Computational and Structural Biotechnology Journal 19 (2021) 5292–5308

COMPUTATIONAL AND STRUCTURAL BIOTECHNOLOGY JOURNAL

journal homepage: www.elsevier.com/locate/csbj

# An Ebola, *Neisseria and Trypanosoma* human protein interaction census reveals a conserved human protein cluster targeted by various human pathogens

Shishir K Gupta [a,b,1], Alicia Ponte-Sucre [c,d,1,2], Elena Bencurova [a], Thomas Dandekar [a,e,2,*]

[a] *Functional Genomics & Systems Biology Group, Department of Bioinformatics, Biocenter, Am Hubland, University of Würzburg, 97074 Würzburg, Germany*
[b] *Evolutionary Genomics Group, Center for Computational and Theoretical Biology, University of Würzburg, 97078 Würzburg, Germany*
[c] *Laboratorio de Fisiología Molecular, Instituto de Medicina Experimental, Escuela Luis Razetti, Universidad Central de Venezuela, Caracas, Venezuela*
[d] *Medical Mission Institute, Hermann-Schell-Str. 7, 97074 Würzburg, Germany*
[e] *EMBL Heidelberg, BioComputing Unit, Meyerhofstraße 1, 69117 Heidelberg, Germany*

ARTICLE INFO

*Article history:*
Received 19 March 2021
Received in revised form 14 September 2021
Accepted 15 September 2021
Available online 16 September 2021

*Keywords:*
Blood-borne pathogen
Interactome
Protein–protein interaction
Life cycle
Hub proteins
Antibiotic targeting
Pathway analysis
Ebola virus
*Neisseria meningitidis*
Human African Trypanosomiasis

ABSTRACT

*Filovirus ebolavirus* (ZE; *Zaire ebolavirus*, Bundibugyo ebolavirus), *Neisseria meningitidis* (NM), and *Trypanosoma brucei* (Tb) are serious infectious pathogens, spanning viruses, bacteria and protists and all may target the blood and central nervous system during their life cycle. NM and Tb are extracellular pathogens while ZE is obligatory intracellular, targetting immune privileged sites. By using interactomics and comparative evolutionary analysis we studied whether conserved human proteins are targeted by these pathogens. We examined 2797 unique pathogen-targeted human proteins. The information derived from orthology searches of experimentally validated protein–protein interactions (PPIs) resulted both in unique and shared PPIs for each pathogen. Comparing and analyzing conserved and pathogen-specific infection pathways for NM, TB and ZE, we identified human proteins predicted to be targeted in at least two of the compared host-pathogen networks. However, four proteins were common to all three host-pathogen interactomes: the elongation factor 1-alpha 1 (EEF1A1), the SWI/SNF complex subunit SMARCC2 (matrix-associated actin-dependent regulator of chromatin subfamily C), the dolichyl-dipho sphooligosaccharide--protein glycosyltransferase subunit 1 (RPN1), and the tubulin beta-5 chain (TUBB). These four human proteins all are also involved in cytoskeleton and its regulation and are often addressed by various human pathogens. Specifically, we found (i) 56 human pathogenic bacteria and viruses that target these four proteins, (ii) the well researched new pandemic pathogen SARS-CoV-2 targets two of these four human proteins and (iii) nine human pathogenic fungi (yet another evolutionary distant organism group) target three of the conserved proteins by 130 high confidence interactions.



## 1. Introduction

About $10^{13}$ human cells coexist with $10^{14}$ bacterial, fungal, and protozoan cells, i.e., thousands of microbial species. This means that we live in "close intimacy" with a wide variety of microbes. Some of them are capable of causing us illness or death once the "eco-system" is broken and an immunocompromised situation arises, as is the case for *Neisseria meningitidis* [1]. For the aims of this manuscript we will name it as an opportunistic pathogen. On the other hand, dedicated pathogens do not require the host to be immunocompromised or injured before invading it. Their specialized mechanisms for crossing cellular and biochemical barriers and for eliciting specific responses from the host organism contribute to the survival and multiplication of the pathogen, as is the case of Ebola virus and *Trypanosoma* sp. The host responses may even contribute to pathogen survival and multiplication, as happens for Ebola virus and *Trypanosoma* sp. [1–5].

*Abbreviations:* PPI, protein–protein interaction; PPIs, protein–protein interactions; DIP, database of interacting proteins; HLA, human leucocyte antigen; **ZE**, *for* Zaire ebolavirus, *Bundibugyo ebolavirus*; **NM**, *Neisseria meningitidis*; Tb, for *Trypanosoma brucei;*.

* Corresponding author at: Functional Genomics & Systems Biology Group, Department of Bioinformatics, Biocenter, Am Hubland, University of Würzburg, 97074 Würzburg, Germany.
*E-mail address:* dandekar@biozentrum.uni-wuerzburg.de (T. Dandekar).

[1] Equal contributing first authors.
[2] Equal contributing senior authors.

https://doi.org/10.1016/j.csbj.2021.09.017

Furthermore, (i) viruses, including Ebola virus, only replicate inside cells, thus being intracellular pathogens; (ii) the Gram-negative bacterium *N. meningitidis* is an asymptomatic colonizer (harmless commensal) of the human nasopharynx, mostly using propionic acid available outside of cells in that environment. However, to cause disease it must traverse the blood-cerebrospinal fluid (B-CSF) barrier. This has lead researchers to postulate transcellular passage (and, implicitly, cell entry); (iii) the parasitic protist *T. brucei* is not intracellular in either of its two hosts, tse-tse flies (insect stage) or man/cattle (bloodstream stage); however, the constantly changing specific factor variant surface glycoprotein (VSGs) of *T. brucei* constitutes a complicating issue. How trypanosomes cross the brain parenchyma and persist behind an intact blood–brain barrier (BBB), how their growth in the brain is controlled and how the unicellular trypanosomes exit the brain and cause a relapse are still questions to be answered. Additional details can be found in Supplementary file 1.

Cohabitation between human populations at risk and their animals - as pathogen reservoirs - is common, and their everyday contact may even permit disease transmission [6]. Thus, ecological and evolutionary hints, together with disease dynamics, are fundamental for understanding and analysing the temporal, organizational and spatial interactions, even within a single pathogen. That is, hours to months, cellular to ecosystem levels, and local to pandemic spread of diseases [6]. Additionally, some pathogens circulate between individuals of a single species, while others flow among multiple hosts, need insect vectors, or subsist in reservoirs. Hence, we compare in this first part of our study *Filovirus ebolavirus* (ZE for *Zaire ebolavirus*, Bundibugyo ebolavirus), *Neisseria meningitidis* (NM for *N. meningitidis*) and *Trypanosoma brucei* (Tb for *T. gambiense* and *rhodesiense*, *T. brucei sp.*).

It is fundamental and may be attractive to elucidate defense strategies common among hosts when confronted with these three diverse pathogen types - whether opportunistic or professional - expressing a miriad of host-pathogen interactions. Interactions and strategies may differ, and yet investigating those commonly utilized would help further define how the immune response reacts to pathogens. However, we realized during the course of this study that the tools we have at our disposal are only *in silico* and to uncover immune reactions to pathogens would require a large-scale study with many organisms and many experiments testing human immune responses. Instead, to validate our findings from these three very different organisms, we can with our *in silico* approach only systematically determine how far always the same human proteins are targeted by a wide variety of pathogens while any functional conclusions need experimental data. For both parts of our study we used a host-pathogen protein–protein interactions (HP-PPIs) framework [7]. On a large scale, HP-PPIs systematically provide a global view to understand the basis of infectious diseases while at small-scale it is possible to identify conserved interactions in different host-pathogen pairs. PPIs conservation reflects the different levels of interactions paving the route to infection and host defense response. The aim, design and setting of the study were to seek direct information on the proteins and protein interactions in pathogens targeting human proteins from different kingdoms to reveal potential conserved host interactions as well as specific or only partly shared protein interactions.

However, only a systematic approach is here powerful, hence we rechecked and could then confirm the most central four human proteins commonly targeted in all three pathogens by a systematic census over human pathogenic organisms, comparing bacteria, viridae and fungi. The available host-pathogen protein–protein interaction databases support conserved targeting of these four human proteins by various human pathogens.

## 2. Methods

The flow diagram (Fig. 1) illustrates all the pipeline steps.

### 2.1. Proteomes and amino acid sequences

The reference proteomes of human (reference proteome accession: UP000005640) and *Bundibugyo ebolavirus* (reference proteome accession: UP000143891) were retrieved from Uniprot database [8]. The reference proteomes of *Trypanosoma brucei brucei* TREU927 (genome assembly ASM244v1, RefSeq assembly accession: GCF_000002445.1) and *Neisseria meningitidis* MC58 (genome assembly ASM880v1, RefSeq assembly accession: GCF_000008 805.1) were retrieved from RefSeq database [9].

### 2.2. Inferring protein–protein interactions based on interologs

The experimentally determined host-pathogen PPIs from PHISTO database [10] and PPIs present in the DIP (Database of Interacting Proteins) database [11] were used as a template to reconstruct interologs based PPIs networks [12] of human-ZE, human-Tb and human-NM. The stand-alone InParanoid program version 4.1 was used to assign proteins from the three pathogens to orthology groups. It uses the reciprocal best blast hit strategy to identify the clusters of orthologs and paralogs [13]. Only the seed ortholog pair whose Bootstrap value was 100% was considered to increase prediction confidence.

### 2.3. Pruning PPIs with domain-domain interactions

The amino acid sequences of identified non-redundant DIP-based interologs of host-pathogen PPIs were extracted and domains were assigned to the sequences using Pfam version 29.0 [14]. The information of interacting protein domains was prepared using the three databases Domine [15], DIMA 3.0 [16] and IDDI database [17]. They consider the information of the Pfam families suggested to interact based on structural information present in the protein complexes available in Protein Data Bank (PDB) [18]. This list was used to parse the DIP-based interologs of PPIs using customized Perl scripts, and only those interolog [12,19] interactions consisting of at least one true domain-domain interaction pair as listed in the chosen databases were retained.

### 2.4. Host-pathogen PPI networks

The host-pathogen PPIs predicted from both the PHISTO and DIP-based interactions were merged and redundancies were removed. Cytoscape version 2.8.1 [20] was used to visualize all three host-pathogen PPI networks. The network hubs were determined using the cyto-Hubba plugin [21]. The networks were further curated by extensive manual effort, and the connectors of proteins related to the human immune system were retrieved to generate three immune subnetworks. Immune subnetwoks were further analysed with Cytoscape to identify the central host proteins targeted by all three selected pathogens.

### 2.5. Gene-Ontology (GO) annotation and pathway Over-Representation analysis (ORA)

The protein sequences of pathogen interacting host proteins were extracted from all three established host-pathogen PPIs and resulting proteins were functionally annotated by the BLAST2GO software suite version 2.4.1 [22]. Fisher's exact test was used to identify significantly over-represented GO terms. The pathway over-representation analysis of conserved human proteins that

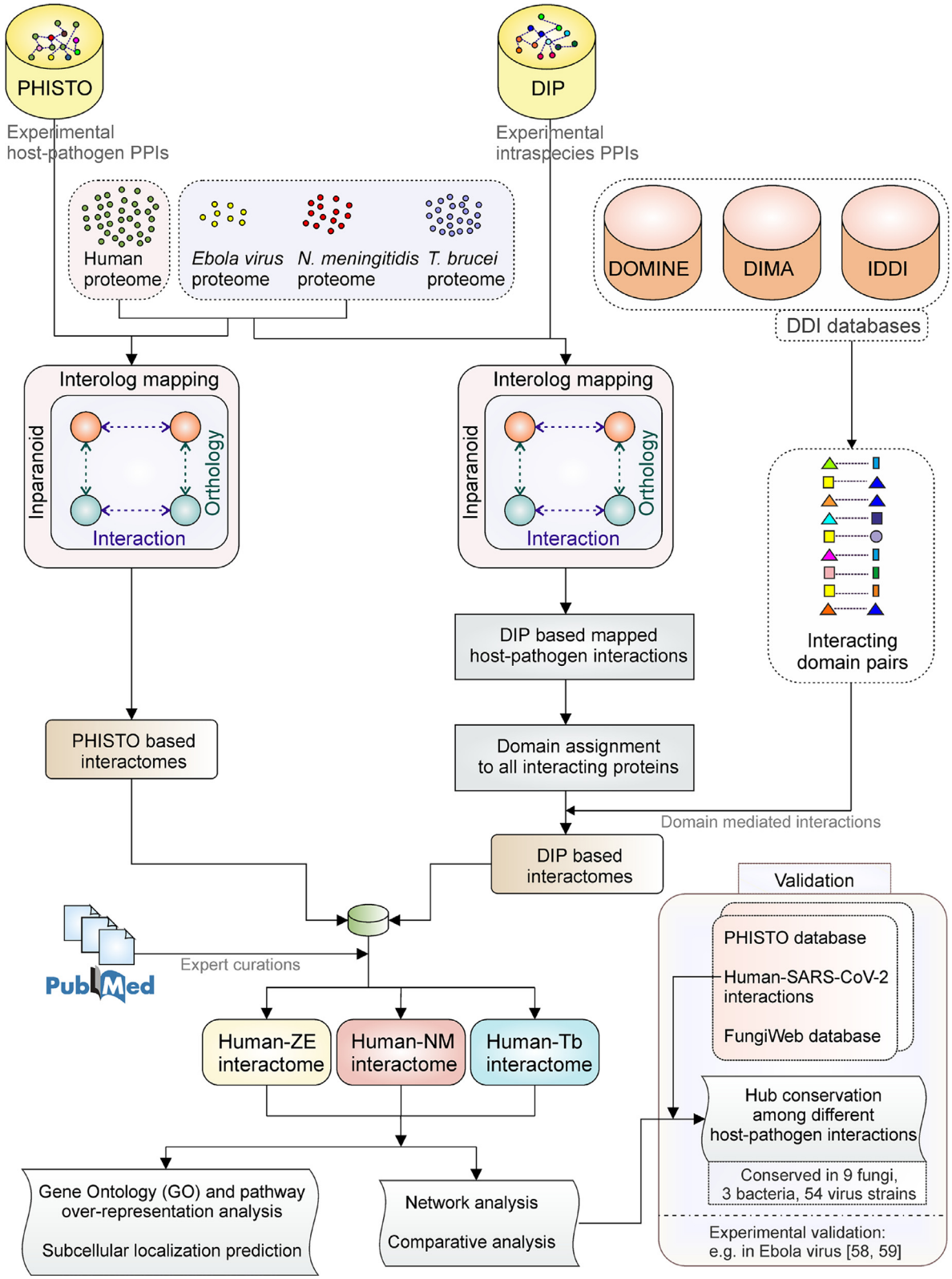



**Fig. 1. Schematic representation of our different analysis steps in a flow diagram.** The pipeline can be used for analysis of interactions and human protein targeting by any pathogen. Here we apply it for: (step 1) the triple comparison *Ebola virus*, *Neisseria meningitidis* and *Trypanosoma brucei* and their protein interactions with human host proteins and (step 2, part at bottom of the diagram) extensive validation of the conservation of interactions comparing different human pathogenic organisms and kingdoms (bacteria, viridae and fungi).

were targeted by at least two pathogens was performed using the pathway analysis tool at InnateDB database [23]. The p-value was calculated using hypergeometric tests, and Benjamini-Hochberg adjustment was used for multiple test correction.

### 2.6. Subcellular localization

The subcellular localization of immune-related human proteins was determined with extended version of KnowPredsite [24] available at UniLoc server (bioapp.iis.sinica.edu.tw/UniLoc/), a knowledge-based classifier for protein subcellular localization [25]. The prediction server KnowPredsite II is capable of predicting single and multiple localization sites for both the eukaryotic and prokaryotic proteins.

### 2.7. Analysis of conservation of host-pathogen protein interactions and protein targets

All host-pathogen protein–protein pairs and interactions from the HPIDB database [26] were examined and conservation of identified interactions as well as conserved targeting of proteins by all three pathogens were accessed. We determined conservation of interactions using the interolog approach [19]. We further assessed the conservation of targeting by further organisms regarding the identified common host proteins. For this validation we collected the interaction data from the PHISTO database [10], human-SARS-CoV-2 interactomes [27,28], and FungiWeb database (https://fungiweb.bioapps.biozentrum.uni-wuerzburg.de/php/fungiweb.php) and analysed the targeting of identified conserved host proteins comparing different pathogens.

## 3. Results

### 3.1. Reconstruction of host-pathogen interaction networks

Although primary interactions between host cells and pathogens are usually diverse, pathogens use several mechanisms to signal their presence or to invade host cells. As these host-pathogen PPIs mainly occur due to biophysical properties of the involved protein molecules, they do not completely disrupt host protein networks. They can, however, effectively modulate host signaling and interactome [29,30]. Of note, critical proteins from the host interactome [31] which enable the pathogen to switch host systems [32] are preferentially targeted. We reconstructed and analyzed HP-PPI (host-pathogen protein–protein interaction) networks comparing the virus *Zaire ebolavirus*, (Bundibugyo ebolavirus, ZE), the prokaryotic bacterium *Neisseria meningitidis* (NM) and the eukaryotic protist (*Trypanosoma brucei gambiense* and *rhodesiense,* (Tb). We mapped physical PPIs between the human host and each pathogen, using orthology relations and host-pathogen interaction databases. Briefly, to reconstruct these HP-PPI networks, we first collected each pathogen and human genome and their corresponding protein-coding genes. Next, we created the respective networks using PHISTO and DIP database as templates of experimentally verified PPIs (see material and methods section). Due to their structural features, exposed protein domains have higher probabilities to be involved in PPIs. These protein domains are often conserved [33] and used by many pathogens to invade and disseminate in the host [34,35]. Therefore, we used a stringent interaction mapping pipeline that refines contact predictions by considering known domain-domain interactions. This approach produces high-quality interactome data with low false-discovery rates [36–39]. The list of all obtained interactions is given in Supplementary file 2.

To further characterize protein functions in these host-pathogen PPI networks we used Gene Ontology (GO) annotation. Fig. 2 displays the number of genes involved in the interactome for each pathogen and its counterpart in humans (panels a-d). Although each HP-PPI network express unique characteristics, the overall profile of gene type involved in the interaction is similar for all three classes of organism analyzed. That is, most host proteins involved in the interaction are expressed either in the plasma membrane, or relate to DNA or RNA functions, or are mitochondrial proteins. Furthermore, with increasing complexity (from viruses to lower eukaryotes) of the invading pathogen, an increasing number of proteins seem to be involved in the host-pathogen interaction for all compartments or functions; of note, most pathogen proteins participating in the interaction are not well characterized in their function. We hence focussed in the second part of our investigation more on the conservation of the targeting of the well characterized human proteins by various pathogens.

Fig. 2g and 2 h compare the number and percentage of human genes related to each category participating in the interactome for each disease. Fig. 2i and 2j illustrate the number and percentage of pathogen genes related to each category involved in the interactome for NM and Tb. The results indicate that the overall profile of protein type contributing to the interaction is similar and does not depend on which organism class the pathogen belongs, thus suggesting a high conservation involved in these host-pathogen interactions.

### 3.2. Unique and shared human proteins in host-pathogen networks

Pathogens tend to interact with host network hubs hijacking them for their own profit [40]. However, the host network should remain ‘‘robust”, to prevent catastrophic failure of host cell metabolism. The reverse would hinder pathogen survival and proliferation and kill the host [41]. We thus analyzed proteins participating in the interactome and represented in the HP-PPI networks. Fig. 3 summarizes numbers for unique and shared targeted human proteins by all three pathogens. Supplementary file 3 lists pathogen targeted human proteins in reconstructed HP-PPI networks. A total of 2797 unique human proteins were targeted by the three analyzed pathogens. Eukaryotic cells herein represented by trypanosomes were found to achieve the highest numbers of specific host interactions (1414); this was followed by *Neisseria* (753) and Ebola (57), respectively. To dissect paths involved in each host-pathogen interaction, we further performed pathway analysis and classification, collected by approaches such as Netpath [42], BioCarta [43], KEGG [44] and Reactome [45] database classification.

Analysis of the human-Tb HP-PPI network using the Netpath database identifies EGFR, thymic stromal lymphopoietin (TSLP) and androgen receptor pathways, while data collected from Reactome point towards host defense mechanisms, immune signalling, cell senescence, and response to stress as top categories involved. According to the BioCarta database, further signalling relies mainly on thrombospondin-1 (TSp-1) induced apoptosis, extracellular-signal-regulated kinases/FYN proto-oncogene, src family tyrosine kinase/Tyrosine kinase (Erk/Fyn/Tyr) activation and EGF signalling. These results are further supported by KEGG which identifies NOD-like receptor signalling, antigen processing, and trypanosomal defense pathways, as central for the HP-PPI network. Finally, InnateDB classifies innate responses against trypanosomes as relying on IL-2 and IL-12 as well as osteopontin and atypical NF-kB mediated processes (Supplementary file 4).

Despite targeting common proteins, host defense details vary for each disease. For example, interacting host proteins shared between NM and ZE include those associated with mRNA stability, gene expression control, and mRNA splicing (reactome classification). Remarkably, Netpath reveals again androgen receptor pathways as significant, and TNFalpha response pathways as involved in host response against these two pathogens.

According to the reactome database, ZE and Tb, despite differing greatly in complexity and size, trigger shared specific human defense systems centering on eukaryotic translation, elongation,

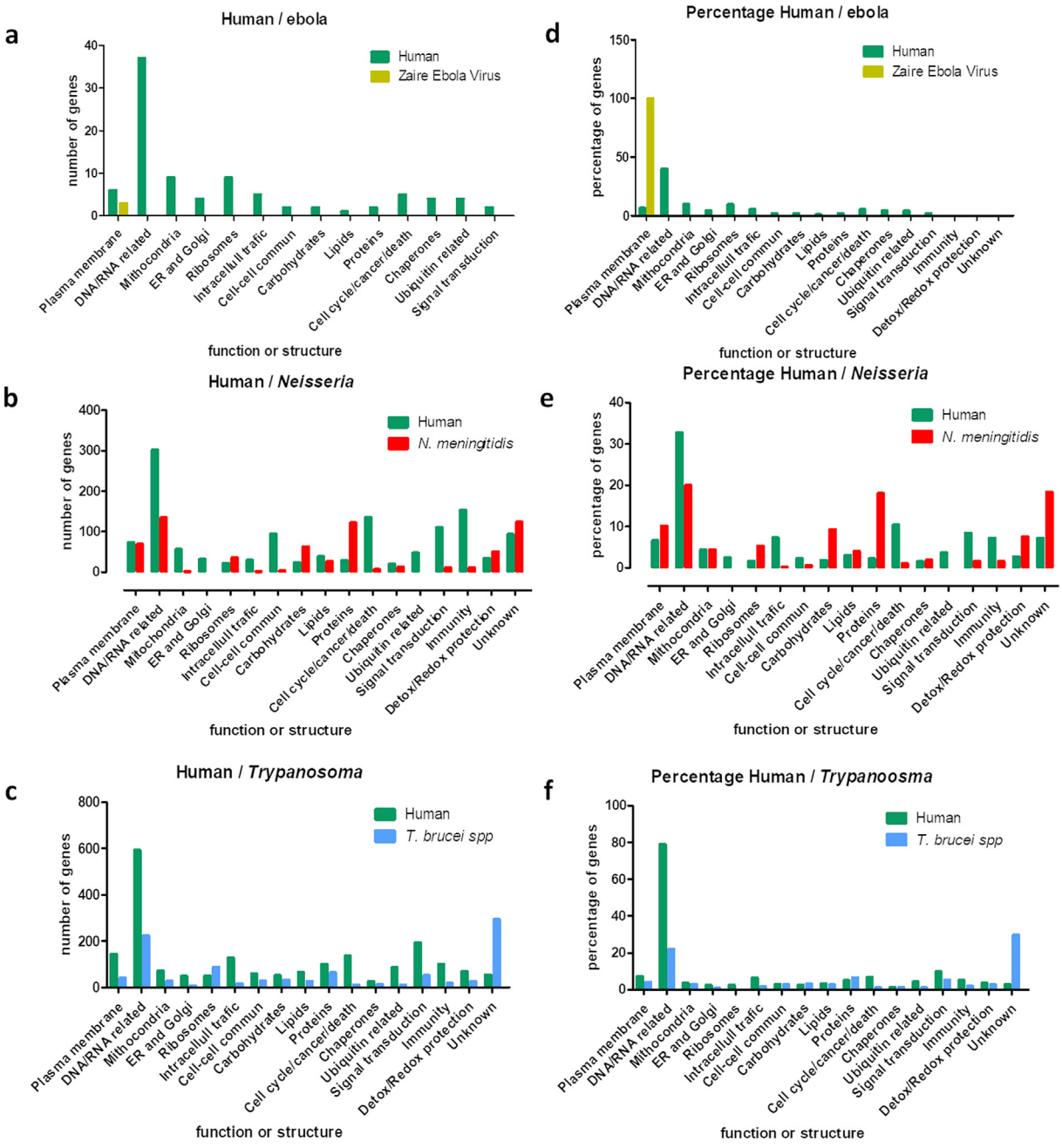



**Fig. 2. Comparison of host-pathogen protein–protein interactions (HP-PPI) in different organisms.** Number of genes involved in the HP-PPI network for ebola (panel a), *Neisseria* (panel b) and *Trypanosoma* (panel c) and its counterpart in humans; **(d, e, f)** Percentage of genes involved in the HP-PPI network for ebola (panel d), *Neisseria* (panel e) and *Trypanosoma* (panel f) and its counterpart in humans. **(g, h)** Number of human genes (panel g) resp. percentage of genes (panel h) involved in the HP-PPI network for each disease (H-human, E – ebola, N- *Neisseria*, T – *Trypanosoma*). **(i,j)** Number (panel i) resp. percentage (panel j) of pathogen genes related to their function or structure (H-human, E – ebola, N- *Neisseria*, T – *Trypanosoma*).

life cycle, and infection, involving ribosomes and PID-BardI signalling, modulated by Rb and p53 proteins according to KEGG. Additionally, Netpath shows a shared TSLP pathway response and again the TNFalpha response.

As a conclusion from this triple comparison, our data suggest that commonly targeted host proteins include host defense and differentiation/maturation pathways. Additionally, the two bigger HP-PPI networks (NM, Tb) also involve reprogramming of host-cell senescence and apoptosis; the two smaller pathogens (NM, ZE) influence mRNA; and Tb and ZE impact protein translation and specific host defense pathways.

### *3.3. Characterization of pathogen proteins and their role in human-pathogen interaction*

We further reviewed disease characteristics for ZE, NM and Tb. Detailed information is listed and explained in Supplementary file 5 (Comparison Table) and Supplementary file 1 (Text file including details for each pathogen and disease), including general responses occurring during the infection processes of all three pathogens. A thorough assessment of the summarized data points provides functional similarities and differences occurring during infection processes. The collected biological information on immune-

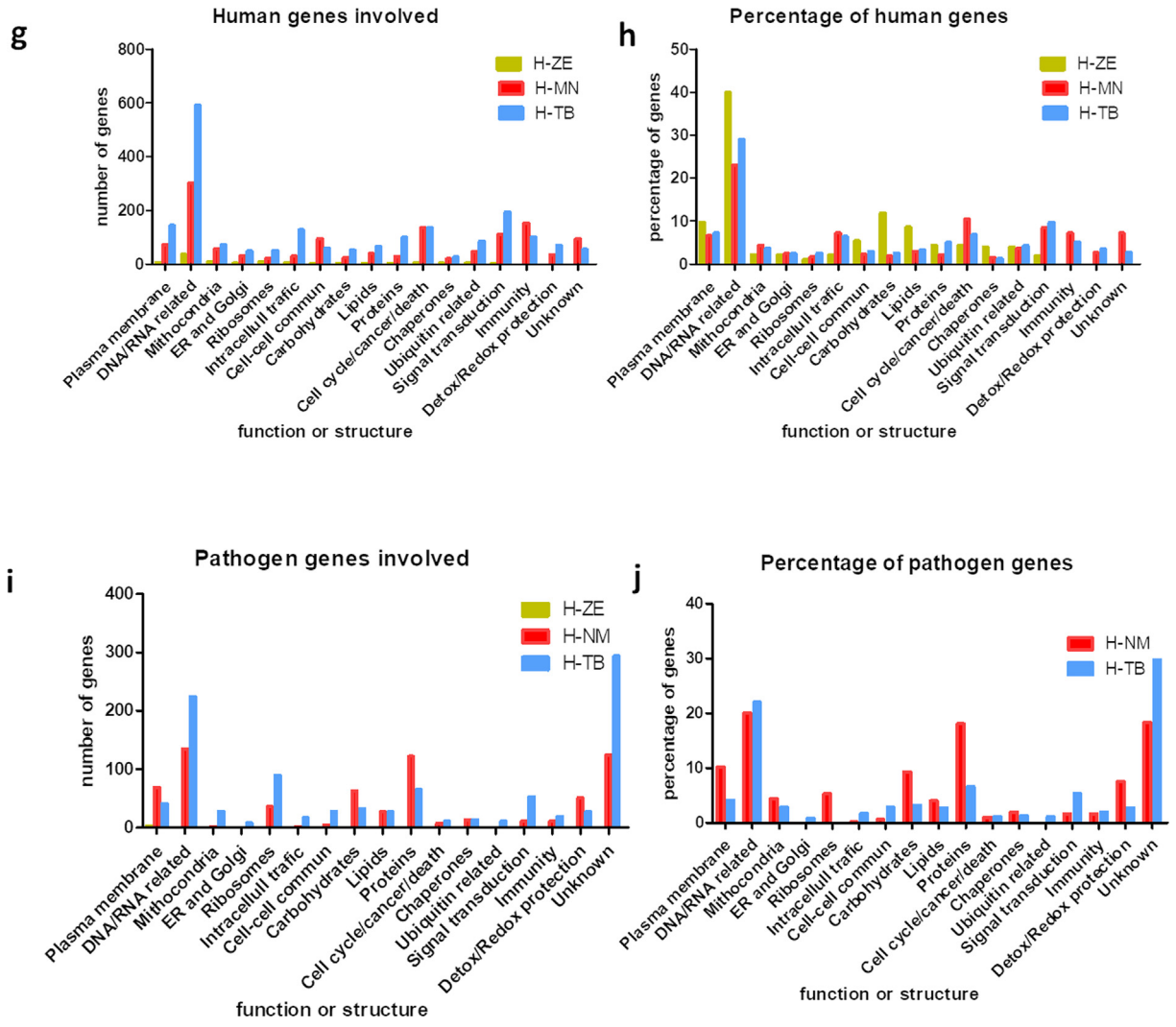



**Fig. 2** (*continued*)

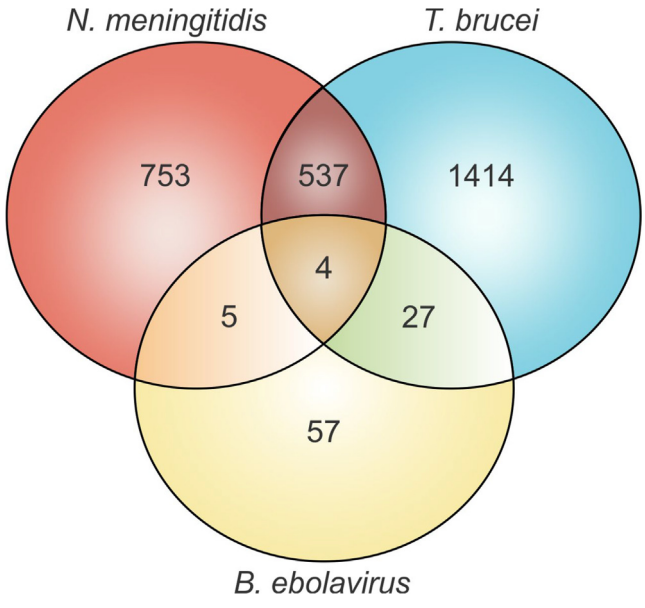



**Fig. 3. Unique and common targeted human proteins by three pathogens: *Neisseria meningitidis*, *Trypanosoma brucei* and *Zaire ebolavirus*.** See Supplementary file 3 for annotation of involved proteins.

related and infection-associated proteins was used for extensive individual curations of the reconstructed preliminary HP-PPI networks. This literature-driven analysis was crucial to address the next phase of our investigation directed towards the identification of infection-related HP-PPIs, based on these curated subnetworks and comparisons between host and pathogen interactions for each organism.

Identification of specific infection-related HP-PPIs and their association with data presented in Supplementary file Figure 1 to 3; Supplementary data 4 till 12 (comparing and describing disease characteristics for each pathogen) allowed the description of useful interaction examples.

For example, the Ebola virus glycoprotein ZE-GP, enables viral entry into target cells [46]. It also antagonizes host cell protein tetherin (bone marrow stromal antigen 2, BST2, also known as tetherin, see Table 1) antiviral activity. This plasma membrane molecule usually restricts virus release from infected cells. It is not clear how ZE-GP antagonizes tetherin. However, this result emphasizes the interaction existing between both molecules and highlights that understanding how Ebola virus offsets innate immune system antiviral effectors might help to define novel antiviral intervention targets. BST2 plays a key role in cellular antiviral response against several enveloped virus including HIV-1, thus preventing subsequent infections [47,48]. BST2 has recently

**Table 1**
Host defense associated human proteins involved in physical interaction with ZE, TB and NM proteins.

| | ZE | NM | Tb | Human protein (annotation) |
|---|---|---|---|---|
| **Extracellular** | | IGHA2 5U | IGHA2 3U | Ig alpha-2 chain C region |
| | | LTBP4 3U, 1C | LTBP4 3U, 1C | Latent-transforming growth factor beta-binding protein 4 |
| | | TINAGL1 11U, 2C | TINAGL1 3U, 2C | Tubulointerstitial nephritis antigen-like (Glucocorticoid-inducible protein 5,(Oxidized LDL-responsive gene 2 protein) (Tubulointerstitial nephritis antigen-related protein,TIN Ag-related protein,TIN-Ag-RP) |
| | | IL18BP 1C | IL18BP 2U, 1C | Interleukin-18-binding protein (IL-18BP, Tadekinig-alfa) |
| | | IGKC 10U | IGKC 2U | Ig kappa chain C region (Fragment) |
| | | IGHG1 2U, 1C | IGHG1 1C | Ig gamma-1 chain C region |
| | | CD74 7U, 2C | CD74 5U, 2C | HLA class II histocompatibility antigen gamma chain (HLA-DR antigens-associated invariant chain, Ia antigen-associated invariant chain, Ii, p33) |
| **Plasma membrane** | BST2 1U | | | Bone marrow stromal antigen 2 (HM1.24 antigen, Tetherin, CD antigen CD317) |
| | | MPEG1 1C | MPEG1 1C | Macrophage-expressed gene 1 protein (Macrophage gene 1 protein, Mpg-1) |
| | | IGHA2 5U | IGHA2 3U | Ig alpha-2 chain C region |
| | | IL2RB 1U | IL2RB 1U | Interleukin-2 receptor subunit beta (IL-2 receptor subunit beta, IL-2R subunit beta, IL-2RB, High affinity IL-2 receptor subunit beta,p70-75, p75, CD antigen CD122) |
| | | BCAP31 1U | BCAP31 1U | B-cell receptor-associated protein 31 (BCR-associated protein 31, Bap31, 6C6-AG tumor-associated antigen, Protein CDM, p28) |
| | | IFI27 1U | IFI27 1U | Interferon alpha-inducible protein 27, mitochondrial (p27, Interferon alpha-induced 11.5 kDa protein, Interferon-stimulated gene 12a protein, ISG12(a) |
| | | NKTR 1U | NKTR 1U | NK-tumor recognition protein (NK-TR protein) (Natural-killer cells cyclophilin-related protein, [Includes: Putative peptidyl-prolyl cis-trans isomerase (PPIase, EC 5.2.1.8, Rotamase)] |
| | | SIGLEC11 1U, 1C | SIGLEC11 1U, 1C | Sialic acid-binding Ig-like lectin 11 (Sialic acid-binding lectin 11, Siglec-11) |
| | | MCEMP1 1U | MCEMP1 2U | Mast cell-expressed membrane protein 1 |
| | | MS4A1 2U | MS4A1 1U | B-lymphocyte antigen CD20 (B-lymphocyte surface antigen B1, Bp35, Leukocyte surface antigen Leu-16, Membrane-spanning 4-domains subfamily A member 1, CD antigen CD20) |
| | | CD74 7U,2C | CD74 5U, 2C | HLA class II histocompatibility antigen gamma chain (HLA-DR antigens-associated invariant chain, Ia antigen-associated invariant chain, Ii, p33) |
| | | TINAGL1 11U,2C | TINAGL1 3U,2C | Tubulointerstitial nephritis antigen-like (Glucocorticoid-inducible protein 5, Oxidized LDL-responsive gene 2 protein, OLRG-2) (Tubulointerstitial nephritis antigen-related protein) (TIN Ag-related protein) (TIN-Ag-RP) |
| | | HLA-A 1U | HLA-A 2U | HLA class I histocompatibility antigen, A-29 alpha chain (Aw-19, MHC class I antigen A*29) |
| | | HLA-B 3U 1C | HLA-B 1C | HLA class I histocompatibility antigen, B-42 alpha chain (MHC class I antigen B*42) |
| | | HLA-E 3U | HLA-E 1U | HLA class I histocompatibility antigen, alpha chain E (MHC class I antigen E) |
| | | HLA-DRA 1U, 1C | HLA-DRA 1U, 1C | HLA class II histocompatibility antigen, DR alpha chain (MHC class II antigen DRA) |
| | | HLA-DRB1 2U, 2C | HLA-DRB1 2C | HLA class II histocompatibility antigen, DRB1-3 chain (Clone P2-beta-3, MHC class II antigen DRB1*3) |
| | | HLA-DRB5 1U, 1C | HLA-DRB5 1C | HLA class II histocompatibility antigen, DRB1-3 chain (Clone P2-beta-3, MHC class II antigen DRB1*3) |
| | | IGSF6 1C | IGSF6 1C | Immunoglobulin superfamily member 6 (IgSF6) (Protein DORA) |
| **Cytoplasm** | BST2 1U | | | Bone marrow stromal antigen 2 (HM1.24 antigen, Tetherin, CD antigen CD317) |
| | | PSME2 2C | PSME2 2U, 2C | Proteasome activator complex subunit 2 (11S regulator complex subunit beta) (REG-beta) (Activator of multicatalytic protease subunit 2, Proteasome activator 28 subunit beta, PA28b, PA28beta) |
| | | TRAF6 1U | TRAF6 6U | TNF receptor-associated factor 6 (EC 6.3.2.-) (E3 ubiquitin-protein ligase TRAF6, Interleukin-1 signal transducer, RING finger protein 85) |
| | | TP53 2U, 1C | TP53 5U, 1C | Cellular tumor antigen p53 (Antigen NY-CO-13, Phosphoprotein p53, Tumor suppressor p53) |
| | | PSMA7 1U | PSMA7 1U | Proteasome subunit alpha type-7 (EC 3.4.25.1) (Proteasome subunit RC6-1, Proteasome subunit XAPC7) |
| | | TGFB1i1 1C | TGFB1i1 1C | Transforming growth factor beta-1-induced transcript 1 protein (Androgen receptor coactivator 55 kDa protein, Androgen receptor-associated protein of 55 kDa, Hydrogen peroxide-inducible clone 5 protein, Hic-5) |
| | | PSMD3 2C | PSMD3 2C | 26S proteasome non-ATPase regulatory subunit 3 (26S proteasome regulatory subunit RPN3, 26S proteasome regulatory subunit S3, Proteasome subunit p58) |
| | | PAAF1 1U | PAAF1 3U | Proteasomal ATPase-associated factor 1 (Protein G-16, WD repeat-containing protein 71) |
| **Cytoskeleton** | | TGFB1i1 1C | TGFB1i1 1C | Transforming growth factor beta-1-induced transcript 1 protein (Androgen receptor coactivator 55 kDa protein, Androgen receptor-associated protein of 55 kDa, Hydrogen peroxide-inducible clone 5 protein, Hic-5) |
| | | PAAF1 1U | | Proteasomal ATPase-associated factor 1 (Protein G-16, WD repeat-containing protein 71) |
| **Lysosome** | | HLA-DRB1 2U,2C | HLA-DRB1 2C | HLA class II histocompatibility antigen, DRB1-3 chain (Clone P2-beta-3, MHC class II antigen DRB1*3) |
| | | HLA-DRB5 1U, 1C | HLA-DRB5 1C | HLA class II histocompatibility antigen, DR beta 5 chain (DR beta-5) (DR2-beta-2, (Dw2, MHC class II antigen DRB5) |
| **Nucleus** | PCNA 1U | | PCNA 3U | Proliferating cell nuclear antigen (Cyclin) |
| | | TRAF6 1U | TRAF6 6U | TNF receptor-associated factor 6 (EC 6.3.2.-) (E3 ubiquitin-protein ligase TRAF6, Interleukin-1 signal transducer, RING finger protein 85) |
| | | PSME2 2C | PSME2 2U, 2C | Proteasome activator complex subunit 2 (11S regulator complex subunit beta) (REG-beta) (Activator of multicatalytic protease subunit 2, Proteasome activator 28 subunit beta, PA28b, PA28beta) |
| | | PAAF1 1U | PAAF1 3U | Proteasomal ATPase-associated factor 1 (Protein G-16, WD repeat-containing protein 71) |
| | | TBRG1 1U, 1C | TBRG1 1U, 1C | Transforming growth factor beta regulator 1 (Nuclear interactor of ARF and Mdm2) |
| | | PBX2 1C | PBX2 1U, 1C | Pre-B-cell leukemia transcription factor 2 (Homeobox protein PBX2) (Protein G17) |
| | | PSMD3 2C | PSMD3 2C | 26S proteasome non-ATPase regulatory subunit 3 (26S proteasome regulatory subunit RPN3, 26S proteasome regulatory subunit S3, Proteasome subunit p58) |
| | | PSMA7 1U | PSMA7 1U | Proteasome subunit alpha type-7 (EC 3.4.25.1) (Proteasome subunit RC6-1, Proteasome subunit XAPC7) |
| | | TGFB1i1 1C | TGFB1i1 1C | Transforming growth factor beta-1-induced transcript 1 protein (Androgen receptor coactivator 55 kDa protein, Androgen receptor-associated protein of 55 kDa, Hydrogen peroxide-inducible clone 5 protein, Hic-5) |
| | | TP53 2U, 1C | TP53 5U, 1C | Cellular tumor antigen p53 (Antigen NY-CO-13, Phosphoprotein p53, Tumor suppressor p53) |
| **Endoplasmic reticulum** | | TP53 2U, 1C | TP53 5U, 1C | Cellular tumor antigen p53 (Antigen NY-CO-13, Phosphoprotein p53, Tumor suppressor p53) |
| | | HLA-DRB1 2U, 2C | HLA-DRB1 2C | HLA class II histocompatibility antigen, DRB1-3 chain (Clone P2-beta-3, MHC class II antigen DRB1*3) |
| | | HLA-DRB5 1U, 1C | HLA-DRB5 1C | HLA class II histocompatibility antigen, DR beta 5 chain (DR beta-5) (DR2-beta-2, (Dw2, MHC class II antigen DRB5) |
| | | BCAP31 1U | BCAP31 1U | B-cell receptor-associated protein 31 (BCR-associated protein 31) (Bap31, 6C6-AG tumor-associated antigen, Protein CDM, p28) |
| **Golgi apparatus** | BST2 1U | | | Bone marrow stromal antigen 2 (HM1.24 antigen, Tetherin, CD antigen CD317) |
| | | HLA-DRB1 2U, 2C | HLA-DRB1 2C | HLA class II histocompatibility antigen, DRB1-3 chain (Clone P2-beta-3, MHC class II antigen DRB1*3) |
| | | HLA-DRB5 1U, 1C | HLA-DRB5 1C | HLA class II histocompatibility antigen, DR beta 5 chain (DR beta-5) (DR2-beta-2, (Dw2, MHC class II antigen DRB5) |
| | | BCAP31 1U | BCAP31 1U | B-cell receptor-associated protein 31 (BCR-associated protein 31) (Bap31, 6C6-AG tumor-associated antigen, Protein CDM, p28) |
| **Centriole** | | PAAF1 1U | PAAF1 3U | Proteasomal ATPase-associated factor 1 (Protein G-16, WD repeat-containing protein 71) |

Color code: Inteferon associated proteins, Blue: Antigen, antibody and HLA associated proteins, Green: Cytokine and chemokine associated proteins, Red-Brown: Proteasome associated proteins, Light-brown: Transcription factors associated proteins, Red: Cell recognition associated proteins. Note that EBOV-human immune system interaction concentrates in two proteins, one at the nuclear level (proliferation) and one that can be found in the cytoplasm, Golgi apparatus and plasma membrane. On the other hand, human immune related proteins involved in the interaction with *T. brucei* and *N. meninigiditis* overlap in high proportion. In the table the following codes are used: U for unique, and C for conserved interactions; the numbers correspond to the number of pathogenic proteins in corresponding category interacting with them. Information about the function of these proteins can be found in the Supplementary file 6. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

caught attention as deletion of BST2 cytoplasmic and transmembrane N-terminal domains results in SARS-CoV, SARS-CoV-2, and influenza virus production suppression in a VERO cell line [49]. Hence, this host protein is important for Ebola virus replication and potentially for several corona viruses. BST2 has been implicated in SARS-CoV-2 cellular restriction [50] and SARS-CoV-2 spike protein generally downregulates tetherin to enhance viral spread [51].

A second example related to NM follows. Heptose 1,7 bisphosphate (HBP) actively secreted or released after NM lysis seems to be delivered to the cytoplasm through the endocytic route. HBP is a metabolic intermediate in lipopolysaccharide (LPS) biosynthesis and is highly conserved in Gram-negative bacteria. HBP presence in the cytoplasm is sensed by tumor necrosis factor receptor-associated factor (TRAF)-interacting protein with a forkhead-associated domain (TIFA). During infection, signaling series of events occur including TRAF2/6 [TNF receptor-associated factor 6 (Table 1) oligomerization and ubiquitination. After recruitment and inclusion of additional host factors this molecular signaling leads to the activation of nuclear factor (NF)-kB and proinflammatory gene expression [52,53].

A final example (Supplementary files 1 and 5) highlights that in trypanosomiasis, several genes and pathways associate with the latent phenotype, suggesting an increased activation of both T- and B-cells in patients with latent infection [54]. The data suggest increased lymphocyte activation in HAT patients, relative to latent cases. Addtional information also predict activation of lymphocytes ($p < 0.0001$), activation of T cells ($p = 0.0004$), co-stimulation of T lymphocytes ($p = 0.0013$), and $Ca^{2+}$ flux ($p = 0.0004$) as an indication of lymphocyte activation, as significantly different between acive and latent cases of trypanosomiasis. Using transcriptome data to conduct an exome-wide single nucleotide polymorphism (SNP) association has suggested that SNP in human major histocompatibility locus associates with disease severity, thus suggesting that T cell activation is a determining factor in disease outcome [54].

The overlapping curated HP-PPI subnetworks of interacting proteins belonging to pathogen and host were further used to understand the role of described host defense PPIs. Subnetwork comparison suggested the existence of common human mechanisms to fight against these three pathogens as depicted in Fig. 4-a-c and summarized in Table 1 (detailed interactions are listed in Supplementary file 6). Fig. 4 illustrates how redundancy of host nodes in different HP-PPI networks are fundamental for survival during host-pathogen encounters [55], either to impair or to promote host protection, relying on innate and cellular defense mechanisms [56]. Moreover, a direct mapping of functional processes (Supplementary files 1 and 5), can be built from processes and gene

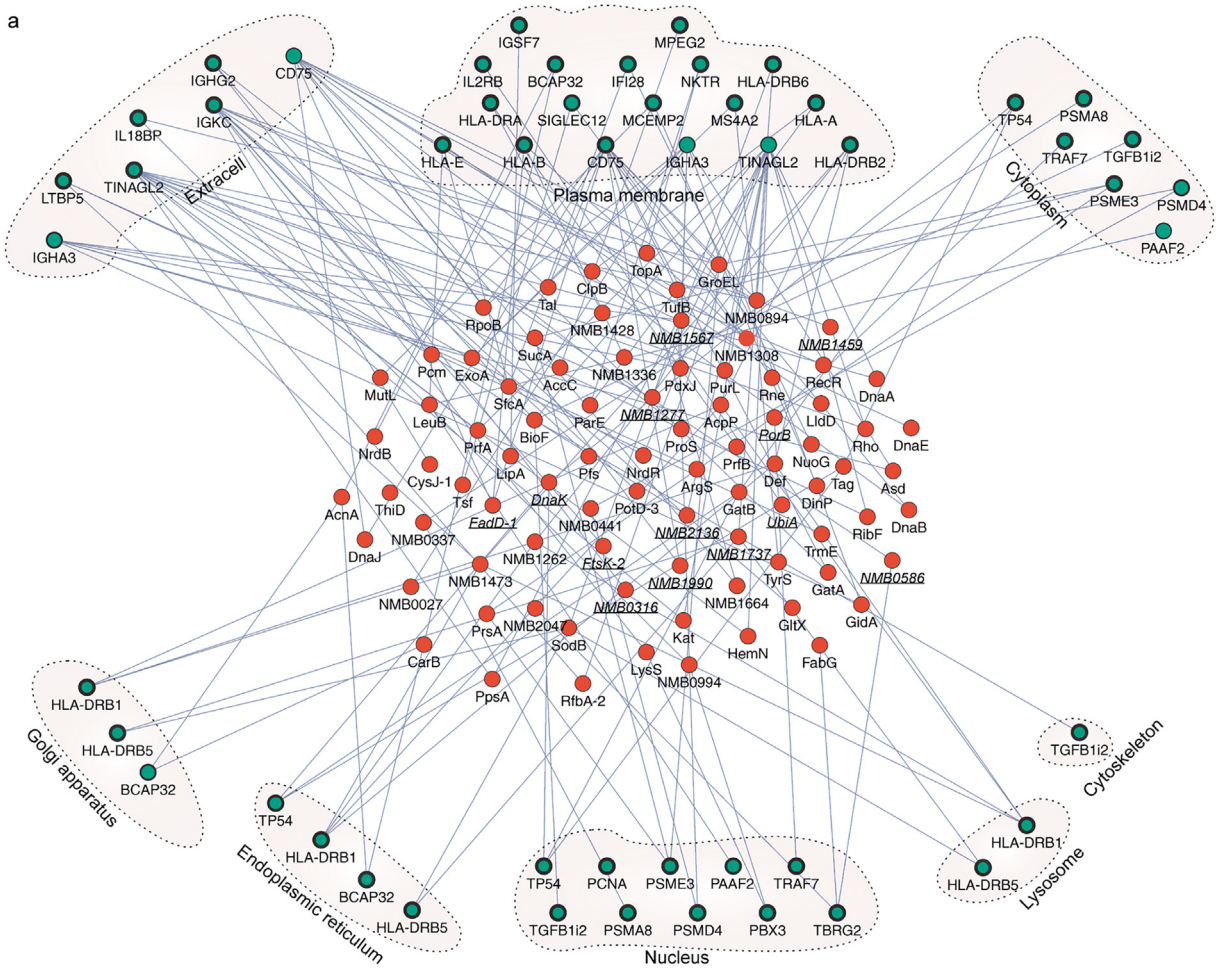
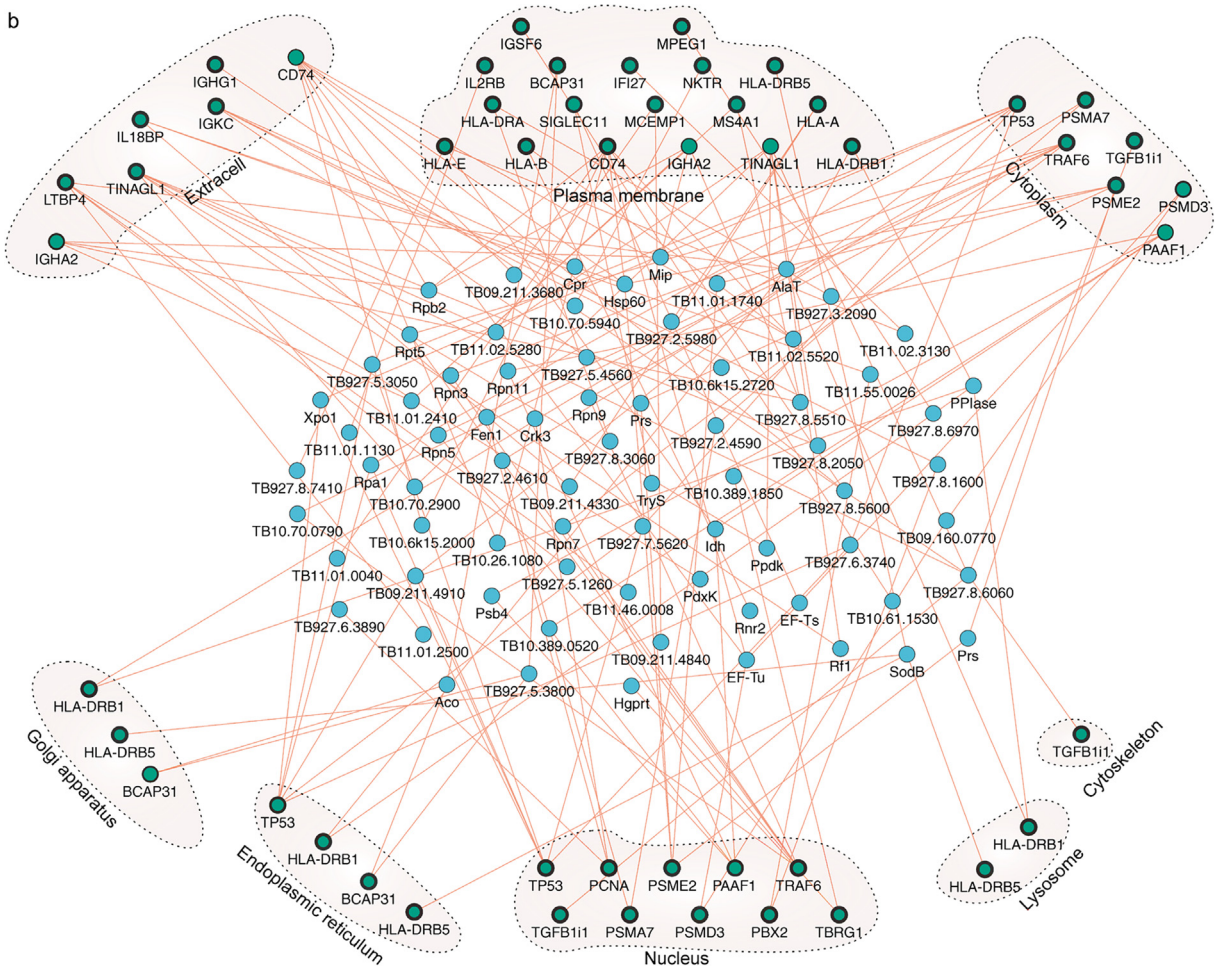



**Fig. 4. An interaction map of the human host with three pathogens comparing protein–protein interactions (PPIs).** The map is based on interologs. The color of the nodes indicates the different organisms: **(a)** *Neisseria meningitidis* = red; **(b)** *Trypanosoma brucei* = blue; **(c)** *Zaire ebolavirus* = yellow, and green = human proteins. Human nodes are shaded in the color clouds and arranged according to their subcellular localization. The border thickness around green human nodes indicates the confidence of localization in different cellular compartments. Membrane proteins of *N. meningitidis* are depicted by underlined and italics text while the *N. meningitidis* membrane proteins with multiple localization is shown in italics. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)



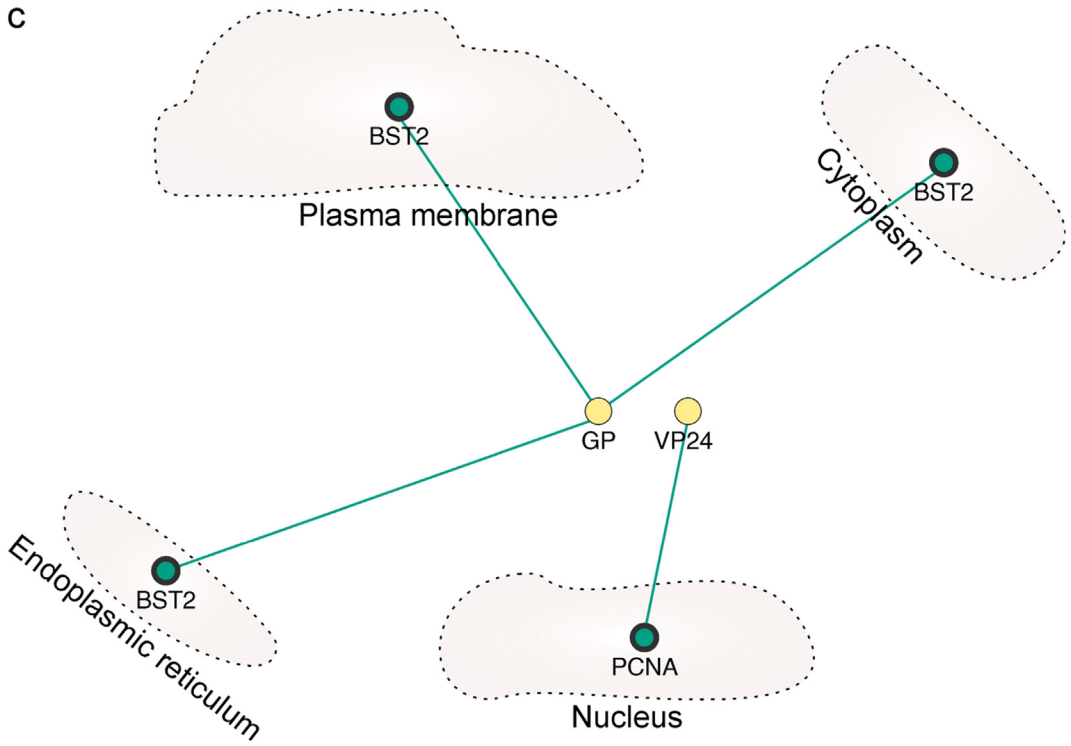




**Fig. 4** (*continued*)

onthology associated with these protein networks (listed with uniprot identifiers and annotations both for human host and pathogen in Supplementary file 6). Additionally, an analysis of strongly connected (protein hubs) or weakly connected proteins can be found in Supplementary file 7. In this way a systematic overview of proteins, annotations and processes is given, appropriately fitting disease characteristics and considering the identified human host and pathogen protein interactions.

Host-pathogen interaction details and individual protein function including host defense-related signalling pathways and compartments are given in Table 1. This table categorizes unique (U) and conserved (C) host proteins involved in the interaction with pathogens. Specific host protection mechanisms against all three pathogens include interferon associated and induced proteins, the HLA system, cytokines and chemokines and proteasome associated proteins. Regulation of this host defense network is achieved by transcription factors (yellow lettering) via cell recognition associated proteins (red lettering). Interestingly, the ZE interaction molecules interact strongly with two human proteins, one is located in the nucleus and related to proliferation) and another one has multiple locations including cytoplasm, Golgi apparatus, and plasma membrane: lipid raft associated protein CD antigen CD317, or stromal bone marrow antigen-2, respectively (see Fig. 4).

### 3.4. Central interactome proteins shared by all pathogens and humans, implications

Our HP-PPI results illustrate that only four host proteins targeted all three HP-PPI networks: the elongation factor 1-alpha 1 (EEF1A1), the SWI/SNF complex subunit SMARCC2, the dolichyl-diphosphooligosaccharide--protein glycosyltransferase subunit 1 (RPN1), and the tubulin beta-5 chain (TUBB). The identified host-pathogen interactions as compared in Fig. 4a-c, suitable for each specific pathogen disease characteristic (supplementary files 1 and 5), share a common denominator: top four hub proteins are exploited by pathogens to gain entry to the host and survive inside it. As illustrated in Fig. 5, (colour codes differentiating all three pathogens), these four central proteins are strongly and well connected in the human interactome (Supplementary Fig. 1). Nor-

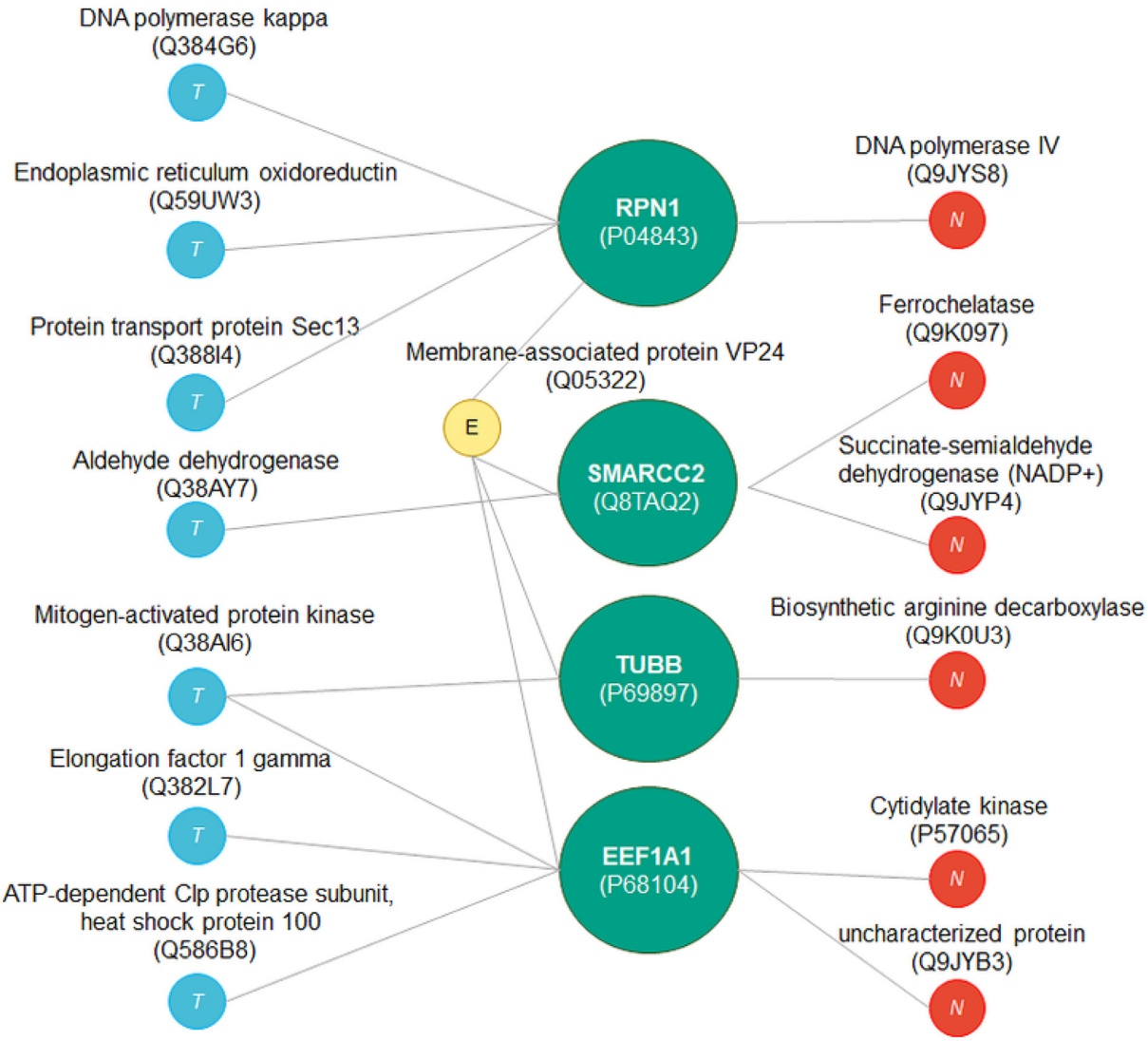



**Fig. 5.** Pathogens target common hubs of human proteins intraspecies PPI network. Only the primary interactors of common hubs are shown. The color of the nodes indicates the different organism's proteins (blue = *Trypanosoma brucei*; red = *Neisseria meningitidis*; yellow = *Zaire ebolavirus*; and green = Human). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.) (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

mally, hub proteins are well conserved [19,57] and hence represent good pathogen targets [29].

Based on these results that suggest a central role of these host hub proteins, we postulate that they may represent potential targets for the design of infection disruption tools and that their manipulation might stop all three infections. For example, EEF1A1 relates to host cell functional changes, so that infection can better exploit cell protein synthesis. SMARCC2 and the TUBB may target cytoskeleton hubs related to infection; their modulation may block all three infections. RPN1 might provide an important target involving glycosyl and lipid metabolism, which all of the three infective agents use.

Table 2 lists these conserved four human proteins as well as their interactions with pathogen proteins, inferred by homology, predicted, putative, or demonstrated by transcription. Involved functions and processes of these proteins are also described in Table 2. Accumulating experimental evidence suggests that these pathogen proteins share these central interactions with the host and exploit this for pathogen entry. In particular for Ebola virus, the interactions have been demonstrated experimentally by high-affinity purification coupled to label-free mass-spectrometry-based approach for VP24 and all four human proteins [58,59]. Moreover, further evidence is available for *Trypanosoma* and even in selected cases for *Neisseria* (Table 2). Altogether, these results suggest that a common, shared protein interaction route could be implicated in pathogen entry.

### 3.5. Large-scale host-pathogen data validate the clade of four host proteins to be well conserved among different organisms in host-pathogen interactions

To validate this unexpected finding further, we analyzed conservation of these four central interactions in large-scale host pathogen databases over animals, fungi, bacteria and viruses and checked participation of all four host proteins in different cellular pathways and networks.

#### 3.5.1. Experimental data confirming the pathogen protein interaction with four conserved human proteins

The conservation of the identified host-pathogen interactions can not be accurately accessed because of the lack of experimentaly validated interspecies interactions between pathogens and host species other than human. The largest collection of interspecies host-pathogen interaction database (HPIDB) contains 94.6% exclusively interactions between pathogens and humans [26] (Supplementary file 8a). The second most common host species in HPIDB is mouse which consists of only 1.1% of interspecies inteactions. The rest of the 64 host species contributes 4.2% of total interactions. This bias towards human data do not allow us to validate interaction conservation comparing multiple host-pathogen interaction pairs. Therefore, we assessed the identified interactions with four central proteins (EEF1A1, RPN1, SMARCC2, TUBB) independently (Table 3). The interaction between all the four central proteins and ZE protein VP24 has been validated by tandem affinity purification experiments [58]. In total, two host-pathogen interactions involving EEF1A1 and NM, one interaction involving RPN1 and NM, two interactions involving SMARCC2 and NM, one interaction involving TUBB and NM were transferred from highly conserved experimentally verified host-pathogen interactions. Similarly, two interaction involving EEF1A1 and Tb and 1 interaction involving TUBB and Tb were transferred from highly conserved experimentally verified host-pathogen interactions. We identified unique interaction between EEF1A1 and Tb, three interactions between RPN1 and Tb, and one interaction between SMARCC2 and Tb by interaspecies interaction transfer and domain-domain interactions. All these mentioned interactions are listed in (Table 3).

#### 3.5.2. Many pathogenic bacteria and virusses target these four conserved human proteins

To understand the importance of these four central proteins, we studied further the conservation of its targeting by different pathogens. We used the interaction data from PHISTO database which comprise human experimental interactions with 97 bacteria, 6 fungi, 11 protozoa and 474 viruses [10] and we examined whether and to which extent these pathogens target the four central proteins identified by our analysis. This assessment revealed alltogether 56 unique pathogens from bacteria and viruses that target these four central proteins to establish the infection (Supplementary file 8b). For instance, EEF1A1 was targeted by 36 unique pathogen species/strains creating 60 host-pathogen interactions (Fig. 6).

This clearly indicates that these four central proteins are among the preferential targets of pathogens.

#### 3.5.3. SARS-CoV-2 targets also two proteins of this conserved protein clade

To examine the importance of our four human proteins further we also checked whether these proteins are targeted by the novel, intensely investigated global human pathogen SARS-CoV-2 by analysing published, experimentally confirmed human-SARS-CoV-2 interactomes [27,28]. We found experimental evidence that out of these four protein targets, EEF1A1 and RPN1 are also targeted by SARS-CoV-2 (Supplementary Fig. 2).

In a recent work, EEF1A1 was shown as a poteintail drug target against the B.1.1.7 variant of SARS-CoV-2 [60]. EEF1A1 can be efficiently targeted by plitidepsin [61]. On the pathway level EEF1A1 and RPN1 both also involves in metabolism of proteins and translation (Supplementary file 8c) and hijacking the host translational machinery is one the survival strategy of SARS-CoV-2 [62].

#### 3.5.4. Various pathogenic fungi target these four conserved human proteins

We further looked at the FungiWeb database to identify the interaction between proteins of nine fungal pathogens (*Aspergillus fumigatus, Candida glabrata, Candida albicans, Coccidioides immitis, Cryptococcus neoformans, Debaryomyces hansenii, Lodderomyces elongisporus, Paracoccidioides brasiliensis, Rhizopus delemar*) with central host targets. Here we identified 130 high confidence interactions with these fungal pathogens and three of the four conserved human proteins: EEF1A1, RPN1 and TUBB (Supplementary file 8d). EEF1A1 was identified as a highly targeted protein as all the nine fungal pathogen targets this protein and all together creates 78 interactions (Supplementary Fig. 3). Pathway analysis of these four proteins (Supplementary file 8c) reveals that all four are involved in cytoskeletal signaling and regulate by this also other central processes such as translation and host cell entry (EEF1A1), protein processing, proteasome and translation (RPN1), chromatin remodeling (SMARCC2), cell cycle and centrosome maturation (TUBB).

Taken together, these three interaction surveys confirm that the four proteins are often targeted by various human pathogens, even if they are far evolutionary apart and from different kingdoms (see discussion).

## 4. Discussion

Pathogens infect and replicate and grow inside the human host. Targeting host proteins improves survival in the host and is selected in evolution. We wanted to better understand specific

**Table 2**
List of proteins identified as shared hubs among the human intraspecies PPI network and ZE, TB and NM proteins.

| Human protein | UniProt ID | Human protein function | Pathogen | Protein from pathogen | Uniprot ID (evidence level) * | Pathogen protein function | Reported experimental validation | Reference |
|---|---|---|---|---|---|---|---|---|
| **Dolichyl-diphosphooligosaccharide–protein glycosyltransferase subunit 1** | **P04843** | Post-translational protein modification | *Trypanosoma brucei brucei* | DNA polymerase kappa | **Q384G6** (predicted) | Error-prone translesion synthesis | N/A | [80]** |
| | | | *Trypanosoma brucei brucei* | Endoplasmic reticulum oxidoreductin | **Q57UW3** (predicted) | Protein folding | N/A | |
| | | | *Trypanosoma brucei brucei* | Protein transport protein Sec13 | **Q388I4**(inferred from homology) | Protein transport from endoplasmic reticulum | N/A | |
| | | | *Neisseria meningitidis* | DNA polymerase IV | **Q9JYS8**(inferred from homology) | DNA replication, error-prone translesion synthesis | N/A | |
| | | | Ebola virus | Membrane-associated protein VP24 | **Q05322** | Host-virus interaction, influencing host immune response | affinity chromatography technology | [59] |
| **SWI/SNF complex subunit SMARCC2** | **Q8TAQ2** | Chromatin remodeling | *Trypanosoma brucei brucei* | Aldehyde dehydrogenase | **Q38AY7** (putative) | Catalysis aldehyde oxidation | N/A | |
| | | | *Neisseria meningitidis* | Succinate-semialdehyde dehydrogenase | **Q9JYP4**(inferred from homology) | Cell grown in presence of Mn | N/A | |
| | | | *Neisseria meningitidis* | Ferrochelatase | **Q9K097** (inferred from homology) | Porphyrin biosynthesis | N/A | |
| | | | Ebola virus | Membrane-associated protein VP24 | **Q05322** | Host-virus interaction, influencing host immune response | tandem affinity purification | [58] |
| **Tubulin beta-5 chain** | **P07437** | The major component of microtubules (cytoskeleton formation) | *Trypanosoma brucei brucei* | Mitogen-activated protein kinase | **Q26802** (transcription) | survival of *Trypanosoma* | N/A | [78]** |
| | | | *Neisseria meningitidis* | Biosynthetic arginine decarboxylase | **Q9K0U3** (inferred from homology) | Agmatine biosynthesis | N/A | [77,85]** |
| | | | Ebola virus | Membrane-associated protein VP24 | **Q05322** | Host-virus interaction, influencing host immune response | tandem affinity purification | [58] |
| **Elongation factor 1-alpha 1** | **P68104** | Transcription regulation | *Trypanosoma brucei brucei* | Elongation factor 1 gamma | **Q382L7** (predicted) | Transcription regulation | N/A | |
| | | | *Trypanosoma brucei brucei* | ATP-dependent Clp protease subunit, heat shock protein 100 (HSP100) | **Q586B8** (inferred from homology) | Regulating mitochondrial DNA replication | N/A | |
| | | | *Trypanosoma brucei brucei* | Mitogen-activated protein kinase | **Q26802** (transcription) | survival of *Trypanosoma* | N/A | [78]** |
| | | | *Neisseria meningitidis* | Cytidylate kinase | **P57065**(inferred from homology) | Nucleic acid biosynthesis | N/A | |
| | | | *Neisseria meningitidis* | Uncharacterized protein | **Q9JYB3** | N/A | N/A | |
| | | | Ebola virus | Membrane-associated protein VP24 | **Q05322** | Host-virus interaction, influencing host immune response | tandem affinity purification | [58] |

N/A – not available, *Uniprot accessed on July 28, 2020. **potential indirect evidence.

**Table 3**
Experimental data confirming pathogen proteins targeting the four central proteins.

| Host | Pathogen | Validation | Reference |
|---|---|---|---|
| EEF1A1 | VP24 | Experimentally validated | [58] |
| EEF1A1 | NMB1663 | Conserved in EEF1A1(Human)-FTT_1214c(*Francisella tularensis*) | [86] |
| EEF1A1 | Cmk | Conserved in EEF1A1(Human)-KCY(*Bacillus anthracis*) | [56] |
| EEF1A1 | TB11.01.4750 | Intraspecies and domain-domain interaction based | [87] |
| EEF1A1 | TB10.6 k15.2790 | Conserved in EEF1A1(Human)-US3(*Human herpesvirus 1 Strain 17*) | [58] |
| EEF1A1 | TB927.2.5980 | Conserved in EEF1A1(Human)-CLPB(*Bacillus anthracis*) | [86] |
| RPN1 | VP24 | Experimentally validated | [58,86] |
| RPN1 | DinB | Conserved in RPN1(Human)-DPO4(*Yersinia pestis*) | [86] |
| RPN1 | TB11.01.0040 | Intraspecies and domain-domain interaction based | [87] |
| RPN1 | TB10.61.2630 | Intraspecies and domain-domain interaction based | [87] |
| RPN1 | TB927.8.4890 | Intraspecies and domain-domain interaction based | [87] |
| SMARCC2 | VP24 | Experimentally validated | [58] |
| SMARCC2 | HemH | Conserved in SMARCC2 (Human)-HemH1(*Bacillus anthracis*) | [86] |
| SMARCC2 | GabD | Conserved in SMARCC2 (Human)-GabD(*Bacillus anthracis*) | [86] |
| SMARCC2 | TB10.70.0630 | Intraspecies and domain-domain interaction based | [87] |
| TUBB | VP24 | Experimentally validated | [58] |
| TUBB | SpeA | Conserved in TUBB(Human)-SpeA(*Yersinia pestis*) | [86] |
| TUBB | TB10.6 k15.2790 | Conserved in TUBB(Human)-US3(*Human herpesvirus 1 Strain 17*) | [58] |

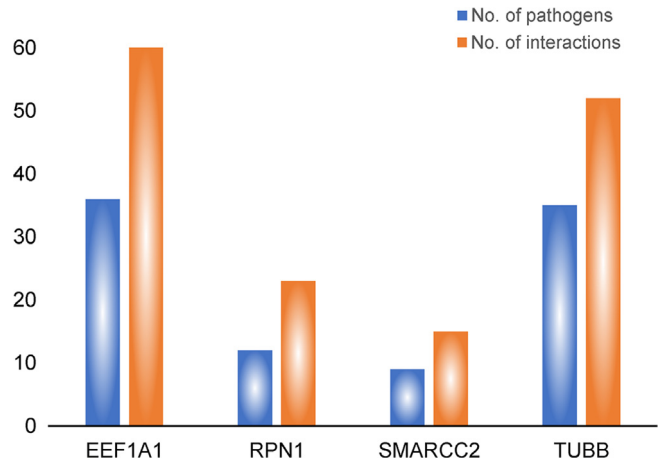



**Fig. 6.** Conserved targeting of the four central proteins by different pathogens. Only experimentally validated interactions are included. For the details of protein interactions (organism name, involved proteins) please see Supplementary file 8b.

adaptations and common pathways shared by three dangerous human pathogens able to infect septically human beings. We compared ZE, NM and Tb and specific but also shared pathways and next test, which proteins of the host these three pathogens target. We hypothesized that this common ability of all three pathogens to be blood borne would allow to reveal ready interpretable shared host pathways attacked and found indeed four human host proteins shared as common targets by all three pathogens.

This result was then extensively counter-checked by a battery of organisms combining various data from bacterial and viral host-pathogen interaction databases, latest results from SARS-CoV-2 research as well as databases on human proteins targeted by fungi. All these data confirm: Yes, the four human proteins are well conserved and used for direct targeting by many human pathogens and we can identify the human pathways they are involved in, suggesting a central role in the infection process for the human host often suppressed by the pathogen.

### *4.1. A common defense network revealed from the three blood borne pathogens*

#### *4.1.1. ZE interacts with the complement cascade*

Cholesterol-enriched lipid raft (plasma membrane) microdomains are a main gateway used by this virus [63,64]. ZE enhances its infectivity by stimulating antibody production and virus-antibody complexes that bind to monocytes and macrophage Fc-receptors [65] and activates the complement cascade to facilitate virus entry into cells such as primate kidney cells [66]. The complex constituted by virus, antibodies and C1, then binds C1q ligands at the immune cell surface, promoting either virus binding to ZE-specific receptors or endocytosis [66].

Once internalized, a cytopathogenic effect of ZE causes annihilation of infected cells. Monocytes/macrophages and dendritic cells (DCs) are ZE early replication sites and disseminate the virus from spleen and lymph nodes to other tissues such as fibroblasts, hepatocytes, adrenal cells and epithelial cells [64]. Two proteins are involved in the ZE-human HP-PPI network as herein described: PCNA and the bone marrow stromal antigen 2 (see Table 1). PCNA is a proliferating cell nuclear antigen (Cyclin), essential for replication and involved in DNA repair, epigenetics and chromatin remodelling [67]. PCNA-interacting ZE proteins VP35 and VP24 are involved in viral particle maturation, immune activity suppression, and inhibition of the tetherin activity, and probably facilitate the evasion process. ZE evades Type-I IFNs responses (IFN-α and IFN-β), thus eliminating their protective role especially of IFN-α [68]. Interferon activates neighboring cells, increases expression of viral restriction factors, e.g. tetherin [68–70] and early (first hours) strong secretion of pro- and anti-inflammatory cytokines via TLR4 stimulation of DCs and macrophages [69]. Hence, ZE host defense interaction facilitates viral proliferation, while the second proteins restricts viral spreading. The implications of this balance should be further analyzed. Downstream effects of antigen-presenting cell dysfunction affect adaptive immunity: massive natural killer (NK) cell loss in peripheral blood, unbalanced maturation signals for DCs with lymphoid depletion and necrosis occur in spleen, thymus and lymph nodes. As no signs of lymphocyte virus infection are detected, this suggests a bystander mechanism of apoptosis, or necrotic cell death [69].

#### *4.1.2. NM and Tb activate common pathways in the host*

NM and Tb both have a blood infection stage, can invade the brain and cause meningitis. Their respective HP-PPI networks targeting human defense mechanisms overlap a lot (see Table 1), targeting antigen presentation, cytokine and chemokine secretion and antibody production. E.g. interleukin-2 receptor subunit beta, TNF receptor-associated factor 6 and interleukin-18-binding protein trigger pathogen-specific sites of thrombo-inflammation (Table 1). Extracellular matrix protein LTBP4, the latent-transforming growth factor-beta (TGF-β)-binding-protein-4, participates only in the human-NM HP-PPI network [71]. In Tb, changing glycoprotein coat or Variant Surface Glycoprotein (VSG) of bloodstream parasites is crucial for evading host antibody activity, and the complement cascade [3,72]. Similarly in NM, high-frequency antigenic variation of other surface antigens, like lipooligosaccharide (LOS), opacity-

associated proteins (Opa) and type IV pili, leads to poor immune control of NM infections [73]. Common to NM and Tb diseases is also the depletion of the complement cascade function. In *Trypanosoma* infections by *Tb gambiense* and *Tb rhodesiense*, reduction in complement activity is caused by their massive activation of complement [74]. Human infection by NM begins by the colonization of the mucosal epithelium in cervix, conjunctiva, fallopian tubes, nasopharynx, rectum and urethra, by bacterial type IV pilin and Opa proteins [75] targeting host cell specialized adhesins.

#### 4.1.3. Conserved host and pathogen-specific pathways

Our census shows that the three evolutionary distant pathogens develop similar strategies for host mitigation, defense and evasion by interaction with the same host proteins (Table 1). Detailed shared and unique PPIs occurring in different compartments are listed in Table 1. Host interactions with pathogen proteins happen in six different cellular compartments (green nodes). NM and Tb target both human immunoglobulin heavy constant gamma-1 (IGHG1) chain, tubulointerstitial nephritis antigen-like protein (TINAGL1), latent-transforming growth factor beta-binding protein-4 (LTBP4) and interleukin-18-binding protein (IL18BP). Host plasma membrane proteins include HLAs (human leukocyte antigen; HLA-B, HLA-DRA, HLA-DRB1 and HLA-DRB5), CD surface protein (CD74), Macrophage-expressed gene 1 protein (MPEG1), immunoglobulin superfamily member-6 (IGSF6) and sialic acid-binding Ig-like lectin-11 (SIGLEC11).

Cytoplasmic host targets for NM and Tb include apoptotic and inflammatory p53 and transforming growth factor beta-1-induced transcript 1 protein (TGFB1i1), which regulates integrin, Wnt and TGFB signaling pathways. Shared host targets are also HLA molecules HLA-DRB1 and HLA-DRB5 in Golgi apparatus and the endoplasmic reticulum. Shared cyto-nuclear interactions include proteasome related proteins (PSME2 and PSMD3), transforming growth factor-beta regulator 1 (TBRG1) and Pre-B-cell leukemia transcription factor 2 (PBX2).

### 4.2. Infection biology of the three pathogen networks: Specific and general targeting of host proteins

The interactions of the four conserved core proteins (Fig. 5) are contrasted by kingdom-typical interactions: eukaryotic trypanosomes target host nuclear or the mitochondrial proteins, while prokaryotic bacteria target often also metabolic enzymes as well as host cell molecules by surface proteins [76], e.g. NM targets host microtubules [77]. Virus-interactions manipulate the host translation or transcription machinery as is the case for PCNA located in the cytoskeleton.

Many protozoan parasites adequately control growth and differentiation processes for survival in the different environments they encounter during their life cycle. One way to do so may be by expressing different sets of mitogen-activated (MAP) kinases for each proliferative stage. For example, a Tb protein kinase containing the signature of extracellular-signal–regulated kinases has been cloned and analyzed. Deletion of Tb MAPK2 results in delayed differentiation and growth inhibition and cell cycle arrest. Tb MAPK2 may thus control the growth of procyclic form trypanosomes in the tsetse midgut [78]. A second MAP kinase, Tb MAPK5, seems to have a regulatory function during bloodstream form differentiation of Tb. As a result, knockout parasites seem to express an increased sensitivity to the stumpy induction factor, and knockout parasites differentiated prematurely in mice and culture [78]. Whether or not a physical interaction exists between any of these MAP kinases and host elongation factor 1-alpha 1 and/or tubulin-5 beta chain needs to be confirmed. The potential physical interaction might relate to the switch on/off of differentiation processes that occur during the blood stage, and/or to guide parasite attachment to the host cell for example during parasite migration to the brain.

Additionally, the human oligosaccharyltransferase, involved in posttranslational protein modification, is localized in the endoplasmic reticulum [88], suggesting a similar localization with Tb putative proteins involved in protein folding and protein transport from the endoplasmic reticulum. Moreover, in Tb, a polymerase K (POLK) gene is present in 10 copies, in tandem repeats, and is not subjected to alternative splicing [89]. In *T. cruzi*, two copies of the POLK gene with differences in their encoded protein sequences outside of their catalytic core regions have been described. TcPOLK overexpression increases *T. cruzi* survival when exposed to $H_2O_2$ treatment. Phylogenetic analysis of TcPolk copies together with its orthologues shows that the Polk of trypanosomatids has diverged early from other eukaryotes counterparts and at least one of the copies is located at the parasite mitochondria and supports *in vitro* DNA synthesis [79] probably involved in error-prone translesion synthesis. Whether or not the predicted Tb POLK described above is involved in error prone translesion synthesis need to be explored as well as its physical interaction with the human dolichyl-diphosphooligosaccharide-protein glycosyltransferase subunit 1.

In Tb, exit from the endoplasmic reticulum is mediated by coat protein II (COPII)-coated vesicles that bud from distinct ribosome-free regions of the ER membrane known as ER exit sites (ERES). Tb success relies as an infecting pathogen on coat protection by the variant surface glycoprotein (VSG). VSG is membrane-bound by a glycolipid (GPI) anchor, attached in the earliest compartment of the secretory pathway, the endoplasmic reticulum. Therefore, its exit from the endoplasmic reticulum is required for its proper expression at the surface of the parasite [80].

The Venn Diagram summarizes all these interactions (Fig. 3). A high proportion of overlap exists between *Neisseria*- and trypanosome-human interactions, and these conserved PPIs focus on host proteins related to immune system signaling, response to stress and reprogramming of host cell senescence. On the contrary, overlapping interactions between ZE and NM focus on mRNA stability and splicing while interactions overlapping between ZE and Tb concern translation and elongation, ribosome function and life cycle. Taken together, we speculate that all these shared PPIs play a role in protein translation. Data presented here were generated by different bioinformatic tools providing theoretical predictions and revealing similar pathogen strategies. In addition to the central interaction clade, most host cell compartments are targeted by the pathogens, with the overlap between host-pathogen interactions exposed best when comparing NM and Tb. These data highlights potential therapeutic targets for the design of agents to impair the infection produced by these pathogens and encourages the design of experiments which can shed light on the mechanisms involved in the conserved host-pathogen PPIs described.

### 4.3. Validation of the conserved four proteins targeted by pathogens

To further validate the four proteins commonly targeted by NM, TB, ZE, all available experimental evidence for these proteins was collected (Table 3), notably for ZE there is extensive direct experimental validation, for NM and Tb we collected circumstantial evidence. Nevertheless, this still could be a rare coincidence from the particular pathogen choice. By three different large-scale data sets we hence did next confirm that the four human proteins found to be targeted comparing the evolutionary very distant pathogens NM, TB, ZE are no chance observation but these four human proteins are very often targeted by human pathogens. Specifically, we found evidence from well-curated databases with experimental data that

(i) 56 human pathogenic bacteria and viruses target these four proteins,
(ii) the well researched new pandemic pathogen SARS-CoV-2 targets two of these four human proteins and
(iii) nine human pathogenic fungi (yet another evolutionary distant organism group) target three of the conserved proteins by 130 high confidence interactions.

Together, all these data confirm that these four central proteins are highly conserved and hence important proteins that are targeted by many human pathogens including bacteria, fungi, viridae and trypanosomes.

### *4.4. Four proteins are highly conserved targets as they are targeted by many human pathogens*

The clade of only four interacting proteins present in the host and very strongly and conserved targeted seems to assure the proper decoding, produce cellular proteins and relocate them to the appropriate place. All four are important regulators of the cytoskeleton as explained in the following: EF1α is undoubtedly important for protein translation. However, EF1α proteins also modulate cytoskeleton activities, exhibit chaperone-like activity, and are key proteins for cell proliferation and cell death in human tumours [81]. EF1α assists replication of many RNA viruses, such as the respiratory syncytialvirus (RSV). In this case, down-regulation of EF1α restricts the expression of viral genomic RNA and the release of infectious virus [82]; in parasites, EF1α has also been implicated in pathogenesis [83] and host cell invasion [84]. The other three proteins also organize the cytoskeleton and we think that cytoskeletal signaling and its organization are both critical for the host defense as well as the success of the pathogen, and hence this results in these highly conserved targets: SWI/SNF complex subunit SMARCC2 is the matrix-associated actin-dependent regulator of chromatin subfamily C. It is targeted by bacteria and viridae but not targeted by fungi. The dolichyl-diphosphooligosaccharide--protein glycosyltransferase subunit 1 (RPN1) transfers glycosides to phospholipids and oligosaccharides. This function is again important in the cytoskeleton. Finally, the tubulin beta-5 chain (TUBB) is an important component of the cytoskeleton. The central functions of the four proteins (Supplementary File 8c) support the idea that many pathogens profit from targeting central host processes via the cytoskeleton and host entry by targeting these four proteins. Hence, this descriptive and *in silico* finding of high conservation of a human protein clade in infection requires experimental follow up to better understand the exact function of the four proteins in human-pathogen interactions and for the infection process. Ultimately, this could open up new therapeutical intervention strategies targeting host proteins pharmacologically which has in addition the promise of preventing fast resistance mutations which evolve from survival selection pressure by targeting pathogen proteins (e.g. antibiotic resistance of numerous pathogens Supplementary File 8c).

## 5. Conclusions

Septic pathogens *Trypanosoma*, *Neisseria* and Ebola virus invade the host in different manners; however, we found four common human proteins which are targeted by each of the pathogens. Experimentally validated direct interactions are available for VP24 from ZE to all four conserved human proteins and there is circumstantial evidence for NM and Tb. Next, we systematically looked at all available human-pathogen database data comparing human pathogenic bacteria, viruses and fungi. We found that indeed these four proteins are often targeted by human pathogens according to the available database information. This includes confirmation by experimental data in several specific cases including two of these proteins to be targeted by SARS-CoV-2. Our results suggest an unexpected human protein clade of four proteins targeted often by human pathogens: EF1α, involved in host cell invasion; tubulin beta-5 chain; cytoscelettal glycosyltransferase RPN1; actin-dependent regulator SMARCC2. All four proteins are involved in cytoskeletal organization and by this in further central human host processes.

## CRediT authorship contribution statement

Conceptuatlization: APS, SG, TD; Data curation and Formal analysis: All; Funding acquisition: APS, TD; Investigation: All; Methodology: SG, EB; Project administration: TD; Software: SG; Supervisions: APS, TD; Validation: All; Visualization: SG, EB; Writing - original draft: APS, SG, TD; Writing - review and editing: All.

## Ethics approval and consent to participate or to publish

Not applicable, this manuscript does not include human participants, human data or human tissue.

## Availability of data and material

The datasets generated and analysed during the current study are all available in the manuscript and its supplementary materials.

## Declaration of Competing Interest

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

## Acknowledgements

This work was supported by the Land Bavaria (contribution to DFG Project number 324392634 – TRR 221/INF). TD acknowledges the support of the Deutsche Forschungsgemeinschaft (Project number 374031971 – TRR 240/INF, platelet effects; and project number 210879364 – TRR 124/B1; fungal interactions). APS acknowledges the support of the Alexander von Humboldt Foundation for return visits to Würzburg University and Missionärztliches Institut, Würzburg. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. We thank Dr. Rapp-Galmiche and Sara Giddins for stylistic and language corrections.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.csbj.2021.09.017.

## References

[1] Goeijenbier M, van Kampen JJ, Reusken CB, Koopmans MP, van Gorp EC. Ebola virus disease: a review on epidemiology, symptoms, treatment and pathogenesis. Neth J Med 2014;72:442–8.
[2] Moole H, Chitta S, Victor D, Kandula M, Moole V, Ghadiam H, et al. Association of clinical signs and symptoms of Ebola viral disease with case fatality: a systematic review and meta-analysis. J Community Hosp Intern Med Perspect 2015;5(4):28406. https://doi.org/10.3402/jchimp.v5.28406.
[3] Ponte-Sucre A. An Overview of Trypanosoma brucei Infections: An Intense Host-Parasite Interaction. Front Microbiol 2016;7:2126.
[4] Rewar S, Mirdha D. Transmission of ebola virus disease: an overview. Ann Glob Health. 2014;80:444–51.

[5] Stijlemans B, Caljon G, Van Den Abbeele J, Van Ginderachter JA, Magez S, De Trez C. Immune Evasion Strategies of Trypanosoma brucei within the Mammalian Host: Progression to Pathogenicity. Front Immunol 2016;7:233.
[6] Welburn SC, Beange I, Ducrotoy MJ, Okello AL. The neglected zoonoses–the case for integrated control and advocacy. Clin Microbiol Infect 2015;21 (5):433–43.
[7] Gupta SK, Osmanoglu, Ö., Srivastava, M., Bencúrová, E., Dandekar, T. Dandekar Pathogen and Host-Pathogen Protein Interactions Provide a Key to Identify Novel Drug Targets. Reference Module in Biomedical Sciences; Systems Medicine: Integrative, Qualit Computat Approach Elsevier; 2020.
[8] The UC. UniProt: the universal protein knowledgebase. Nucleic Acids Res 2017;45:D158–69.
[9] O'Leary NA, Wright MW, Brister JR, Ciufo S, Haddad D, McVeigh R, et al. Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. Nucleic Acids Res 2016;44(D1): D733–45.
[10] Durmus Tekir S, Cakir T, Ardic E, Sayilirbas AS, Konuk G, Konuk M, et al. PHISTO: pathogen-host interaction search tool. Bioinformatics. 2013;29:1357-8.
[11] Salwinski L, Miller CS, Smith AJ, Pettit FK, Bowie JU, Eisenberg D. The Database of Interacting Proteins: 2004 update. Nucleic Acids Res 2004;32: D449–51.
[12] Yu H, Luscombe NM, Lu HX, Zhu X, Xia Y, Han JD, et al. Annotation transfer between genomes: protein-protein interologs and protein-DNA regulogs. Genome Res 2004;14:1107–18.
[13] Sonnhammer EL, Ostlund G. InParanoid 8: orthology analysis between 273 proteomes, mostly eukaryotic. Nucleic Acids Res. 2015;43:D234-9.
[14] Finn RD, Coggill P, Eberhardt RY, Eddy SR, Mistry J, Mitchell AL, et al. The Pfam protein families database: towards a more sustainable future. Nucleic Acids Res 2016;44(D1):D279–85.
[15] Yellaboina S, Tasneem A, Zaykin DV, Raghavachari B, Jothi R. DOMINE: a comprehensive collection of known and predicted domain-domain interactions. Nucleic Acids Res. 2011;39:D730-5.
[16] Luo Q, Pagel P, Vilne B, Frishman D. DIMA 3.0: Domain Interaction Map. Nucleic Acids Res 2011;39:D724–9.
[17] Kim Y, Min B, Yi GS. IDDI: integrated domain-domain interaction and protein interaction analysis system. Proteome Sci 2012;10(Suppl 1):S9.
[18] Berman HM, Battistuz T, Bhat TN, Bluhm WF, Bourne PE, Burkhardt K, et al. The Protein Data Bank. Acta Crystallogr D Biol Crystallogr 2002;58(6):899–907.
[19] Brown KR, Jurisica I. Unequal evolutionary conservation of human protein interactions in interologous networks. Genome Biol 2007;8(5):R95. https://doi.org/10.1186/gb-2007-8-5-r95.
[20] Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 2003;13:2498–504.
[21] Chin CH, Chen SH, Wu HH, Ho CW, Ko MT, Lin CY. cytoHubba: identifying hub objects and sub-networks from complex interactome. BMC Syst Biol 2014;8 (Suppl 4):S11.
[22] Conesa A, Gotz S, Garcia-Gomez JM, Terol J, Talon M, Robles M. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics 2005;21(18):3674–6.
[23] Breuer K, Foroushani AK, Laird MR, Chen C, Sribnaia A, Lo R, et al. InnateDB: systems biology of innate immunity and beyond–recent updates and continuing curation. Nucleic Acids Res 2013;41:D1228-33.
[24] Lin HN, Chen CT, Sung TY, Ho SY, Hsu WL. Protein subcellular localization prediction of eukaryotes using a knowledge-based approach. BMC Bioinf 2009;10(Suppl 15):S8.
[25] Lin H-N, Chen C-T, Sung T-Y, Hsu W-L. UniLoc: A universal protein localization site predictor for eukaryotes and prokaryotes. bioRxiv. 2018:252916.
[26] Ammari MG, Gresham CR, McCarthy FM, Nanduri B. HPIDB 2.0: a curated database for host-pathogen interactions. Database (Oxford). 2016;2016.
[27] Gordon DE, Jang GM, Bouhaddou M, Xu J, Obernier K, White KM, et al. A SARS-CoV-2 protein interaction map reveals targets for drug repurposing. Nature 2020;583(7816):459–68.
[28] Samavarchi-Tehrani P, Abdouni H, Knight JDR, Astori A, Samson R, Lin Z-Y, et al. A SARS-CoV-2 – host proximity interactome. bioRxiv. 2020:2020.09.03.282103.
[29] Durmus Tekir S, Cakir T, Ulgen KO. Infection Strategies of Bacterial and Viral Pathogens through Pathogen-Human Protein-Protein Interactions. Front Microbiol 2012;3:46.
[30] Epperson ML, Lee CA, Fremont DH. Subversion of cytokine networks by virally encoded decoy receptors. Immunol Rev 2012;250(1):199–215.
[31] Schleker S, Trilling M. Data-warehousing of protein-protein interactions indicates that pathogens preferentially target hub and bottleneck proteins. Front Microbiol 2013;4:51.
[32] Nicod C, Banaei-Esfahani A, Collins BC. Elucidation of host-pathogen protein-protein interactions to uncover mechanisms of host cell rewiring. Curr Opin Microbiol 2017;39:7–15.
[33] Rual J-F, Venkatesan K, Hao T, Hirozane-Kishikawa T, Dricot A, Li N, et al. Towards a proteome-scale map of the human protein-protein interaction network. Nature 2005;437(7062):1173–8.
[34] Littler SJ, Hubbard SJ. Conservation of orientation and sequence in protein domain–domain interactions. J Mol Biol 2005;345(5):1265–79.
[35] Aloy P, Ceulemans H, Stark A, Russell RB. The relationship between sequence and interaction divergence in proteins. J Mol Biol 2003;332(5):989–98.
[36] Gupta SK, Srivastava M, Osmanoglu O, Dandekar T. Genome-wide inference of the Camponotus floridanus protein-protein interaction network using homologous mapping and interacting domain profile pairs. Sci Rep 2020;10:2334.
[37] Itzhaki Z, Akiva E, Altuvia Y, Margalit H. Evolutionary conservation of domain-domain interactions. Genome Biol 2006;7:R125.
[38] Bencurova E, Gupta S, Sarukhanyan E, Dandekar T. Identification of Antifungal Targets Based on Computer Modeling. J Fungi (Basel) 2018;4(3):81. https://doi.org/10.3390/jof4030081.
[39] Gupta SK, Gross R, Dandekar T. An antibiotic target ranking and prioritization pipeline combining sequence, structure and network-based approaches exemplified for Serratia marcescens. Gene 2016;591(1):268–78.
[40] Yang H, Ke Y, Wang J, Tan Y, Myeni SK, Li D, et al. Insight into bacterial virulence mechanisms against host immune response via the Yersinia pestis-human protein-protein interaction network. Infect Immun 2011;79 (11):4413–24.
[41] Crua Asensio N, Munoz Giner E, de Groot NS, Torrent BM. Centrality in the host-pathogen interactome is associated with pathogen fitness during infection. Nat Commun 2017;8:14092.
[42] Kandasamy K, Mohan S, Raju R, Keerthikumar S, Kumar GSSAMEER, Venugopal AK, et al. NetPath: a public resource of curated signal transduction pathways. Genome Biol 2010;11(1):R3. https://doi.org/10.1186/gb-2010-11-1-r3.
[43] Nishimura D. BioCarta. Biotech Software & Internet Report: The Computer Software Journal for Scient. 2001;2:117-20.
[44] Kanehisa M, Araki M, Goto S, Hattori M, Hirakawa M, Itoh M, et al. KEGG for linking genomes to life and the environment. Nucleic Acids Res 2008;36: D480–4.
[45] Vastrik I, D'Eustachio P, Schmidt E, Joshi-Tope G, Gopinath G, Croft D, et al. Reactome: a knowledge base of biologic pathways and processes. Genome Biol 2007;8(3):R39. https://doi.org/10.1186/gb-2007-8-3-r39.
[46] Brinkmann C, Nehlmeier I, Walendy-Gnirß K, Nehls J, González Hernández M, Hoffmann M, et al. The Tetherin Antagonism of the Ebola Virus Glycoprotein Requires an Intact Receptor-Binding Domain and Can Be Blocked by GP1-Specific Antibodies. J Virol 2016;90(24):11075–86.
[47] Neil SJD, Zang T, Bieniasz PD. Tetherin inhibits retrovirus release and is antagonized by HIV-1 Vpu. Nature 2008;451(7177):425–30.
[48] Taylor JK, Coleman CM, Postel S, Sisk JM, Bernbaum JG, Venkataraman T, et al. Severe Acute Respiratory Syndrome Coronavirus ORF7a Inhibits Bone Marrow Stromal Antigen 2 Virion Tethering through a Novel Mechanism of Glycosylation Interference. J Virol 2015;89(23):11820–33.
[49] Dolskiy AA, Bodnev SA, Nazarenko AA, Smirnova AM, Pyankova OG, Matveeva AK, et al. Deletion of BST2 Cytoplasmic and Transmembrane N-Terminal Domains Results in SARS-CoV, SARS-CoV-2, and Influenza Virus Production Suppression in a Vero Cell Line. Front Mol Biosci. 2020;7. https://doi.org/10.3389/fmolb.2020.616798.
[50] Martin-Sancho L, Lewinski MK, Pache L, Stoneham CA, Yin X, Pratt D, et al. Functional Landscape of SARS-CoV-2 Cellular Restriction. bioRxiv. 2020.
[51] Stewart H, Johansen KH, McGovern N, Palmulli R, Carnell GW, Heeney JL, et al. SARS-CoV-2 spike downregulates tetherin to enhance viral spread. bioRxiv. 2021.
[52] Zhai Y, Wang C, Jiang Z. Cross-talk between bacterial PAMPs and host PRRs. Natl Sci Rev 2018; 5: 791-2.
[53] Pachathundikandi K, Backert S. Heptose 1,7-Bisphosphate Directed TIFA Oligomerization: A Novel PAMP-Recognizing Signaling Platform in the Control of Bacterial Infections. Gastroenterology 2018;154(4):778–83.
[54] Capewell P, Bucheton B, Clucas C, Ilboudo H, Cooper A, Gorman T-A, et al. T cell activation and the HLA locus associate with latent infections of human African trypanosomiasis. bioRxiv. 2017:184762.
[55] Phizicky EM, Fields S. Protein-protein interactions: methods for detection and analysis. Microbiol Rev 1995;59(1):94–123.
[56] Dyer MD, Murali TM, Sobral BW. Computational prediction of host-pathogen protein-protein interactions. Bioinformatics. 2007;23:i159-66.
[57] Fraser HB, Hirsh AE, Steinmetz LM, Scharfe C, Feldman MW. Evolutionary rate in the protein interaction network. Science 2002;296:750–2.
[58] Pichlmair A, Kandasamy K, Alvisi G, Mulhern O, Sacco R, Habjan M, et al. Viral immune modulators perturb the human molecular network by common and unique strategies. Nature 2012;487(7408):486–90.
[59] García-Dorival I, Wu W, Dowall S, Armstrong S, Touzelet O, Wastling J, et al. Elucidation of the Ebola virus VP24 cellular interactome and disruption of virus biology through targeted inhibition of host-cell protein function. J Proteome Res 2014;13(11):5120–35.
[60] Reuschl AK, Thorne LG, Zuliani-Alvarez L, Bouhaddou M, Obernier K, Hiatt J, et al. Host-directed therapies against early-lineage SARS-CoV-2 retain efficacy against B.1.1.7 variant. bioRxiv 2021.
[61] White KM, Rosales R, Yildiz S, Kehrer T, Miorin L, Moreno E, et al. Plitidepsin has potent preclinical efficacy against SARS-CoV-2 by targeting the host protein eEF1A. Science 2021;371(6532):926–31.
[62] de Breyne S, Vindry C, Guillin O, Conde L, Mure F, Gruffat H, et al. Translational control of coronaviruses. Nucleic Acids Res. 2020; 48: 12502-22.
[63] Bavari S, Bosio CM, Wiegand E, Ruthel G, Will AB, Geisbert TW, et al. Lipid raft microdomains: a gateway for compartmentalized trafficking of Ebola and Marburg viruses. J Exp Med 2002;195:593-602.
[64] Falasca L, Agrati C, Petrosillo N, Di Caro A, Capobianchi MR, Ippolito G, et al. Molecular mechanisms of Ebola virus pathogenesis: focus on cell death. Cell Death Differ 2015;22(8):1250–9.

[65] Furuyama W, Marzi A, Carmody AB, Maruyama J, Kuroda M, Miyamoto H, et al. Fcgamma-receptor IIa-mediated Src Signaling Pathway Is Essential for the Antibody-Dependent Enhancement of Ebola Virus Infection. PLoS Pathog 2016;12:e1006139.
[66] Takada A, Feldmann H, Ksiazek TG, Kawaoka Y. Antibody-dependent enhancement of Ebola virus infection. J Virol 2003;77(13):7539–44.
[67] Moldovan G-L, Pfander B, Jentsch S. PCNA, the maestro of the replication fork. Cell 2007;129(4):665–79.
[68] Kuhl A, Pohlmann S. How Ebola virus counters the interferon system. Zoonoses Public Health. 2012;59(Suppl 2):116–31.
[69] Douglas JL, Gustin JK, Viswanathan K, Mansouri M, Moses AV, Früh K, et al. The great escape: viral strategies to counter BST-2/tetherin. PLoS Pathog 2010;6(5):e1000913.
[70] Le Tortorec A, Willey S, Neil SJD. Antiviral inhibition of enveloped virus release by tetherin/BST-2: action and counteraction. Viruses. 2011;3(5):520–40.
[71] Robertson IB, Horiguchi M, Zilberberg L, Dabovic B, Hadjiolova K, Rifkin DB. Latent TGF-beta-binding proteins. Matrix Biol 2015;47:44–53.
[72] Donelson JE, Hill KL, El-Sayed NM. Multiple mechanisms of immune evasion by African trypanosomes. Mol Biochem Parasitol 1998;91:51–66.
[73] van der Woude MW, Bäumler AJ. Phase and antigenic variation in bacteria. Clin Microbiol Rev 2004;17(3):581–611.
[74] Vincendeau P, Bouteille B. Immunology and immunopathology of African trypanosomiasis. An Acad Bras Cienc. 2006;78(4):645–65.
[75] Coureuil M, Join-Lambert O, Lecuyer H, Bourdoulous S, Marullo S, Nassif X. Pathogenesis of meningococcemia. Cold Spring Harb Perspect Med. 2013;3.
[76] Kánová E, Jiménez-Munguía I, Majerová P, Tkáčová Z, Bhide K, Mertinková P, et al. Deciphering the Interactome of Neisseria meningitidis With Human Brain Microvascular Endothelial Cells. Front Microbiol 2018;9:2294.
[77] Talà A, Cogli L, De Stefano M, Cammarota M, Spinosa MR, Bucci C, et al. Serogroup-specific interaction of Neisseria meningitidis capsular polysaccharide with host cell microtubules and effects on tubulin polymerization. Infect Immun 2014;82(1):265–74.
[78] Domenicali Pfister D, Burkard G, Morand S, Renggli CK, Roditi I, Vassella E. A Mitogen-activated protein kinase controls differentiation of bloodstream forms of Trypanosoma brucei. Eukaryot Cell 2006;5(7):1126–35.
[79] Rajao MA, Passos-Silva DG, DaRocha WD, Franco GR, Macedo AM, Pena SD, et al. DNA polymerase kappa from Trypanosoma cruzi localizes to the mitochondria, bypasses 8-oxoguanine lesions and performs DNA synthesis in a recombination intermediate. Mol Microbiol. 2009;71:185-97.
[80] Kruzel EK, Zimmett GP, 3rd, Bangs JD. Life Stage-Specific Cargo Receptors Facilitate Glycosylphosphatidylinositol-Anchored Surface Coat Protein Transport in Trypanosoma brucei. mSphere. 2017;2.
[81] Abbas W, Kumar A, Herbein G. The eEF1A Proteins: At the Crossroads of Oncogenesis, Apoptosis, and Viral Infections. Front Oncol 2015;5:75.
[82] Wei T, Li D, Marcial D, Khan M, Lin M-H, Snape N, et al. The eukaryotic elongation factor 1A is critical for genome replication of the paramyxovirus respiratory syncytial virus. PLoS ONE 2014;9(12):e114447.
[83] Nandan D, Reiner NE. Leishmania donovani engages in regulatory interference by targeting macrophage protein tyrosine phosphatase SHP-1. Clin Immunol 2005;114(3):266–77.
[84] Matsubayashi M, Teramoto-Kimata I, Uni S, Lillehoj HS, Matsuda H, Furuya M, et al. Elongation factor-1alpha is a novel protein associated with host cell invasion and a potential protective antigen of Cryptosporidium parvum. J Biol Chem 2013;288:34111–20.
[85] Coureuil M, Join-Lambert O, Lécuyer H, Bourdoulous S, Marullo S, Nassif X. Mechanism of meningeal invasion by Neisseria meningitidis. Virulence. 2012;3(2):164–72.
[86] Dyer MD, Neff C, Dufford M, Rivera CG, Shattuck D, Bassaganya-Riera J, et al. The human-bacterial pathogen protein interaction networks of Bacillus anthracis, Francisella tularensis, and Yersinia pestis. PLoS ONE 2010;5(8):e12089.
[87] Xenarios I, Rice DW, Salwinski L, Baron MK, Marcotte EM, Eisenberg D. DIP: the database of interacting proteins. Nucleic Acids Res 2000;28:289–91.
[88] Yan Q, Lannarz WJ. Oligosaccharyltransferase: a complex multisubunit enzyme of the endoplasmic reticulum. Biochem Biophys Res Commun . 1999;266(3):684–9. https://doi.org/10.1006/bbrc.1999.1886.
[89] Passos-Silva DG et al. Overview of DNA Repair in Trypanosoma cruzi, Trypanosoma brucei, and Leishmania major. J Nucleic Acids. 2010(840768). https://doi.org/10.4061/2010/840768.