bioRxiv preprint doi: https://doi.org/10.1101/147868; this version posted June 8, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

An integrative method to unravel the host-parasite interactome: an orthology-based approach

Yesid Cuesta-Astroz[1,2] ¶, Alberto Santos[4] ¶, Guilherme Oliveira[1,3*], Lars J. Jensen[4*]

[1] Centro de Pesquisas René Rachou (CpqRR), Fundação Oswaldo Cruz (FIOCRUZ), Belo Horizonte, Brazil

[2] Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais (UFMG), Belo Horizonte, Brazil

[3] Instituto Tecnologico Vale, Belém, PA, Brazil

[4] Novo Nordisk Foundation Center for Protein Research, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark.

¶These authors have contributed equally to this work.

* Corresponding author

E-mail: lars.juhl.jensen@cpr.ku.dk; guilherme.oliveira@itv.org

Abstract

The study of molecular host–parasite interactions is essential to understand parasitic infection and adaptation within the host system. As well, prevention and treatment of

bioRxiv preprint doi: https://doi.org/10.1101/147868; this version posted June 8, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

infectious diseases require clear understanding of the molecular crosstalk between parasites and their hosts. As yet, experimental large–scale identification of host–parasite molecular interactions remains challenging and the use of *in silico* predictions becomes then necessary. Here, we propose a computational integrative approach to predict host–parasite protein–protein interaction (PPI) networks resulting from the infection of human by 12 different parasites. We used an orthology-based method to transfer high–confidence intra–species interactions obtained from the STRING database to the corresponding inter–species protein pairs in the host–parasite system. To reduce the number of spurious predictions, our approach uses either the parasites predicted secretome and membrane proteins or only the secretome depending on whether they are uni- or multicellular respectively. Besides, the host proteome is filtered for proteins expressed in selected cellular localizations and tissues supporting the parasites growth. We evaluated the inferred interactions by analyzing the enriched biological processes and pathways in the predicted networks and their association to known parasitic invasion and evasion mechanisms. The resulting PPI networks were compared across parasites to identify common mechanisms that may define a global pathogenic hallmark. The predicted PPI networks can be visualized and downloaded at http://orthohpi.jensenlab.org.

Author Summary

A protein-protein interaction (PPI) network is a collection of interactions between proteins from one or more organisms. Host–parasite PPIs are key to understanding the biology of different parasitic diseases, since predicting PPIs enable to know more about the parasite invasion, infection and persist. Our understanding of PPIs between host and parasites is still very limited, as no many systematic experimental studies have so far been performed. Efficacy of treatments for parasitic diseases is limited and in many cases parasites evolve resistance. Thus, there is an urgent need to develop novel drugs or vaccines for these neglected diseases, and thus interest in the functions and interactions of proteins associated with parasitism processes. Here we developed an *in silico* method to shed light on the

bioRxiv preprint doi: https://doi.org/10.1101/147868; this version posted June 8, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

interactome in twelve human parasites by combining an orthology based strategy and integrating, domain–domain interaction data, sub-cellular localization and to give a spatial context, we only took in to account those human tissues that support the parasite´s tropism. Here we show that is possible to identify relevant interactions across different parasites and their human host and that these interactions are well supported based on the biology of the parasites.

Introduction

Parasitic diseases result in millions of deaths each year. For example, *Plasmodium falciparum* is responsible for most of the malaria deaths with an estimated 214 million cases and 438,000 deaths worldwide in 2015 [1]. Around 7 million people worldwide are infected with *Trypanosoma cruzi*, which causes Chagas disease and results in life-long morbidity and disability as well as more than 7,000 deaths per year [1]. Another highly prevalent disease, Leishmaniasis that accounts for 20 to 30 thousand deaths a year and is caused by protozoan parasites of the *Leishmania* genus [1]. Revealing the mechanisms of parasitic infection, evasion of the host's immune system and survival are essential to identify vaccines and therapies against these neglected diseases. One last parasitic disease, which is relevant in this work is Schistosomiasis that is a neglected parasitic disease mainly caused by five species of the genus *Schistosoma*. This parasite has a prevalence estimated in 200 million people worldwide [2]. The number of treatments for schistosomiasis is limited and their efficacy as well; in many cases parasites develop resistance. Thus, there is an urgent need to develop novel drugs or vaccines.

Typically, parasites have complex life cycles with several morphological stages and infect many host cells and tissues. For that, the parasites display a resourceful capacity to live in heterogeneous environmental conditions (intra and extra–cellular parasites) and to the pressure from the hosts' immunological response [3]. For examples, extracellular parasites

bioRxiv preprint doi: https://doi.org/10.1101/147868; this version posted June 8, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

remodel tissues in order to migrate and evade the immune system [4]. Likewise, intracellular parasites shape cellular processes and remodel host's cells to adjust their niche during infection [5]. The manipulation of these processes and pathways happens through molecular interactions that the parasites use in their own advantage.

The study of molecular host–parasite interactions is essential to understand parasitic infection, local adaptation within the host and pathogenesis, in addition these complex interactions can be described as a network [6]. Pathogens affect their hosts partly by interacting with host cells proteins, which defines a molecular interplay between the parasite survival mechanisms and the host defense and metabolic systems [7]. Understanding this molecular crosstalk can provide insights into specific interactions that could be targeted to avoid the parasite's pathogenic consequences [8]. Intra–species protein–protein interactions (PPIs) have been studied in depth and there exist large datasets containing experimentally or computationally predicted interactions [9,10]. However, the number of available datasets providing host–pathogen PPIs is limited and challenged by the intrinsic difficulties of analyzing simultaneously the host and pathogen systems in high-throughput experiments.

Thus, host–pathogen PPIs have mainly been computationally predicted using distinct strategies such as approaches based on sequence [8,11–14], structure [15,16] and gene expression [17]. Homology–based prediction is one of the most common approaches to predict host–pathogen PPIs. These approaches have been extensively used to infer intra–species interactions [10,11,18] as well as host–pathogen PPIs [13–15,19]. These methods assume that interactions between proteins in one species can be transferred to homolog proteins in another species (interologs). Using experimental PPIs and interologs were predicted 3,090 interactions between *P. falciparum* and *H. sapiens*, these interactions were grouped based on their biological, metabolic and cellular processes [15]. Homology based methods are also combined with other approaches in order to asses the quality of the interactions obtained, these additional approaches are for example random forest classifiers and interactions are subsequently filtered for parasite specific characteristics [14].

bioRxiv preprint doi: https://doi.org/10.1101/147868; this version posted June 8, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

In this work, we developed an *in silico* method to unravel the host–parasite interactome across 12 human parasites, namely *Trypanosoma brucei, Trypanosoma cruzi, Trichinella spiralis, Schistosoma mansoni, Giardia lamblia, Plasmodium falciparum, Cryptosporidium hominis, Cryptosporidium parvum, Leishmania braziliensis, Leishmania mexicana, Leishmania donovani and Leishmania infantum*. Our method is based on orthology and its strengths reside in 1) incorporating high–confidence intra–species interactions, which are not limited to known physical interactions but complemented with predictions and interactions mined from the scientific literature [10], 2) using fined grained orthology assignments instead of simple sequence similarity and 3) including parasite–specific biological context such as lifestyle (uni- or multicellular) and tissue infection. The objective of integrating multiple layers of information is to reduce the number of falsely predicted interactions, increase the reliability and thereby provide a better understanding of the parasite's molecular mechanisms. Finally, we propose that this approach can be applied to any host–pathogen system to predict relevant molecular interactions and define the context in which they unfold.

Materials and methods

Proteome filtering: adding parasite-specific biological context

The first step in the prediction pipeline is to filter both the host and parasite proteomes according to the specific characteristics of the studied parasite (Fig 1A). Firstly, for the interaction to happen, proteins in the parasite need to be secreted or membrane proteins depending on the type of parasite (unicellular or multicellular). When analyzing parasites such as helminths (multi–cellular and extracellular) we used the soluble secretome. Conversely, in case of unicellular parasites, we used both the soluble secretome and membrane proteins. To identify soluble and membrane proteins, we used different available bioinformatics tools to predict subcellular localization (Fig 1B). SignalP [20] was used to identify classical secretory proteins. Proteins predicted to be secreted were scanned for the

bioRxiv preprint doi: https://doi.org/10.1101/147868; this version posted June 8, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

presence of mitochondrial sequences by TargetP [21] and transmembrane helices by the transmembrane identification based on hidden Markov model (TMHMM) method [22]. As well, we filtered host proteins to consider only those located on the cellular membrane and extracellular space by using the COMPARTMENTS database, which provides high-confidence information on cellular localization of proteins (confidence score > 3) [23].

Fig 1. Method workflow. A) The method input consists of the host's and parasite's proteomes and specific characteristics of the parasite's attributes (uni-multicellular) and tropism (tissues). Our approach filters the host proteome according to the specified list of tissues using TISSUES database and limits it to only cellular membrane and extracellular proteins extracted from the COMPARTMENTS database. The parasite's proteome is filtered to only membrane or secreted proteins depending on the specified parasite attribute. The next step in the pipeline identifies in eggNOG database, orthologous proteins for both the host's and the parasite's filtered proteomes. The workflow continues with the extraction of intra–species protein–protein interactions from STRING database for all the orthologous proteins. These intra–species interactions are then transferred to the host–parasite system as long as the interacting proteins have an orthologous protein in host and in the parasite. **B)** Workflow for soluble secretome prediction. MT: mitochondrial, TM: transmembrane. **C)** A simple scheme of how the orthology transfer is implemented.

Additionally, our method included context on the specific tissue tropism of the studied parasite by limiting the predictions to host proteins expressed in tissues relevant in the parasite's life cycle. For this purpose, we used the TISSUES database [24], which provides protein profiles of tissue expression. Protein–tissue associations in this database are also scored in similar way to the COMPARTMENTS database, which allowed us to use only high-confidence associations (confidence score > 3).

bioRxiv preprint doi: https://doi.org/10.1101/147868; this version posted June 8, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Orthologs identification and PPIs transfer

Host–parasite PPIs were predicted using orthology-based transfer. This approach relies on cross-species data integration to predict inter-species protein–protein interactions. Conserved intra-species interactions from multiple organisms, namely interologs, are transferred to the host–parasite system when orthologous proteins exist in these species. For example, an intra–species interaction between protein A and protein B is transferred if the host and the parasite have orthologs to A and B (Fig 1C).

To obtain intra-species PPIs, we used the STRING database [10]. STRING provides PPIs from a variety of sources and evidence types, which expands the list of high–confidence PPIs while not limiting it to only known physical interactions. One of the integrated sources in this database is based on homology transfer, however, these transferred interactions, interologs, were omitted in our approach (STRING file including the distinction: direct *vs.* interologs). In this file, each of the evidence channels (Neighborhood, Gene fusion, Co–occurrence, Co–expression, Experiments, Databases and Text-mining) is scored independently so we combined the scores using Eq.(1).

$$Eq.(1)\ score_{combined} = 1 - \frac{1}{(1-p)^{(N-1)}} * \prod_{i=1}^{N}(1 - score_i)$$

, where *i* is the different evidence channels (Neighborhood, Gene Fusion, Co–occurrence, Co–expression, Experiments, Databases and Text-mining) and *p* is the prior probability of two proteins being linked (p = 0.063). We only used high-confidence interactions ($\mathrm{score}_{\mathrm{combined}} > 0.7$).

For each of the interactions from STRING, we used the fine–grained orthologs function in eggNOG database [25] to identify orthologous proteins in human and in the different parasites. We transferred an interaction as long as it involved proteins that had an orthologous protein in the parasite and another one in human, and were in the filtered proteome. Several metrics were collected to facilitate the analyses of the interactions

bioRxiv preprint doi: https://doi.org/10.1101/147868; this version posted June 8, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

predicted: maximum confidence score transferred, maximum confidence score transferred from the Experiments channel in STRING, species from which the interactions are transferred and the eggNOG non-supervised orthologous groups (NOGs) they belong to, and also the number of transferred interactions corrected by the number of paralogs (duplicates).

Domain–domain and linear motif–domain annotations

To know which interactions may be physical — rather than only functional associations — we annotated our interaction predictions with domain–domain interaction predictions from iPfam [26] and 3did [27], and linear motif–domain interactions from ELM database [28]. These databases provide predictions based on structural information from the Protein Data Bank [29]. Protein domains were predicted using Pfam scan, which combines the HMMER tool [30] and the domain models from Pfam version 29 [31]. Linear motif–domain interactions are predicted using the regular expressions provided in the ELM database.

Network analysis

Once we obtained the predicted host–parasite PPI networks, we used the topology of the network to identify relevant proteins, which may play critical roles in the host–parasite crosstalk. We used betweenness centrality to pinpoint the proteins whose targeting would most disrupt this communication [32].

To identify key biological processes enriched in the predicted host–parasite PPI networks, we performed functional enrichment analysis using Gene Ontology biological processes (omitting the inferred evidence codes) [33] and Reactome pathways annotations [34]. We focused only on the host proteins predicted to interact given the limited functional annotation for the studied parasites. The functional enrichment was performed using Fisher's exact test and correction for multiple testing (FDR correction Benjamini–Hochberg (BH); 0.05). The enrichment was calculated using as background only the filtered proteome.

bioRxiv preprint doi: https://doi.org/10.1101/147868; this version posted June 8, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

To overcome the lack of functional characterization of the studied parasites, we also investigated the functional classification of COGs [35] provided by the eggNOG database [25]. These annotations classify COGs into functional categories that can be used to characterize the proteins grouped in these clusters (S6 Fig). These categories were transferred to the parasite proteins in the network and when the category was "Function unknown" we assigned as putative functions the categories of their interaction partners in human (S1 Table).

Web interface

To provide access to the predictions generated with our approach, we developed a web interface for OrthoHPI (http://orthohpi.jensenlab.org/). This web site provides interactive, predicted host–parasite PPI networks, built with d3.js (https://d3js.org/), which allow the user to easily navigate the full networks as well as the tissue–specific ones. The nodes in these networks represent parasite and human proteins and their sizes correspond to their betweenness centrality in the network. The edges between nodes show predicted molecular host–parasite interactions and are weighted using the maximum score transferred from STRING database. The predictions can be downloaded in tab separated values file format (tsv) or in Graph Modeling Language format (gml), which is compatible with Cytoscape [36].

Results and discussion

We applied our integrative orthology-based approach to predict the host–parasite PPIs networks for 12 different parasites (Table 1; Fig 1). The method returned a total of 10,096 interactions for 6,039 proteins (5,265 host proteins, 770 parasite proteins) being *T. cruzi* and *G. lamblia* the largest and the smallest predicted PPIs networks, respectively (Table 2; Fig 2A). The overview network of all 12 parasites (Fig 2A) shows that there are predicted interactions common to all parasites and parasite–specific ones. This distinction allowed us to investigate global, parasite– or genus–specific mechanisms. Our prediction method

bioRxiv preprint doi: https://doi.org/10.1101/147868; this version posted June 8, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

transferred most of the STRING high-confidence intra–species interactions from model organisms (*M. musculus*, *D. rerio*, *D. melanogaster*, *X. tropicalis*, etc.) (Fig 2B).

Tax_id	Species	E/I	U/M	Tissues/Cell type	Proteome size	SS (%)	MP (%)	Input
5691	*Trypanosoma brucei*	E	U	Skin, blood, lymph	8,747	534 (6,1)	330 (3,7)	SS+MP
353153	*Trypanosoma cruzi*	I	U	Skin, heart, blood, eye, macrophages	19,602	1,623 (8,2)	1,423 (7,2)	SS+MP
6334	*Trichinella spiralis*	E	M	Intestine, brain, heart, muscle, lung	16,380	1,338 (8,1)	488 (2,9)	SS
6183	*Schistosoma mansoni*	E	M	Skin, liver, intestine, lung, blood	11,770	659 (5,6)	252 (2,1)	SS
184922	*Giardia lamblia*	E	U	Intestine	6,502	361 (5,5)	251 (3,8)	SS+MP
5833	*Plasmodium falciparum*	I	U	Skin, blood, liver, erythrocytes, hepatocytes	5,429	280 (5,1)	358 (6,5)	SS+MP
353151	*Cryptosporidium hominis*	E	U	Intestine	3,885	244	137	SS+MP
353152	*Cryptosporidum parvum*	E	U	Intestine	3,805	273	156	SS+MP
420245	*Leishmania braziliensis*	I	U	Skin, nose, mouth, blood	8,160	208	225	SS+MP
929439	*Leishmania mexicana*	I	U	Skin, nose, mouth, blood	8,147	194	225	SS+MP
5661	*Leishmania donovani*	I	U	Skin, liver, spleen, blood, bone marrow	8,032	212	223	SS+MP
435258	*Leishmania infantum*	I	U	Skin, liver, spleen, blood, bone marrow	8,150	211	233	SS+MP

Table 1. Human parasites analyzed. E: Extracellular, I: Intracellular, Cell type: host cell of an intracellular parasite. Tissues: different tissues associated with the parasite´s tropism.

bioRxiv preprint doi: https://doi.org/10.1101/147868; this version posted June 8, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Proteome size: sequences in the predicted proteome. SS: Soluble secretome, MP: Membrane proteins (%: is the percentage in relation to the overall proteome.), U: unicellular, M: multicellular.

Fig 2. Predicted interactome. A) This panel shows all the predicted interactions into a single network for all the parasites and how the different parasite proteins are distributed unraveling host–parasite–specific interactions, genus–specific or shared among several parasite species. **B)** In this figure we present the species that contribute the most in the PPI transfer. Most of the interologs transferred to the host–parasite system corresponds to model organisms, which account for most of the high–confidence PPI in STRING database.

The resulting host–parasite PPIs networks were analyzed to identify central proteins in the networks, tissue-specific connections and enriched biological processes and pathways. The lack of experimentally validated host–parasite interactions that could be used as a gold–standard impedes validating the quality of the predicted interactions. Instead, we validated the plausibility of the network by looking at the known functions in which the involved proteins participate. The analyses are divided into two: a case study focusing on the in-depth examination of the human–*S. mansoni* predicted interactome, and a global study of the common mechanisms targeted by the studied parasites and the shared human interactors.

Name	Parasite proteins	Host proteins	Parasite nodes	Host nodes	Total nodes	Total edges	Avg. degree
C. parvum	397	1003	13	60	77	80	1.039
T. cruzi	4056	6000	156	991	1147	3062	2.67
P. falciparum	977	7513	83	747	830	1099	1.324

bioRxiv preprint doi: https://doi.org/10.1101/147868; this version posted June 8, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

L. braziliensis	294	4879	41	245	286	405	1.416
G. lamblia	658	1041	11	18	29	23	0.79
C. hominis	324	1003	5	20	25	20	0.8
S. mansoni	318	3503	69	429	498	581	1.16
L. infantum	325	8090	57	755	812	1164	1.433
T. brucei	1038	1824	101	265	366	528	1.442
L. donovani	310	8090	55	576	631	954	1.512
T. spiralis	705	4317	128	739	867	1389	1.6
L. mexicana	305	4879	51	420	471	791	1.677

Table 2. Human–parasite PPI networks. This table contains the number of host and parasite nodes, the number of interactions and average degree of the predicted interactomes.

bioRxiv preprint doi: https://doi.org/10.1101/147868; this version posted June 8, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Case study: Human – *S. mansoni* interactome

Network topology analysis. In this analysis, we use the topological characteristics of the predicted human–*S. mansoni* PPI network to identify central proteins. Network centrality helps prioritizing proteins that have a relevant role in the network, which may translate into biologically relevant roles as well. In the human–*S. mansoni* interactome network (Fig 3) the nodes with the highest centrality were alpha-2 macroglobulin (Smp_089670), heat shock protein 70 (Smp_049550), laminin subunit beta 1 (Smp_148790) and the cell adhesion molecule Smp_171460 (Fig 4). A literature search revealed that these parasite proteins are indeed involved in crucial biological processes needed for invasion and survival of the parasite: protease activity regulation, inhibition of blood coagulation, cell adhesion and migration [37]. Our results predict that the *S. mansoni* protein Alpha-2 macroglobulin (Smp_089670) interacts with host proteins involved in matrix extracellular organization such as SERPINE1 or metalloproteases (MMP3, MMP13, MMP8 and MMP1). Interactions with the extracellular matrix (ECM) components influence several biological processes and ultimately the fate of the host cell [38]. Parasites encounter the ECM as a barrier and they deploy several mechanisms to overcome it such as cell adhesion, induction of ECM degradation and regulation of immune response [4,39,40]. For example, Alpha-2 macroglobulin is known to inhibit the predicted four proteinases by a 'trapping' mechanism that limits the access of the proteinases to their substrates [37] . Laminin subunit beta 1 (Smp_148790) and Smp_171460 are involved in cell adhesion and have interacting partners belonging to the metalloprotease, collagenase and laminin families.

Fig 3. Human–*S. mansoni* predicted interactome. The layout of the nodes in the network (ellipses: human proteins, diamonds: *S. mansoni* proteins) uses Markov Clustering to group and locate the nodes. The tools used to visualize the networks in this article are: Cytoscape [36] and the clusterMaker app [41].

bioRxiv preprint doi: https://doi.org/10.1101/147868; this version posted June 8, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Fig 4. Topology analysis. This network shows the selected central nodes in the Human–*S. mansoni* predicted interactome (ellipses: human proteins, diamonds: S mansoni proteins). The central nodes are calculated using betweenness centrality and highlighted in the figure with a higher size.

The analysis highlighted as well a human protein central in the predicted interactome, PLEC protein, which strengthens cells and tissues by acting as a cross-linking element of the cytoskeleton [42]. *S. mansoni* interaction with this protein may compromise cell and tissue integrity (cell junction) and may be used to migrate across different tissues. This alteration of cell junction and tissue remodeling has been found in some helminths parasites [43].

bioRxiv preprint doi: https://doi.org/10.1101/147868; this version posted June 8, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Domain and linear motif analysis. We annotated the interacting proteins in the predicted human–*S. mansoni* PPI networks with domain and linear motif predictions from Pfam scan and ELM database respectively [28,30]. Around 20% of the proteins in the network (94 human proteins, 16 parasite proteins) contained interacting domains or linear motifs (Fig 5). The domain–domain or linear motif–domain interactions were integrated from iPfam, 3did and ELM databases [26–28]. The most prevalent domain in the parasite proteins is the 'Papain family cysteine protease' (predicted in Smp_019030, Smp_139160 and Smp_193000). This family of proteins has already been validated as drug targets for many parasites including *P. falciparum* and the African and American trypanosomes [44]. Another relevant conserved sequence is the 'Calcineurin (PP2B)-docking motif LxvP' present in the parasite central node Smp_171460 and the putative alkaline phosphatase Smp_171460. Calcineurin signaling manipulation provides influence over several physiological contexts such as cell adhesion, immune response, stress and post–translationally controlled events [45,46]. Similarly, several human proteins in the network contain kinase domains (AKT1, CDK1, AURKB, TAOK1, ROCK2 and STK10), which would affirm the delivered leverage of the parasite of the host cellular signaling [47–50].

Fig 5. Domain and linear–motif analysis. This network shows in dark green those nodes that have a supporting domain or linear–motif predicted to interact according to iPfam, 3did or ELM. Interacting domains or linear–motifs were not found for the light green nodes. (ellipses: human proteins, diamonds: *S mansoni* proteins)

bioRxiv preprint doi: https://doi.org/10.1101/147868; this version posted June 8, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Tissue–specific network analysis. The human–*S. mansoni* interaction network is not static and the tissue expression data used to filter the host proteome can be used also to give context to the predicted molecular associations. Modeling spatial context into the predicted network allowed identification of relevant tissue-specific interactions through the parasite's life cycle (S1-S5 Figs). For instance, in the blood-specific network (S1 Fig), we identified several interactions between alpha 2 macroglobulin and coagulation factors. These interactions could result in inhibition of the coagulation factors, which may facilitate the migration of the parasite through the broken tissue/vessels [37]. Equally compelling is the predicted lungs-specific interaction between the host protein FZD1, previously implicated in the inflammatory response in the alveolus [51], and the parasite protein Smp_151400 (Fig S2).

Specific characteristics of the *S. mansoni* life cycle are clearly displayed by the number of tissue-specific interactions. In this life cycle, skin is the penetration tissue and this event is happens rapidly (approximately 1 hour) [52]. According to our results, the method predicted the fewest number of interactions in this tissue. The next stage in the life cycle after skin involves migration through the bloodstream and lungs to the liver where the parasite develops into an adult male/female schistosome [53]. The number of predicted interactions in lung and liver is higher than in skin. This increase in the number of interactions could be explained by the development of the adult parasite and the energetic requirements associated with the tropism across different host tissues. The pathogenesis of human schistosomiasis initiates when the parasite's eggs are embedded in the lung, liver or intestine, which induces inflammation, granulomas, fibrosis and subsequent morbidities [53].

bioRxiv preprint doi: https://doi.org/10.1101/147868; this version posted June 8, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Functional analysis

Common pathways across parasites. The twelve parasites included in our study are strikingly divergent in terms of phylogenetic context and their biology, for example distinct forms of invasion that leads to differences in parasite's tropism and the invasion of different sets of tissues within the host (Table 1). However, when combining all the predicted PPI networks, we found some common core pathways and biological processes targeted in the host by all the studied parasites. Indeed, it has been previously shown that evolutionary distinct parasites can target the same pathways as a result of convergent evolution, for example in *Arabidopsis* pathogens, such as the bacterium *Pseudomonas syringae* and the eukaryote *Hyaloperonospora arabidopsidis* [54]. Our analysis across the studied parasites of the targeted host pathways confirms this. We used annotations from both biological process GO terms and Reactome pathways in human to get an overview of the shared pathways targeted by the studied parasites.

In the analysis, we found terms enriched and common across 10 interactomes related to metabolism, protein folding and transport, and biosynthesis (Fig 6A). These targeted pathways have been identified already as crucial to *P. falciparum* and other Apicomplexa parasites, for instance the use of the host's amino acid synthesis and modulation of the host metabolic network [55,56]. This host metabolism dependence has also been seen in other parasites such as in Trypanosomatids [57].

bioRxiv preprint doi: https://doi.org/10.1101/147868; this version posted June 8, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Pathways linked to lifestyle. Different lifestyles (intra- and extracellular parasites) may define specific targeted pathways in the host. We analyzed the common processes and pathways enriched in the intracellular and extracellular parasites specifically. The enrichment performed in the extracellular species networks did not return any lifestyle-specific targeted mechanisms, probably consequence of the common behavior of the intra– and extracellular parasites in the studied interactomes. However, the specific lifestyle requirements of the intracellular parasites (*T. cruzi*, *L. braziliensis*, *L. infantum*, *L. donovani*, *L. mexicana* and *P. falciparum*) revealed 11 exclusive biological processes related to exploitative mechanisms to acquire nutrients from the cytosol of host cells (i.e. glycerophospholipid biosynthetic process and glycerolipid metabolic process) carrying out a type of sustainability interactions. Eukaryotic pathogens are able to synthesize a number of nutrients required for growth *de novo,* however it is more advantageous to conserve energy and take host derived resources [55,58]. The enriched processes are related to the acquisition of host lipids necessary for the parasites to assemble a large amount of new membranes during replication within host cells. In terms of host immune response to parasitic infection we obtained the GO term common to intracellular parasites: antigen processing and presentation of exogenous peptide antigen via MHC class I. This term indicate that innate sensing of parasites is important for the induction of pro-inflammatory responses aimed at controlling infection [58,59].

Reactome pathways annotation provided 2 additional relevant pathways exclusive to intracellular parasites: scavenging by class F receptors and scavenger receptors (SR). These pathways include SR-A, MARCO and CD36 proteins, which belong to the innate immune defense by acting as pattern-recognition receptors, in particular against bacterial pathogens [60].

Scavenger receptors (SR) are expressed on macrophages and dendritic cells and intracellular pathogens such as hepatitis C virus or *P. falciparum* utilize the SR to gain entry into host cells [60]. The scavenger receptor MARCO is involved in *L. major* infection in macrophages [61] and the influence of other SR on *Leishmania* macrophage infection has

bioRxiv preprint doi: https://doi.org/10.1101/147868; this version posted June 8, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

been demonstrated [62,63]. Scavenger receptors are also involved in the formation and maturation of parasitophorous vacuole in intracellular parasites [64]. These vacuoles contain intracellular parasites initially inside the host cell and are thought to be rich in nutrients essential for parasite survival [65]. The formation of these vacuoles involves a host cell membrane invagination process, which is related to the reactome pathway Scavenging by Class F Receptors and also to the endocytic vesicle membrane [34].

***Leishmania*-specific pathways.** Across the *Leishmania* species, we identified genus-specific biological processes related to lipid metabolic process and more specifically to fatty acid metabolism. These processes may be associated with how *Leishmania* species deplete membrane cholesterol and disrupt lipid rafts in host macrophages during the invasion process [65]. Therapies based on restoration of cholesterol levels and raft-associated proteins have been explored as promising strategies [61].

In species that cause mucocutaneous leishmaniasis (*L. braziliensis, L. mexicana*), we found 2 specific processes: regulation of acetyl-CoA biosynthetic process from pyruvate and regulation of acyl-CoA biosynthetic process. Regarding visceral leishmania (*L. donovani, L. infantum*) we found 28 specific GO terms associated with relevant biological process in the host-parasite interaction such as: regulation of defense response to virus by virus (GO:0050690), regulation of translation in response to stress (GO:0043555), glycosylceramide catabolic process (GO:0046477) among others.

In terms of Reactome pathways in *Leishmania* 4 pathways were specific. Pathways such as signal transduction (R-HSA-1227986), immune system (R-HSA-1236977), Metabolism (R-HSA-1483255) to be more specific in the phosphatidylinositol metabolism, which is a component of the host cell membrane, and the other pathway was also involved in synthesis of plasma membrane intrinsic proteins (R-HSA-1660499). In mucocutaneous leishmania we obtained only 1 pathway specific, which is regulation of pyruvate dehydrogenase (PDH) complex (R-HSA-204174). Regarding visceral leishmaniasis we found 22 specific pathways related with response to stress, metabolism of proteins, disease (cancer) among others.

bioRxiv preprint doi: https://doi.org/10.1101/147868; this version posted June 8, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Helminth-specific pathways. In helminths (*S. mansoni, T.spiralis*) we identified 18 specific biological process, these were involved in coagulation such as hemostasis, blood coagulation and regulation of body fluid levels among others. Several helminth parasites imbibe host blood, including the hookworms, the flukes and the major livestock nematode parasites [66]. Blood coagulation is triggered by different pathways, which are targeted by blood-feeding parasites to inhibit coagulation and prolong blood flow [66]. In helminths, the inhibition of blood coagulation is related with proteinases that facilitate invasion of host tissues and digest host proteins. Additionally parasite proteinases help pathogens evade the host immune response and prevent blood coagulation [67].

Helminths-specific enriched Reactome pathways targeted host mechanisms: Sphingolipid metabolism and glycosphingolipid metabolism, which are known to have a key role in the interaction host–helminths. The sphingolipids and glycosphingolipids are a class of lipids that serve as integral components of eukaryotic cell membranes and act as signaling molecules in many cellular functions and play crucial roles in the regulation of pathobiological processes [68]. In helminths lipids may have a role in pathogenesis by helping the larvae to survive in the host tissue [68].

Common targeted host proteins across interactomes

In this analysis, we studied the common host proteins to all or most of the studied parasites (Fig 6B), which could help identify possible general therapeutic targets. In all the inferred interactomes, GANAB (neutral alpha glucosidase AB) and P4HB (protein disulfide isomerase) proteins were predicted to interact with parasite proteins.

Fig 6. **Targeted host processes and proteins. A)** This figure shows some of the common and most relevant biological processes (GO BP) targeted by the parasites. **B)** The studied parasites interact with several common proteins some of which have been already linked to

bioRxiv preprint doi: https://doi.org/10.1101/147868; this version posted June 8, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

pathogenic activity. Both the targeted processes and pathways and these common host interacting proteins may open new therapeutic possibilities.

GANAB protein is related with the alteration of eosinophil proteome [69]. Strikingly, eosinophils are important mediators of allergies, asthma and adverse dug reactions and are also related to the host defense mechanisms against helminth infections, which are characterized by eosinophilia [70–72]. Straub et al., 2011 [69] conducted a comparative proteomic analysis of eonisophils (healthy vs hypereosinophilia from acute fascioliasis) and GANAB was one of the four proteins significantly upregulated in the *Fasciola* patient.

Protein disulfide isomerase (P4HB) has a key role in the internalization of some pathogens [73]. For example, during the host invasion process of *L. chagasi* increased levels of P4HB were found to induced phagocytosis in the promastigote phase and inhibition of expression of this gene reduced phagocytosis [74]. The role of P4HB has also been associated to other pathogens such as HIV, dengue virus or rotavirus [75–77]. In dengue virus, P4HB was linked to reduction of β1 and β3 integrins allowing for the entry of the virus [76] and in MA104 cells, thiol blockers and P4HB inhibitors decreased the entry of rotavirus [77]. These evidences show that GANAB and P4HB are relevant proteins not only in our 12 interactomes, but also in other host-pathogen systems, which confirm these proteins as possible hallmarks of the host-parasite interaction. Other host proteins common in several parasite interactomes (Fig 7) such as HYOU1, PDIA3, HSPA9, CTSS, etc., have been already found deregulated upon pathogenic infection [78–83].

The OrthoHPI web resource

We developed a web resource (http://orthohpi.jensenlab.org) to provide all the predicted interactomes for all the studied parasites (Fig 7). The aim of OrthoHPI is to facilitate the analysis of the predicted host–parasite PPI networks, provide easy access to the full and

bioRxiv preprint doi: https://doi.org/10.1101/147868; this version posted June 8, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

tissue–specific networks and produce visual, navigable and interactive networks easily manipulated by users.

Fig 7. The OrthoHPI web resource. The web site (http://orthohpi.jensenlab.org) allows you to visualize the predicted interactomes for all the studied parasites in three simple steps. 1) Select parasite, 2) Choose to visualize the full network or a tissue–specific one and 3) Visualize the interactive network and download the interactions.

Conclusion

Our method showed molecular crosstalk between *S. mansoni* and human proteins enriched in biological processes and tissue-specific interactions essential in the life cycle of the parasite. Moreover, this approach highlighted mechanisms and specific components that may be alternative druggable targets. All these findings indicate the functional relevance of the predicted host-pathogen PPIs. An advantage of this method is that it has clear biological basis (orthology), is scalable and can be applied to many different host–pathogen systems.

Our method also could be used to assign functions to parasite´s hypothetical proteins, which are interacting with known proteins, assuming that clustered proteins tend to have similar functions and functionally related proteins can interact with each other [84–86]. This application would be useful in parasites genomics, considering that a large number of parasite´s proteins are annotated as hypothetical proteins, and our networks provides a useful resource for annotation of those proteins.

Despite the limitations that a computational method may have, our predictions provide an enriched list of host-parasites interactions, and till experimental limitations to predict host-parasite interactions are resolved, the present study or set of interactomes can efficiently serve as a theoretical base for experiments aiming to identify potential drug targets and to know more about the mechanisms to infect, migrate and persist within the host.

bioRxiv preprint doi: https://doi.org/10.1101/147868; this version posted June 8, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

References

1. WHO | World Health Statistics 2015. WHO. World Health Organization; 2016;

2. Adenowo AF, Oyinloye BE, Ogunyinka BI, Kappo AP. Impact of human schistosomiasis in sub-Saharan Africa. Brazilian J Infect Dis. 2015;19: 196–205. doi:10.1016/j.bjid.2014.11.004

3. Cuesta-Astroz Y, Oliveira FS de, Nahum LA, Oliveira G. Helminth secretomes reflect different lifestyles and parasitized hosts. Int J Parasitol. 2017; doi:10.1016/j.ijpara.2017.01.007

4. Piña-Vázquez C, Reyes-López M, Ortíz-Estrada G, de la Garza M, Serrano-Luna J. Host-Parasite Interaction: Parasite-Derived and -Induced Proteases That Degrade Human Extracellular Matrix. J Parasitol Res. Hindawi Publishing Corporation; 2012;2012: 1–24. doi:10.1155/2012/748206

5. Nelson MM, Jones AR, Carmen JC, Sinai AP, Burchmore R, Wastling JM. Modulation of the host cell proteome by the intracellular apicomplexan parasite Toxoplasma gondii. Infect Immun. American Society for Microbiology (ASM); 2008;76: 828–44. doi:10.1128/IAI.01115-07

6. Dallas T, Park AW, Drake JM. Predicting cryptic links in host-parasite networks. Koella J, editor. PLoS Comput Biol. 2017;13: e1005557. doi:10.1371/journal.pcbi.1005557

7. Salzet M, Capron A, Stefano GB, Capron A, Dessaint JP, Capron A, et al. Molecular Crosstalk in Host–Parasite Relationships: Parasitol Today. Elsevier; 2000;16: 536–540. doi:10.1016/S0169-4758(00)01787-7

8. Taylor CM, Fischer K, Abubucker S, Wang Z, Martin J, Jiang D, et al. Targeting protein-protein interactions for parasite control. PLoS One. Public Library of Science; 2011;6: e18381. doi:10.1371/journal.pone.0018381

bioRxiv preprint doi: https://doi.org/10.1101/147868; this version posted June 8, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

9. Stark C, Breitkreutz BJ, Chatr-Aryamontri A, Boucher L, Oughtred R, Livstone MS, et al. The BioGRID Interaction Database: 2011 update. Nucleic Acids Res. 2011;

10. Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, Huerta-Cepas J, et al. STRING v10: protein-protein interaction networks, integrated over the tree of life. Nucleic Acids Res. 2015;43: D447-52. doi:10.1093/nar/gku1003

11. Yu H, Luscombe NM, Lu HX, Zhu X, Xia Y, Han J-DJ, et al. Annotation transfer between genomes: protein-protein interologs and protein-DNA regulogs. Genome Res. Cold Spring Harbor Laboratory Press; 2004;14: 1107–18. doi:10.1101/gr.1774904

12. Kotelnikova E, Kalinin A, Yuryev A, Maslov S. Prediction of protein-protein interactions on the basis of evolutionary conservation of protein functions. Evol Bioinform Online. Libertas Academica; 2007;3: 197–206.

13. Zhou H, Gao S, Nguyen N, Fan M, Jin J, Liu B, et al. Stringent homology-based prediction of H. sapiens-M. tuberculosis H37Rv protein-protein interactions. Biol Direct. BioMed Central; 2014;9: 5. doi:10.1186/1745-6150-9-5

14. Wuchty S. Computational Prediction of Host-Parasite Protein Interactions between P. falciparum and H. sapiens. Borrmann S, editor. PLoS One. 2011;6: e26960. doi:10.1371/journal.pone.0026960

15. Lee S-A, Chan C, Tsai C-H, Lai J-M, Wang F-S, Kao C-Y, et al. Ortholog-based protein-protein interaction prediction and its application to inter-species interactions. BMC Bioinformatics. BioMed Central; 2008;9 Suppl 12: S11. doi:10.1186/1471-2105-9-S12-S11

16. Davis FP, Barkan DT, Eswar N, McKerrow JH, Sali A. Host-pathogen protein interactions predicted by comparative modeling. Protein Sci. 2007;16: 2585–2596. doi:10.1110/ps.073228407

17. Reid AJ, Berriman M. Genes involved in host-parasite interactions can be revealed by

bioRxiv preprint doi: https://doi.org/10.1101/147868; this version posted June 8, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

their correlated expression. Nucleic Acids Res. 2013;41: 1508–1518. doi:10.1093/nar/gks1340

18. Murakami Y, Mizuguchi K. Homology-based prediction of interactions between proteins using Averaged One-Dependence Estimators. BMC Bioinformatics. 2014;15: 213. doi:10.1186/1471-2105-15-213

19. Ramakrishnan G, Srinivasan N, Padmapriya P, Natarajan V. Homology-Based Prediction of Potential Protein-Protein Interactions between Human Erythrocytes and Plasmodium falciparum. Bioinform Biol Insights. Libertas Academica; 2015;9: 195–206. doi:10.4137/BBI.S31880

20. Petersen TN, Brunak S, von Heijne G, Nielsen H. SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat Methods. 2011;8: 785–6. doi:10.1038/nmeth.1701

21. Emanuelsson O, Nielsen H, Brunak S, von Heijne G. Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. J Mol Biol. 2000;300: 1005–1016. doi:10.1006/jmbi.2000.3903

22. Krogh a, Larsson B, von Heijne G, Sonnhammer EL. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J Mol Biol. 2001;305: 567–580. doi:10.1006/jmbi.2000.4315

23. Binder JX, Pletscher-Frankild S, Tsafou K, Stolte C, O'Donoghue SI, Schneider R, et al. COMPARTMENTS: unification and visualization of protein subcellular localization evidence. Database. 2014;2014: bau012-bau012. doi:10.1093/database/bau012

24. Santos A, Tsafou K, Stolte C, Pletscher-Frankild S, O'Donoghue SI, Jensen LJ. Comprehensive comparison of large-scale tissue expression datasets. PeerJ Inc.; 2015; doi:10.7287/peerj.preprints.1072v1

25. Huerta-Cepas J, Szklarczyk D, Forslund K, Cook H, Heller D, Walter MC, et al. eggNOG 4.5: a hierarchical orthology framework with improved functional annotations

bioRxiv preprint doi: https://doi.org/10.1101/147868; this version posted June 8, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

for eukaryotic, prokaryotic and viral sequences. Nucleic Acids Res. Oxford University Press; 2016;44: D286-93. doi:10.1093/nar/gkv1248

26. Finn RD, Miller BL, Clements J, Bateman A. iPfam: a database of protein family and domain interactions found in the Protein Data Bank. Nucleic Acids Res. Oxford University Press; 2014;42: D364-73. doi:10.1093/nar/gkt1210

27. Mosca R, Céol A, Stein A, Olivella R, Aloy P. 3did: a catalog of domain-based interactions of known three-dimensional structure. Nucleic Acids Res. 2014;42: D374–D379. doi:10.1093/nar/gkt887

28. Dinkel H, Van Roey K, Michael S, Kumar M, Uyar B, Altenberg B, et al. ELM 2016—data update and new functionality of the eukaryotic linear motif resource. Nucleic Acids Res. 2016;44: D294–D300. doi:10.1093/nar/gkv1291

29. 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: 899–907.

30. Finn RD, Clements J, Eddy SR. HMMER web server: interactive sequence similarity searching. Nucleic Acids Res. Oxford University Press; 2011;39: W29-37. doi:10.1093/nar/gkr367

31. 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: D279-85. doi:10.1093/nar/gkv1344

32. Lu L, Zhang M. Edge Betweenness Centrality. Encyclopedia of Systems Biology. New York, NY: Springer New York; 2013. pp. 647–648. doi:10.1007/978-1-4419-9863-7_874

33. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet. NIH Public Access; 2000;25: 25–9. doi:10.1038/75556

bioRxiv preprint doi: https://doi.org/10.1101/147868; this version posted June 8, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

34. Croft D, O'Kelly G, Wu G, Haw R, Gillespie M, Matthews L, et al. Reactome: a database of reactions, pathways and biological processes. Nucleic Acids Res. Oxford University Press; 2011;39: D691-7. doi:10.1093/nar/gkq1018

35. Tatusov RL. The COG database: new developments in phylogenetic classification of proteins from complete genomes. Nucleic Acids Res. 2001;29: 22–28. doi:10.1093/nar/29.1.22

36. Christmas R, Avila-Campillo I, Bolouri H, Schwikowski B, Anderson M, Kelley R, et al. Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. AACR Educ B. 2005;2005: 12.

37. Protasio A V, Dunne DW, Berriman M. Comparative study of transcriptome profiles of mechanical- and skin-transformed Schistosoma mansoni schistosomula. PLoS Negl Trop Dis. 2013;7: e2091. doi:10.1371/journal.pntd.0002091

38. Aumailley M, Gayraud B. Structure and biological activity of the extracellular matrix. J Mol Med. Springer-Verlag; 1998;76: 253–265. doi:10.1007/s001090050215

39. Silva-Almeida M, Carvalho LO, Abreu-Silva AL, Souza CS, Hardoim DJ, Calabrese KS. Extracellular matrix alterations in experimental Leishmania amazonensis infection in susceptible and resistant mice. Vet Res. 2012;43: 10. doi:10.1186/1297-9716-43-10

40. Geurts N, Opdenakker G, Van den Steen PE. Matrix metalloproteinases as therapeutic targets in protozoan parasitic infections. Pharmacol Ther. 2012;133: 257–279. doi:10.1016/j.pharmthera.2011.11.008

41. Morris JH, Apeltsin L, Newman AM, Baumbach J, Wittkop T, Su G, et al. clusterMaker: a multi-algorithm clustering plugin for Cytoscape. BMC Bioinformatics. 2011;12: 436. doi:10.1186/1471-2105-12-436

42. Liu CG, Maercker C, Castañon MJ, Hauptmann R, Wiche G. Human plectin: organization of the gene, sequence analysis, and chromosome localization (8q24).

bioRxiv preprint doi: https://doi.org/10.1101/147868; this version posted June 8, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Proc Natl Acad Sci U S A. 1996;93: 4278–83.

43. Su C, Cao Y, Kaplan J, Zhang M, Li W, Conroy M, et al. Duodenal helminth infection alters barrier function of the colonic epithelium via adaptive immune activation. Infect Immun. American Society for Microbiology (ASM); 2011;79: 2285–94. doi:10.1128/IAI.01123-10

44. Selzer PM (Paul M. Antiparasitic and antibacterial drug discovery□: from molecular targets to drug candidates. Wiley-VCH; 2009.

45. Nygren PJ, Scott JD. Regulation of the phosphatase PP2B by protein-protein interactions. Biochem Soc Trans. NIH Public Access; 2016;44: 1313–1319. doi:10.1042/BST20160150

46. Naderer T, Dandash O, McConville MJ. Calcineurin is required for Leishmania major stress response pathways and for virulence in the mammalian host. Mol Microbiol. Blackwell Publishing Ltd; 2011;80: 471–480. doi:10.1111/j.1365-2958.2011.07584.x

47. Doerig C, Billker O, Haystead T, Sharma P, Tobin AB, Waters NC. Protein kinases of malaria parasites: an update. Trends Parasitol. 2008;24: 570–7. doi:10.1016/j.pt.2008.08.007

48. Zhang VM, Chavchich M, Waters NC. Targeting protein kinases in the malaria parasite: update of an antimalarial drug target. Curr Top Med Chem. 2012;12: 456–72.

49. Luckhart S, Pakpour N, Giulivi C. Host-pathogen interactions in malaria: cross-kingdom signaling and mitochondrial regulation. Curr Opin Immunol. NIH Public Access; 2015;36: 73–9. doi:10.1016/j.coi.2015.07.002

50. Semblat J-P, Doerig C. PAK in pathogen-host interactions. Cell Logist. Taylor & Francis; 2012;2: 126–131. doi:10.4161/cl.20222

51. Li Y, Shi J, Yang J, Ma Y, Cheng L, Zeng J, et al. A Wnt/β-catenin negative feedback

bioRxiv preprint doi: https://doi.org/10.1101/147868; this version posted June 8, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

loop represses TLR-triggered inflammatory responses in alveolar epithelial cells. Mol Immunol. 2014;59: 128–135. doi:10.1016/j.molimm.2014.02.002

52. WILSON RA, Christensen N ⁄rnbjerg, CRABTREE JE, WILSON RA, Dillon GP, Feltwell T, et al. The saga of schistosome migration and attrition. Parasitology. Cambridge University Press; 2009;136: 1581–1592. doi:10.1017/S0031182009005708

53. Merrifield M, Hotez PJ, Beaumier CM, Gillespie P, Strych U, Hayward T, et al. Advancing a vaccine to prevent human schistosomiasis. Vaccine. 2016;34: 2988–2991. doi:10.1016/j.vaccine.2016.03.079

54. Mukhtar MS, Carvunis A-R, Dreze M, Epple P, Steinbrenner J, Moore J, et al. Independently evolved virulence effectors converge onto hubs in a plant immune system network. Science. NIH Public Access; 2011;333: 596–601. doi:10.1126/science.1203659

55. Olszewski KL, Morrisey JM, Wilinski D, Burns JM, Vaidya AB, Rabinowitz JD, et al. Host-parasite interactions revealed by Plasmodium falciparum metabolomics. Cell Host Microbe. 2009;5: 191–9. doi:10.1016/j.chom.2009.01.004

56. Xu T, Ping J, Yu Y, Yu F, Yu Y, Hao P, et al. Revealing parasite influence in metabolic pathways in Apicomplexa infected patients. BMC Bioinformatics. 2010;11: S13. doi:10.1186/1471-2105-11-S11-S13

57. Coustou V, Biran M, Breton M, Guegan F, Rivière L, Plazolles N, et al. Glucose-induced remodeling of intermediary and energy metabolism in procyclic Trypanosoma brucei. J Biol Chem. American Society for Biochemistry and Molecular Biology; 2008;283: 16342–54. doi:10.1074/jbc.M709592200

58. Swann J, Jamshidi N, Lewis NE, Winzeler EA. Systems analysis of host-parasite interactions. Wiley Interdiscip Rev Syst Biol Med. Wiley-Blackwell; 2015;7: 381–400.

bioRxiv preprint doi: https://doi.org/10.1101/147868; this version posted June 8, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

doi:10.1002/wsbm.1311

59. Binns D, Dimmer E, Huntley R, Barrell D, O'Donovan C, Apweiler R. QuickGO: a web-based tool for Gene Ontology searching. Bioinformatics. Oxford University Press; 2009;25: 3045–6. doi:10.1093/bioinformatics/btp536

60. Areschoug T, Gordon S. Scavenger receptors: role in innate immunity and microbial pathogenesis. Cell Microbiol. Blackwell Publishing Ltd; 2009;11: 1160–1169. doi:10.1111/j.1462-5822.2009.01326.x

61. GOMES IN, PALMA LC, CAMPOS GO, LIMA JGB, DE ALMEIDA TF, DE MENEZES JPB, et al. The scavenger receptor MARCO is involved in *Leishmania major* infection by CBA/J macrophages. Parasite Immunol. 2009;31: 188–198. doi:10.1111/j.1365-3024.2009.01093.x

62. Kropf P, Fuentes JM, Fähnrich E, Arpa L, Herath S, Weber V, et al. Arginase and polyamine synthesis are key factors in the regulation of experimental leishmaniasis in vivo. FASEB J. 2005;19: 1000–2. doi:10.1096/fj.04-3416fje

63. Corcoran L, Vremec D, Febbraio M, Baldwin T, Handman E. Differential regulation of CD36 expression in antigen-presenting cells: Oct-2 dependence in B lymphocytes but not dendritic cells or macrophages. Int Immunol. 2002;14: 1099–104.

64. Okuda K, Tong M, Dempsey B, Moore KJ, Gazzinelli RT, Silverman N. Leishmania amazonensis Engages CD36 to Drive Parasitophorous Vacuole Maturation. Zamboni DS, editor. PLOS Pathog. Public Library of Science; 2016;12: e1005669. doi:10.1371/journal.ppat.1005669

65. Walker DM, Oghumu S, Gupta G, McGwire BS, Drew ME, Satoskar AR. Mechanisms of cellular invasion by intracellular parasites. Cell Mol Life Sci. NIH Public Access; 2014;71: 1245–63. doi:10.1007/s00018-013-1491-1

66. KNOX DP. Proteinase inhibitors and helminth parasite infection. Parasite Immunol. Blackwell Publishing Ltd; 2007;29: 57–71. doi:10.1111/j.1365-3024.2006.00913.x

bioRxiv preprint doi: https://doi.org/10.1101/147868; this version posted June 8, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

67. Dzik JM. Molecules released by helminth parasites involved in host colonization. Acta Biochim Pol. 2006;53: 33–64. doi:20061201 [pii]

68. Heung LJ, Luberto C, Del Poeta M. Role of sphingolipids in microbial pathogenesis. Infect Immun. American Society for Microbiology (ASM); 2006;74: 28–39. doi:10.1128/IAI.74.1.28-39.2006

69. Straub C, Burnham JP, White AC, Pazdrak K, Sanchez C, Watanabe LC, et al. Altered eosinophil proteome in a patient with hypereosinophilia from acute fascioliasis. Clin Vaccine Immunol. American Society for Microbiology (ASM); 2011;18: 1999–2002. doi:10.1128/CVI.05373-11

70. Meeusen EN, Balic A. Do eosinophils have a role in the killing of helminth parasites? Parasitol Today. 2000;16: 95–101.

71. Abraham D, Leon O, Schnyder-Candrian S, Wang CC, Galioto AM, Kerepesi LA, et al. Immunoglobulin E and eosinophil-dependent protective immunity to larval Onchocerca volvulus in mice immunized with irradiated larvae. Infect Immun. 2004;72: 810–7.

72. O'Connell AE, Hess JA, Santiago GA, Nolan TJ, Lok JB, Lee JJ, et al. Major basic protein from eosinophils and myeloperoxidase from neutrophils are required for protective immunity to Strongyloides stercoralis in mice. Infect Immun. American Society for Microbiology (ASM); 2011;79: 2770–8. doi:10.1128/IAI.00931-10

73. Ali Khan H, Mutus B. Protein disulfide isomerase a multifunctional protein with multiple physiological roles. Front Chem. Frontiers Media SA; 2014;2: 70. doi:10.3389/fchem.2014.00070

74. Santos CXC, Stolf BS, Takemoto PVA, Amanso AM, Lopes LR, Souza EB, et al. Protein disulfide isomerase (PDI) associates with NADPH oxidase and is required for phagocytosis of Leishmania chagasi promastigotes by macrophages. J Leukoc Biol. 2009;86: 989–998. doi:10.1189/jlb.0608354

bioRxiv preprint doi: https://doi.org/10.1101/147868; this version posted June 8, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

75. Ryser HJ, Levy EM, Mandel R, DiSciullo GJ. Inhibition of human immunodeficiency virus infection by agents that interfere with thiol-disulfide interchange upon virus-receptor interaction. Proc Natl Acad Sci U S A. 1994;91: 4559–63.

76. Wan S-W, Lin C-F, Lu Y-T, Lei H-Y, Anderson R, Lin Y-S. Endothelial cell surface expression of protein disulfide isomerase activates β1 and β3 integrins and facilitates dengue virus infection. J Cell Biochem. 2011;113: n/a-n/a. doi:10.1002/jcb.24037

77. Calderon MN, Guerrero CA, Acosta O, Lopez S, Arias CF. Inhibiting Rotavirus Infection by Membrane-Impermeant Thiol/Disulfide Exchange Blockers and Antibodies against Protein Disulfide Isomerase. Intervirology. 2012;55: 451–464. doi:10.1159/000335262

78. Chiu H-C, Hannemann H, Heesom KJ, Matthews DA, Davidson AD. High-Throughput Quantitative Proteomic Analysis of Dengue Virus Type 2 Infected A549 Cells. Thiel V, editor. PLoS One. Public Library of Science; 2014;9: e93305. doi:10.1371/journal.pone.0093305

79. Cortés A, Sotillo J, Muñoz-Antoli C, Fried B, Esteban JG, Toledo R. Altered Protein Expression in the Ileum of Mice Associated with the Development of Chronic Infections with Echinostoma caproni (Trematoda). PLoS Negl Trop Dis. Public Library of Science; 2015;9: e0004082. doi:10.1371/journal.pntd.0004082

80. Pando-Robles V, Oses-Prieto JA, Rodríguez-Gandarilla M, Meneses-Romero E, Burlingame AL, Batista CVF. Quantitative proteomic analysis of Huh-7 cells infected with Dengue virus by label-free LC–MS. J Proteomics. 2014;111: 16–29. doi:10.1016/j.jprot.2014.06.029

81. Rinaldi M, Dreesen L, Hoorens PR, Li RW, Claerebout E, Goddeeris B, et al. Infection with the gastrointestinal nematode Ostertagia ostertagi in cattle affects mucus biosynthesis in the abomasum. Vet Res. BioMed Central; 2011;42: 61. doi:10.1186/1297-9716-42-61

bioRxiv preprint doi: https://doi.org/10.1101/147868; this version posted June 8, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

82. Morias Y, Abels C, Laoui D, Van Overmeire E, Guilliams M, Schouppe E, et al. Ly6C-Monocytes Regulate Parasite-Induced Liver Inflammation by Inducing the Differentiation of Pathogenic Ly6C+ Monocytes into Macrophages. Loke P, editor. PLOS Pathog. Public Library of Science; 2015;11: e1004873. doi:10.1371/journal.ppat.1004873

83. Manque PA, Probst CM, Probst C, Pereira MCS, Rampazzo RCP, Ozaki LS, et al. Trypanosoma cruzi infection induces a global host cell response in cardiomyocytes. Infect Immun. American Society for Microbiology (ASM); 2011;79: 1855–62. doi:10.1128/IAI.00643-10

84. Goldberg DS, Roth FP. Assessing experimentally derived interactions in a small world. Proc Natl Acad Sci U S A. National Academy of Sciences; 2003;100: 4372–6. doi:10.1073/pnas.0735871100

85. Guo X, Xu Y, Bian G, Pike AD, Xie Y, Xi Z. Response of the mosquito protein interaction network to dengue infection. BMC Genomics. 2010;11: 380. doi:10.1186/1471-2164-11-380

86. Vazquez A, Flammini A, Maritan A, Vespignani A. Global protein function prediction from protein-protein interaction networks. Nat Biotechnol. 2003;21: 697–700. doi:10.1038/nbt825

Supporting information

S1 Table. Putative functional annotation of parasite proteins. We assigned eggNOG functional annotations from host COGs to parasite proteins with no “Function unknown” in the COGs they belonged to.

S1 Fig. Blood–specific network. Human–*S. mansoni* predicted interactions located in blood.

bioRxiv preprint doi: https://doi.org/10.1101/147868; this version posted June 8, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

S2 Fig. Lung–specific network. Human–*S. mansoni* predicted interactions located in the lungs.

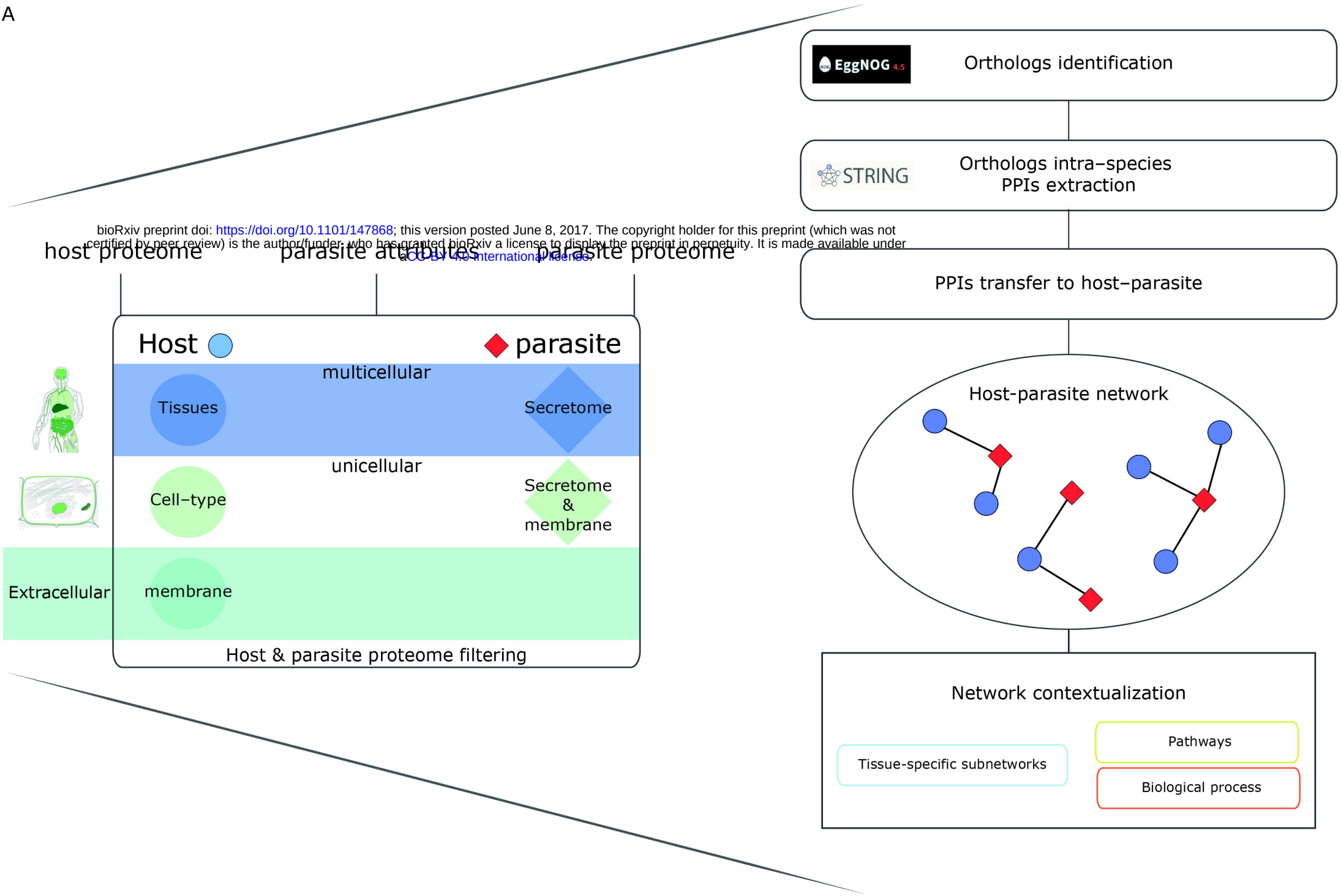
S3 Fig. Liver–specific network. Human–*S. mansoni* predicted interactions located in the liver.

S4 Fig. Skin–specific network. Human–*S. mansoni* predicted interactions located in the lungs.

S5 Fig. Intestine–specific network. Human–*S. mansoni* predicted interactions located in the lungs.

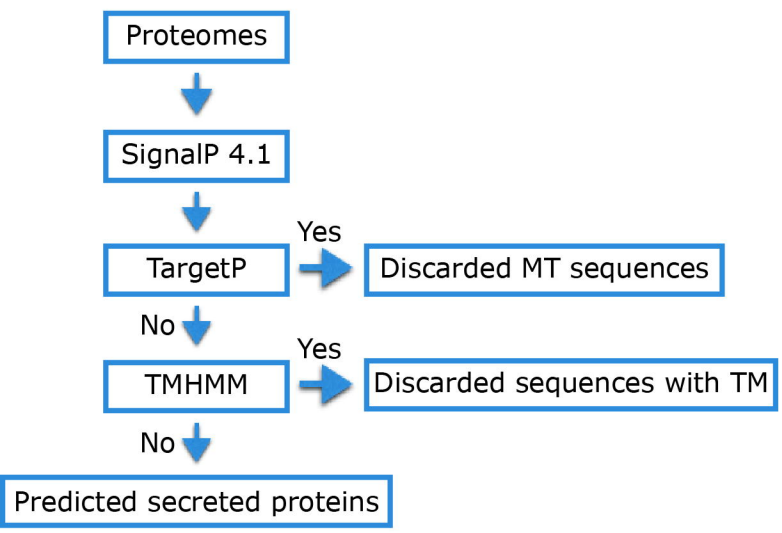
S6 Fig. Intra-species eggNOG groups functional annotation transferred. Functions annotated in eggNOG to the most contributing clusters of orthologs groups (COGs) used in the orthology transfer method.

A

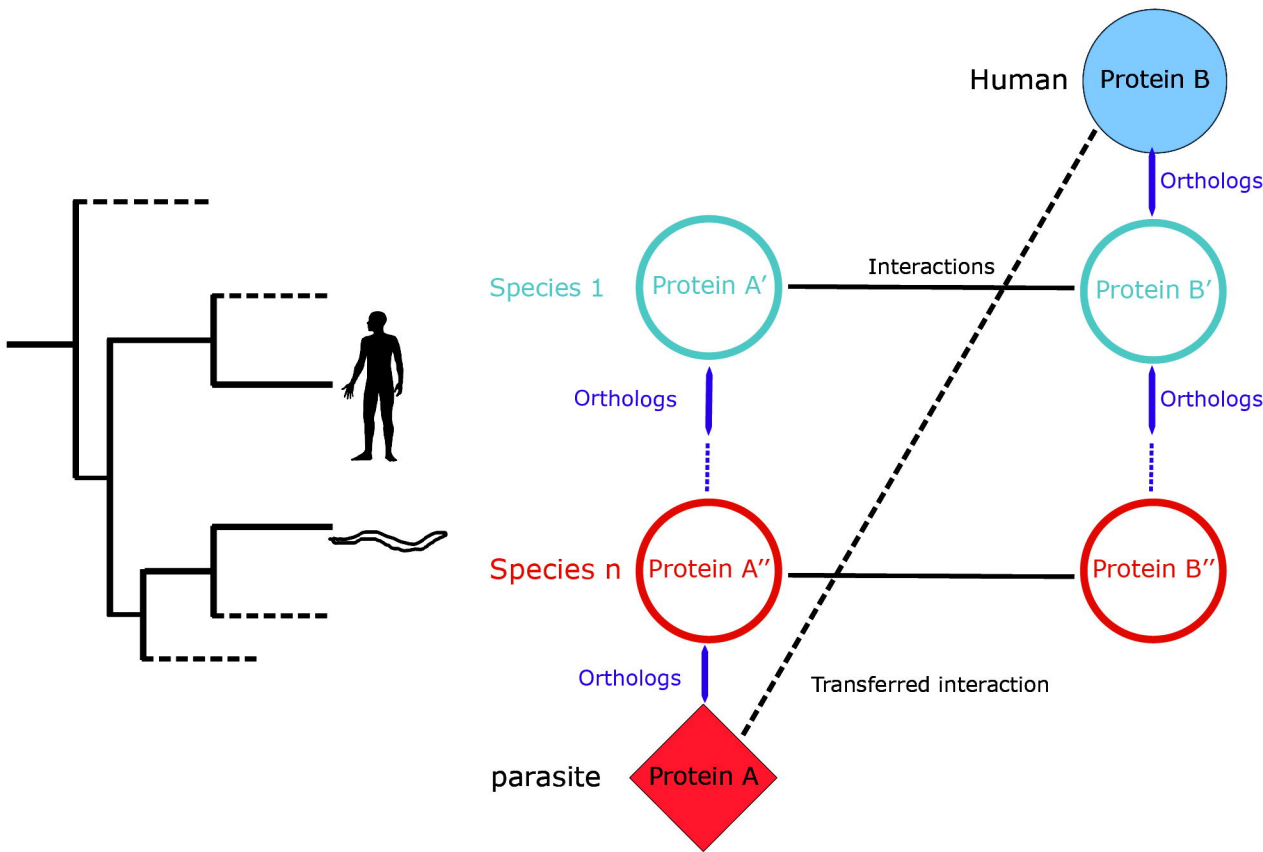


bioRxiv preprint doi: https://doi.org/10.1101/147868; this version posted June 8, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

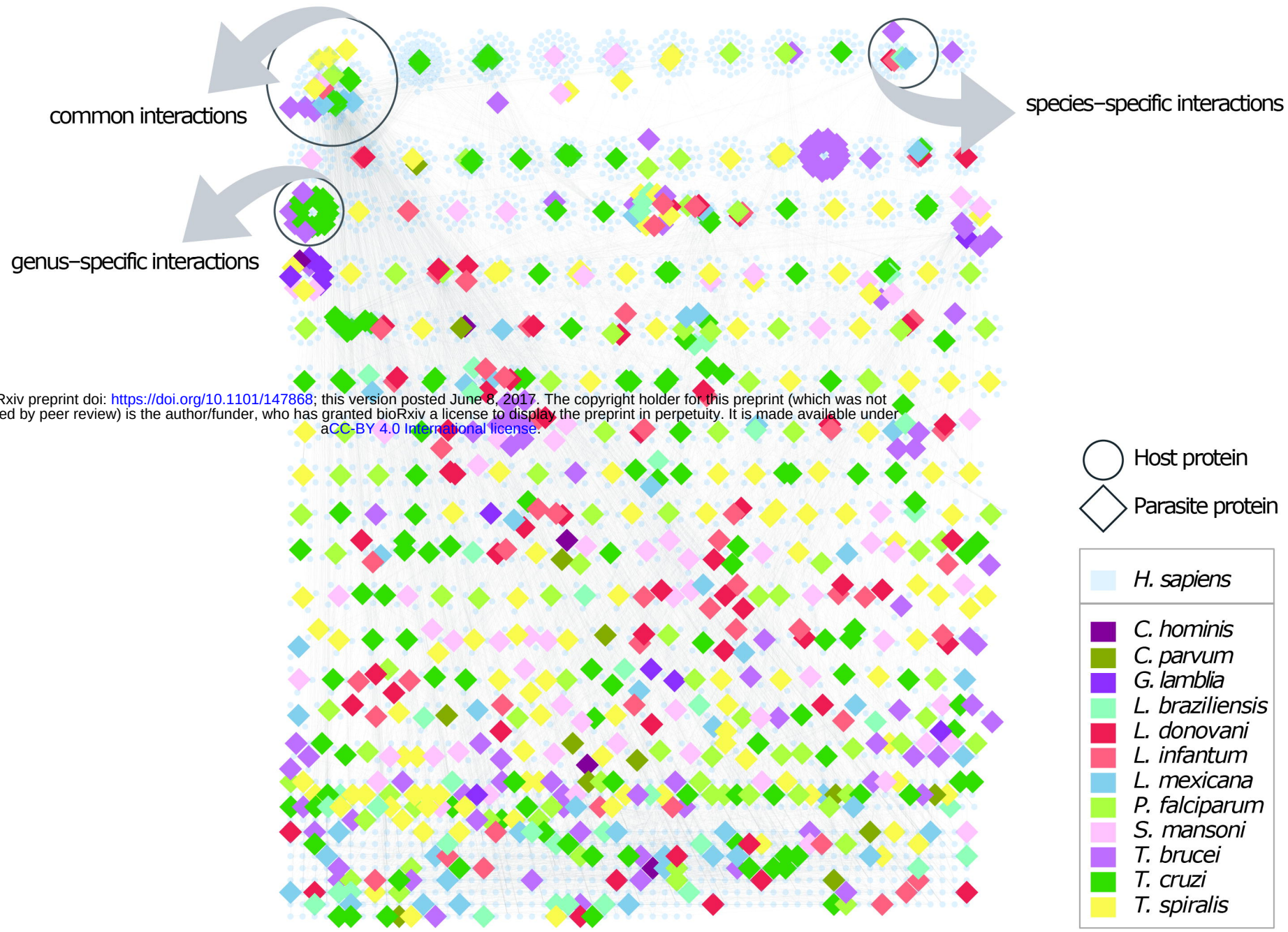
B



C

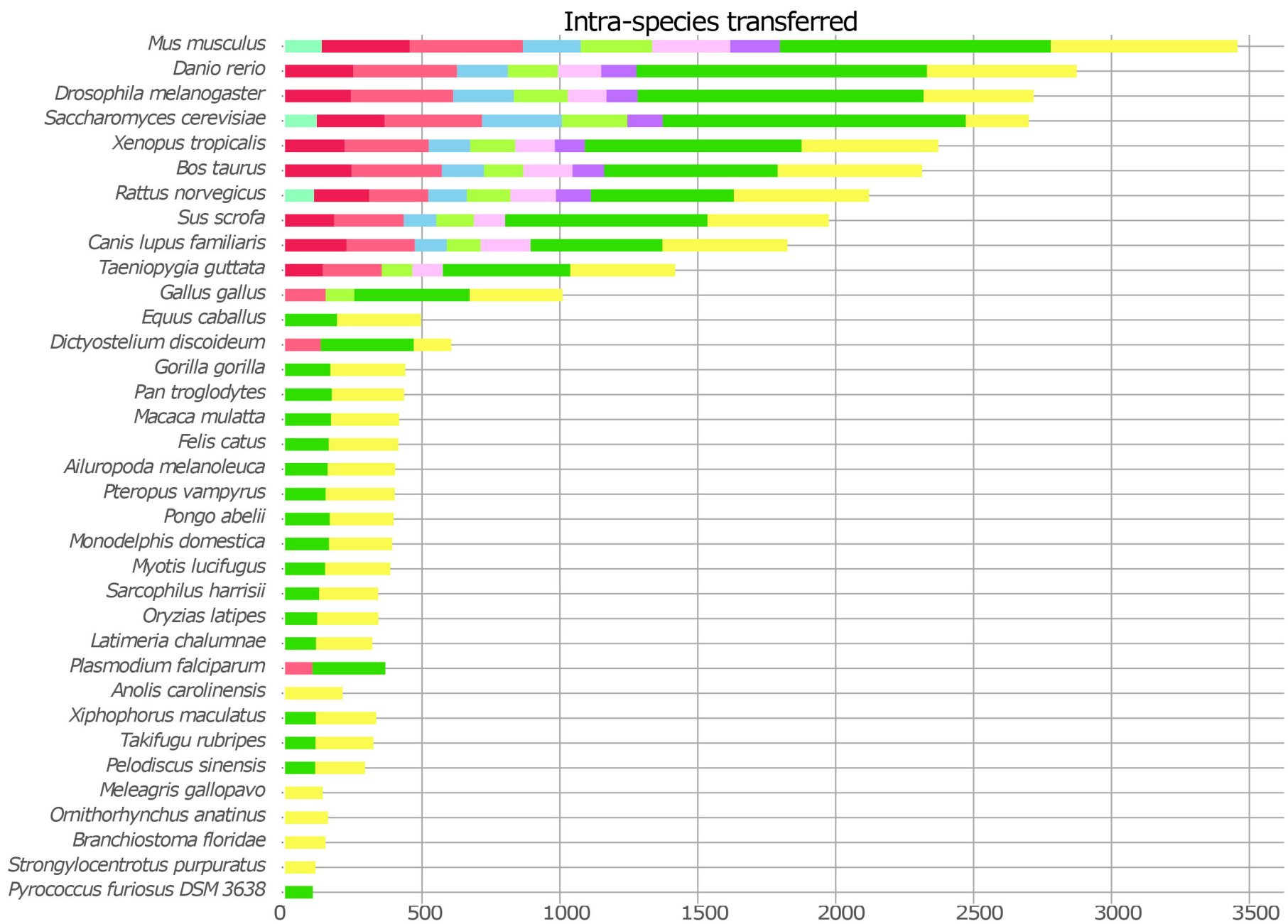


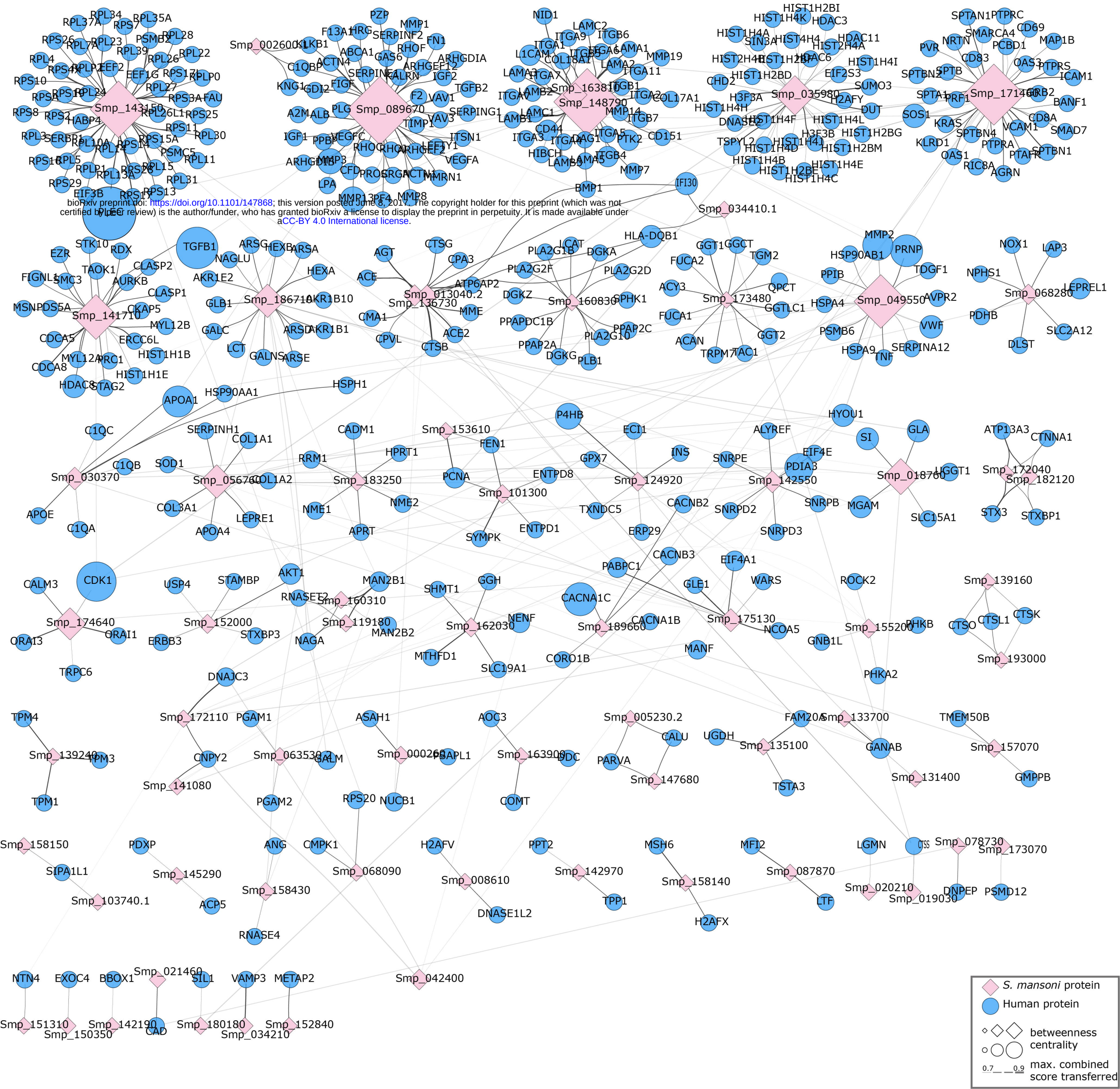
A



bioRxiv preprint doi: https://doi.org/10.1101/147868; this version posted June 8, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

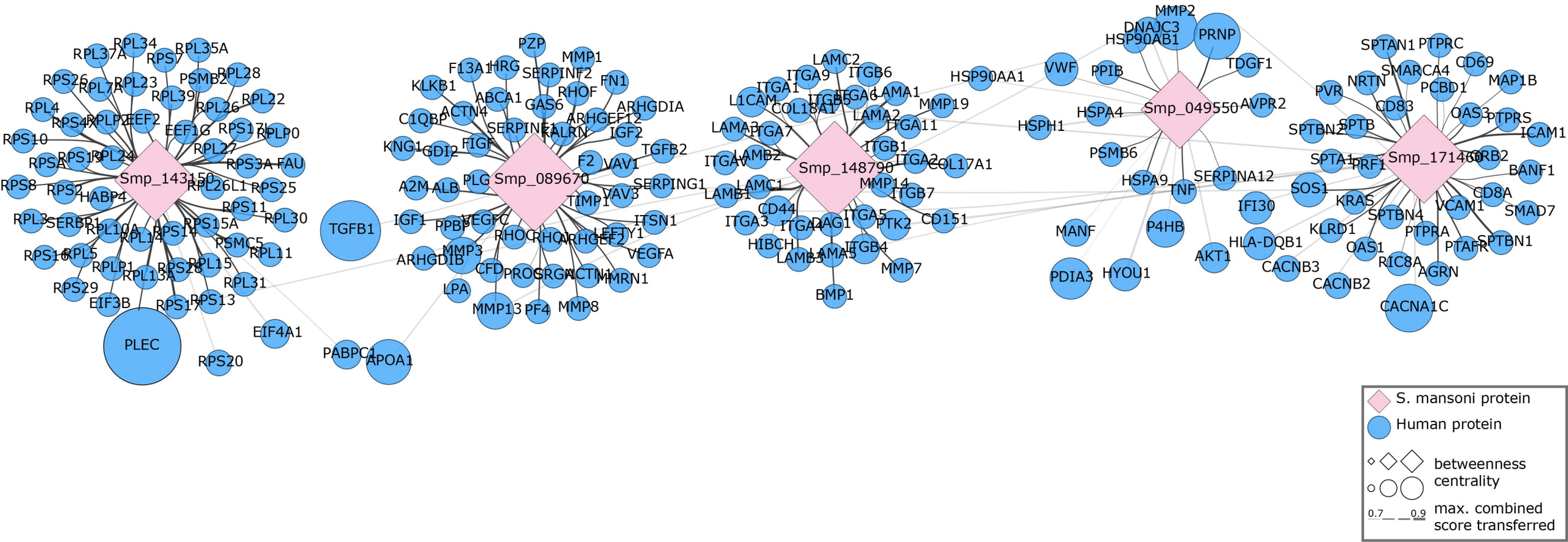
B



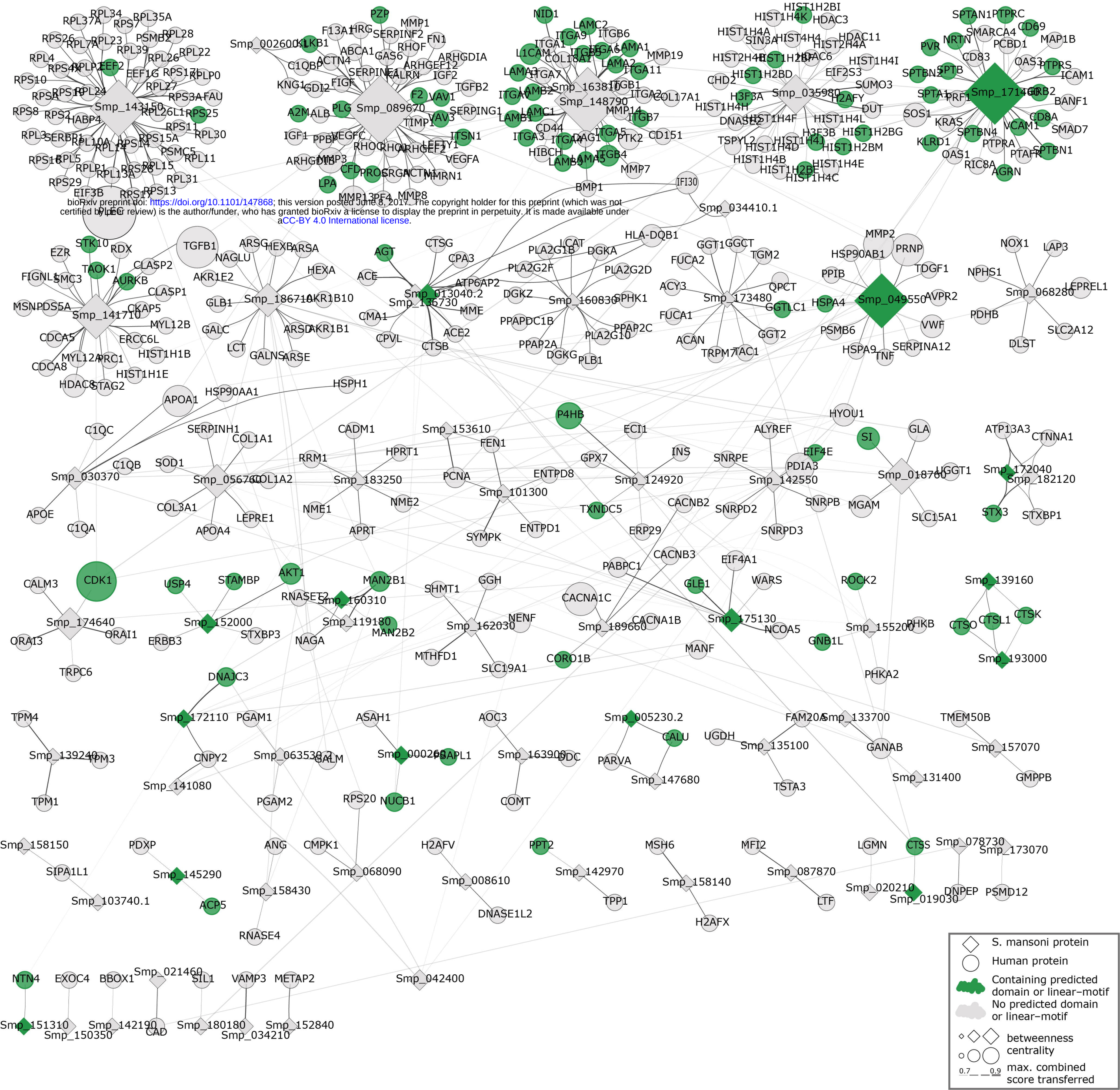


bioRxiv preprint doi: https://doi.org/10.1101/147868; this version posted June 8, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Figure 4



Smp_143150
Smp_089670
Smp_148790
Smp_049550
Smp_171460
PLEC
TGFB1
PABPC1
APOA1
RPL34
RPL37A
RPS7
RPL35A
RPS26
RPL7A
RPL23
PSMB2
RPL28
RPL39
RPL26
RPL22
RPL4
RPS4
RPLP2
EEF2
EEF1G
RPS17L
RPLP0
RPS10
RPL27
RPSA
RPS19
RPL24
RPS3A
FAU
RPS8
RPS2
HABP4
RPL26L1
RPS25
RPS11
RPL3
SERBP1
RPL10A
RPS14
RPS15A
RPL30
RPL14
PSMC5
RPS16
RPL5
RPL11
RPLP1
RPL13A
RPS28
RPL15
RPL31
RPS29
EIF3B
RPS17
RPS13
EIF4A1
RPS20
PZP
MMP1
F13A1
HRG
SERPINF2
FN1
KLKB1
ABCA1
RHOF
GAS6
ARHGDIA
C1QBP
ACTN4
SERPINF1
ARHGEF12
KALRN
IGF2
KNG1
GDI2
FIGF
TGFB2
F2
VAV1
A2M
ALB
PLG
SERPING1
VAV3
TIMP1
IGF1
PPBP
VEGFC
ITSN1
RHOC
RHOA
ARHGEF2
LEFTY1
VEGFA
ARHGDIB
MMP3
CFD
PROS
SRGA
ACTN1
MMRN1
LPA
MMP13
PF4
MMP8
LAMC2
ITGA9
ITGB6
ITGA1
ITGB5
ITGA6
LAMA1
L1CAM
COL18A1
MMP19
LAMA3
ITGA7
LAMA2
ITGA11
ITGAV
LAMB2
ITGB1
ITGA2
COL17A1
LAMC1
MMP14
LAMB1
ITGB7
CD44
ITGA3
ITGA4
DAG1
ITGA5
PTK2
CD151
HIBCH
LAMB3
LAMA5
ITGB4
MMP7
BMP1
HSP90AA1
VWF
PPIB
MMP2
DNAJC3
HSP90AB1
PRNP
TDGF1
HSPA4
AVPR2
HSPH1
PSMB6
HSPA9
TNF
SERPINA12
IFI30
MANF
P4HB
HLA-DQB1
AKT1
PDIA3
HYOU1
SPTAN1
PTPRC
CD69
SMARCA4
NRTN
PCBD1
MAP1B
PVR
CD83
OAS3
PTPRS
SPTBN2
SPTB
ICAM1
SPTA1
PRF1
GRB2
BANF1
SOS1
CD8A
KRAS
VCAM1
SMAD7
SPTBN4
KLRD1
PTPRA
SPTBN1
OAS1
PTAFR
RIC8A
AGRN
CACNB3
CACNB2
CACNA1C
S. mansoni protein
Human protein
betweenness centrality
0.7
0.9
max. combined score transferred

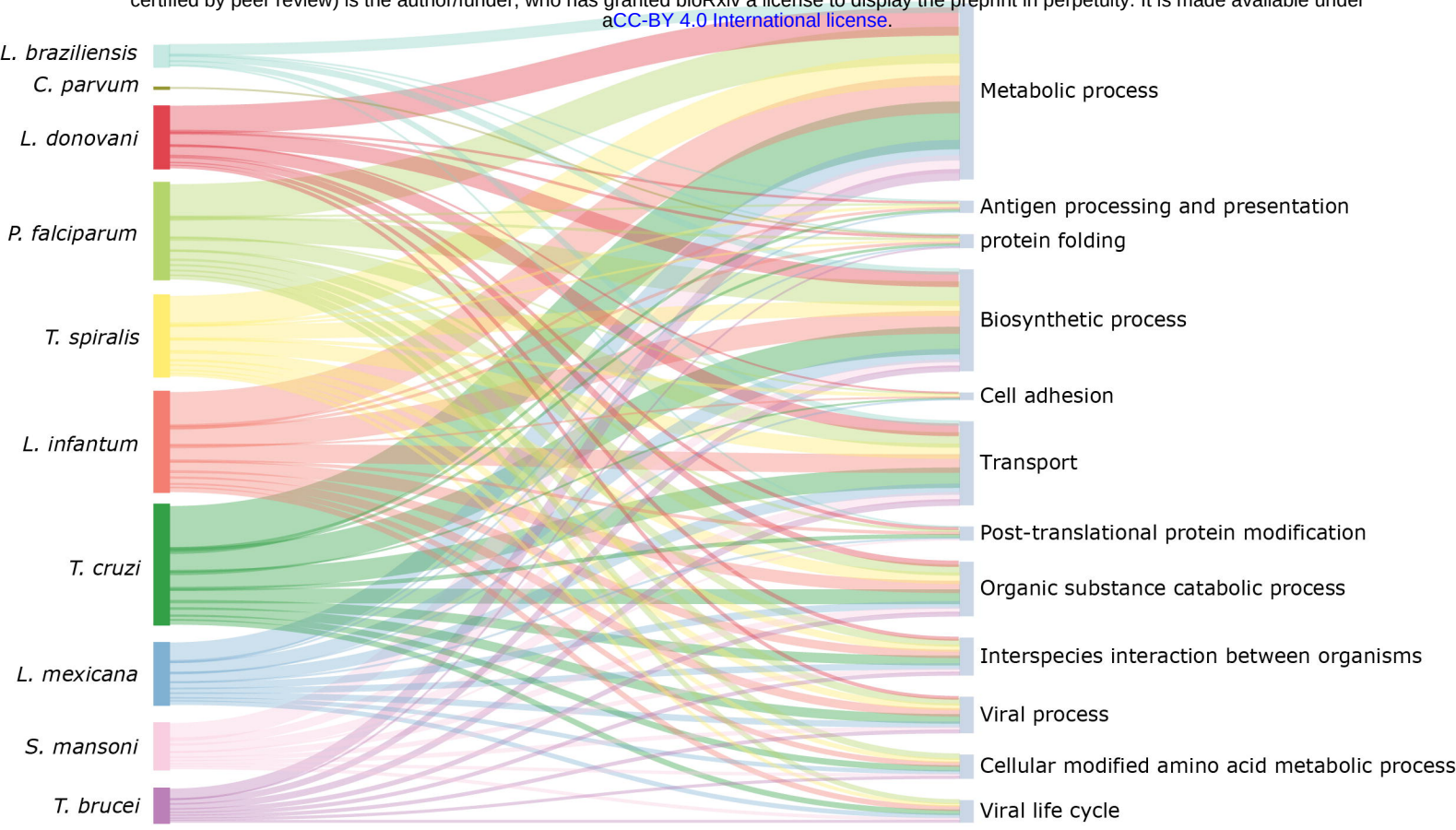


bioRxiv preprint doi: https://doi.org/10.1101/147868; this version posted June 8, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Figure 7

A

bioRxiv preprint doi: https://doi.org/10.1101/147868; this version posted June 8, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.



B

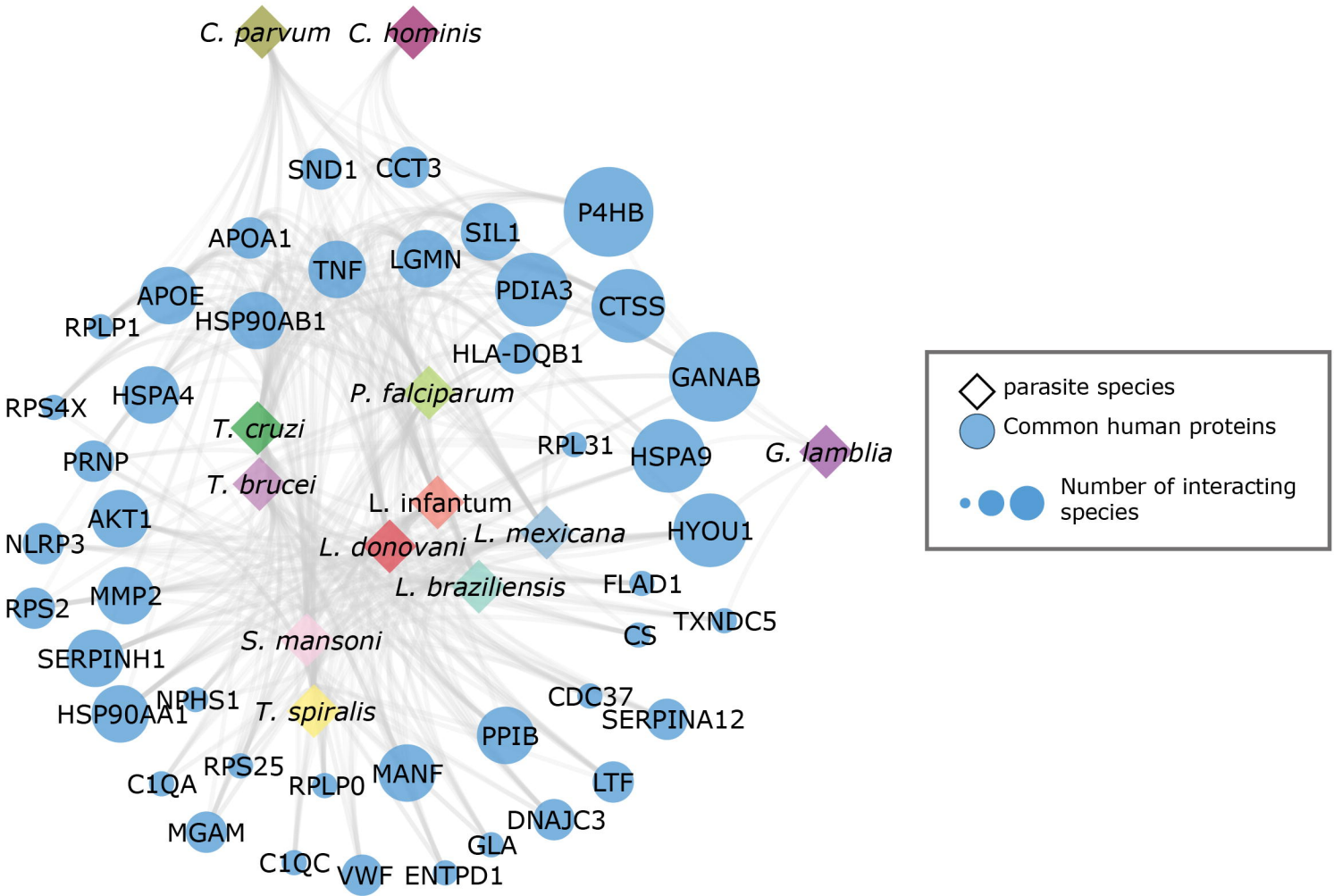
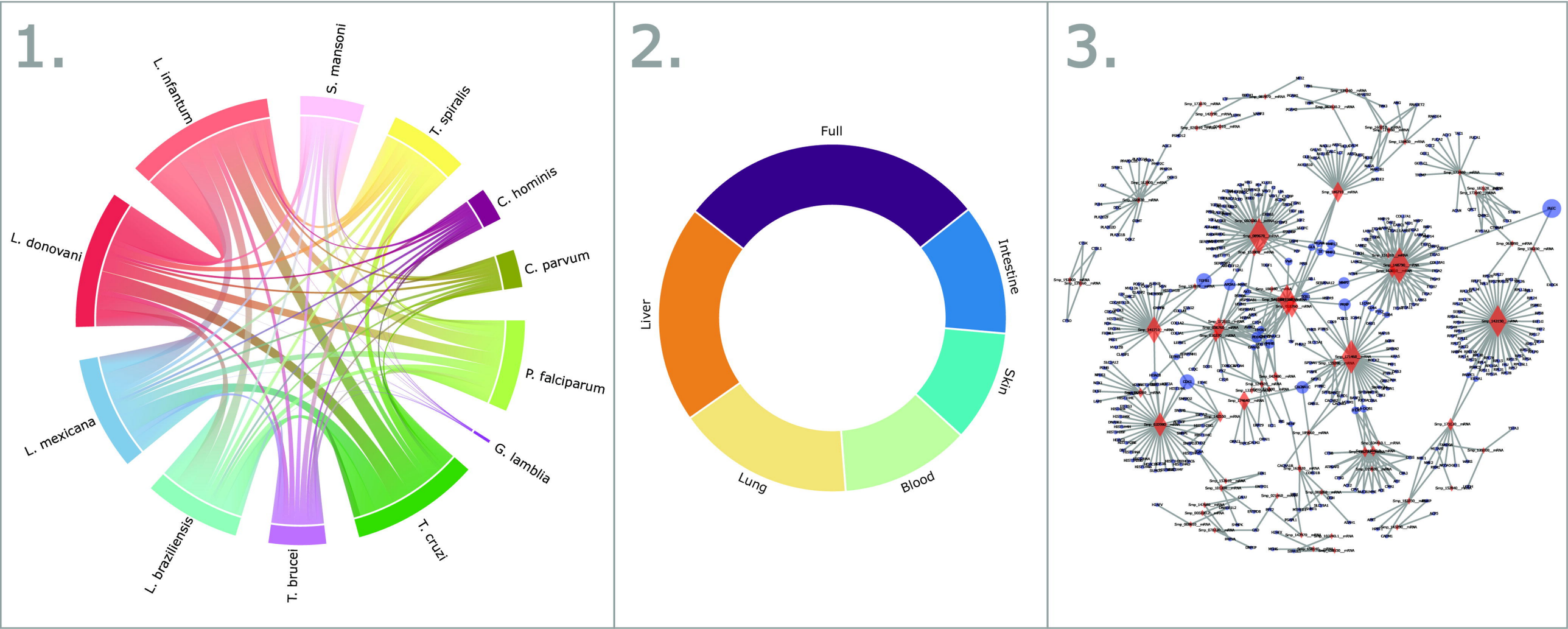


Figure 6

OrthoHPI