the topology of the network [65, 66]. Betweenness centrality enables the identification of nodes that act as bridges (bottlenecks) between other nodes in the network and hence are important in influencing the communication between nodes within the network [67]. Viruses through the evolutionary processes have evolved strategies that allow them to efficiently adapt to the scale-free nature of the human interactome and via targeted attacks interact with essential proteins (hubs, bottlenecks) that are highly connected, allowing them to influence multiple functions and pathways simultaneously [68].

Viral-mediated perturbations of the host's interactome that lead to the generation of complex diseases involve the interaction of viral proteins with the host's PPIs network, therefore merging the virus–human PPI and the human disease PPI networks can enable to understand viral-induced pathogenesis [54, 56, 68, 69]. In the context of NDs, the virus–host PPI network can be merged with the ND-related PPI networks, creating an integrated virus–host–ND PPI network allowing to identify viral-mediated pathogenic mechanisms in NDs (see Figure 2 for a schematic representation).

Resources and computational analysis of the virus–host–disease PPI network

The analysis of virus–host–disease interactions can involve several approaches with the most commonly used bioinformatics pipeline approach involving the following steps: (i) reconstruction and visualization of the virus–host–disease PPI network, (ii) topological network analysis and (iii) Gene ontology (GO) [70] and pathway enrichment analysis. There are several reviews that have extensively described databases and tools that can be implemented to investigate these central steps [71, 72] . In this section, we summarize some of the recourses that can be utilized at each step for the investigation of virus–host–disease PPIs, which are also illustrated in Figure 3.

Reconstruction and visualization of the virus–host–disease PPI network

Data for the reconstruction of the virus–host PPI network can be collected from several databases such as VirHostNet 2.0 [73], Viruses.STRING [74], PHISTO [75], VirusMentha [76] and HPIDB 3.0 [77]. The main approaches utilized for the reconstruction of disease PPI networks are using disease genes obtained from gene to disease association databases, such as MalaCards [78]
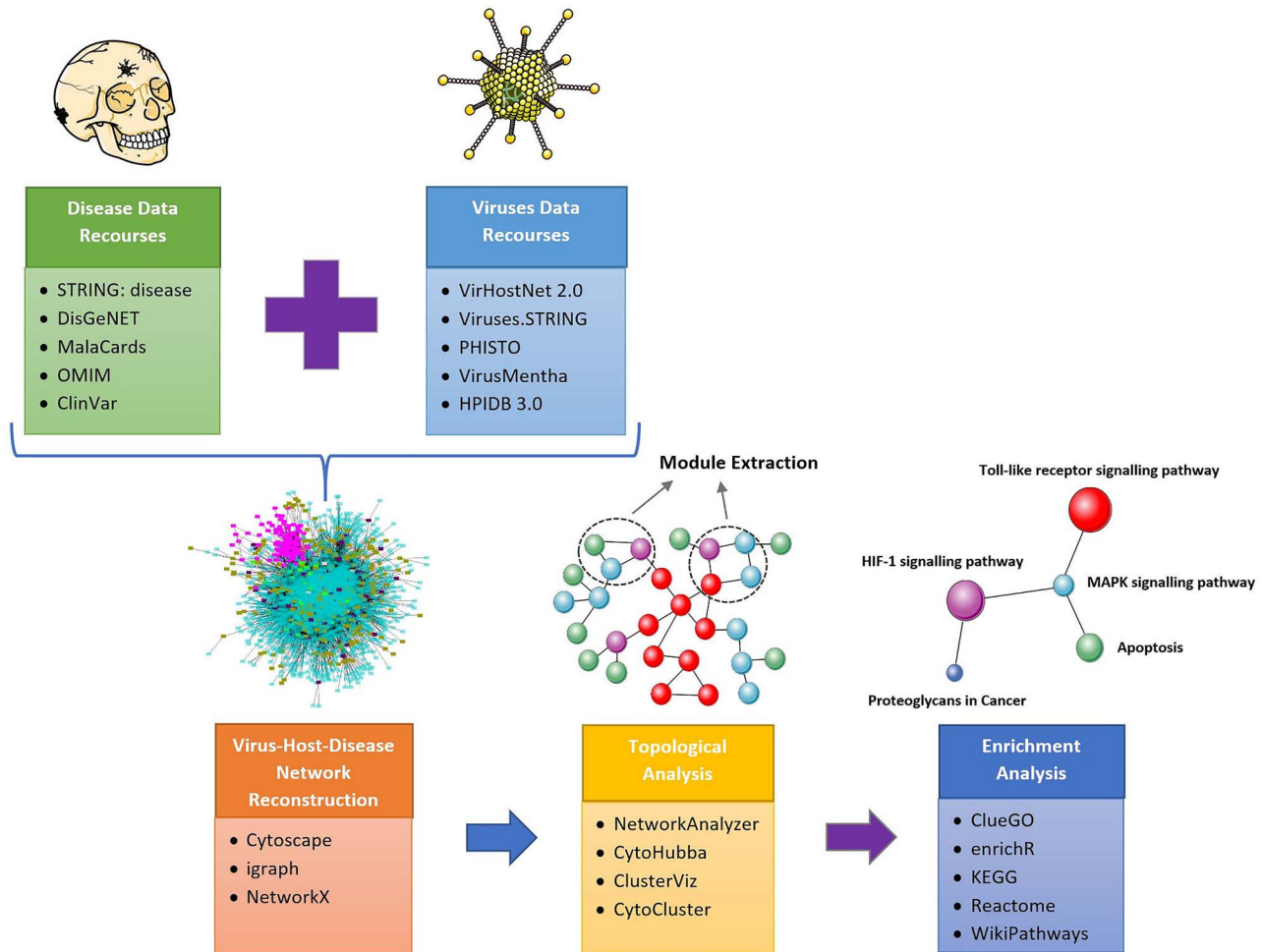
Figure 3. Illustration of the most commonly used approach for the investigation of virus–host–disease PPIs highlighting recourses and tools that can be utilized at each step. The first step involves the collection of data, and virus–host PPIs can be collected from several databases such as VirHostNet 2.0 [73], Viruses.STRING [74], PHISTO [75], VirusMentha [76] and HPIDB 3.0 [77]. Disease data can be collected from databases such as the STRING: disease query app in Cytoscape [81], MalaCards [78], DisGeNET [79], OMIM [237] and ClinVar [238]. The second step involves the visualization of the integrated network, which can be performed with Cytoscape [85], the igraph package in python and R [86] and the NetworkX package in python [87]. The third step involves network topological analysis that can be performed with several plugins offered by Cytoscape such as NetworkAnalyzer [88] and CytoHubba [89], as well as clustering analysis apps, such as CytoCluster [112] and ClusterViz [113]. The final step involves enrichment analysis that can be performed either on the whole network or on subnetworks using the ClueGO [96] plugin in Cytoscape [98] and enrichR [105], an R interface to the Enrichr database [106, 107]. Pathway enrichment analysis can be performed form several databases including the KEGG [99], Reactome [100] and WikiPathways [101]. Figure contains illustrations obtained from Servier medical art (https://smart.servier.com/), provided free and licensed under the Creative Commons Attribution 3.0 Unported License.

and DisGeNET [79], or using gene expression data from disease patients obtained from the NCBI GEO database [80], and then converting these genes into the corresponding protein products. Another approach is to use the *STRING: disease* query app in Cytoscape [81], which provides directly disease-associated proteins instead of genes, therefore overcoming the problem of having to convert genes into proteins. The *STRING: disease* data are obtained from the DISEASES database [82] that obtains gene to disease associations from automatic text mining, manually curated databases such as UniProt Knowledgebase (UniProtKB), genome-wide association studies and cancer mutation data. The different types of evidence are then unified by assigning them a confidence score, with 5 stars been the highest confidence and 1 star the lowest. Virus–host and human intra-species PPI data can also be obtained from the PISCQUIC Web Service Client app in Cytoscape, an integration tool that provides data from multiple databases [83]. Visualization and integration of the virus–host

and disease PPI networks can be performed with several open source software such as Gephi [84], Cytoscape [85], the igraph package in python and R [86] and the NetworkX package in python [87].

Topological and enrichment analysis of the virus–host–disease PPI network

Cystoscope offers numerous plugins, such as NetworkAnalyzer [88] and CytoHubba [89] for topological network analysis. Topological properties allow to identify the most important interactions within the network. The most common properties used for the investigation of virus-host interactions are: node degree, degree centrality, betweenness centrality, shortest path and modularity score [62, 63, 67, 90, 91].

GO and pathway enrichment analysis of pathogen–host PPIs have been used in several studies to identify common

and unique infection strategies, as well as pathogen-mediated pathogenic mechanisms that might lead to disease [58, 63, 91–93]. Therefore, enrichment analysis can be used to identify highly targeted biological pathways by viral proteins, providing insight into the mechanisms of pathogen obstruction of the host's immune system and cellular processes [94]. For example, GO enrichment analysis can be performed to identify functional terms that are significantly targeted by pathogens in the three ontologies terms: biological process, molecular functions and cellular components [95]. Functional analysis of the viral-targeted human proteins can be performed via the ClueGO [96] and BINGO [97] plugins in Cytoscape [98]. The ClueGO app also allows to perform pathway enrichment analysis from several databases including the Kyoto Encyclopedia of Genes and Genomes (KEGG) [99], Reactome [100] and WikiPathways [101]. Other available pathway enrichment analysis web tools are the KEGG Orthology-Based Annotation System [102], the Database for Annotation, Visualization and Integrated Discovery [103, 104] and enrichR [105], an R interface to the Enrichr database [106, 107].

In addition, the PathwayConnector web tool [108], which was created by our group, can be used to create a complementary network of the KEGG enrichment analysis results, allowing to identify missing nodes (pathways) that interact with the enriched pathways. GO and pathway enrichment analysis can be performed either on the whole network or in clusters of sub-networks. Clustering algorithms such as MCODE [109], the community clustering algorithm [110] and K-means clustering can be used for subnetwork identification, which can be performed with several clustering apps in Cytoscape such as ClusterMaker [111], CytoCluster [112] and ClusterViz [113].

Computational analysis of virus–host–ND PPIs using a knowledge-based approach

Although several studies have investigated the role of viral-mediated perturbations in the generation of human diseases based on PPIs networks [54, 56, 61, 90], there is lack of studies that focus specifically on viral-induced pathogenic mechanisms in the generation of NDs. A recent study by Costa Sa et.al. [114] has investigated the mechanisms of viral-induced pathogenesis in NDs, but at the transcriptomic level by comparing the gene expression profiles of AD or PD patients with three viral infection datasets and was able to identify shared pathways between these NDs and viruses. However, despite the lack of studies that explore virus–ND PPIs, there are several studies that have investigated NDs based on PPI networks. Although NDs encompass a large group of diseases that share as a common pathological component the degeneration of neurons, there is heterogenicity in the subset of neurons, anatomical structures and pathological abnormalities that are affected in each of these diseases [115, 116]. Nonetheless, these diseases share several common pathological mechanisms including protein misfolding and aggregation [117], apoptosis [118], impaired bioenergetics [119], neuroinflammation [120], oxidative stress and decreased antioxidant activity [121].

Therefore, NDs are a group of pathologically distinct but related diseases and two different directions have been used to investigate NDs using PPI networks by either focusing on a specific ND or analyzing a group of NDs, with the majority considering a specific disease [122–124]. The few studies that have investigated a group of NDs have primarily concentrated in identifying common molecular mechanisms by identifying direct protein/genes commonalities and pathways among NDs and indirect network relationships between NDs via topological association by identifying common modules in these diseases [125–128]. There is also lack of specific ND databases with NeuroDNet being one of the few available databases that contains information for genes and SNPs associated with 13 NDs and also allows the reconstruction and analysis of these NDs networks using PPI, regulatory and Boolean networks; however, it was last updated in 2016 [129]. In the next part of our paper, we aim to describe a simple methodology that integrates approaches for the investigation of NDs based on PPIs and elements from the investigation of virus–host PPI networks. Our approach aims to incorporate knowledge-based information about a group of NDs with the aim to focus in the molecular pathology of viral-mediated perturbations in MS disease.

The case of multiple sclerosis

MS is a complex chronic immune-mediated demyelinating disease of the CNS, characterized by the presence of both white and grey mater lesions in the CNS, accompanied by immune cell infiltration of autoreactive T and B cells [130, 131]. Although the etiology of MS pathogenesis still remains unknown, two major opposing hypotheses have been formulated on its origin. The 'outside-in' model, which suggests that an unknown trigger leads to the activation of peripheral lymphocyte cells that migrate into the CNS via the BBB though cell adhesion molecules [132, 133]. The peripheral immune cells then lead to the destruction of oligodendrocytes and the release of myelin antigens causing the activation of resident microglia cells and the subsequent formation of autoreactive T cells [132, 133]. The alternative hypothesis is the 'inside-out' model, where an initiating event within the CNS leads to the release of myelin antigens, causing the activation of resident microglia cells and the recruitment of immune cells [132–134]. The CNS-derived myelin antigens are then drain out of the CNS into the periphery inducing a secondary adaptive immune response in peripheral tissues [132–134]. Evidences for and against both models have been reviewed [130, 132–134]; however, it still remains undetermined if the primary initiator of MS disease is oligodendrocyte degeneration or peripheral immune system dysfunction, which are caused by an unknown trigger. As already mentioned above, viral infections can act as triggers under both the 'outside-in' and the 'inside-out' models of MS pathogenesis.

In the following section, we aim to identify possible mechanisms of viral-induced pathogenesis in MS disease, using a theoretical-driven integrative network-based bioinformatics pipeline approach, with the main steps illustrated in Figure 4. In this paper, we focus on the ND component of the disease by comparing MS with a group of NDs that are also associated with viral infections. In addition, we will briefly explore the autoimmune component of MS. Our pipeline includes the following approaches: (i) reconstruction and visualization of the integrated virus–host–MS PPI network, (ii) topological and knowledge-based subnetwork identification, (iii) pathway and GO enrichment analysis, (iv) a filtering process of the enriched results, (v) comparison and identification of direct common pathological pathways in NDs and autoimmune diseases, (vi) construction of the ND-based MS-enriched KEGG pathway-to-pathway network, (vii) topological analysis, (viii) similarity calculations and clustering analysis and (ix) tissue-specificity enrichment analysis.
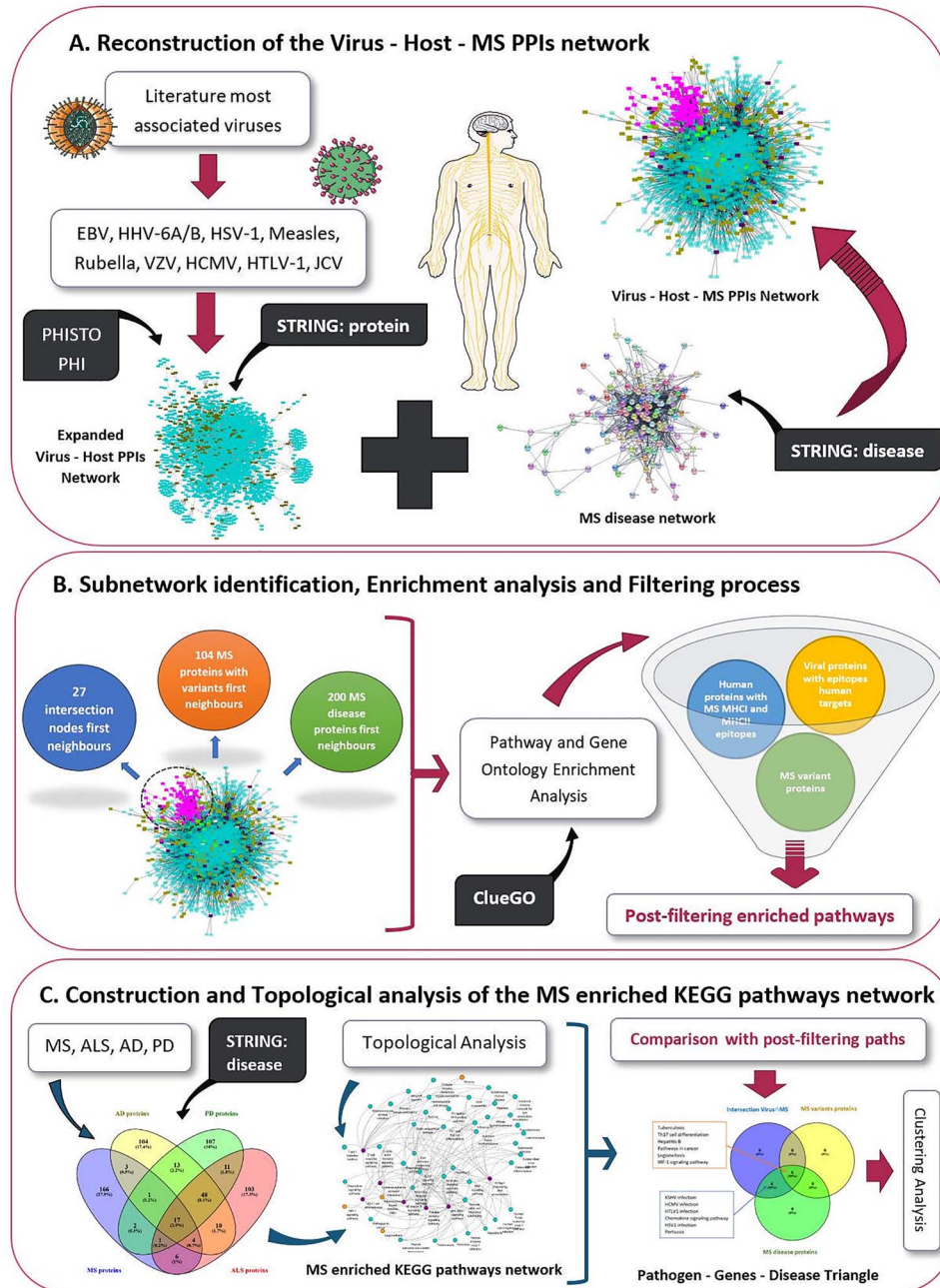
Figure 4. Schematic representation of the methodology applied in this paper to investigate the interaction between virus-host-MS PPIs using a network-based approach with the aim to identify viral-mediated pathogenic mechanisms that might be involved in the development of MS. (A) Reconstruction of the virus-host-MS PPIs network. (B) Subnetwork identification,enrichment analysis and filtering process of the enriched results. (C) Construction and topological analysis of the ND-based MS enriched KEGG pathway to pathway network and comparison with the post filtering enriched pathways obtained from each subnetwork. Figure contains illustrations obtained from Servier medical art (https://smart.servier.com/), provided free and licensed under the Creative Commons Attribution 3.0 Unported License.

Methods

Reconstruction and visualization of the integrated virus–host–MS PPI network

A thorough literature review analysis has been performed to identify viruses highly associated with MS disease (Table 1). Eleven viruses have been identified. From them, HERV-W was excluded due to lack of PPI data. Virus–host PPI data for the 10 selected viruses were collected from PHISTO database, which is one of the most comprehensive databases for pathogen–human host PPI data and contains only experimentally detected interactions that are imported from several PPI databases [75, 135]. For the selected viruses, a unified virus–host PPI network was constructed and visualized in Cytoscape [85]. The characteristics of the 10 viruses and the number of virus–host PPIs for each of the viruses included are indicated in Table 2. The StringAPP [81] was then used to expand the interactome of the viral-targeted human proteins and create edges between the human proteins, in order to create an expanded interconnected virus–host PPI network. The unified virus–host PPIs network contained 2101

Table 2. Viruses–human host PPI network data

Virus species	Family	Virus genome	No. of strains	No. of PHIs	No. of viral proteins	No. of human proteins
HSV-1	Herpesviridae	dsDNA	6	777	66	604
Varicella-zoster virus (HHV-3)	Herpesviridae	dsDNA	2	6	5	4
EBV	Herpesviridae	dsDNA	3	4644	148	1198
HCMW	Herpesviridae	dsDNA	4	110	32	90
HHV-6A	Herpesviridae	dsDNA	2	3	3	3
HHV-6B	Herpesviridae	dsDNA	1	4	3	4
JCV	Polyomaviridae	dsDNA	1	51	5	49
Rubella virus	Togaviridae	+ssRNA	3	60	4	59
Human T-cell leukemia virus type 1 (HTLV-1)	Retroviridae	+ssRNA-RT	5	195	15	178
Measles *Morbillivirus* virus	Paramyxoviridae	-ssRNA	6	529	14	487
		Total	33	6379	295	2676

different viral-targeted human proteins, which were expanded by maximum of 10 additional interactors using the *STRING: protein* query, and the confidence cut-off for the interactions between the human proteins was set at 0.8. The confidence cut-off score determines the nature and quality of the supporting evidence of the interaction between proteins and ranges from 0 (low) to 1.0 (high), with interactions with high confidence score being more likely to be true positives. Expansion of the 2101 viral-targeted human resulted in the import of an additional 238 human proteins. Then, the expanded interconnected virus–host PPIs network, containing 2339 human proteins and 295 viral proteins, was integrated with the MS disease PPI network that contained the 200 disease-associated proteins with the highest disease score obtained from the *STRING: disease* query of the StringApp. Since we are interested in constructing an integrated virus–host–MS disease PPI network, we choose *STRING: disease* to obtain the disease-associated proteins for MS because it provides directly proteins instead of genes and it has a unified disease scoring scheme for the different types of evidences. In addition, because the virus–host PPI network was constructed using *STRING: protein*, the protein identifiers are synchronized, allowing the virus–host and MS disease PPI networks to be merged. Merging of the two networks resulted in the construction of an integrated virus–host–MS PPI network of 2807 nodes and 33 741 edges.

Topological and knowledge-based subnetwork identification in the virus–host–MS PPI network

Due to the size and the complexity of the integrated virus–host–MS PPI network, it is important to identify subnetworks that reveal important virus–host interactions that could directly or indirectly affect disease proteins. Three subnetworks were isolated that account for pathogen–genes–MS disease interactions:

(i) The first neighbors of the intersection nodes between the expanded virus–host PPI network and the MS disease PPI network: merging of the expanded virus–host PPI network with the MS disease network indicated 27 intersection nodes, which are MS disease-related proteins, and 21 nodes are also human viral targets. This subnetwork includes the 27 intersection nodes and their first neighbors.

(ii) The first neighbors of MS-related proteins: the second subnetwork includes the 200 MS disease-associated proteins and their first neighbors.

(iii) The first neighbors of human proteins with MS-related variants: the MS-related variants subnetwork was isolated by mapping on the integrated virus–host–MS PPI network, MS disease-associated variant proteins, which were collected from the DisGeNET database [79]. This mapping revealed 104 MS variant proteins on the network, which were then selected with their first neighbors to highlight the third subnetwork. Based on the local impact hypothesis, viral host targets of the associated virus are located in proximity of disease susceptibility genes within the network, and the expression pattern of such genes changes significantly [56, 65, 136]. Therefore, viral proteins that directly interact with MS disease-associated variant proteins or indirectly by affecting another protein within the same pathway of the disease variant can lead to a shift from virus–host equilibrium state to disease disequilibrium state, causing viral perturbations.

The number of viral and MS disease proteins included in each subnetwork can be found in Table 3.

Enrichment analysis of the three subnetworks

GO and pathway enrichment analysis was performed for the human proteins contained in each of the three subnetworks (Table 3) using the ClueGO plugin in Cytoscape [96]. The enrichment analysis for each subnetwork included the following three databases: (i) KEGG database pathways [99], (ii) Reactome database pathways [100] and (iii) Gene Ontology Immune System Processes (GO ISP), keeping only the terms with significant *P*-value ≤ 0.05 (corrected with Bonferroni step-down).

Similarity and clustering analysis of viral proteins–KEGG pathways interactions

To identify clusters of viral proteins that interact with similar KEGG pathways, Reactome pathways and GO ISP, and vice versa, we used the vegan package in R [137] to measure the Jaccard similarity index. Then, we performed agglomerative hierarchical clustering using the factoextra package in R [138].

Reconstruction and visualization of the MS-enriched KEGG pathway-to-pathway network

NDs share several common pathological mechanisms; therefore, common disease proteins shared between all NDs could reflect

Table 3. Contents of the topological and knowledge-based subnetworks

Subnetworks inclusion criteria	Total proteins in subnetwork	No. of viral proteins	No. of MS disease proteins	No. of human proteins for enrichment analysis
Intersection nodes first neighbors	615	43	89	572
MS disease proteins first neighbors	699	43	200	656
MS variants proteins first neighbors	681	69	131	612
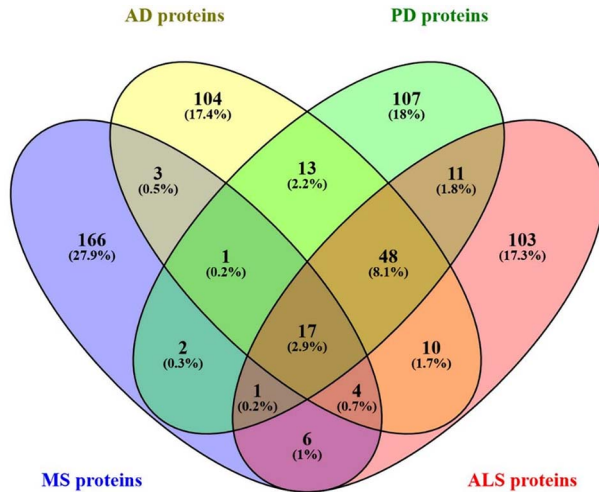
Figure 5. Comparison results of the top 200 highest disease-associated proteins between the four NDs: ALS, MS, PD and AD.

direct commonalities of pathological mechanisms and pathways affected, whereas unique disease-associated proteins might possibly reflect more specific disease phenotype. We selected AD (DOID: 10652), ALS (DOID:332) and PD (DOID:14330) to be compared to MS (DOID:2377) for the following reasons: (i) like MS, their development has also been associated with various viral infections, as indicated in Table 1; (ii) similar to MS, they affect the CNS unlike other NDs that affect the peripheral nervous system; (iii) they are also predominantly sporadic but also have familial forms unlike other NDs like Huntington's disease and spinocerebellar ataxia, which are hereditary and (iv) most importantly these four NDs share several common pathological mechanisms. The *STRING: disease* query of the StringApp in Cytoscape was used to collect disease-associated proteins for the four NDs. The top 200 most disease-associated proteins with the highest disease score for each disorder were collected, with confidence cut-off set at 0.8. We used Venny 2.1 [139], to compare the associated proteins among the four NDs to identify the common disease-associated proteins between all four NDs and those that are unique for MS compared to this specific group of NDs (Figure 5).

We identified 17 common disease proteins between all four NDs, 166 unique to MS and 17 that are shared between MS and some of the other NDs. Enrichment analysis was performed for each group of proteins with the ClueGO app [96] in Cytoscape using the following three databases: (i) KEGG database [99], (ii) Reactome database [100] and (iii) GO ISP, keeping only terms with significant P-value ≤0.05 (corrected with Bonferroni step-down). The enriched KEGG pathway results obtained from the 17 common ND proteins were then entered into the PathwayConnector web tool [108] to discover a complementary network of pathways that interact with the common ND pathways (Figure 6).

Moreover, PathwayConnector was used to reconstruct and visualize the MS-enriched KEGG pathways network. Topological analysis was performed using the CytoHubba app of Cytoscape [89] with the aim to identify hubs and bottlenecks nodes in the MS-enriched KEGG pathways network. The ClusterMaker app [111] was then used to perform community clustering [110] on the MS-enriched KEGG pathways network, which detects clusters of nodes based on their connectivity, allowing to identify community clusters of pathways that interact with the hub–bottleneck nodes.

The autoimmune component of MS disease

MS is a complex disease having components from both ND and autoimmune diseases. Similar to NDs, the pathogenesis of several autoimmune diseases has also been associated with several infectious agents, including viruses [140, 141]. Interestingly, EBV, which is the strongest environmental factor in the development of MS, has also been associated with other NDs (Table 1), but also with several autoimmune diseases [142–145]. So how is one virus able to cause different disease phenotypes? Therefore, to identify commonalities between MS and other autoimmune diseases, we used a similar approach as in the comparison of MS with a group of NDs that were also associated with viral infections. We selected systemic lupus erythematosus (SLE) (DOID: 9074), rheumatoid arthritis (RA) (DOID: 7148) and type 1 diabetes (DOID: 9744) to be compared to MS (DOID:2377) because (i) they are non-nervous system autoimmune diseases and (ii) their development has also being associated with several viral infections, as described in Table 4. Enrichment analysis was then performed for each group of proteins using the KEGG database. We then compared the obtained statistically significant KEGG-enriched results with those identified in the comparison of MS with the three NDs, with the aim to identify common pathways between the two major components that characterize MS pathophysiology.

Tissue-specificity enrichment analysis of virus–disease interactions

The final 12 KEGG pathways, 52 GO ISP and 5 Reactome pathways are targeted by 80 viral proteins from eight viral species. We used TissueEnrich web application (http://tissueenrich.gdcb.iastate.edu/) [156] to perform tissue-specific gene enrichment analysis to identify possible tissues where the virus–disease interaction could manifest leading to the dysregulation of the identified pathways. To perform the analysis, we used the human RNA-seq data from the GTEx database [157], which is the largest dataset of normal tissue expression, containing 56 different tissues. However, in the TissueEnrich application samples from sub-tissues are grouped together, i.e. different brain areas are grouped under brain, resulting into 29 human tissues. We considered genes as tissue specific, using the 'tissue enriched' parameter, which defines genes as tissue specific if they have at least five times

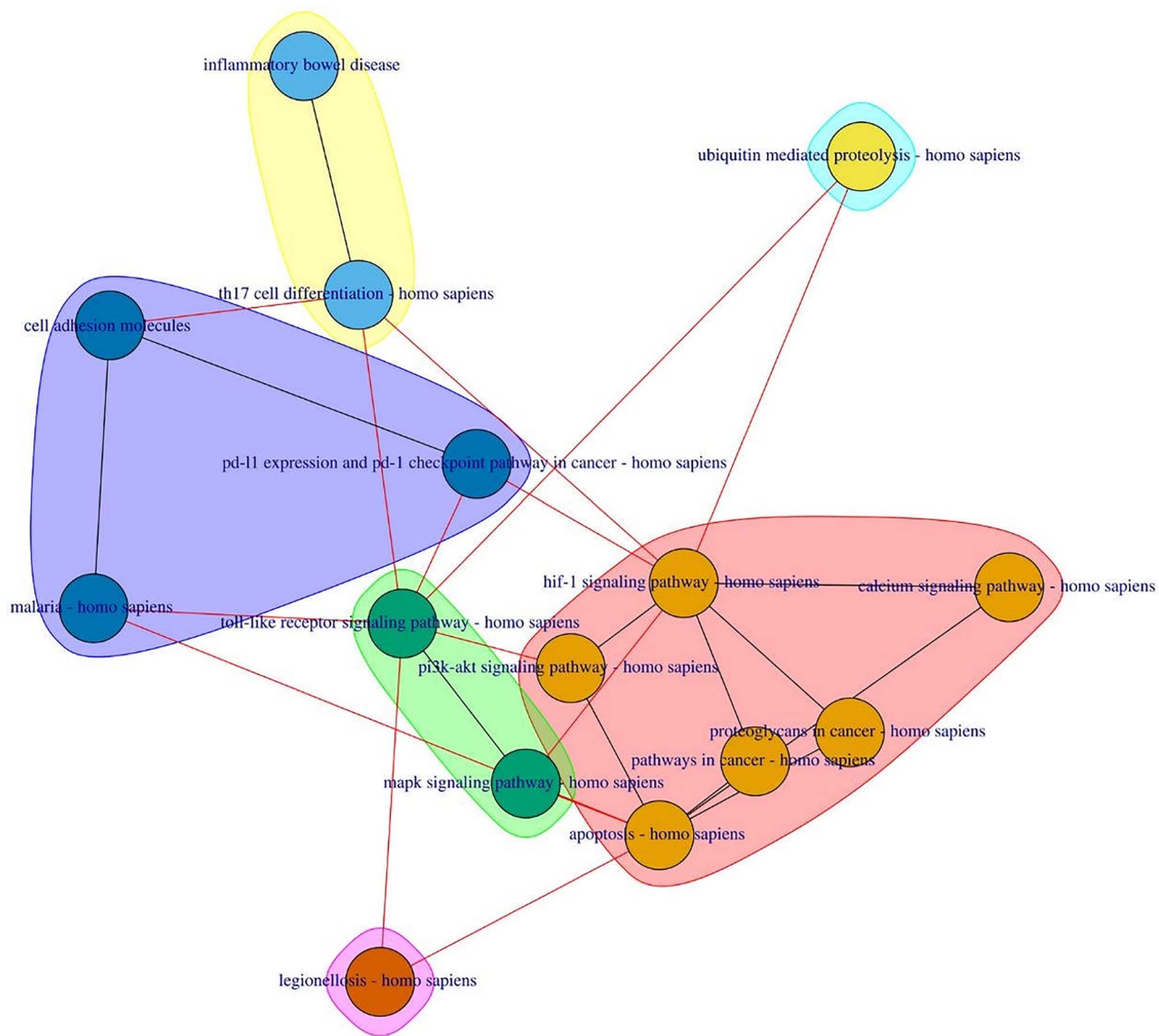## clustering using: edge betweenness algorithm (clusters:6)

Figure 6. Complementary network of the four common NDs KEGG pathways (IBD, HIF-1 signaling pathway, malaria and legionellosis) and their interactions with the 11 complementary nodes/pathways, created using PathwayConnector.

Table 4. Viruses associated with autoimmune diseases

| Autoimmune disease	Viruses	References
SLE	EBV, HCMV, HTLV-1, HIV-1, human endogenous retroviruses (HERVs), parvovirus B19	[142, 143, 145–150]
RA	EBV, parvovirus B19, HIV, hepatitis B virus, hepatitis C virus, alphaviruses, HTLV-1, rubella virus, HCMV	[143–145, 150, 151]
Type 1 diabetes	Enteroviruses (coxsackie B virus), EBV, HERVs, mumps virus, rubella virus, hepatitis C virus, HCMV, rotavirus	[150, 152–155]

% terms per group
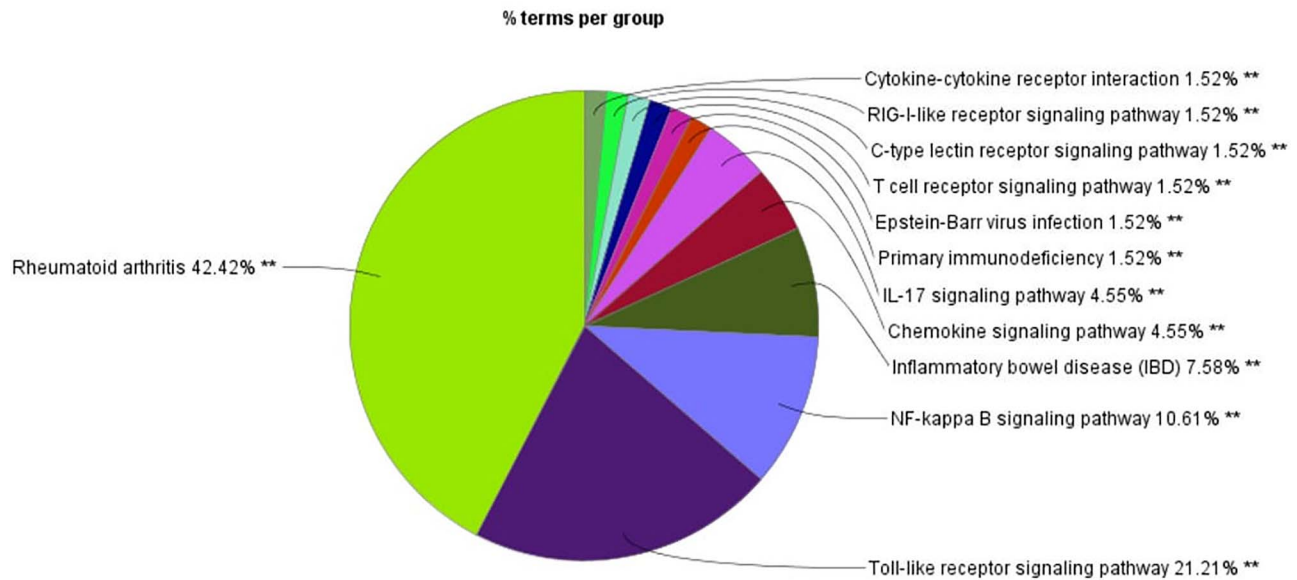
Figure 7. Enriched KEGG pathway analysis results of the 166 unique MS disease proteins, obtained using the ClueGO app in Cystoscope, with the pathways classified into groups and the percentage indicating the number of terms in each group.

higher expression levels in a particular tissue compared to all other tissues. We also measured the statistical significance of the tissue-specific genes using the fold-change test.

Results

Enrichment analysis results of the MS unique, MS shared and ND common disease proteins

The enrichment analysis of the 17 common ND-associated proteins indicated that only microglia cell activation is a statistically significant common GO ISP, no Reactome pathways were found and four KEGG pathways were identified: inflammatory bowel disease (IBD), malaria, HIF-1 signaling pathway and legionellosis pathways. Complementary network reconstruction of the four KEGG pathways indicated that they interact with 11 complementary pathways (Figure 6). Enrichment analysis of the 166 MS unique disease-associated proteins revealed 49 KEGG pathways (Figure 7), 38 Reactome pathways and 188 GO ISP. The 49 KEGG pathways are either infectious disease or immune-related pathways (see Supplementary File 1 available online at https://academic.oup.com/bib), thus confirming the association between pathogens and their interaction with the immune system in the development of MS. Interestingly, 42.4% of the KEGG pathways belong to the Rheumatoid arthritis group. The GO ISP enrichment analysis results indicated that 92% of the significant terms belong to the group of positive regulation of lymphocyte proliferation, which is important for the development of adaptive immune responses against pathogens in the CNS [158]. The Reactome results indicated that the majority of the terms belong to interleukin (IL) signaling groups, to the chemokine receptors bind chemokines group and the diseases of the immune system group. The enriched group results of the Reactome and GO IPS can be found in Supplementary File 1 available online at https://academic.oup.com/bib.

In addition, the enrichment analysis results of the 17 disease-associated proteins that are shared between MS and some of the other NDs resulted in 2 Reactome pathways, the IL-10 signaling and IL-4 and IL-13 signaling pathways, 1 GO ISP, the cytokine secretion involved in immune responses and 7 KEGG pathways.

Comparison between MS ND-derived pathways versus MS autoimmune-derived pathways

We identified 45 common disease-associated proteins between MS and the three autoimmune diseases SLE, RA and type I diabetes, 102 unique to MS and 53 disease-associated proteins that are shared between MS and some of the other autoimmune diseases. The enrichment analysis of the 45 common disease-associated proteins resulted in the identification of 47 statistically significant KEGG pathways, which belong to nine groups, as indicated in Figure 8. The KEGG pathway groups are either infectious diseases, autoimmune diseases or immune system components, which highlights the association between infectious diseases and the immune system in the development of autoimmune diseases [159]. In addition, 23.75 and 15.0% of the pathways belong to the malaria and IBD groups, respectively. These pathways have also been identified in our comparison of MS with the three NDs (AD, ALS and PD), as common ND pathways. In addition, Legionellosis was also identified in both the common MS ND-derived pathways and the common MS autoimmune-derived pathways, whereas the HIF-1 signaling pathway was only identified in the MS common ND-derived pathways. Therefore, malaria, IBD and legionellosis could represent possible crossroad points between the MS ND component and the MS autoimmune disease component.

The enrichment results of the 53 disease-associated proteins that are shared between MS and some of the other autoimmune diseases resulted in 31 statistically significant KEGG pathways, with 48.78% of the pathways belonging to the EBV infection group, as indicated in Figure 9. Similar to MS, EBV infection has been associated as an environmental factor in the development of the other three autoimmune diseases SLE, RA and type 1 diabetes [142–145, 150]. Therefore, the results confirm that EBV infection is a shared pathogenic mechanism in these diseases,
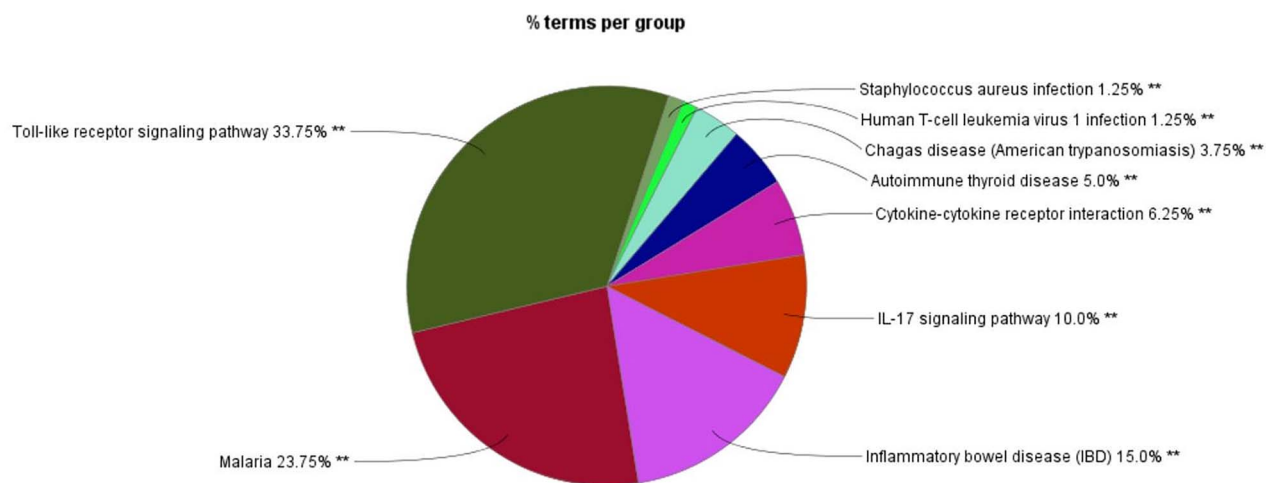
% terms per group

Figure 8. KEGG enrichment analysis results of the 45 common disease proteins between MS, SLE, RA and type 1 diabetes, with the pathways classified into groups and the percentage indicating the number of terms in each group.
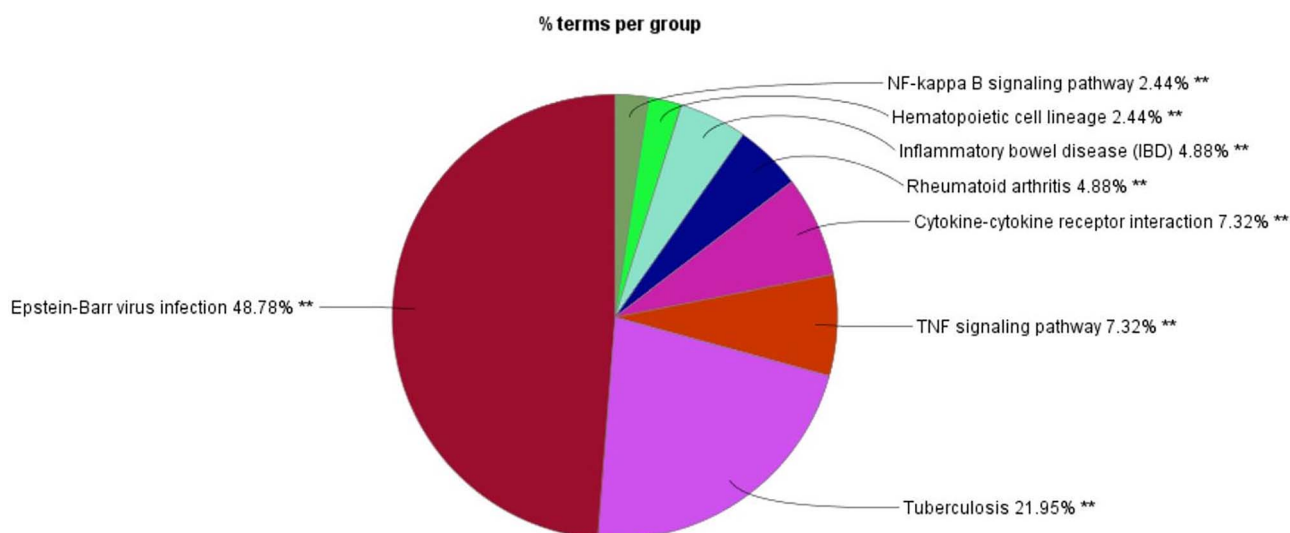
% terms per group

Figure 9. KEGG enrichment analysis results of the 53 disease proteins that are shared between MS and some of the other autoimmune diseases (SLE, RA, type 1 diabetes), with the pathways classified into groups and the percentage indicating the number of terms in each group.

but the results also indicate that the pathogenic effects of EBV in these diseases are possibly mediated through different biological mechanisms.

Topological and clustering analysis of the MS-enriched KEGG pathways network

Reconstruction and topological analysis of the MS-enriched KEGG pathways network allowed to identify the most important bottleneck and hub pathways within the network. Topological analysis revealed seven hub–bottleneck nodes/pathways that can act as a bridge of communication between the rest of the pathways in the network. These high centrality nodes were identified by using the average score of the top 10 bottleneck, degree and closeness centrality pathway results obtained via the CytoHubba app [89]. We also examined the topological characteristics of the four common ND pathways in the MS-enriched KEGG pathways network, which indicated that malaria, HIF-1 signaling pathway and legionellosis pathways are non-hub–non-bottleneck nodes, whereas the IBD pathway acts as a hub–non-bottleneck node.

Clustering analysis of the MS-enriched KEGG pathways network using the community clustering algorithm [110] allowed to identify four clusters of community nodes that interact with the seven hub–bottleneck nodes. The clustering results indicate that the majority of viral infectious disease pathways are clustered together with the hub–bottleneck nodes Toll-like receptor signaling pathway and JAK-STAT signaling pathway (Figure 10A), whereas few viral infectious disease pathways are located in cluster c, with the hub–bottleneck nodes Th17 cell differentiation and NF-kappa B signaling pathways (Figure 10C).

Filtering process of the topological and knowledge-based subnetwork enrichment results

The enrichment analysis of the human proteins contained in each subnetwork (Table 3) revealed multiple significant GO ISP and pathways (KEGG and Reactome). A filtering process (Figure 11) was performed on the enriched terms for each of the three subnetworks in an effort to isolate the most relevant GO ISP and pathways. The selection criteria are listed below:
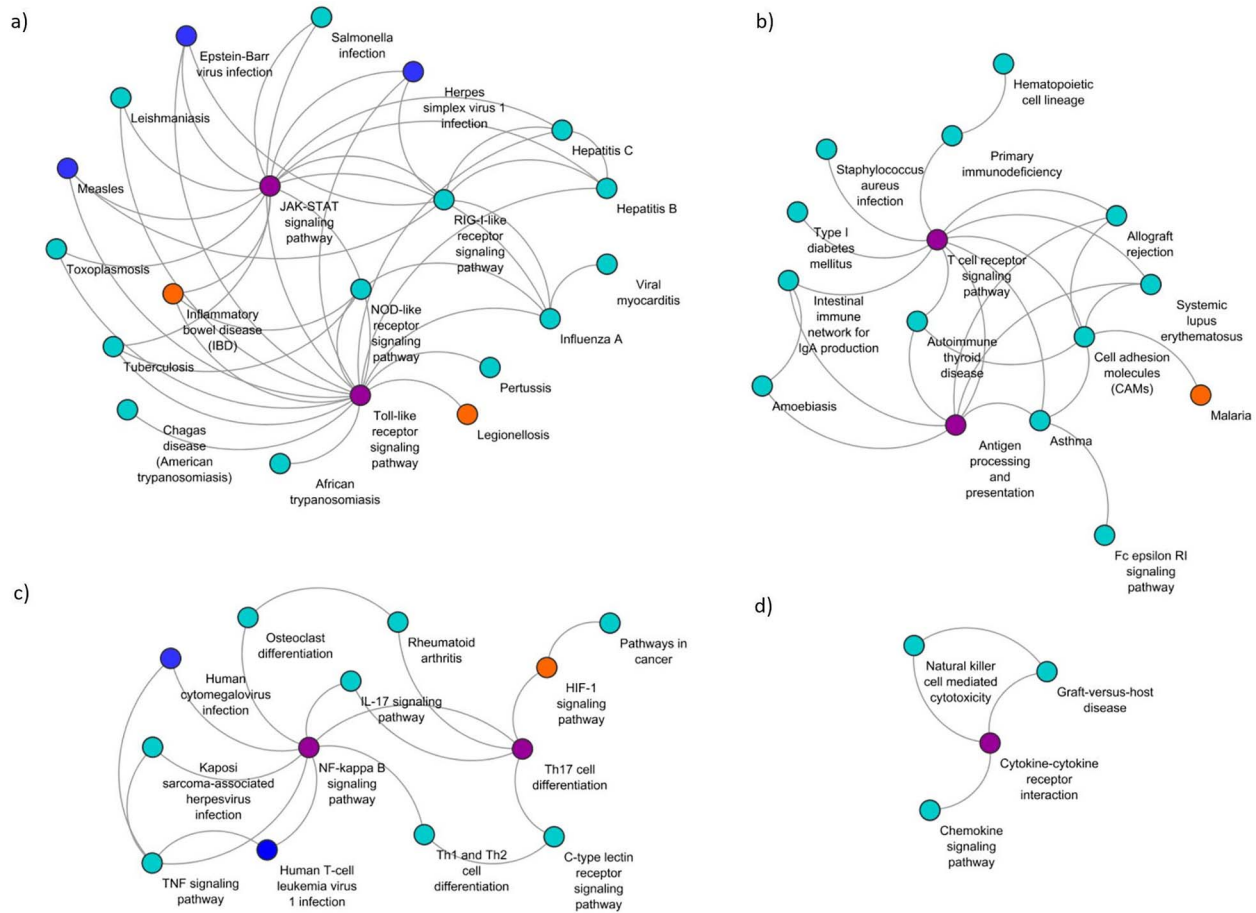
Figure 10. Community clustering of the MS-enriched KEGG pathways network resulted in the formation of four community clusters (a-d). Orange color nodes represent the four common ND pathways between ALS, MS, PD and AD. Hub–bottleneck nodes are represented in color purple, and blue color nodes represent infectious disease pathways of some of the viruses we included for the reconstruction of our virus–host–MS PPIs network.

(i) Criterion 1: Only pathways and GO ISP containing human proteins that are targets of immunogenic viral proteins were selected. To identify the immunogenic viral proteins, immune epitope data for the selected viruses that are found in humans during viral infection were obtained from the Immune Epitope Database (IEDB) [160].

(ii) Criterion 2: The results obtained from the first filtering process were then filtered by using immune epitope data found in humans with MS disease. Immune epitope data found in MS disease individuals, which have MHCI- or MHCII-restricted antigen recognition, meaning that T cells will only respond to the antigens from this protein only when they are bound to the relevant MHC molecules, were obtained from the IEDB [160]. The epitope data contained few epitopes against viral proteins where their human targets were selected as in criterion 1, but the majority were self-epitopes against human proteins which were also selected.

(iii) Criterion 3: Finally, the filtering results obtained were filtered to select pathways containing MS variants proteins collected previously from the *DisGeNET database* [79].

This three-criteria filtering process was applied on the intersection subnetwork and the MS disease protein subnetwork, whereas for the MS variant protein subnetwork only the first two criteria were applied. The filtering process aims to account for the immunogenicity, autoimmunity and genetic components of MS disease as the final pathways obtained contain immunogenic viral proteins and their human targets, host proteins with self-epitopes and MS disease variants proteins.

Identification of post-filtering enrichment results that are also MS-enriched terms

The filtering process significantly reduced the number of enriched pathways (KEGG and Reactome) and GO ISP for the three subnetworks, allowing possibly to isolate the most relevant pathways. More specifically, the enriched KEGG pathways before the filtering process for the intersection subnetwork that contains nodes that are both human viral targets and MS-related proteins were 94, for the MS disease proteins and MS variants protein subnetworks 91 and 94, whereas after the filtering process 36, 34, 28 KEGG pathways remained, respectively. The number of enriched Reactome pathways and GO ISP for each subnetwork before and after the filtering process can be found in Supplementary File 2 available online at https://academic.oup.com/bib. Comparison of the post-filtering enrichment results for each of the three subnetworks with the MS unique, MS shared and common ND pathways that were previously found, allowed to identify and select enriched terms from the three subnetworks that are also MS disease-related terms.
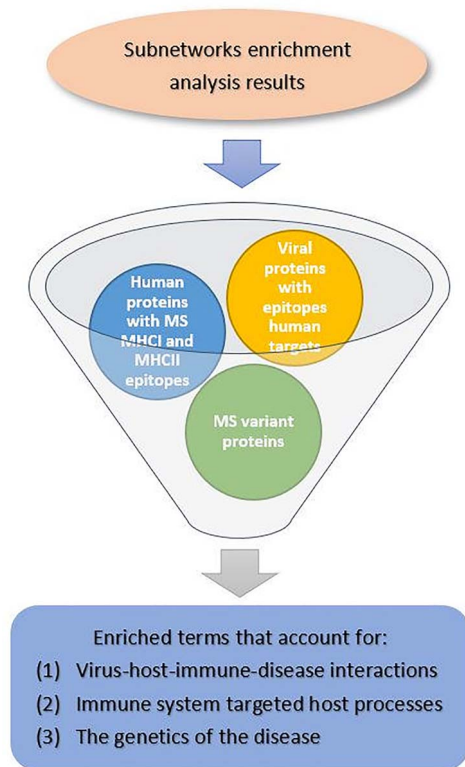
Figure 11. Filtering process applied on the enriched analysis results of the three subnetworks.

Comparison of the enriched pathways between the three subnetworks

The three subnetworks represent the triangle to pathogenesis by accounting the importance of three elements, pathogen–genes–disease interactions, and their role in the pathogenesis of MS disease (Figure 12A). To identify the KEGG pathways, Reactome pathways and GO ISP that fall within the pathogen–genes–disease triangle, we compared the post-filtering enrichment results that are also MS disease-related terms between the three subnetworks, which resulted in the identification of 12 KEGG pathways shown in Figure 12B, 5 Reactome pathways and 52 GO ISP, which can be found in Supplementary File 2 available online at https://academic.oup.com/bib.

Analysis of the complementary network of the final 12 KEGG pathways

Pathway-to-pathway network reconstruction of the 12 KEGG pathways using the PathwayConnector tool indicated that only Th17 cell differentiation pathway, HIF-1 signaling pathway and pathways in cancer functionally interact, whereas the rest 9 KEGG pathways are disconnected with each other. A complementary network was therefore created by using the missing pathway approach to identify proximal pathways that interact with the 12 KEGG pathways, which led to the identification of 7 complementary pathways, shown in Table 5. Comparison of the complementary pathways with the MS disease-related pathways showed that four out of the seven pathways are also MS unique pathways, whereas three of the complementary pathways, namely TGF-beta signaling, Calcium signaling and MAPK signaling pathway, do not belong to the MS disease pathway terms (Table 5).

Topological analysis of the MS-enriched KEGG pathways network indicated seven hub–bottleneck nodes that act as a bridge of communication within the network, five of these hub–bottleneck nodes are found in the complementary network, of which four are complementary nodes and one is part of the 12 KEGG pathways. A schematic visualization (Figure 13) illustrates the complementary network of the 12 KEGG pathways.

Similarity and clustering results of viral proteins–KEGG pathways interactions

The final 12 KEGG pathways are targeted by 67 viral proteins from eight viral species, more specifically EBV, HCMV, HHV-6A, HHV-6B, HSV-1, HTLV-1, Measles and Rubella, with the majority being EBV strains. The final step of our analysis involved identifying clusters within the 67 viral proteins that target similar KEGG pathways and clusters within the 12 KEGG pathways that are targeted by similar viral proteins. Similarity analysis was performed using the Jaccard similarity index and hierarchical clustering analysis for both dendrograms was performed using the average distance (also called mean) method as it had the highest correlation coefficient value in both cases. The clustering dendrogram in Figure 14 indicates the presence of 16 clusters of viral proteins based on pathway target similarity with the 12 KEGG pathways. The clustering dendrogram of the 12 KEGG pathways and a heatmap plot indicating the interactions between the 67 viral proteins with the 12 KEGG pathways can be found in Supplementary File 2 available online at https://academic.oup.com/bib. The clustering dendrograms and the heatmap plots of the 5 Reactome pathways and the 52 GO ISP indicating the interactions between their viral targets can also be found in Supplementary File 2 available online at https://academic.oup.com/bib.

Tissue-specificity analysis results of virus–disease interactions

We performed tissue-specific gene enrichment analysis on two groups: (i) the 200 MS disease-associated proteins and (ii) on the 1771 human protein targets of the 80 viral proteins that interact with our identified pathways (12 KEGG pathways, 52 GO ISP and 5 Reactome pathways). The aim was to identify common tissues between the two groups that might represent the virus–disease interaction areas where the identified eight viral species might exert their pathogenic effects in MS. The 200 MS disease-associated proteins showed statistically significant tissue-specific enrichment in 13 tissues, indicated in Figure 15, using the GTEx dataset, based on the fold-change test. Figure 15 contains the spleen, which is part of the lymphatic system, the nerve and brain tissues, which represent the two major pathophysiological characteristics of MS, that of neuron degeneration in the brain and immune system activation. We also observe multiple tissues related to the lower gastrointestinal (GI) system (colon, liver, small intestine, pancreas), which is consistent with the clinical manifestation of MS disease, as almost two-thirds of MS patients have at least one chronic GI-associated symptom, including irritable bowel syndrome and constipation [161]. The results also include the breast and cervix uteri, which are female reproductive organs; females have increased risk compared to men in developing MS [162]. In addition, female MS patients can have breast discomfort or pain and patients undertaking immunomodulatory drugs have an increased risk of cervical cancer [163, 164]. Other tissues identified are the muscle, skin
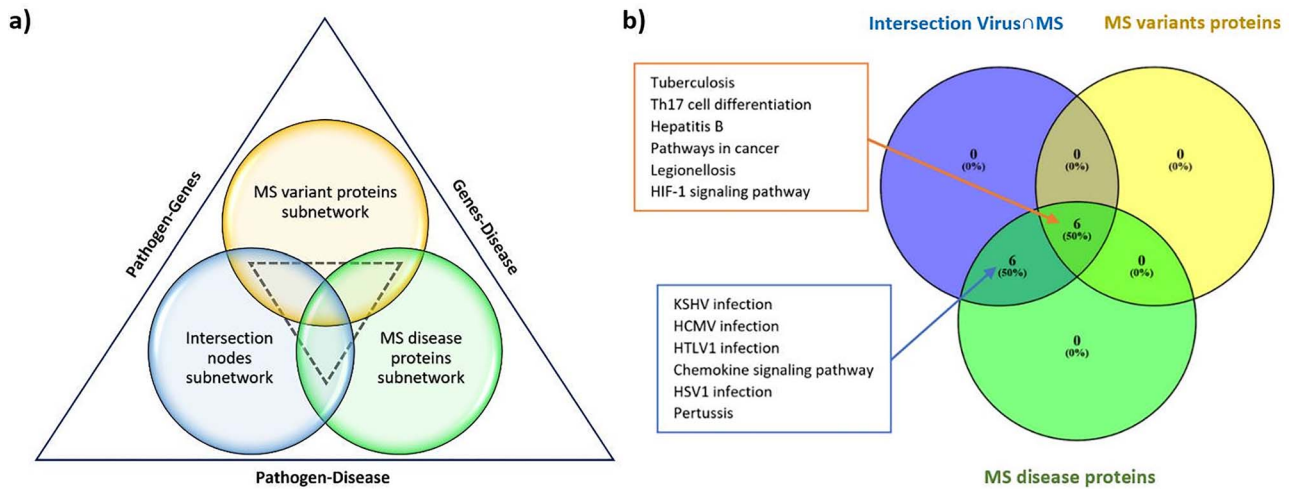
Figure 12. (**A**) Pathogen–genes–disease triangle. (**B**) Comparison of the post-filtering KEGG pathways enriched results that are also MS disease-related terms between the three subnetworks, indicating the 12 KEGG pathways that fall within the pathogen–genes–disease triangle.

Table 5. Comparison of the seven complementary pathways (*) identified in the 12 KEGG pathways complementary network with the MS disease pathways terms

KEGG pathway name	MS unique, MS shared, NDs common	Average *P*-value	Whole network *P*-value <0.05	Whole network *P*-value >0.05
Tuberculosis	MS unique	4.05E−14	5.29E−05	–
Kaposi's sarcoma herpesvirus (KSHV) infection	MS unique	2.87E−17	9.03E−08	–
Th17 cell differentiation	MS unique	3.00E−14	4.08E−05	–
HCMV infection	MS unique	6.57E−16	7.79E−10	–
Hepatitis B	MS unique	6.02E−22	1.27E−13	–
HTLV-1 infection	MS unique	1.65E−15	3.98E−17	–
Chemokine signaling pathway	MS unique	1.02E−09	–	8.09E−01
HSV-1 infection	MS unique	4.42E−27	5.68E−28	–
Pathways in cancer	MS unique	9.92E−14	3.16E−05	–
Pertussis	MS shared	1.40E−06	–	6.86E−01
Legionellosis	MS unique, MS shared, NDs common	6.31E−11	7.15E−08	–
Hypoxia-inducible factor 1(HIF-1) signalling pathway	NDs common	1.03E−04	3.05E−02	–
Transforming growth factor beta (TGF-beta) signalling pathway*	–	1.01E−02	–	6.05E−02
T-cell receptor signaling pathway* Calcium signaling pathway*	MS unique	4.71E−12 –	7.08E−04 –	– 1.00E+00
Janus kinase (JAK)-signal transducer and activator of transcription (STAT) pathway *	MS unique	8.74E−10	–	3.90E−01
Mitogen-activated protein kinase (MAPK) signalling pathway*	–	1.97E−05	–	2.10E−01
Nuclear factor (NF)-kappa B signalling pathway*	MS unique	5.47E−17	1.32E−08	–
Toll-like receptor signaling pathway*	MS unique	8.53E−19	6.67E−09	–

Average *P*-value obtained from the three subnetworks and whole network *P*-values of the 12 KEGG pathways and the 7 complementary pathways (*).

and artery, which are also consistent with the clinical manifestation of the disease, as MS patients have spasticity symptoms due to impaired signaling between the nerves and the muscles, altered sensation and itching, and vascular symptoms, such as changes in major neck arteries [165–167]. The kidney tissue is also included; some MS patients can develop kidney problems due to urinary incontinence, which is a frequent symptom in MS [168]. In addition, there is bidirectional crosstalk between the kidney and the brain, with patients with chronic kidney disease frequently developing neurological disorders [169]. Moreover, the kidney is involved in the secretion of several hormones including the active form of vitamin D, and interestingly, low levels of vitamin D are associated with increased MS risk [170, 171].

The tissue-specific gene enrichment analysis of the 1771 human proteins targeted by the 80 viral proteins identified 14 tissues, shown in Figure 16, where the interaction of these viral proteins with the host could be manifested. Comparison with the 13 enriched tissues of the 200 MS disease proteins indicated nine common tissues (spleen, nerve, liver, kidney, cervix uteri, artery, muscle, skin and brain) that could represent possible
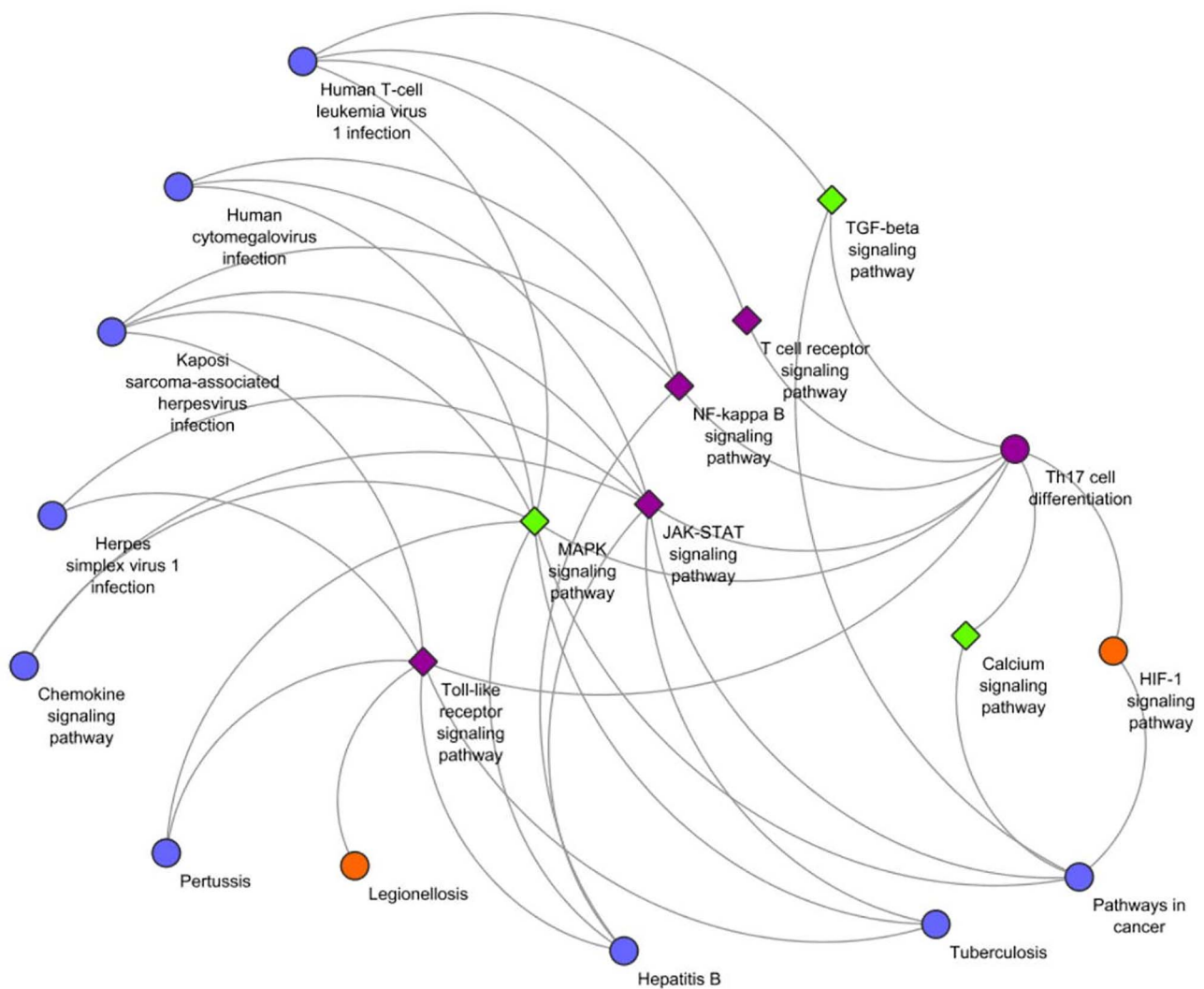
Figure 13. Schematic visualization of the complementary network for the 12 KEGG pathways, which are indicated in circular nodes, and the 7 complementary nodes shown in diamond shape. The orange nodes represent nodes that are also part of the four common ND pathways and purple nodes are hub–bottleneck nodes that act as a bridge of communication between nodes/pathways in the MS disease KEGG pathways network. The green nodes are complementary nodes, which are not MS disease-related pathway terms (MS unique, MS shared or NDs common).

interaction areas where the virus-induced dysregulation of the identified pathways could occur. The 80 viral proteins belong to 18 viral strains from eight species (EBV, HCMV, HHV-6A, HHV-6B, HSV-1, HTLV-1, Measles and Rubella), whose viral tropism varies. Viral tropism defines the ability of different viral species or viral strains to infect different cell types or tissues, which is an important factor that defines the tissue specificity of virus–host interactions [172]. Evidences described below confirm the interaction of these eight viral species with these nine tissues and indicates which species interacts with each of these tissues.

EBV, HCMV, HHV-6A, HHV-6B and HSV-1 belong to the herpesvirus family and are neurotropic viruses that can infect neuronal cells and the brain and are associated with various neurological disorders [49, 173]. Measles, Rubella and HTLV-1 can also be neuropathogenic and can cause demyelinating diseases in certain cases [49, 174–176]. EBV, HSV and HHV-6 have been detected in higher frequency in the CNS of MS patients compared to controls [177]. In addition, all members of the herpesvirus family can case mild liver disease [178] and members of the *betaherpesvirus*, HCMV and HHV-6 can establish latent infection in the kidneys [179]. Liver and kidney involvement has also been found during Measles infection [180, 181]. Rubella infection can also cause acute liver failure in rare occasions [182]. HSV-1 can cause oral skin herpes and eczema herpeticum [183]. Acute infections with measles and rubella viruses involve a characteristic skin rash, and individuals infected with HTLV-1 can develop several dermatological symptoms [184–186]. Infectious mononucleosis, caused by EBV, can also be presented with a rash and an enlarged spleen [187]. HTLV-1 proviral DNA has been found in the liver, spleen, brain and in the blood *in vivo* [188]. Dissemination of the measles virus includes the skin, spleen and the circulatory system [189]. HSV-1 and HCMV can target endothelial cells and cause atherosclerosis, a cardiovascular disease that involves narrowing of the arteries [190], and DNA from EBV, HSV and HCMV has been found in arteriosclerotic plaques [191].HTLV-1 is also associated with the development of coronary artery diseases, and there is evidence of increased seroprevalence of HTLV-1 in individuals with cardiovascular diseases [192, 193]. Muscle pain is also a common symptom in several viral infections, such as measles, and certain viruses can
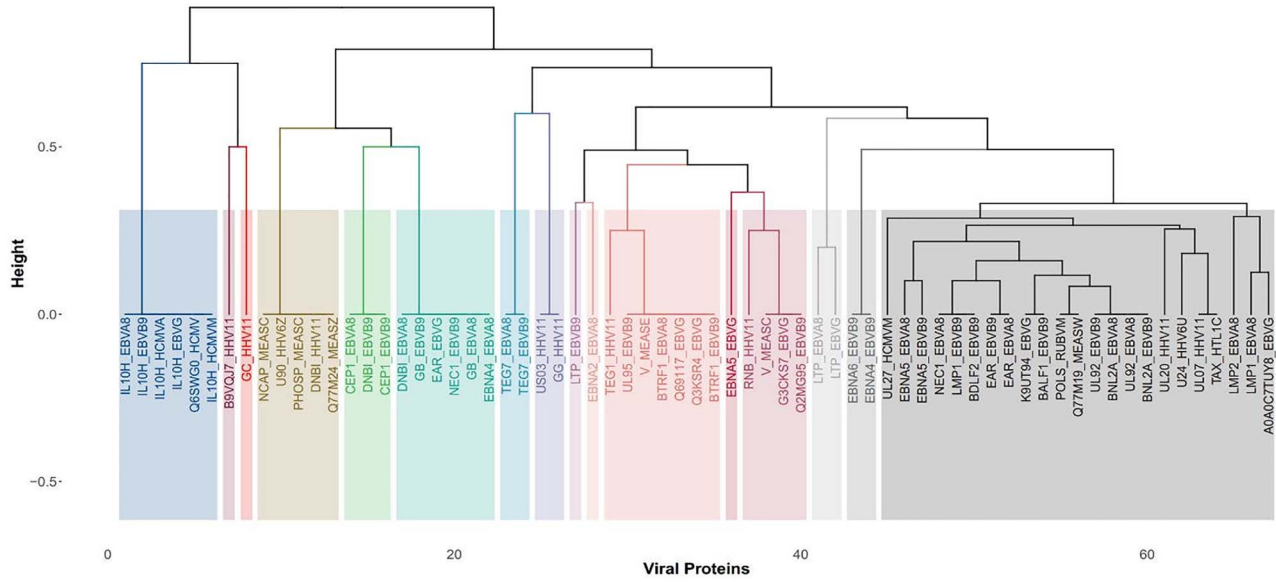
Figure 14. Clustering dendrogram of the 67 viral proteins from 8 viral species EBV, HCMV, HHV-6A, HHV-6B, HSV-1, HTLV-1, Measles and Rubella based on target similarity of the final 12 KEGG pathways.
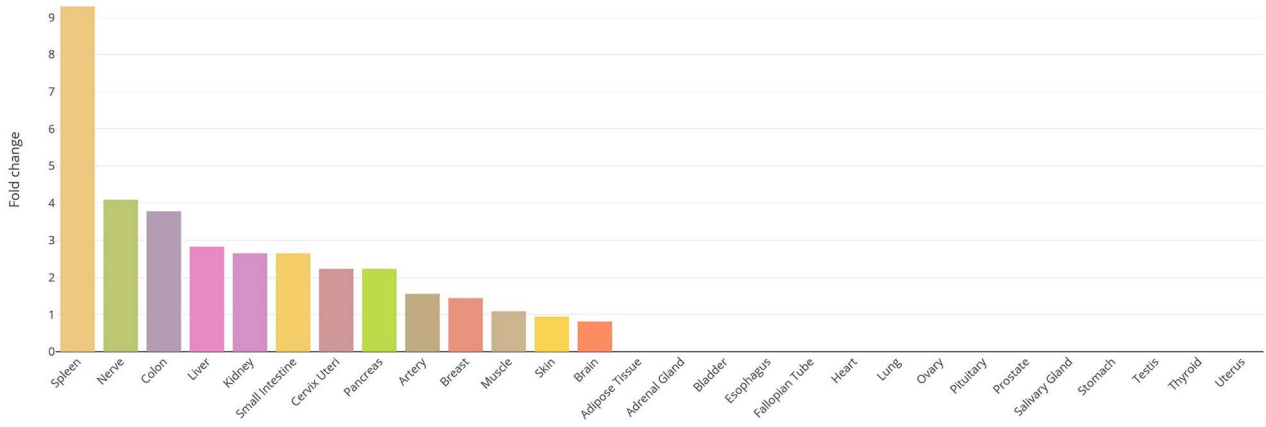
Figure 15. Tissue-specific gene enrichment analysis results of the 200 MS disease-associated proteins, using the GTEx dataset.
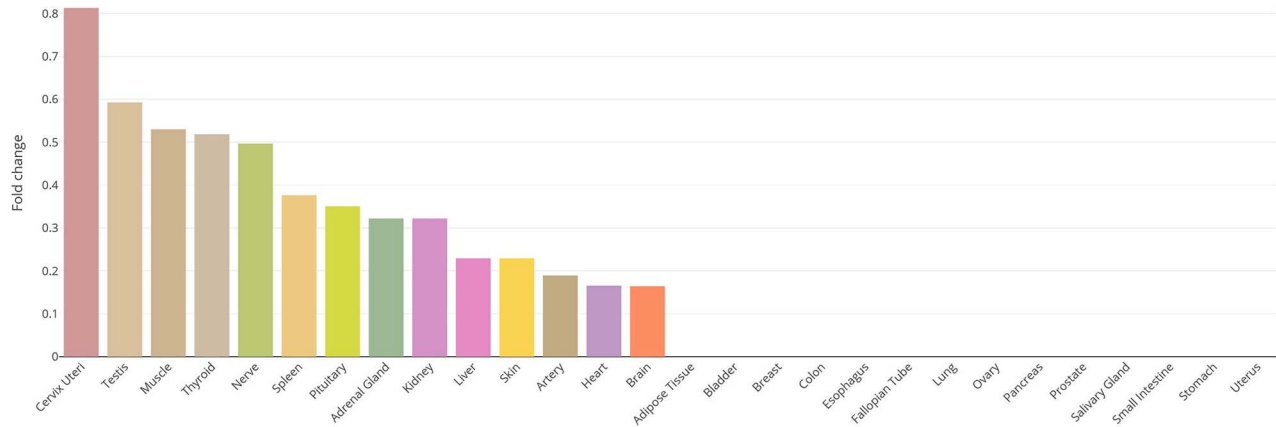
Figure 16. Tissue-specific gene enrichment analysis results of the human targets of the 80 viral proteins that interact with the 12 KEGG pathways, 52 GO ISP and 5 Reactome pathways, using the GTEx dataset.

also infect muscles cells, such as HCMV that can infect smooth muscle cells [194]. Moreover, HLTV-1 can cause inflammatory myopathies in some individuals [195]. Several of the included viruses can also infect the cervix, for example HSV-1 can cause genital herpes, HHV-6 can infect cervical epithelial cells and both EBV and HCMV have also been detected in the uterine cervix [196–198]. Furthermore, both EBV and HTLV-1 are oncogenic viruses, and evidence indicates that coinfection of the cervix by any of this two viruses with human papillomavirus confers increased risk for the development of cervical cancer, suggesting the existence of synergistic oncogenic effects [199–202]. The presence of measles virus and rubella viruses was also detected in the uterine cervix [203–205].

Validation of the contribution of the 12 KEGG pathways in MS

Our analysis revealed 12 KEGG pathways by which 8 viral species associated with the development of MS might exert their viral-mediated pathogenic effects. In this section, we evaluate the contribution of the identified pathways in MS disease.

First, we performed pathway enrichment analysis on the 200 MS disease-associated proteins, using the same parameters as before, which led to the identification of 71 statistically significant KEGG pathways (see Table 1 of Supplementary File 3 available online at https://academic.oup.com/bib). The enrichment results of the MS disease proteins indicated the presence of all 12 KEGG pathways (Figure 12B) in the 71 KEGG pathways, therefore confirming the involvement of our 12 identified pathways in MS disease. In addition, the 71 MS-enriched KEGG pathways contained several infectious disease pathways, including 5 out of the 8 viral species that were identified in our analysis to interact with the 12 KEGG pathways, namely EBV, HTLV-1, HCMV, HSV-1 and Measles. However, KEGG database does not contain pathways for the other three viral species (HHV6A, HHV6B and Rubella); hence their contribution in MS disease cannot be excluded.

From the identified 12 KEGG pathways, 8 are infectious disease pathways and viral proteins from the 8 viral species can interact with these pathways even though the pathway might be of another infectious disease. For example, Q77M19_MEASW protein of the Measles virus targets human proteins associated with each of the eight infectious disease pathways even though none of them is the Measles infectious disease pathway. This is because viruses from different families or species within the same family can have common and unique infection strategies; therefore, viruses that interact with several similar human targets that participate in an infectious disease pathway might be also able to exert the same perturbation that causes a certain comorbidity. These 8 infectious disease pathways include HTLV-1, HCMV and HSV-1 viruses which have been associated with the development of MS, as shown in Table 1, and were also included in our reconstructed virus-host-MS PPI network.

The 12 KEGG pathways also include the Chemokine signaling pathway, which is involved in the recruitment of leukocytes in inflamed or infected tissues [206] and the Th17 cell differentiation pathway. Activation of Th17 cells also participates in the recruitment of neutrophils and macrophage via the production of proinflammatory cytokines [207]. Both chemokines and Th17 cells have been shown to play an important role in the immunopathogenesis of MS as they participate in the recruitment of pathogenic immune system cells in the CNS [208–210]. Viruses to escape host immune system detection and destruction have developed an array of immune evasion strategies

including interfering with cellular components of the immune response and immune effector functions [43]. Our analysis indicated that viruses target human proteins from these immune system pathways, therefore by interfering with these pathways viruses can possibly lead to the dysregulation of immune system process, which in turn propagates the emergence of autoimmune responses.

We also identified the HIF-1 signaling pathway, which is involved in the regulation of hypoxia responses. Inflammation can exacerbate hypoxia and vice versa, and the 'hypoxia–inflammation cycle' has been hypothesized to play a role in MS progression [211]. HIF-1 alpha can also be induced by a wide range of infectious agents, and its expression influences infectious disease outcome [212]. In addition, increased expression of HIF-1 alpha is found in several cancers, and interestingly several oncogenic viruses, including EBV and HTLV-1, can increase the expression of HIF-1 [213, 214]. Pathways in cancer was also one of the 12 KEGG pathways, and the pathway-to-pathway network reconstruction indicated that Th17 cell differentiation pathway, HIF-1 signaling pathway and pathways in cancer are the only pathways that functionally interact with each other from the 12 KEGG pathways. Immune system responses play an essential role in cancer as they are involved in the elimination of malignant cells. MS patients undertaking immunomodulatory disease-modifying therapies (DMTs) have increased risk in developing certain types of cancer [215]. Further investigation is therefore warranted to understand the contribution of the interaction between the identified viruses and the pathway-to-pathway interactions between Th17 cell differentiation pathway, HIF-1 signaling pathway and pathways in cancer in the development of MS, but also how the effects of DMTs on these pathways might contribute to the emergence of cancer comorbidities.

Discussion

In this study, we performed a theoretical-driven network-based bioinformatics pipeline approach that utilizes various databases and bioinformatics tools presented in the first part of this paper with the aim to identify viral-mediated pathogenic mechanisms that may be associated with the development and/or progression of MS disease. This disease is characterized by both a neurodegenerative and an autoimmune component. Viral infections have been associated with both the development of NDs and autoimmune diseases. In this paper, we investigated viral-mediated perturbations in MS disease, by focusing on the neurodegenerative component of the disease, by comparing MS disease-associated proteins with three other NDs (ALS, PD and AD) that have been also associated with viral infections in order to identify MS unique, MS shared and ND common pathways. We identified four common ND KEGG pathways that may represent common pathological mechanisms in this group of NDs: the IBD, HIF-1 signaling pathway, malaria and legionellosis pathways. Interestingly, all four NDs have been associated with IBD, which is characterized by chronic inflammation in the gastrointestinal gut, and one main hypothesis is that abnormal brain–gut interactions might be involved in its pathogenesis [216–220]. Accumulating evidence suggests that microbiota dysbiosis facilitates IBD genesis and the development and/or progression of several NDs, including PD, AD, ALS and MS via the perturbation of the brain–gut axis [221, 222]. For example, microbiota-induced pathological mechanisms can cause increased BBB permeability, augment neuroinflammation and affect the production of several neurotransmitters produced by gut microbiota, such as dopamine that is dysregulated in PD [223, 224]. Interestingly, evidences suggest

that commensal microbiota can influence viral infections via direct or indirect mechanisms as they can suppress or promote certain viral infections [225, 226]. On the other hand, infection with several viruses can lead to microbiota dysbiosis [225], which could also represent another possible pathological mechanism by which viruses could lead to NDs. Therefore, future work will aim to investigate the interaction between viruses and commensal microbiota and its possible contribution in the development or progression of NDs using omics data analysis. The HIF-1 signaling pathway on the other hand is involved in the regulation of molecular signals during hypoxia, infection and the stimulation of pro-inflammatory signals [212]. Hypoxia and the HIF-1 signaling pathway have been associated with the development and progression of wide range of diseases, with hypoxia being a common characteristic in NDs [227]. Malaria and legionellosis are infectious pathways, thus highlighting the possible association between pathogens and NDs.

Statistically significant enrichment analysis of the 166 unique MS disease proteins also confirmed the association between pathogens and their interaction with the immune system in the development of MS. In addition, the KEGG enrichment results indicate that 42.4% of the pathways belong to the Rheumatoid arthritis group, suggesting the existence of common pathogenic mechanisms between MS and Rheumatoid arthritis [228–230]. The GO ISP enrichment analysis results also indicated that 92% of the MS unique significant terms belong to the group of positive regulation of lymphocyte proliferation. The balance between positive regulation signals that mediate the activation and proliferation of lymphocytes and negative regulation signals that induce their termination is essential in order to avoid autoimmune disorders, such as MS [231]. Homeostasis in the regulation of immune responses can be disrupted when immune self-tolerance fails or when the immune system fails to control chronic viral infections, resulting in continuous stimulation of lymphocytes, and hence chronic positive activation of immune responses [231]. Therefore, the GO ISP results indicate that homeostasis in the regulation of immune process in MS is dysregulated, and there is a shift towards positive activation of lymphocytes that possibly contributes to MS pathogenesis.

Topological analysis of the MS-enriched KEGG pathways network allowed to identify seven hub–bottleneck nodes (Figure 17) that can act as a bridge of communication between the rest of the pathways in the network. Pathogens are known to target hubs and bottlenecks proteins in order to exert systemic infectious effects [63, 68, 91]; therefore, hub–bottleneck nodes can possibly act as disease communicator nodes bridging the communication between infectious pathways and the rest of the MS-enriched KEGG pathways network. Subsequently, any dysregulation in a pathway that interacts with a hub–bottleneck or dysregulation of a hub–bottleneck node directly can lead to a cascade of events causing systemic dysregulation of a subset of pathways, resulting possibly in the development of MS. In order to identify communities of infectious disease pathways that interact with the identified hub–bottleneck pathways, we performed community clustering analysis of the MS-enriched KEGG pathways network that indicated four clusters (Figure 10). More specifically, the community clustering results showed that EBV, HSV-1 and Measles viral infection pathways, which have being linked as a risk factor for the development of MS interact with the Toll-like receptor signaling pathway and JAK-STAT signaling pathway hub–bottleneck nodes (Figure 10A). Whereas, HCMV and HTLV-1 that are also associated with MS form a different community cluster and interact with the NF-kappa B signaling

pathway and Th17 cell differentiation pathway hub–bottleneck nodes (Figure 10C).

Our methodology also involved the reconstruction, subnetwork identification and pathway enrichment analysis of the integrated virus–host–MS PPI network, where after applying our filtering process and isolating pathways that are also MS disease-enriched terms, we were able to identify 12 enriched KEGG pathways, (Figure 12B), 52 GO ISP and 5 Reactome pathways (Supplementary File 1 available online at https://academic.oup.com/bib). The final list of GO ISP and pathways obtained through our pipeline approach accounts for pathogen–genes–disease interactions as they are pathways that contain MS variants proteins, MS disease proteins and virus–host PPIs. Furthermore, via our filtering process, the final results account for the immunogenicity, autoimmunity and genetic components of MS disease. Finally, via the filtering process and the comparison with other NDs by obtaining only pathways that are associated with MS unique, MS shared and NDs common pathways, we were able to isolate the most relevant MS disease pathways from a large pool of enriched pathway results. The specificity of our processes in extracting relevant MS disease pathways by which viruses might exert their pathogenic effects is analyzed in Supplementary File 4 available online at https://academic.oup.com/bib.

To investigate pathway-to-pathway interactions of the final 12 KEGG pathways, a complementary network was created as it can provide possible pathways of communication of the 12 KEGG pathways with other pathways, particularly MS disease-related pathways and hub–bottleneck disease communicator nodes. The complementary network, shown in Figure 13, indicates seven proximal complementary nodes with JAK-STAT signaling pathway, Toll-like receptor signaling pathway, NF-kappa B signaling pathway and T-cell receptor signaling complementary nodes being also MS disease-related pathways and hub–bottleneck disease communicator nodes in the MS-enriched KEGG pathways network. All the 12 KEGG pathways directly interact with one or more of the hub–bottleneck disease communicator nodes; therefore by interacting with disease communicator nodes, they can exert systemic effects within the MS-enriched KEGG pathways network (Figure 13). In addition, the three complementary nodes, Calcium signaling pathway, MAPK signaling pathway and TGF-beta signaling pathway that are not MS disease-related pathways interact with the Th17 cell differentiation pathway, which is part of the 12 final KEGG pathways and a disease communicator node in the MS-enriched KEGG pathways network; therefore, they might indirectly contribute to the development of MS (Figure 13). The Th17 cell differentiation pathway acts as a key disease communicator node in the complementary network of the final 12 KEGG pathways as it interacts with all seven of the complementary nodes of which four are also disease communicator nodes. Moreover, it interacts with the HIF-1 signaling pathway which was identified through our analysis as one of the four common ND pathways. This makes the Th17 cell differentiation pathway a key viral-mediated pathogenic mechanism and a possible therapeutic target for MS disease. The relapsing–remitting MS phase of MS involves the infiltration of CD4+ T cells with Th1 and Th17 proinflammatory phenotypes [208, 209], thus highlighting the role of the Th17 cell differentiation pathway in the immunopathogenesis of MS. Th17 cells are also involved in the immunopathogenesis of multiple other autoimmune diseases including IBD and Rheumatoid arthritis [232]. In addition, interferon beta, which is used as a first line of treatment for MS patients at the
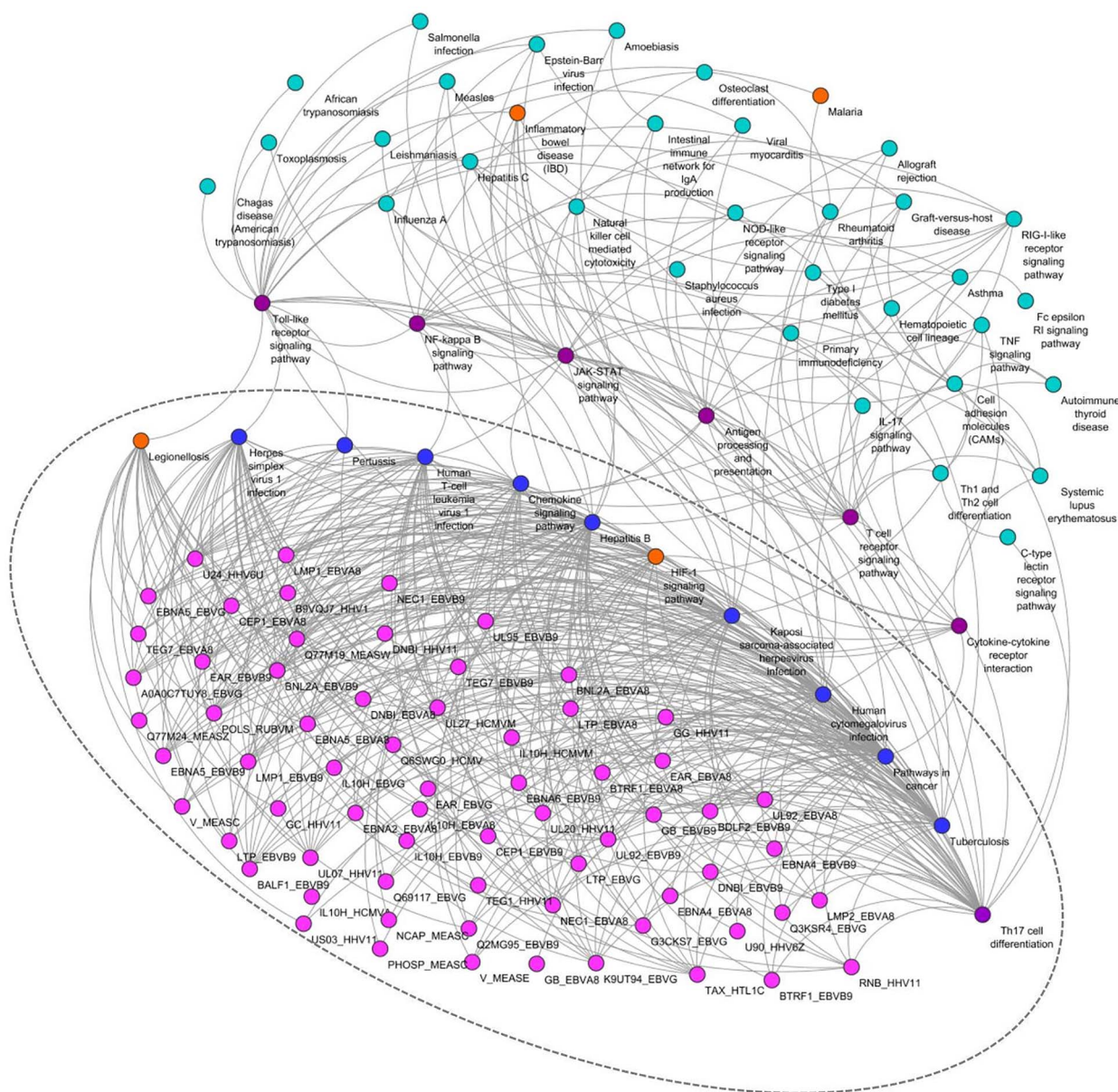
Figure 17. Schematic illustration of the possible viral-mediated pathogenic mechanisms obtained through our pipeline approach indicating how the resulting 67 viral proteins (lilac color) targeting the 12 identified KEGG pathways in the MS-enriched KEGG pathways network, which in turn interact with the hub–bottleneck disease communicator nodes (purple color), can exert systemic effects within the network and lead to the development of MS. The ND common pathways are indicated in orange.

relapsing–remitting MS phenotype, has shown to exert its pharmacological effects by inhibiting indirectly Th17 cell differentiation [233–235].

The final 12 KEGG pathways, 52 GO ISP and 5 Reactome pathways are targeted by 80 viral proteins from 8 viral species (EBV, HCMV, HHV-6A, HHV-6B, HSV-1, HTLV-1, Measles and Rubella). Therefore, viruses by targeting the identified pathways in the MS-enriched KEGG pathways network, illustrated in Figure 17, which in turn interact with the hub–bottleneck disease communicator pathways, may be able to exert systemic pathogenic effects, thus leading to the development of MS. By performing similarity and clustering analysis, we were able to identify the group of pathways (KEGG or Reactome) and GO ISP that can be targeted by each viral protein, but also clusters of viral proteins

from different species that can target similar pathways. The clustering analysis also indicates viral proteins that can possibly exert similar effects in the MS disease pathways network by affecting the same group of pathways.

Tissue-specific gene enrichment analysis led to the identification of nine common tissues (spleen, nerve, liver, kidney, cervix uteri, artery, muscle, skin and brain) that could represent possible interaction areas where the virus–MS disease interactions could occur. The cervix uteri had the highest fold change value in the tissue-enrichment analysis of the human targets of the 80 viral proteins, and literature evidence indicates that all of the 8 viral species have been detected in the cervix, including the oncogenic viruses EBV and HTLV-1. Female MS patients undertaking immunomodulatory drugs have increased

risk of cervical cancer [163, 164]. Therefore, infection of the cervix uteri by these viruses in association with other cofactors particularly immunosuppressive drugs that are known to cause viral reactivation could contribute to increased cervical cancer risk in MS. Our analysis also indicated that KEGG pathway-to-pathway interactions between Th17 cell differentiation pathway, HIF-1 signaling pathway and pathways in cancer that are targeted by the identified eight viral species and are part of our final 12 KEGG pathways might contribute to cancer comorbidity in MS. Therefore, further investigation is warranted to understand how the interaction of viruses, particularly known oncogenic viruses, with these three pathways in combination with immunomodulatory drugs might contribute to the development of cervical cancer.

Finally, we briefly explored the autoimmune component of MS by comparing disease-associated proteins between MS and three other autoimmune diseases SLE, RA and type 1 diabetes that are also associated with viral infections. All of these diseases are associated with EBV infection, and the enrichment analysis of the 53 MS disease-associated proteins that are shared with some of these other autoimmune diseases indicated that 48.78% of the shared KEGG pathway enriched results belonging to the EBV infection group. These results confirm the role of EBV infection in the pathogenesis of these four autoimmune diseases, but also indicate that the pathogenic effects of EBV are possibly mediated through different biological mechanisms in these diseases. This finding might provide a possible explanation on how the same virus might be able to cause different disease phenotypes.

Pipeline limitations

One important limitation of our study is the imbalance of available experimental virus–host PPIs data, as more studied viruses like EBV have more available data, which accounts for 72.8% of the virus–host PPI data included in our integrated virus–host–MS PPI network, than less studied viruses, such as HHV-6A and HHV-6B. A current approach used to overcome this limitation is to use computationally predicted virus–host PPIs obtained from machine learning algorithms, but such methods are not without drawbacks due to the high false-positive and false-negative rates [236]. Another limitation inherent to pathway enrichment analysis is that the standard approach is to select only statistically significant pathways with P-value ≤0.05. However, pathways that are not revealed through the current analysis as statistically significant could also be important as viral proteins also interact with human protein targets in these pathways and therefore might be able to cause disease effects via non-statistically significant pathways. Furthermore, in this paper, we explore the effect of several viruses, with different cellular and tissue tropism, in the development of MS using an integrated heterogeneous network. Another limitation of our approach is that it does not account for tissue-specific interactions of these viruses with the host. However, viral tissue-specific interactions with the host can be investigated when having a single virus or viruses that infect the same cell types or tissues, which is not the case here.

Conclusion

In this paper, we initially reviewed databases and tools that can be utilized to investigate viral-mediated perturbations of the host's interactome that lead to the generation of complex diseases, such as NDs. We then presented our integrative network-based bioinformatics pipeline approach that accounts for pathogen–genes–disease PPIs with the aim to identify possible viral-mediated pathogenic mechanisms in MS disease.

Comparison between disease-associated proteins of four NDs (ALS, MS, PD and AD) associated with viruses confirmed the role of pathogens in the development of NDs and led to the identification of the HIF-1 signaling pathway as a possible common pathogenic mechanism in NDs. Reconstruction and topological analysis of the MS-enriched KEGG pathways network that includes MS unique, MS shared and ND common pathways enabled us to identify seven hub–bottleneck nodes that can act as disease communicator nodes and exert systemic effects.

Through the reconstruction of the virus–host–MS PPIs network and the application of our methodology, we were also able to identify 12 KEGG pathways, 52 GO ISP and 5 Reactome pathways that are targeted by 80 viral proteins from 8 viral species. These viruses might exert their viral-mediated pathogenic mechanisms by interacting with these 12 KEGG pathways, which in turn interact with the hub–bottleneck disease communicator nodes, allowing them to exert systemic effects, hence affecting several MS disease pathways.

Finally, our analysis highlights the Th17 differentiation pathway, which is part of the 12 underlined KEGG pathways as a key viral-mediated pathogenic mechanism and a possible therapeutic target for MS disease as it is a central hub–bottleneck disease communicator node in the MS disease-enriched KEGG pathways network. To further elucidate the role of these pathways in MS disease, computational modeling can be used to determine if the viral proteins that target a specific pathway can shift the pathway into a disequilibrium state causing viral-mediated perturbations in the host interactome.

Key Points

- NDs are chronic degenerative neurological diseases, and currently, there are no effective pharmacotherapies for their treatment.
- Although several environmental and genetic factors have been implicated to their development, the exact underlying mechanisms are still unclear.
- Viral infections have been associated with several NDs, with virus–host PPIs representing a key pathogenic mechanism that can lead to the generation of perturbations within the human interactome.
- This work illustrates how network-based approaches used for the investigation of NDs and virus–host PPIs can be integrated to identify viral-mediated pathogenic mechanisms in MS.

Supplementary data

Supplementary data are available online at *Briefings in Bioinformatics*.

Data availability

The data used in this article are derived from sources in the public domain.

Funding

References

1. Migliore L, Coppedè F. . Genetics, environmental factors and the emerging role of epigenetics in neurodegenerative diseases. Mutat. Res. Fundam. Mol. Mech, 2009;**667**:82–97.

2. Patrick KL, Bell SL, Weindel CG, *et al*. Exploring the "multiple-hit hypothesis" of neurodegenerative disease: bacterial infection comes up to bat. *FCIMB* 2019;**9**:138.

3. De Chiara G, Marcocci ME, Sgarbanti R, *et al*. Infectious agents and neurodegeneration. *Mol Neurobiol* 2012;**46**:614–38.

4. Alam M, Alam Q, Kamal M, *et al*. Infectious agents and neurodegenerative diseases: exploring the links. *Curr Top Med Chem* 2017;**17**:1390–9.

5. Zhou L, Miranda-Saksena M, Saksena NK. Viruses and neurodegeneration. *Virol J* 2013;**10**:172.

6. Sochocka M, Zwolińska K, Leszek J. The infectious etiology of Alzheimer's disease. *Curr Neuropharmacol* 2017;**15**:996–1009.

7. Itzhaki RF, Dobson CB, Shipley SJ, *et al*. The role of viruses and of APOE in dementia. *Ann N Y Acad Sci* 2004;**1019**:15–8.

8. Readhead B, Haure-Mirande JV, Funk CC, *et al*. Multiscale analysis of independent Alzheimer's cohorts finds disruption of molecular, genetic, and clinical networks by human herpesvirus. *Neuron* 2018;**99**:6464–82.e7.

9. Carbone I, Lazzarotto T, Ianni M, *et al*. Herpes virus in Alzheimer's disease: relation to progression of the disease. *Neurobiol Aging* 2014;**35**:122–9.

10. Kuhlmann I, Minihane AM, Huebbe P, *et al*. Apolipoprotein e genotype and hepatitis C, HIV and herpes simplex disease risk: a literature review. *Lipids Health Dis* 2010;**9**:8.

11. Broccolo F, Fusetti L, Ceccherini-Nelli L. Possible role of human herpesvirus 6 as a trigger of autoimmune disease. *Sci World J* 2013;**2013**.

12. Dolei A. The aliens inside us: HERV-W endogenous retroviruses and multiple sclerosis. *Mult Scler* 2018;**24**:42–7.

13. Dolei A, Uleri E, Ibba G, *et al*. The aliens inside human DNA: HERV-W/MSRV/syncytin-1 endogenous retroviruses and neurodegeneration. *J Infect Dev Ctries* 2015;**9**: 577–87.

14. Lee SM, Morcos Y, Jang H, *et al*. HTLV-1 induced molecular mimicry in neurological disease. *Mol Mimicry Infect Autoimmune Dis* 2006;**296**:125–36.

15. Oger J. HTLV-1 infection and the viral etiology of multiple sclerosis. *J Neurol Sci* 2007;**262**:100–4.

16. Jarius S, Eichhorn P, Franciotta D, *et al*. The MRZ reaction as a highly specific marker of multiple sclerosis: re-evaluation and structured review of the literature. *J Neurol* 2017;**264**:453–66.

17. Vanheusden M, Stinissen P, 't Hart BA, *et al*. Cytomegalovirus: a culprit or protector in multiple sclerosis? *Trends Mol Med* 2015;**21**:16–23.

18. Tselis A. Evidence for viral etiology of multiple sclerosis. *Semin Neurol* 2011;**31**:307–16.

19. Stoner GL. Implications of progressive multifocal leukoencephalopathy and JC virus for the etiology of MS. *Acta Neurol Scand* 1991;**83**:20–33.

20. Wortman MJ, Lundberg PS, Dagdanova AV, *et al*. Opportunistic DNA recombination with Epstein-Barr virus at sites of control region rearrangements mediating JC virus Neurovirulence. *J Infect Dis* 2016;**213**:1436–43.

21. Jang H, Boltz DA, Webster RG, *et al*. Viral parkinsonism. *Biochim Biophys Acta Mol Basis Dis* 2009;**1792**:714–21.

22. Olsen LK, Dowd E, McKernan DP. A role for viral infections in Parkinson's etiology? *Neuronal Signal* 2018;**2**:NS20170166.

23. Le BX, Wang X, Xiang Y, *et al*. The association between infectious burden and Parkinson's disease: A case-control study. *Parkinsonism Relat Disord* 2015;**21**:877–81.

24. Nielsen NM, Rostgaard K, Hjalgrim H, *et al*. Poliomyelitis and Parkinson disease. *JAMA* 2002;**287**:1650–1.

25. Xue YC, Feuer R, Cashman N, *et al*. Enteroviral infection: the forgotten link to amyotrophic lateral sclerosis? *Front Mol Neurosci* 2018;**11**:63.

26. Bowen LN, Tyagi R, Li W, *et al*. HIV-associated motor neuron disease: HERV-K activation and response to antiretroviral therapy. *Neurology* 2016;**87**:1756–62.

27. Alfahad T, Nath A. Retroviruses and amyotrophic lateral sclerosis. *Antivir Res* 2013;**99**:180–7.

28. Nielsen TR, Rostgaard K, Nielsen NM, *et al*. Multiple sclerosis after infectious mononucleosis. *Arch Neurol* 2007;**64**:72–5.

29. Robinson RL, Shahida S, Madan N, *et al*. Transient parkinsonism in West Nile virus encephalitis. *Am J Med* 2003;**115**:252–3.

30. Kanagarajan K, Ganesh S, Alakhras M, *et al*. West Nile virus presenting as cerebellar ataxia and fever: case report. *South Med J* 2003;**96**:600–1.

31. Sejvar JJ. Clinical manifestations and outcomes of West Nile virus infection. *Viruses* 2014;**6**:606–23.

32. Jang H, Boltz D, Sturm-Ramirez K, *et al*. Highly pathogenic H5N1 influenza virus can enter the central nervous system and induce neuroinflammation and neurodegeneration. *Proc Natl Acad Sci U S A* 2009;**106**:14063–8.

33. Hamaue N, Ogata A, Terado M, *et al*. Brain catecholamine alterations and pathological features with aging in Parkinson disease model rat induced by Japanese encephalitis virus. *Neurochem Res* 2006;**31**:1451–5.

34. Cerna F, Mehrad B, Luby JP, *et al*. St. Louis encephalitis and the substantia nigra: MR imaging evaluation. *Am J Neuroradiol* 1999;**20**:1281–3.

35. Xue YC, Ruller CM, Fung G, *et al*. Enteroviral infection leads to transactive response DNA-binding protein 43 pathology in vivo. *Am J Pathol* 2018;**188**:2853–62.

36. Mackenzie IRA, Rademakers R. The role of transactive response DNA-binding protein-43 in amyotrophic lateral sclerosis and frontotemporal dementia. *Curr Opin Neurol* 2008;**21**:693–700.

37. Küry P, Nath A, Créange A, *et al*. Human endogenous retroviruses in neurological diseases. *Trends Mol Med* 2018;**24**:379–94.

38. Li W, Lee MH, Henderson L, *et al*. Human endogenous retrovirus-K contributes to motor neuron disease. *Sci Transl Med* 2015;**7**:307ra153.

39. Lin WR, Shang D, Itzhaki RF. Neurotropic viruses and Alzheimer disease: interaction of herpes simplex type I virus and apolipoprotein E in the etiology of the disease. *Mol Chem Neuropathol* 1996;**28**:135–41.

40. Wozniak MA, Shipley SJ, Combrinck M, *et al*. Productive herpes simplex virus in brain of elderly normal subjects and Alzheimer's disease patients. *J Med Virol* 2005;**75**: 300–6.

41. Moir RD, Lathe R, Tanzi RE. The antimicrobial protection hypothesis of Alzheimer's disease. *Alzheimers Dement* 2018;**14**:1602–14.

42. van den Pol AN. Viral infections in the developing and mature brain. *Trends Neurosci* 2006;**29**:398–406.

43. Vossen MTM, Westerhout EM, Söderberg-Nauclér C, *et al*. Viral immune evasion: a masterpiece of evolution. *Immunogenetics* 2002;**54**:527–42.

44. Calistri A, Munegato D, Carli I, *et al*. The ubiquitin-conjugating system: multiple roles in viral replication and infection. *Cell* 2014;**3**:386–417.

45. Loureiro J, Ploegh HL. Antigen presentation and the ubiquitin-proteasome system in host-pathogen interactions. *Adv Immunol* 2006;**92**:225–305.

46. Deleidi M, Isacson O. Viral and inflammatory triggers of neurodegenerative diseases. *Sci Transl Med* 2012;**47**:121ps3.

47. Chen WW, Zhang X, Huang WJ. Role of neuroinflammation in neurodegenerative diseases (review). *Mol Med Rep* 2016;**13**:3391–6.

48. Kristensson K. Potential role of viruses in neurodegeneration. *Mol Chem Neuropathol* 1992;**16**:45–58.

49. Ludlow M, Kortekaas J, Herden C, *et al*. Neurotropic virus infections as the cause of immediate and delayed neuropathology. *Acta Neuropathol* 2016;**131**:159–84.

50. Stojdl DF, Lichty BD, Ten Oever BR, *et al*. VSV strains with defects in their ability to shutdown innate immunity are potent systemic anti-cancer agents. *Cancer Cell* 2003;**4**:263–75.

51. Church JA. CCR5 deficiency increases risk of symptomatic West Nile virus infection. *Pediatrics* 2007;**203**:35–40.

52. Wicker JA, Whiteman MC, Beasley DWC, *et al*. A single amino acid substitution in the central portion of the West Nile virus NS4B protein confers a highly attenuated phenotype in mice. *Virology* 2006;**349**:245–53.

53. Hou Y, Dan X, Babbar M, *et al*. Ageing as a risk factor for neurodegenerative disease. *Nat Rev Neurol* 2019;**15**:565–581.

54. Navratil V, de Chassey B, Combe CR, *et al*. When the human viral infectome and diseasome networks collide: towards a systems biology platform for the aetiology of human diseases. *BMC Syst Biol* 2011;**21**:5:13.

55. Brito AF, Pinney JW. Protein-protein interactions in virus-host systems. *Front Microbiol* 2017;**8**:1557.

56. Gulbahce N, Yan H, Dricot A, *et al*. Viral perturbations of host networks reflect disease etiology. *PLoS Comput Biol* 2012;**8**:e1002531.

57. Barabási AL, Albert R. Emergence of scaling in random networks. *Science* 1999;**286**:509–12.

58. Durmuş S, Ülgen K. Comparative interactomics for virus–human protein–protein interactions: DNA viruses versus RNA viruses. *FEBS Open Bio* 2017;**7**:96–107.

59. Franzosa EA, Xia Y. Structural principles within the human-virus protein-protein interaction network. *Proc Natl Acad Sci* 2011;**108**:10538–43.

60. Daugherty MD, Malik HS. Rules of engagement: molecular insights from host-virus arms races. *Annu Rev Genet* 2012;**46**:677–700.

61. Rachita HR, Nagarajaram HA. Viral proteins that bridge unconnected proteins and components in the human PPI network. *Mol BioSyst* 2014;**10**:2448–58.

62. Calderwood MA, Venkatesan K, Xing L, *et al*. Epstein-Barr virus and virus human protein interaction maps. *Proc Natl Acad Sci U S A* 2007;**104**:7606–11.

63. Dyer MD, Murali TM, Sobral BW. The landscape of human proteins interacting with viruses and other pathogens. *PLoS Pathog* 2008;**4**:e32.

64. Wuchty S, Siwo G, Ferdig MT. Viral organization of human proteins. *PLoS One* 2010;**5**:e11796.

65. Barabási AL, Oltvai ZN. Network biology: understanding the cell's functional organization. *Nat Rev Genet* 2004;**5**:101–13.

66. Jeong H, Mason SP, Barabási AL, *et al*. Lethality and centrality in protein networks. *Nature* 2001;**411**:41–2.

67. Gursoy A, Keskin O, Nussinov R. Topological properties of protein interaction networks from a structural perspective. *Biochem Soc Trans* 2008;**36**:1398–403.

68. Meyniel-Schicklin L, De Chassey B, André P, *et al*. Viruses and interactomes in translation. *Mol Cell Proteomics* 2012;**11**:M111.014738.

69. Uetz P, Dong YA, Zeretzke C, *et al*. Herpesviral protein networks and their interaction with the human proteome. *Science* 2006;**311**:239–42.

70. Ashburner M, Ball CA, Blake JA, *et al*. Gene ontology: tool for the unification of biology. *Nat Genet* 2000;**25**:25–9.

71. Pan A, Lahiri C, Rajendiran A, *et al*. Computational analysis of protein interaction networks for infectious diseases. *Brief Bioinform* 2016;**17**:517–26.

72. Friedel CC. Computational analysis of virus-host interactions. *Virus Host Interact Methods Protoc* 2013;**1064**:115–30.

73. Guirimand T, Delmotte S, Navratil V. VirHostNet 2.0: surfing on the web of virus/host molecular interactions data. *Nucleic Acids Res* 2015;**43**:D583–7.

74. Cook HV, Doncheva NT, Szklarczyk D, *et al*. Viruses. STRING: a virus-host protein-protein interaction database. *Viruses* 2018;**10**:519.

75. Durmuş Tekir S, Çakir T, Ardiç E, *et al*. PHISTO: pathogen-host interaction search tool. *Bioinformatics* 2013;**29**:1357–8.

76. Calderone A, Licata L, Cesareni G. Virus Mentha: a new resource for virus-host protein interactions. *Nucleic Acids Res* 2015;**43**:D588–92.

77. Ammari MG, Gresham CR, McCarthy FM, *et al*. HPIDB 2.0: a curated database for host-pathogen interactions. *Database (Oxford)* 2016;**2016**:baw103.

78. Rappaport N, Twik M, Plaschkes I, *et al*. MalaCards: an amalgamated human disease compendium with diverse clinical and genetic annotation and structured search. *Nucleic Acids Res* 2017;**45**:D877–D887.

79. Piñero J, Queralt-Rosinach N, Bravo À, *et al*. DisGeNET: a discovery platform for the dynamical exploration of human diseases and their genes. *Database* 2015;**2015**:bav028.

80. Edgar R, Domrachev M, Lash AE. Gene expression omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res* 2002;**30**:207–10.

81. Doncheva N, Morris J, Gorodkin J, *et al*. Cytoscape stringApp: network analysis and visualization of proteomics data. *Cytoscape StringApp Netw Anal Vis Proteomics Data* 2018;**18**:623–32.

82. Pletscher-Frankild S, Pallejà A, Tsafou K, *et al*. DISEASES: text mining and data integration of disease-gene associations. *Methods* 2015;**74**:83–9.

83. del-Toro N, Dumousseau M, Orchard S, *et al*. A new reference implementation of the PSICQUIC web service. *Nucleic Acids Res* 2013;**41**:W601W601–6.

84. Bastian M, Heymann S, Jacomy M. Gephi: an open source software for exploring and manipulating networks. In: *Third International AAAI Conference on Weblogs and Social Media*, ICWSM 2009, p. 17–20. AAAI Publications, San Jose, California, USA.

85. Shannon P, Markiel A, Ozier O, *et al*. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* 2003;**13**:2498–504.

86. Csardi G, Nepusz T. The igraph software package for complex network research. *InterJ Complex Syst* 2006;**1695**:1–9.

87. Hagberg AA, Schult DA, Swart PJ. Exploring network structure, dynamics, and function using NetworkX. In: *Proceedings of the 7th Python in Science Conference*, (SciPy 2008), p.11–15. Pasadena, CA USA.

88. Assenov Y, Ramírez F, Schelhorn SESE, *et al*. Computing topological parameters of biological networks. *Bioinformatics* 2008;**24**:282–4.

89. Chin CH, Chen SH, Wu HH, *et al*. cytoHubba: identifying hub objects and sub-networks from complex interactome. *BMC Syst Biol* 2014;**8**:S11.

90. De Chassey B, Navratil V, Tafforeau L, *et al*. Hepatitis C virus infection protein network. *Mol Syst Biol* 2008;**4**:230.

91. Tekir SD, Çakir T, Ülgen KÖ. Infection strategies of bacterial and viral pathogens through pathogen-human protein-protein interactions. *Front Microbiol* 2012;**3**:46.

92. Pichlmair A, Kandasamy K, Alvisi G, *et al*. Viral immune modulators perturb the human molecular network by common and unique strategies. *Nature* 2012;**487**:486–90.

93. Zhao Z, Xia J, Tastan O, *et al*. Virus interactions with human signal transduction pathways. *Int J Comput Biol Drug Des* 2011;**4**:83–105.

94. Zhou H, Jin J, Wong L. Progress in computational studies of host-pathogen interactions. *J Bioinforma Comput Biol* 2012;**11**:1230001.

95. Davis FP, Barkan DT, Eswar N, *et al*. Host-pathogen protein interactions predicted by comparative modeling. *Protein Sci* 2007;**16**:2585–96.

96. Bindea G, Mlecnik B, Hackl H, *et al*. ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. *Bioinformatics* 2009;**25**:1091–3.

97. Maere S, Heymans K, Kuiper M. BiNGO: a Cytoscape plugin to assess overrepresentation of Gene Ontology categories in biological networks. *Bioinformatics* 2005;**21**:3448–9.

98. Saito R, Smoot ME, Ono K, *et al*. A travel guide to Cytoscape plugins. *Nat Methods* 2012;**9**:1069–76.

99. Kanehisa MKEGG. Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res* 2000;**28**:27–30.

100. Jassal B, Matthews L, Viteri G, *et al*. The reactome pathway knowledgebase. *Nucleic Acids Res* 2020;**48**(D1):D498–D503.

101. Slenter DN, Kutmon M, Hanspers K, *et al*. WikiPathways: a multifaceted pathway database bridging metabolomics to other omics research. *Nucleic Acids Res* 2018;**46**: D661–7.

102. Xie C, Mao X, Huang J, *et al*. KOBAS 2.0: a web server for annotation and identification of enriched pathways and diseases. *Nucleic Acids Res* 2011;**39**:W316–22.

103. Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* 2009;**4**:44–57.

104. Huang DW, Sherman BT, Lempicki RA. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res* 2009;**37**:1–13.

105. Jawaid W. enrich R: provides an R Interface to 'Enrichr'. *R Package version* version 3.0. 2020, CRAN, (https://cran.r-project.org/web/packages/enrichR/index.html) 2019.

106. Kuleshov MV, Jones MR, Rouillard AD, *et al*. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res* 2016;**44**:W90–7.

107. Chen EY, Tan CM, Kou Y, *et al*. Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinformatics* 2013;**14**:128.

108. Minadakis G, Zachariou M, Oulas A, *et al*. PathwayConnector: finding complementary pathways to enhance functional analysis. *Bioinformatics* 2019;**35**:889–91.

109. Bader GD, Hogue CW An automated method for finding molecular complexes in large protein interaction networks. *BMC Bioinformatics*. 2003;**4**:2.

110. Su G, Kuchinsky A, Morris JH, *et al*. GLay: community structure analysis of biological networks. *Bioinformatics* 2010;**26**:3135–7.

111. Morris JH, Apeltsin L, Newman AM, *et al*. ClusterMaker: a multi-algorithm clustering plugin for Cytoscape. *BMC Bioinformatics* 2011;**12**:436.

112. Li M, Li D, Tang Y, *et al*. Cytocluster: a cytoscape plugin for cluster analysis and visualization of biological networks. *Int J Mol Sci* 2017;**18**:1880.

113. Wang J, Zhong J, Chen G, *et al*. ClusterViz: a Cytoscape APP for cluster analysis of biological network. *IEEE/ACM Trans Comput Biol Bioinform* 2015;**12**:815–22.

114. Costa Sa AC, Madsen H, Brown JR. Shared molecular signatures across neurodegenerative diseases and herpes virus infections highlights potential mechanisms for maladaptive innate immune responses. *Sci Rep* 2019;**9**: 8795.

115. Dugger BN, Dickson DW. Pathology of neurodegenerative diseases. *Cold Spring Harb Perspect Biol* 2017;**9**:a028035.

116. Przedborski S, Vila M, Jackson-Lewis V. Neurodegeneration: what is it and where are we? *J Clin Invest* 2003;**111**:3–10.

117. Ross CA, Poirier MA. Protein aggregation and neurodegenerative disease. *Nat Med* 2004;**10**:S10–7.

118. Friedlander RM. Apoptosis and caspases in neurodegenerative diseases. *N Engl J Med*. 2003;**348**:1365–75.

119. Golpich M, Amini E, Mohamed Z, *et al*. Mitochondrial dysfunction and biogenesis in neurodegenerative diseases: pathogenesis and treatment. *CNS Neurosci Ther* 2017;**23**:5–22.

120. Glass CK, Saijo K, Winner B, *et al*. Mechanisms underlying inflammation in neurodegeneration. *Cell* 2010;**140**:918–34.

121. Islam MT. Oxidative stress and mitochondrial dysfunction-linked neurodegenerative disorders. *Neurol Res* 2017; **39**:73–82.

122. Liu Z-P, Wang Y, Zhang X-S, *et al*. Identifying dysfunctional crosstalk of pathways in various regions of Alzheimer's disease brains. *BMC Syst Biol* 2010;**4**:S11.

123. Cervantes-Gracia K, Husi H. Integrative analysis of multiple sclerosis using a systems biology approach. *Sci Rep* 2018;**8**:5633.

124. Kalathur RKR, Pedro Pinto J, Sahoo B, *et al*. HDNetDB: a molecular interaction database for network-oriented investigations into Huntington's disease. *Sci Rep* 2017;**7**:5216.

125. Li P, Nie Y, Yu J. An effective method to identify shared pathways and common factors among neurodegenerative diseases. *PLoS One* 2015;**10**:e0143045.

126. Limviphuvadh V, Tanaka S, Goto S, *et al*. The commonality of protein interaction networks determined in neurodegenerative disorders (NDDs). *Bioinformatics* 2007;**23**: 2129–38.

127. Nguyen TP, Caberlotto L, Morine MJ, *et al*. Network analysis of neurodegenerative disease highlights a role of toll-like receptor signaling. *Biomed Res Int* 2014;**2014**:686505.

128. Goñi J, Esteban FJ, de Mendizábal NV, *et al*. A computational analysis of protein-protein interaction networks in neurodegenerative diseases. *BMC Syst Biol* 2008;**2**:52.

129. Vasaikar SV, Padhi AK, Jayaram B, *et al*. NeuroDNet - an open source platform for constructing and analyzing neurode-generative disease networks. *BMC Neurosci* 2013;**14**:3.

130. Hemmer B, Kerschensteiner M, Korn T. Role of the innate and adaptive immune responses in the course of multiple sclerosis. *Lancet Neurol* 2015;**14**:406–19.

131. Hernández-Pedro NY, Espinosa-Ramirez G, De La Cruz VP, *et al*. Initial immunopathogenesis of multiple sclerosis: innate immune response. *Clin Dev Immunol* 2013;**2013**:413465.

132. Sen MK, Almuslehi MSM, Shortland PJ, *et al*. Revisiting the pathoetiology of multiple sclerosis: has the tail been wagging the mouse? *Front Immunol* 2020;**11**:572186.

133. Stys PK. Pathoetiology of multiple sclerosis: are we barking up the wrong tree? *F1000Prime Rep* 2013;**5**:20.

134. Stys PK, Zamponi GW, Van Minnen J, *et al*. Will the real multiple sclerosis please stand up? *Nat Rev Neurosci* 2012;**13**:507–14.

135. Cesur MF, Durmuş S. Systems biology modeling to study pathogen–host interactions. *Methods Mol Biol* 2018;**1734**:97–112.

136. Goh KI, Cusick ME, Valle D, *et al*. The human disease network. *Proc Natl Acad Sci U S A* 2007;**104**:8685–90.

137. Oksanen J, Blanchet FG, Friendly M, *et al*. Package 'vegan'. A community ecology package. *Community Ecol Package CRAN*, https://CRAN.R-project.org/package=vegan 2019.

138. Kassambara A, Mundt. factoextra: Extract and Visualize the Results of Multivariate Data Analyses. Package Version 1.0.7. R Packag. version 2020;CRAN, https://CRAN.R-project.org/package=factoextra 2020.

139. Oliveros JC. Venny. An interactive tool for comparing lists with Venn diagrams. http://bioinfogp.cnb.csic.es/tools/venny/index.html 2007–2015.

140. Hussein HM, Rahal EA. The role of viral infections in the development of autoimmune diseases. *Crit Rev Microbiol* 2019;**45**:394–412.

141. Sfriso P, Ghirardello A, Botsios C, *et al*. Infections and autoimmunity: the multifaceted relationship. *J Leukoc Biol* 2010;**87**:385–95.

142. Draborg AH, Duus K, Houen G. Epstein–Barr virus and systemic lupus erythematosus. *Clin Dev Immunol* 2012;**2012**:370516.

143. Toussirot É, Roudier J. Epstein-Barr virus in autoimmune diseases. *Best Pract Res Clin Rheumatol* 2008;**22**:883–96.

144. Lossius A, Johansen JN, Torkildsen Ø, *et al*. Epstein-Barr virus in systemic lupus erythematosus, rheumatoid arthritis and multiple sclerosis-association and causation. *Viruses* 2012;**4**:3701–30.

145. Draborg AH, Duus K, Houen G. Epstein–Barr virus in systemic autoimmune diseases. *Clin Dev Immunol* 2013;**2013**:535738.

146. Nelson P, Rylance P, Roden D, *et al*. Viruses as potential pathogenic agents in systemic lupus erythematosus. *Lupus* 2014;**23**:596–605.

147. Akimoto M, Matsushita K, Suruga Y, *et al*. Clinical manifestations of human T lymphotropic virus type I-infected patients with systemic lupus erythematosus. *J Rheumatol* 2007;**34**:1841–8.

148. Rigante D, Esposito S. Infections and systemic lupus erythematosus: binding or sparring partners? *Int J Mol Sci* 2015;**16**17331–43.

149. Pan Q, Liu Z, Liao S, *et al*. Current mechanistic insights into the role of infection in systemic lupus erythematosus. *Biomed Pharmacother* 2019;**117**:109122.

150. Harley JB, Chen X, Pujato M, *et al*. Transcription factors operate across disease loci, with EBNA2 implicated in autoimmunity. *Nat Genet* 2018;**50**699–707.

151. Franssila R, Hedman K. Viral causes of arthritis. *Best Pract Res Clin Rheumatol* 2006;**20**1139–57.

152. Filippi CM, Von Herrath MG. Viral trigger for type 1 diabetes: pros and cons. *Diabetes* 2008;**57**2863–71.

153. Levet S, Medina J, Joanou J, *et al*. An ancestral retroviral protein identified as a therapeutic target in type-1 diabetes. *JCI Insight* 2017;**2**:e94387.

154. van der Werf N, Kroese FGM, Rozing J, *et al*. Viral infections as potential triggers of type 1 diabetes. *Diabetes Metab Res Rev* 2007;**23**169–83.

155. Antonelli A. Hepatitis C virus infection and type 1 and type 2 diabetes mellitus. World. *J Diabetes* 2014;**5**586–600.

156. Jain A, Tuteja. Tissue enrich: tissue-specific gene enrichment analysis. *Bioinformatics* 2019;**35**1966–7.

157. Ardlie KG, DeLuca DS, Segrè AV, *et al*. The genotype-tissue expression (GTEx) pilot analysis: multitissue gene regulation in humans. *Science (80-)* 2015;**348**648–60.

158. Libbey JE, Fujinami RS. Adaptive immune response to viral infections in the central nervous system. *Handb Clin Neurol* 2014;**123**225–47.

159. Arango M-T, Shoenfeld Y, Cervera R, *et al*. Chapter 19. Infection and autoimmune diseases. Anaya JM, Shoenfeld Y, Rojas-Villarraga A, et al., editors. Bogota (Colombia): El Rosario University Press, *Autoimmunity From Bench to Bedside* 2013:303–20.

160. Vita R, Overton JA, Greenbaum JA, *et al*. The immune epitope database (IEDB) 3.0. *Nucleic Acids Res* 2015;**43**D405–12.

161. Levinthal DJ, Rahman A, Nusrat S, *et al*. Adding to the burden: gastrointestinal symptoms and syndromes in multiple sclerosis. *Mult Scler Int* 2013;**2013**:319201.

162. Harbo HF, Gold R, Tintora M. Sex and gender issues in multiple sclerosis. *Ther Adv Neurol Disord* 2013;**66**237–48.

163. Doosti R, Togha M, Moghadasi AN, *et al*. Evaluation of the risk of cervical cancer in patients with multiple sclerosis treated with cytotoxic agents: a cohort study. *Iran J Neurol* 2018;**17**64–70.

164. Marrie RA, Reider N, Cohen J, *et al*. A systematic review of the incidence and prevalence of cancer in multiple sclerosis. *Mult Scler J* 2015;**21**294–304.

165. Warren SA, Turpin KVL, Pohar SL, *et al*. Comorbidity and health-related quality of life in people with multiple sclerosis. *Int J MS Care* 2009;**11**:6–16.

166. Pelizzari L, Jakimovski D, Laganà MM, *et al*. Five-year longitudinal study of neck vessel cross-sectional area in multiple sclerosis. *Am J Neuroradiol* 2018;**39**:1703–9.

167. Jakimovski D, Zivadinov R, Pelizzari L, *et al*. Lipoprotein(a) levels are associated with the size of extracranial arteries in multiple sclerosis. *J Vasc Res* 2020;**57**:16–23.

168. Zecca C, Riccitelli GC, Disanto G, *et al*. Urinary incontinence in multiple sclerosis: prevalence, severity and impact on patients' quality of life. *Eur J Neurol* 2016;**23**: 1228–34.

169. Lu R, Kiernan MC, Murray A, *et al*. Kidney-brain crosstalk in the acute and chronic setting. *Nat Rev Nephrol* 2015;**11**:707–19.

170. Sintzel MB, Rametta M, Reder AT. Vitamin D and multiple sclerosis: a comprehensive review. *Neurol Ther* 2018;**7**:59–85.

171. Rostami S, Emami-Aleagha MS, Ghasemi-Kasman M, *et al*. Cross-talks between the kidneys and the central nervous system in multiple sclerosis. *Casp J Intern Med* 2018;**9**:206–10.

172. McFadden G, Mohamed MR, Rahman MM, *et al.* Cytokine determinants of viral tropism. *Nat Rev Immunol* 2009;**9**:645–55.

173. Atluri VSR, Hidalgo M, Samikkannu T, *et al.* Synaptic plasticity and neurological disorders in neurotropic viral infections. *Neural Plast* 2015;**2015**:138979.

174. Hotta H. Neurotropic viruses–classification, structure and characteristics. *Nippon rinsho Japanese J Clin Med* 1997;**55**:777–82.

175. Gessain A, Mahieux R. Tropical spastic paraparesis and HTLV-1 associated myelopathy: clinical, epidemiological, virological and therapeutic aspects. *Rev Neurol (Paris)* 2012;**168**:257–69.

176. Tsuru T, Mizuguchi M, Ohkubo Y, *et al.* Acute disseminated encephalomyelitis after live rubella vaccination. *Brain and Development* 2000;**22**:259–61.

177. Sanders VJ, Felisan S, Waddell A, *et al.* Detection of Herpesviridae in postmortem multiple sclerosis brain tissue and controls by polymerase chain reaction. *J Neurovirol* 1996;**2**:249–58.

178. Fingeroth JD. Herpesvirus infection of the liver. *Infect Dis Clin N Am* 2000;**14**:689–719.

179. Whitley RJ. Herpesviruses. In: Baron S (ed). *Medical Microbiology*, 4th edn. Galveston (TX): University of Texas Medical Branch at Galveston.

180. Dinh A, Fleuret V, Hanslik T. Liver involvement in adults with measles. *Int J Infect Dis* 2013;**17**:e1243–4.

181. Chassort A, Coutherut J, Moreau-Klein A, *et al.* Renal dysfunction in adults during measles. *Med Mal Infect* 2015;**45**:165–8.

182. Figueiredo CA, Cordovani NTB, Castrignano SB, *et al.* Acute liver failure associated with rubella virus in a child. *Pediatr Infect Dis J* 2010;**29**:573–4.

183. Liaw FY, Huang CF, Hsueh JT, *et al.* Eczema herpeticum: a medical emergency. *Can Fam Physician* 2012;**58**:1358–61.

184. Dantas L, Netto E, Glesby MJ, *et al.* Dermatological manifestations of individuals infected with human T cell lymphotropic virus type I (HTLV-I). *Int J Dermatol* 2014;**53**:1098–102.

185. Rota PA, Moss WJ, Takeda M, *et al.* Measles. *Nat Rev Dis Primers* 2016;**2**:16049.

186. Rayala BZ, Morrell DS. Common skin conditions in children: skin infections. *FP Essent* 2017;**453**:26–32.

187. Becker JA, Smith JA. Return to play after infectious mononucleosis. *Sports Health* 2014.

188. Pinto DO, DeMarino C, Pleet ML, *et al.* HTLV-1 extracellular vesicles promote cell-to-cell contact. *Front Microbiol* 2019;**6**:232–8.

189. Laksono BM, de Vries RD, McQuaid S, *et al.* Measles virus host invasion and pathogenesis. *Viruses* 2016;**88**:210.

190. Bansal A, Biswas D. Role of herpesviruses in coronary artery disease. *Heart India* 2018;**6**:75–80.

191. Ibrahim AI, Obeid MT, Jouma MJ, *et al.* Detection of herpes simplex virus, cytomegalovirus and Epstein-Barr virus DNA in atherosclerotic plaques and in unaffected bypass grafts. *J Clin Virol* 2005;**32**:29–32.

192. Farid Hosseni R, Jabbari F, Shabestari M, *et al.* Human T lymphotropic virus type I (HTLV-I) is a risk factor for coronary artery disease. *Iran J Basic Med Sci* 2013;**16**:217–20.

193. Abolbashari S, Ghayour-Mobarhan M, Ebrahimi M, *et al.* The role of human T-lymphotropic virus (HTLV) in cardiovascular diseases: a review of literature. *ARYA Atheroscler* 2018;**14**:183–7.

194. Reeves MB, Coleman H, Chadderton J, *et al.* Vascular endothelial and smooth muscle cells are unlikely to be major sites of latency of human cytomegalovirus in vivo. *J Gen Virol* 2004;**85**:3337–41.

195. Matsuura E, Yoshimura A, Nozuma S, *et al.* Clinical presentation of axial myopathy in two siblings with HTLV-1 associated myelopathy/tropical spastic paraparesis (HAM/TSP). *BMC Neurol* 2015;**15**:18.

196. Chen M, Popescu N, Woodworth C, *et al.* Human herpesvirus 6 infects cervical epithelial cells and transactivates human papillomavirus gene expression. *J Virol* 1994;**68**:1173–8.

197. Koffa M, Koumantakis E, Ergazaki M, *et al.* Association of herpesvirus infection with the development of genital cancer. *Int J Cancer* 1995;**63**:58–62.

198. Vranic S, Cyprian FS, Akhtar S, *et al.* The role of Epstein-Barr virus in cervical cancer: a brief update. *Front Oncol* 2018;**8**:113.

199. Guidry JT, Scott RS. The interaction between human papillomavirus and other viruses. *Virus Res* 2017;**231**:139–47.

200. Shi Y, Peng SL, Yang LF, *et al.* Co-infection of Epstein-Barr virus and human papillomavirus in human tumorigenesis. *Chin J Cancer* 2016;**35**:16.

201. Blas MM, Alva IE, Garcia PJ, *et al.* Association between human papillomavirus and human T-lymphotropic virus in indigenous women from the Peruvian Amazon. *PLoS One* 2012;**7**:e44240.

202. Lôpo SS, Oliveira PM, Santana IU, *et al.* Evidence of a higher prevalence of HPV infection in HTLV-1-infected women: a cross-sectional study. *Rev Soc Bras Med Trop* 2012;**45**:305–8.

203. Heimann A, Scanlon R, Gentile J, *et al.* Measles cervicitis: report of a case with cytologic and molecular biologic analysis. *Acta Cytol* 1992;**36**:727–30.

204. Ghamales IA, Napolitano PG, Rosa C. Vaginal mucositis in measles. *Infect Dis Obstet Gynecol* 1995;**2**:279–81.

205. Seppälä M, Vaheri A. Natural rubella infection of the female genital tract. *Lancet* 1974;**1**:46–7.

206. Nourshargh S, Alon R. Leukocyte migration into inflamed tissues. *Immunity* 2014;**41**:694–707.

207. Dong C. TH17 cells in development: an updated view of their molecular identity and genetic programming. *Nat Rev Immunol* 2008;**8**:337–48.

208. Jadidi-Niaragh F, Mirshafiey A. Th17 cell, the new player of neuroinflammatory process in multiple sclerosis. *Scand J Immunol* 2011;**74**:1–13.

209. Weiner HL. A shift from adaptive to innate immunity: a potential mechanism of disease progression in multiple sclerosis. *J Neurol* 2008;**255**:3–11.

210. Cheng W, Chen. Chemokines and chemokine receptors in multiple sclerosis. *Mediat Inflamm* 2014;**2014**:659206.

211. Yang R, Dunn JF. Multiple sclerosis disease progression: contributions from a hypoxia–inflammation cycle. *Mult Scler J* 2018;**25**:1715–1718.

212. dos Santos SA, de Andrade Júnior DR. HIF–1alpha and infectious diseases: a new frontier for the development of new therapies. *Rev Inst Med Trop Sao Paulo* 2017;**59**:e92.

213. Patiar S, Harris AL. Role of hypoxia-inducible factor-1α as a cancer therapy target. *Endocr Relat Cancer* 2006;**13**:1715–75.

214. Cuninghame S, Jackson R, Zehbe I. Hypoxia-inducible factor 1 and its role in viral carcinogenesis. *Virology* 2014;**456-457**:370–83.

215. Melamed E, Lee. Multiple sclerosis and cancer: the Ying-Yang effect of disease modifying therapies. *Front Immunol* 2020;**10**:2954.

216. Chen CH, Lin CL, Kao CH. Irritable bowel syndrome is associated with an increased risk of dementia: a nationwide population-based study. *PLoS One* 2016;**11**:e0144589.

217. Lai SW, Liao KF, Lin CL, *et al*. Irritable bowel syndrome correlates with increased risk of Parkinson's disease in Taiwan. *Eur J Epidemiol* 2014;**29**:57–62.

218. Wijnands JMA, Zhu F, Kingwell E, *et al*. Five years before multiple sclerosis onset: phenotyping the prodrome. *Mult Scler J* 2019;**25**:1092–101.

219. Rowin J, Xia Y, Jung B, *et al*. Gut inflammation and dysbiosis in human motor neuron disease. *Phys Rep* 2017;**5**: e13443.

220. Mulak A, Bonaz B. Brain-gut-microbiota axis in Parkinson's disease. *World J Gastroenterol* 2015;**21**:10609–20.

221. Sinagra E, Utzeri E, Morreale GC, *et al*. Microbiota-gut-brain axis and its affect inflammatory bowel disease: pathophysiological concepts and insights for clinicians. *World J Clin Cases* 2020;**8**:1013–25.

222. Dinan TG, Cryan JF. The microbiome-gut-brain Axis in health and disease. *Gastroenterol Clin N Am* 2017;**46**: 77–89.

223. Roy Sarkar S, Banerjee S. Gut microbiota in neurodegenerative disorders. *J Neuroimmunol* 2019;**328**:98–104.

224. Cryan JF, Dinan TG. Mind-altering microorganisms: the impact of the gut microbiota on brain and behaviour. *Nat Rev Neurosci* 2012;**13**:701–12.

225. Li N, Ma WT, Pang M, *et al*. The commensal microbiota and viral infection: a comprehensive review. *Front Immunol* 2019;**10**:1551.

226. Robinson CM, Pfeiffer JK. Viruses and the Microbiota. *Annu. Rev. Virol.* 2014;**1**:55–6969.

227. Bhatia D, Ardekani MS, Shi Q, *et al*. Hypoxia and its emerging therapeutics in neurodegenerative, inflammatory and renal diseases. *Hypoxia Hum Dis* 2017.

228. Becker J, Winthrop KL. Update on rheumatic manifestations of infectious diseases. *Curr Opin Rheumatol* 2010;**22**:72–7.

229. Arleevskaya MI, Kravtsova OA, Lemerle J, *et al*. How rheumatoid arthritis can result from provocation of the immune system by microorganisms and viruses. *Front Microbiol* 2016;**7**:1296.

230. Yinshi Yue. Microbial infection and rheumatoid arthritis. *J Clin Cell Immunol* 2013;**4**:174.

231. Vigan S, Perreau M, Pantaleo G, *et al*. Positive and negative regulation of cellular immune responses in physiologic conditions and diseases. *Clin Dev Immunol* 2012;**2012**:485781.

232. Han L, Yang J, Wang X, *et al*. Th17 cells in autoimmune diseases. *Front Med* 2015;**9**:10–9.

233. Ramgolam S, Markovic-Plese S. Interferon-beta inhibits Th17 cell differentiation in patients with multiple sclerosis. *Endocr Metab Immune Disord Drug Targets* 2012;**10**:161–7.

234. Zhang X, Markovic-Plese S. Interferon beta inhibits the Th17 cell-mediated autoimmune response in patients with relapsing-remitting multiple sclerosis. *Clin Neurol Neurosurg* 2010;**112**:641–5.

235. Durelli L, Conti L, Clerico M, *et al*. T-helper 17 cells expand in multiple sclerosis and are inhibited by interferon-β. *Ann Neurol* 2009;**65**:499–509.

236. Nourani E, Khunjush F, Durmus S. Computational approaches for prediction of pathogen-host protein-protein interactions. *Front Microbiol* 2015;**6**:94.

237. Amberger J, Bocchini CA, Scott AF, *et al*. McKusick's online Mendelian inheritance in man (OMIM®). *Nucleic Acids Res* 2009;**37**:D793–6.

238. Landrum MJ, Chitipiralla S, Brown GR, *et al*. Clin Var: improvements to accessing data. *Nucleic Acids Res* 2020;**48** (D1) :D835–44.