UNIVERSIDAD REGIONAL AMAZÓNICA IKIAM

FACULTAD DE CIENCIAS DE LA VIDA

INGENIERÍA EN BIOTECNOLOGÍA

Analysis of the Protein-Protein Interaction Network of the Parasite-Vector-Human System in Chagas Disease

Proyecto de investigación previo a la obtención del título de:

INGENIERO EN BIOTECNOLOGÍA

AUTOR: FRANCISCO JAVIER MENDOZA PROAÑO

TUTOR: M.Sc. MARCO ANDRÉS VITERI YÁNEZ

COTUTOR: M.Sc. MOISÉS RUBÉN GUALAPURO GUALAPURO

Napo-Ecuador

2024

DECLARACIÓN DE DERECHO DE AUTOR, AUTENTICIDAD Y RESPONSABILIDAD

Tena, 5 de marzo de 2024

De mi consideración:

Yo, **Francisco Javier Mendoza Proaño** con documento de identidad **0603617135**, declaro que los resultados obtenidos en la investigación que presento en este documento final, previo a la obtención del título **Ingeniero en Biotecnología** son absolutamente inéditos, originales, auténticos y personales.

En virtud de lo cual, el contenido, criterios, opiniones, resultados, análisis, interpretaciones, conclusiones, recomendaciones y todos los demás aspectos vertidos en el presente documento son de mi autoría y de mi absoluta responsabilidad.

Por la favorable atención a la presente, suscribo de usted,

Atentamente,

Francisco Javier Mendoza Proaño
CI: 0603617135

AUTORIZACIÓN DE PUBLICACIÓN EN EL REPOSITORIO INSTITUCIONAL

Tena, 5 de marzo de 2024

Yo, **Francisco Javier Mendoza Proaño** con documento de identidad **0603617135**, en calidad de autor y titular de los derechos morales y patrimoniales del trabajo de titulación: **Analysis of the Protein-Protein Interaction Network of the Parasite-Vector-Human System in Chagas Disease**, de conformidad con el Art. 114 del CÓDIGO ÓRGANICO DE LA ECONOMÍA SOCIAL DE LOS CONOCIMIENTOS, CREATIVIDAD E INNOVACIÓN, reconozco a favor de la Universidad Regional Amazónica Ikiam una licencia gratuita, intransferible y no exclusiva para el uso no comercial de la obra, con fines estrictamente académicos.

Así mismo autorizo a la Universidad Regional Amazónica Ikiam para que realice la publicación de este trabajo de titulación en el Repositorio Institucional de conformidad a lo dispuesto en el Art. 144 de la Ley Orgánica de Educación superior.

Atentamente,

Francisco Javier Mendoza Proaño
CI: 0603617135

CERTIFICADO DE DIRECCIÓN DE TRABAJO DE TITULACIÓN

Certifico que el trabajo de titulación: **"Analysis of the Protein-Protein Interaction Network of the Parasite-Vector-Human System in Chagas Disease"** en la modalidad de: artículo, fue realizado por **Francisco Javier Mendoza Proaño**, bajo mi dirección.

El mismo ha sido revisado en su totalidad y analizado por la herramienta de verificación de similitud de contenido; por lo tanto, cumple con los requisitos teóricos, científicos, técnicos, metodológicos y legales establecidos por la Universidad Regional Amazónica Ikiam, para su entrega y defensa.

Tena, 5 de marzo de 2024

Firmado electrónicamente por:
MARCO ANDRES VITERI YANEZ

……

Marco Andrés Viteri Yánez

C.I: 1718163759

MOISES RUBEN GUALAPURO GUALAPURO

Digitally signed by MOISES RUBEN GUALAPURO GUALAPURO
Date: 2024.03.05 15:12:26 -06'00'

……

Moisés Rubén Gualapuro Gualapuro

C.I: 1003124284

AGRADECIMIENTOS

A mi mamá, hermanas y tías, por todo su amor y apoyo durante mi periodo de estudio, por su paciencia y sobre todo por nunca dejarme renunciar.

A mi tutor Marco y cotutor Moisés, que fueron mis ejemplos, amigos y guías, quienes, con su disciplina, su compromiso y su dedicación hicieron que este trabajo sea posible.

DEDICATORIA

A mi mamá Bethi, mis hermanas Melissa y Paulina, mis tías Gladis y Martha, a mi primo Pedro, por darme su ejemplo y aconsejarme a que termine mis estudios, darme un sostén incondicional de afecto y cariño.

A mis amigos Paulina, Milena, Tania, Mahelet, María Fernanda, Ana Cristina, Emily, Betzabeth, Johana, Emilio y Bryan, por su insistencia y apoyo emocional durante la escritura de este trabajo, por sus consejos y sobre todo por apoyarme en todo, dándome fuerzas y ánimos.

A mis nuevos amigos, que estuvieron conmigo en mi última etapa de escritura, que me respaldaron en cada una de mis dificultades y me ayudaron a sobreponerme.

A mi Khalessi, por ser mi pilar en cada noche de desvelo, por su incontable compañía, por nunca dejarme solo y motivarme a escribir. Gracias por darme el coraje y la fuerza para enfrentar cualquier desafío.

ÍNDICE GENERAL

ÍNDICE DE FIGURAS

ÍNDICE DE TABLAS

RESUMEN

Trypanosoma cruzi es la causa etiológica de la Enfermedad de Chagas, una de las enfermedades más críticas en América Latina y considerada una de las 13 enfermedades tropicales más desatendidas a nivel mundial. En los últimos años, los datos ómicos han aumentado sustancialmente en calidad y cantidad. En particular, los de proteómica y transcriptómica se pueden utilizar durante el desarrollo de redes biológicas, donde se producen interacciones proteína-proteína (PPI). En el sistema parásito-vector-humano de la Enfermedad de Chagas, se producen interacciones específicas entre proteínas del parásito (*T. cruzi*), del vector (*Rhodnius prolixus*) y del humano (*Homo sapiens*). Este estudio tuvo como objetivo comprender el sistema parásito-vector-humano de la enfermedad de Chagas a través del análisis automatizado de redes de interacción proteína-proteína. Para ello, se construyó una red PPI basada en suposiciones, luego se identificaron los nodos de interés (componentes conectados, centralidad, medidas comunitarias y peso de unión entre proteínas) y se realizó un análisis de enriquecimiento. En este trabajo se analizaron 101 602 genes, y tras ser filtrados y purificados se obtuvieron 985 genes: 501 nodos y 5 796 interacciones. La identificación de genes fundamentales, incluidos *TNF*, *IL6* y *GADPH*, subraya la importancia de nodos específicos en la red. Estos genes se destacan como posibles objetivos terapéuticos con sus funciones pronunciadas e interacciones sólidas. Su participación sugiere contribuciones críticas al panorama molecular de la enfermedad de Chagas, ofreciendo información valiosa sobre sus mecanismos de comprensión.

Palabras clave: *Trypanosoma cruzi*, interacción proteína-proteína, sistema parásito-vector-humano, objetivos terapéuticos.

ABSTRACT

Trypanosoma cruzi is the etiological cause of Chagas Disease, one of the most critical diseases in Latin America and considered one of the 13 most neglected tropical diseases worldwide. In recent years, omics data has increased substantially in quality and quantity. In particular, those from proteomics and transcriptomics can be used during the development of biological networks, where protein-protein interactions (PPIs) occur. In the parasite-vector-human system in Chagas Disease, specific interactions occur between proteins of the parasite (*T. cruzi*), the vector (*Rhodnius prolixus*), and the human (*Homo sapiens*). This study aimed to understand the parasite-vector-human system of Chagas disease through automated analysis of protein-protein interaction networks. To do this, a PPI network was built based on assumptions, then the nodes of interest were identified (connected components, centrality, community measures, and binding weight between proteins), and an enrichment analysis was performed. In this work, 101 602 genes were analyzed, and after being filtered and purified, 985 genes were obtained: 501 nodes and 5 796 interactions. Identifying pivotal genes, including *TNF*, *IL6*, and *GADPH*, underscores the significance of specific nodes in the network. These genes stand out as potential therapeutic targets with their pronounced roles and robust interactions. Their involvement suggests critical contributions to the molecular landscape of Chagas disease, offering valuable insights into its understanding mechanisms.

Keywords: *Trypanosoma cruzi*, protein-protein interaction, parasite-vector-human system, therapeutic targets.

1. INTRODUCTION

1.1. Protozoan parasites: genus Trypanosoma

Protozoa are non-photosynthetic organisms included in a subgroup of protists [1]. Likewise, on Earth, there are more than 15 000 species of protozoa, of which the most medically significant classes are Flagellata (*Leishmania*, *Trypanosoma*, *Giardia* y *Trichomonas*), Infusoria (*Balantidium*), Sarcodina (*Amobea*) y Sporozoa (*Plasmodium*, *Sarcocystis*, *Cryptosporidium*, *Toxoplasma*, *Babesia* e *Isospora*) (Table 1) [2–6].

Table 1. The main characteristics of Protozoa are of significant medical interest.

Classes of major medical significance	Main clades	Features	Reference
Flagellata	*Leishmania, Trypanosoma*, *Giardia* and *Trichomonas*	Their shape is elongated or spherical, move using flagella, and reproduction is asexual and sexual. It lives in aquatic environments, such as lakes, rivers, and oceans, or even in moist soils and the digestive tracts of animals. One of its leading representatives is *T. cruzi*, the causative agent of Chagas disease.	[2,5–8]
Infusoria	*Balantidium*	Their shape and size vary, but they are commonly oval or elongated, move using cilia that cover their surface, and reproduction is asexual and sexual. It inhabits aquatic environments, especially in stagnant water rich in organic matter. One of the leading causes of diseases of this type is *Paramecium* sp., which can be a problem in swimming pools or untreated water systems. However, it does not usually cause diseases in humans.	[2,3,5,9,10]
Sarcodina	*Amoeba*	Their shape is usually amoeboid; that is, they have a variable shape, move using pseudopodia (temporary extensions of the cell membrane), and reproduction is asexual and sexual. Lives in aquatic and terrestrial environments, such as humid soil. One of the leading causes of diseases of this type is *Entamoeba histolytica*; it is the	[2,3,5,11,12]

		causative agent of amoebiasis, an intestinal disease.	
Sporozoa	*Plasmodium*, *Sarcocystis*, *Cryptosporidium* , *Toxoplasma*, *Babesia* and *Isospora*	Their shape is It is spherical or elongated in shape, lack typical locomotion structures such as flagella, cilia, or pseudopodia since they are obligate intracellular parasites and spend their entire life inside the host cell, and reproduction is asexual and sexual. Being obligate intracellular parasites, their habitat is the interior of the host organisms. One of the leading causes of diseases of this type is *P. falciparum*, the causative agent of malaria in humans.	[2,3,5,6,13,14]

Made by: Francisco Javier Mendoza Proaño.

In this sense, trypanosomes are unicellular, ubiquitous, flagellated organisms that belong to the class Kinetoplastea, which is part of the supergroup of the Excavata (parasites with a single nucleus, mitochondria, and flagellum with mitochondrial DNA called kinetoplast) and to the order Trypanosomatida. Moreover, their development is preponderantly limited to a single host species, although some members of this order may have more than one host species during their life cycle [3,4,7,8,15–18].

On the other hand, the family Trypanosomatidae harbors 19 genera, which are protozoans and are mostly monoxenous (in some cases, they are also dixenous (syn. digenetic) and parasitize different species of animal and plant) (Table 2) [8,19–21]. For example, Phytomonas is a dixenous trypanosomatid whose main vectors are phytophagous insects, parasitizing mainly plants [20]. In this sense, within this family, one of the most recognized genera is *Trypanosoma*, which infects almost all vertebrate classes and is transmitted by different species of reduviid blood-sucking insects, from *Rhodnius*, *Triatoma*, and *Panstrongylus* genera [16,17,19,22]. In this regard, some of the primary diseases transmitted by these vectors in the Americas are leishmaniasis (caused by *Leishmania* spp.) and Chagas Disease (or American trypanosomiasis, caused by *T. cruzi*) (Figure 1), which primarily affect people from low-income countries in Central and South America, and are therefore considered neglected tropical diseases [8,19,23].

Table 2. Genera of the family Trypanosomatidae and its invertebrates host that are currently recognized, as modified from [19].

Genus		Invertebrate host	Geographic distribution	Reference
Monoxenous	*Angomononas*	*Chrysomya putoria*	Africa and South America	[24,25]
		Chrysomya megacephala	America (except Canada and Alaska), Africa, and southern Asia	[26]
		Lucilia cuprina	Australia, New Zealand Africa, and southern North America	[27]
		Ornidia obesa	Eastern South America	[28]
		Zelus leucogrammus	Western South America	[29]
	Blastocrithidia	*Euschistus servus*	Central America and North America	[30]
		Gerris remiges (Aquarius remigis)	Southwestern North America	[27]
	Blechomonas	*Ceratophyllus* spp.	Worldwide distribution (tropical climates)	[31]
		Chaetopsylla spp.	Worldwide distribution (tropical climates)	[32]
		Ctenophthalmus spp.	Worldwide distribution (tropical climates)	[33]
		Ctenocephalides spp.	Worldwide (tropical climates)	[34–36]
		Monopsyllus sciurorum	Central Asia	[37]
		Nosopsyllus fasciatus	Europe	[38]
		Nycteridopsylla spp.	Worldwide distribution (tropical climates)	[39]
		Paraceras melis	Europe	[40]
		Pulex irritans	Worldwide distribution (tropical climates)	[41]
	Crithidia	*Bombus hortorum*	Europe, North Asia, Oceania and America	[42]
		Bombus muscorum	Europe and Asia	[42]
		Bombus terrestris	Central and southern Europe, North Africa	[42]
		Culex spp.	Southern North America and Northern South America	[43]
	Herpetomonas	*Musca domestica*	Worldwide distribution (tropical climates)	[44]
	Kentomonas	*Sarcophaga* (*sensu lato*) sp.	Worldwide distribution (tropical climates)	[45]

	Leptomonas	*Calocorisca altiplana*	America	[46]
		Camptischium clavipes	Central and South America	[47]
		Collaria oleosa	South America	[48]
		Dysdercus spp.	Worldwide distribution (tropical climates)	[49]
		Hyalymenus sp.	Worldwide distribution (tropical climates)	[50]
		Jadera aeola aeola	Central and South America	[51]
		Prepops cf. *accinctus*	Central and South America	[52]
	Lotmaria	*Apis mellifera*	Europe, the Middle East, and Africa	[53]
	Novymonas	*Niesthrea vincentii*	South America	[54]
	Paratrypanosoma	*Culex pipiens*	North America, Europe, West Africa, and Central Asia	[43]
	Sergeia	*Culicoides festivipennis*	Western Palaearctic	[55,56]
		Culicoides truncorum	Western Palaearctic	[55]
	Strigonomonas	*Aedes vexans*	North America, Southern South America, Europe, Asia, North Africa, and Oceania	[57]
		Oncopeltus sp.	Worldwide distribution (tropical climates)	[58]
	Wallaceina	*Calocoris sexguttatus*	Europe	[59]
		Nabis brevis	Europe, and Asia	[60,61]
		Nabis flavomarginatus	Europe, Central Asia, and North America	[60]
	Zelonia	*Simulium* (*Morops*) *dycei*	Australia, Asia, and North America	[62,63]
		Ricolla simillima	Central and South America	[46,63]
Dixenous	*Endotrypanum*	*Phlebotomus* spp.	Worldwide distribution (tropical and Mediterranean climates)	[64]
	Leishmania	*Forcipomyia* (*Lasiohelea*) spp.	Worldwide distribution (tropical climates)	[65]
		Lutzomyia spp.		
		Phlebotomus spp.	Worldwide distribution (tropical and Mediterranean climates)	[64]
	Phytomonas	*Nezara viridula*	Europe, Asia, Africa, and America	[66]
		Oncopeltus fasciatus	Southern North America, Northern Central America and the Caribbean	[67]
		Phthia picta	America and the Caribbean	[68]
	Porcisia	Vector unknown	No information	
	Trypanosoma	*Glossina* spp.	Africa	[69,70]
		Rhodnius prolixus	Central America and northern South America	[71]
		Triatoma infestans	South America	[8,72]

Made by: Francisco Javier Mendoza Proaño.

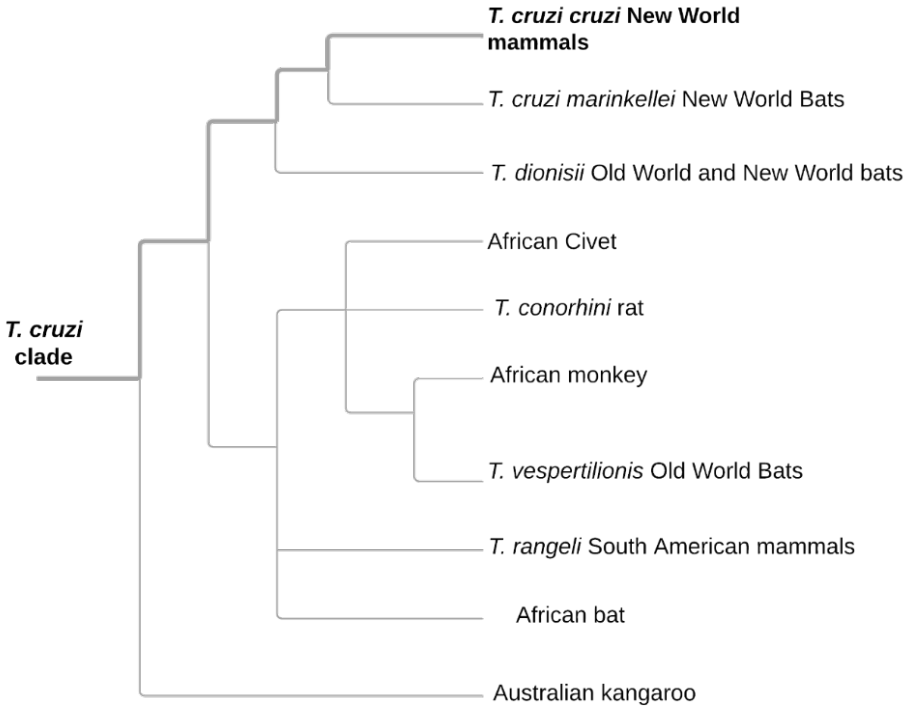

Figure 1. The phylogenetic tree of *T. cruzi* clade, modified from [73,74].
Made by: Francisco Javier Mendoza Proaño.

1.2. Neglected Tropical Disease: Chagas Disease

T. cruzi is the causative agent of Chagas Disease, one of the most critical diseases in Latin America and considered one of the 13 most neglected tropical diseases in the world [17,75–78]. Likewise, according to the World Health Organization (WHO) morbidity estimates, it ranks first among parasitic diseases in the Americas [76,78].

Increasing globalization has introduced this parasite to other continents; however, spraying campaigns in rural areas have considerably reduced the population of the transmitting insects during the last 25 years, reducing the number of infected people from 30 to 8 million [21,75–77,79–81]. Between 20 % and 30 % of those infected may develop Chagas Disease symptomatically, with the potential to be fatal [78]. These individuals may experience cardiomyopathies, digestive mechasyndromes, or both conditions [76].

Likewise, there is a need to generate trypanocidal drugs with greater efficacy, such as benznidazole or nifurtimox, which is a safe and effective alternative with a wide range of anti-protozoa action [76,82,83]. On the other hand, this disease is considered an opportunistic infection in immunocompromised persons, as well as those with human

immunodeficiency virus (HIV) [73,75,76,78,83].

Chagas disease is a public health problem in Ecuador and many other Latin American countries. This disease is endemic in 21 countries, where it affects more than 6 million people, and nearly 70 million live in areas exposed to it. Likewise, it is estimated that 30 000 new cases are generated annually due to all forms of transmission, 12 000 deaths on average, and 8 600 newborns who have contracted the disease during pregnancy [84–87].

In Ecuador, from 2013 to Epidemiological Week 50 of 2023, the Epidemiological Surveillance Subsystem of the Ministry of Public Health (SIVE-MSP, for its acronym in Spanish) records more than 842 confirmed cases of Chagas disease [84,85,88,89]. Despite this, in Ecuador, it is difficult to know with certainty the epidemiological situation, the repercussions, and the situation regarding the prevention and control of this disease. In addition, the provinces with the highest number of cases until 2019 were El Oro with 104, Guayas with 64, and Loja with 60 [84].

1.3. Life cycle and mechanism of action of *T. cruzi*

T. cruzi has a highly complex life cycle, and more than 100 species of Triatominae (Hemiptera: Reduviidae) have been identified that may act as vectors for its transmission. These insects are widely distributed throughout Latin America and the southern United States, especially in rural areas [75,90].

The infection cycle begins when an insect (vector) feeds on the blood of an infected mammal, ingesting trypomastigotes and amastigotes of the parasite. These are transformed into replicative epimastigotes, and after 3 to 4 weeks, the infective meta-cyclic trypomastigotes are found in the feces of the vector without having divided in its hindgut [75,76,91–93]. Wastes contaminated with parasites can be transmitted to a new host if there is contact with the new host's oral and nasal mucosal membranes and other exposed surfaces [75,94,95].

There are other ways by which the parasite can enter the host, such as blood transfusions. For this reason, methods of screening blood donors for Chagas disease have been implemented. However, this form of transmission is less common than that mentioned above [75,76,78,93,96]. In addition, the parasite can also be transmitted congenitally. According to several studies, carrier pregnant women have a transmission

incidence of 1 % and 10 % [75,97–99]. Finally, other transmission forms exist, such as ingesting food or beverages contaminated with trypomastigotes, laboratory accidents during experiments involving this parasite, and organ transplants [75,76,78,93,100–103].

This disease is composed of two phases: the acute phase, which lasts 6 to 8 weeks, during which many infected patients appear healthy once the disease subsides. Standard methods can detect no organ damage during this phase, so diagnosis is based on parasitological or serological tests. This is followed by the chronic phase, also known as the indeterminate form, in which patients remain indefinitely and may last a lifetime. After several years in this phase, approximately 20 % to 35 % of infected patients may develop irreversible lesions in the autonomic nervous system of the colon and esophagus, and the heart and peripheral nervous system (see figure 2) [76,104–107].

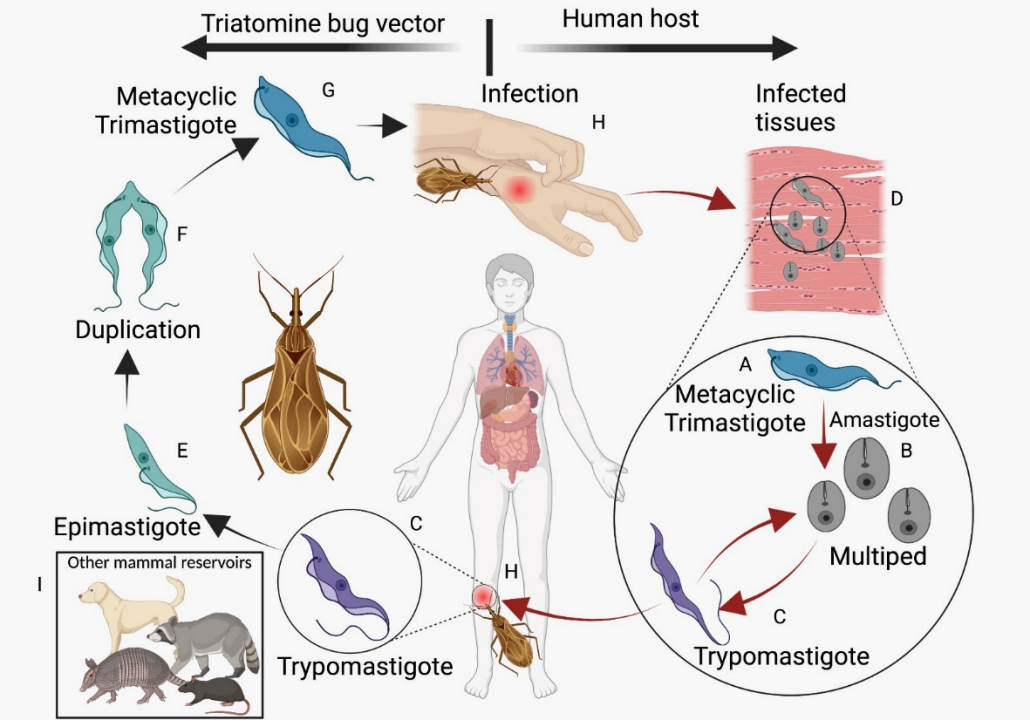

Figure 2. Life cycle and mechanism of action of *T. cruzi*. The triatomine feeds on the host (H), excreting feces that harbor the metacyclic trypomastigotes (A); these transform into amastigotes (B). Amastigotes will multiply by binary fission in the affected tissue, becoming blood trypomastigotes (C) lyse the cells to enter the bloodstream and infect new cells that will transform back into amastigotes (D). Blood trypomastigotes will become epimastigotes (E) in the midgut of the vector, then multiply (F) and migrate to the rectum to transform into metacyclic trypomastigotes (G) and be released with the feces of the triatomine (H). There are other reservoirs of mammals, such as armadillos, dogs, raccoons or rats (I). Adapted from [108].
Made by: Francisco Javier Mendoza Proaño.

1.4. Host immune regulation in the presence of the pathogen

The interaction of the PVH system will be affected by different physiological responses central to the modulation of autophagy, which is a pathway that comprises the degradation of macromolecules, cellular structures, and cell death. For the same reason, the intracellular infection of the parasite would affect the mobilization of calcium and endocytosis of PI3K signaling (activator of tyrosine kinase signals) [109]. On the other hand, during the parasite infection, it will be affected by some components of the immune system, such as the physical barrier, intestinal barrier, abiotic aspects that intervene in the immune response, intestinal microbiota, or survival and adaptation mechanisms of the parasite [110].

Therefore, to avoid the aforementioned, *T. cruzi*, to guarantee its survival inside the host, would avoid different physiological processes such as phagolysome, the expression of virulence factors, direct immunomodulation, and establishment in sites of latency. For example, trypomastigotes evade nucleated cells through different mechanisms, whether the target cell is phagocytosed or not phagocytosed. On the other hand, innate effector cells such as macrophages (M1) are activated by gamma interferon, increasing the expression of nitric oxide synthase and the production of nitric oxide to destroy the parasite. However, macrophages can acquire alternative phenotypes (M2) that will reduce oxide production, increasing the persistence of *T. cruzi*. On the other hand, the reaction protein crosses C9 of the parasite, allowing it to evade the microbicidal activity of macrophages, escaping from the phagolysosome to the cytoplasm, and thus promoting the dephosphorylation of the signal transducer and activator of transcription 1 (STAT1) to interfere the transcription of interleukin 12 (IL12) and tumor necrosis factor (TNF), canceling the microbicidal response. Likewise, this parasite is capable of stimulating the secretion of anti-inflammatory cytokines (interleukin 10 (IL10) and transforming growth factor beta (TGFB)), impairing the protective immune response and promoting the expansion of the infection and resistance of the parasite. In addition, it is capable of inhibiting the assembly of the C3 and C5 convertase, and the action of the AGC10 glycoprotein of the parasite decreases the T cell response during the disease as an effect of the negative regulation of interleukin 2 (IL2) [111].

1.5. Protein-Protein Interactions (PPI) and their application in the study of diseases

In recent years, omics data has increased substantially in terms of quality and quantity. In particular, those from proteomics and transcriptomics can be used during the development of biological networks, where protein-protein interactions (PPIs) are the most important and studied [112,113].

The cell is the basic unit of life; it is made up of biomolecules, such as proteins, which are large and complex molecules necessary for the structure, function, and regulation of different biological processes through complex and organized networks of interactions. [114]. These protein interactions are based on the physical binding of 2 or more proteins in response to various perturbations, providing considerable adaptability for cells to more easily adapt to changing environmental conditions [114–116]. Additionally, some of the most used databases to obtain information on the interaction between proteins are BioGRID [117], DIP [118], GeneMANIA [119], HMDAD [120], HPIDb [121], HPRD [122,123], IMEx [124], InnateDB [125], IntAct [126], MatrixDB [127], MINT [128], MPIDB [129], OrthoHPI [130], PHISTO [131], PHI-base [132], PINA [133], Reactome [134] and String [135] same ones that provide information about the interaction systems [136,137].

1.5.1. Parasite-Vector-Human System Interaction Network

The parasite-vector-human system in Chagas Disease, specific interactions occur between proteins of the parasite (*T. cruzi*), vector (*R. prolixus*) and human (*Homo sapiens*) [71]. They are essential for cell invasion, parasite replication and establishment of infection in the host [84–86,138,139]. These play an essential role in the survival and transmission of the parasite through the vector, as well as in the adaptation of the parasite to the vector's environment. Therefore, they are closely related to parasite pathogenicity, host immune response and disease progression [84,85,140,141].

Finally, understanding the characteristics of PPI in the Parasite-Vector-Human System (PVH) is vital to identifying therapeutic targets, developing disease prevention and control strategies, and improving the understanding of the molecular mechanisms underlying Chagas Disease [142–144]. This analyzes required data interaction such as that offered by the Bioinformatics Resources Center (BRC), focused on collecting DataSets of eukaryotic pathogens and invertebrate vectors of infectious diseases, such as TriTrypDB (https://tritrypdb.org) [145], HostDB (https://hostdb.org) and VEuPathDB

(https://veupathdb.org) [146]. Therefore, future research may provide new perspectives for developing more effective therapies and eradicating this disease [142–144].

Consequently, this study aims to understand the parasite-vector-human system of Chagas Disease through an automated analysis of Protein-Protein Interaction networks. Firstly, data on interactions between proteins relevant to Chagas Disease in the parasite-vector-human system were selected. Next, an algorithm was developed using Python software to analyze the collected data and construct a PPI representing the parasite-vector-human system's interactions. Finally, carry out an analysis and evaluation of the constructed PPI, identifying the most relevant nodes and links, as well as the properties and characteristics of the network.

2. METHODS

2.1. Data mining and pathogenicity gene screening

Chagas DB (https://chagasdb.tagc.univ-amu.fr/ (Last updated: May 31, 2023) (S1 Table) is a repository of manually selected information on molecules significantly related to *T. cruzi* infection from different host species. It includes information on genes, proteins, polymorphisms, hormones, or other chemical compounds that show their interaction during infection. For this analysis, only human protein data related to the diseases was selected; feature data was used from them.

On the other hand, protein data for the parasite (*T. cruzi*) and its vector (*R. prolixus*) were obtained from [113,135,147] (S1 Table). Once the genes of interest are received, these will be extracted through the use of regular expressions (re-library) [148]. Next, they will link to the RCSB: PDB page using the Selenium package's navigation drivers (API). These variables are then stored in a "TSV" file to filter and purge the protein names into a single protein naming format [149]. To do this, the Python v3.11.1 packages [150] will be used again to change their name, which was in PDB format [94], to perform UniProt and String format to perform the analysis [147].

2.2. Assumptions for the construction of the PPI of the PVH System

As previously mentioned, the assumptions of the PPIN were based on the criteria used, with some modifications (conversion of protein names and generation of the PPI) [151]. This will help to evaluate, find, and optimize the search for genes that generate pathogenicity with the VPH System, 1) conversion of different protein identifiers present in RCSB: PDB (Released: December 26, 2023) [152], UniProt (Released: May, 2023) [147] and String v12.0, 2) Transformation of pathogen names into a homogeneous and unique format; concerning the String identifier to purge disparities in syntax and nomenclatures and 3) Discard duplicate records of the data to avoid redundancies during the information search [137].

2.3. Construction of the PPI of the PVH System

Once the filtered and purged PVH System data is obtained, it will be linked using the String API v12.0 [135,153] in Python v3.11.1 [154,155] to calculate their interactions to perform comparisons of the attributes of the PPI network. After this, the collected information will be stored in a data frame, and the network will be generated with the help of the NetworkX library; the PPI was built as proposed by [156] with some modifications [137,154,157] using NCBI-specific taxon identifiers 353153, 13249 and 9606 for *T. cruzi* CL Brener/*T. cruzi* (vector), Reduviidae/*R. prolixus* (parasite) and Human/*H. sapiens* (host), respectably [135]. Consequently, a graph containing vector, parasite, and host proteins was generated. According to [137], this network has three types of edges: 1) host and pathogen, 2) pathogen and vector, and 3) intra-organisms interaction.

2.4. Topological Identification of Nodes of Interest for the PPI of the VPH System

The topological analysis of complex networks, such as metabolomics or transcriptional regulation, will help identify essential nodes related to vital functional activity in organisms. They are important when identifying topological features. For this reason, the following variables were analyzed: Connectivity (Connected Components, Network Diameter, Degree of Nodes, Clustering Coefficient, and Shortest Path between 2 Nodes) [158–161], Centrality Measures [162] (Degree Centrality, Closeness Centrality, and Betweenness Centrality) [163–165], existing Communities [166] (Girvan-Newman [167,168], Louvain [169] and Clauset-Newman-Moore [170]) and Main links between Proteins of Interest using the NetworkX package of Python v3. 11.1 [112,137,156,157,171,172].

Next, this section will define the basic notation measures used throughout the paper for connectivity, as they said by [161,173–176]; centrality measures, as mentioned [163], with some modifications [165]; and communities, as they mentioned [167–170,177]. Let $G = (V, E)$ be an unweighted undirected network, where V and E are the conglomerates of all nodes and edges, respectively, that are present in a network. The following nodes represent their cardinality: $|V| = n$ and set of links (edges) $|E| = m$. On the other hand, let us assume the adjacency matrix (A) of size $n \times m$ in the network G, where the entry a_{ij} indicates whether there is a relationship between nodes i and j if $a_{ij} = 1$; at the same time, if $a_{ij} = 0$, there is no relationship. Furthermore, d_{ij} is the geodesic distance

between both nodes (i and j) or the shortest path length.

2.4.1. Connectivity of the PPI of the PVH System

A network does not necessarily have to have a single set of connected nodes; most keep two or more separate sections disconnected. These will be called components; by definition, there is no path between any pair of nodes of the different components. [174].

Likewise, random network theory predicts that a giant component should be observed if the average degree of node k is greater than 1. This indicates that the majority of the nodes when connected to each other, will form a significant component of the network. However, most of these networks do not satisfy the condition k greater than $\ln N$, which suggests that they tend to divide into isolated groups according to the $G(N,p)$ model (N) labeled nodes are connected with probability p, which considers the probability of connection between nodes. Indeed, although some networks show fragmentation, most do not exhibit this characteristic, which is consistent with expectations derived from random network theory [175].

2.4.1.1. Network Diameter

The diameter of a network is the length of the "longest and shortest path" that separates two nodes. The diameter will always be the longest of the shortest paths between each pair of nodes in the network in which there is a path [174,178]. For graph G, the Network Diameter (ND) of any node u, v and $d(u,v)$ length of the shortest path (minimum number of edges) between nodes can be calculated with the following formula [173,174,176].

$$ND = max_{u,v} d(u,v) \qquad (1a)$$

Nevertheless, denotation is also by the maximum shortest path in the network. In this sense, it is the most significant distance recorded between any pair of nodes.

$$d_{max} \qquad (1b)$$

2.4.1.2. Degree of Nodes

The Degree of a Node k is a node's number of adjacent links. In other words, the number of edges that connect to these. However, if they are undirected networks, the degree of this will be the number of interactions with other nodes. [174,178]. For graph G, the Degree of a Node (L) of any node i^{th} can be calculated with the following formula, where the 1/2 factor corrects for the fact that in the sum (2.1) each link is counted twice [175].

$$\mathrm{L} = \frac{1}{2}\sum_{i=1}^{N} k_i \qquad (2)$$

2.4.1.3. Shortest Path Between Nodes

The number of links a node has, also called "steps," refers to the minimum number of interactions required between two different proteins (i, j) in a network to connect. Ergo, the average distance of all pairs of group nodes is calculated [174,178]. For graph G, the Shortest Path Between Nodes (d) of any node i, j can be calculated with the following formula [175].

$$d(i,j) = \begin{cases} if\ d_{ij} = d_{ji} & It\ is\ an\ undirected\ network, \quad (3a) \\ if\ d_{ij} \neq d_{ji} & It\ is\ an\ undirected\ network, \quad (3b) \end{cases}$$

$d_{ij} = d$: If there is a path of length d between i and j, then $A_{ik} \ldots A_{lj} = 1$ $(A_{ik} \ldots A_{lj} = 0, \mathrm{otherwise})$. The number of paths of length d between i and j is:

$$N_{ij}^{d} = A_{ij}^{d} \qquad (3c)$$

2.4.1.4. Clustering Coefficient

The Clustering Coefficient measures the proportion between links of a node in the i-neighborhood (interconnectivity of the neighbors of a vertex) fractioned by the number of edges that can be had between them [178,179]. The paths with length two are counted; the ones that are closed are listed, and the second is divided by the first, obtaining a number between 0 and 1 [161,174]. For graph G, the Clustering Coefficient $(CC\ or\ C_i)$ of any node v with degree k_v and n_v denotes the number of triangles that go through node can be calculated with the following formula [161,174,179].

$$CC(v) = C_i = \frac{2n_v}{k_v(k_v - 1)}, \quad C_{WS} = \frac{1}{n}\sum_{i=1}^{n} C_i \quad (4)$$

2.4.2. Centrality Measures of the PPI of the PVH System

Centrality Measures allow us to determine proteins (nodes) that are central or that are highly connected to others in a network. This will enable us to establish whether these nodes play vital roles during the interaction between proteins, understand the topological characteristics of the network, or identify molecular targets [164,180]. Whereby, within the analysis of the main Node Centrality Measures (Freeman's Centrality Measures), there are Degree Centrality, Closeness Centrality, and Betweenness Centrality [162,164,172].

2.4.2.1. Degree Centrality

The Degree is the most straightforward Centrality measure in which a node has a union with others within the same network that can be represented, where the nodes with a high degree of centrality will be those with the most significant number of connections and will be called "central to the network" [164,174]. For graph G, the Degree Centrality (DC) of any node u can be calculated with the following formula [163,165].

$$DC(u) = |(u, \dots, \dots) \in E| \; or \; DC(u) = \sum_{v \in V \neq u} a_{u,v} \quad (5)$$

2.4.2.2. Closeness centrality

Closeness Centrality is the inverse of remoteness, defined as the average of all the shortest distances between nodes $d(node\ 1, node\ 2)$. That is the score of the sum of the lengths of time in which information from one node can flow to another node or vice versa. For graph G, the Closeness Centrality (CC) of any node u can be calculated with the following formula [163,165,174].

$$CC(u) = \sum_{v \in V \neq v} \frac{1}{d(u, v)} \quad (6)$$

2.4.2.3. Betweenness centrality

Betweenness Centrality measures the degree or proportion to which a node will interact with the information path of a pair of different nodes. Therefore, a node along the shortest path between another pair of nodes is more significant than a node with fewer shortest paths. For graph G, the Closeness Centrality (BC) of any node u can be calculated with the following formula [163,165,174].

$$BC(u) = \sum_{\substack{i \neq u \neq j \\ i,u,j \in V}} \frac{\sigma_{ij}(u)}{\sigma_{ij}} \qquad (7)$$

2.4.3. Communities of the PPI of the PVH System

Communities play essential roles in the understanding of biological processes (e.g., coding of biological functions in cellular networks) and knowledge of diseases (e.g., proteins that are related to the same disease always tend to interrelate with each other) [175].

The detection of communities, sometimes also called "clustering" (it is important not to confuse this term with clustering coefficients), is the set of methods that allows finding groups of nodes, representing them in substructures of nodes with similar characteristics or connection patterns network specific; applying different criteria, e.g., optimization of modularity, defined as a number between -0.5 and 1 (modularity (Q) is a measure that estimates the quality of a network partition in other communities. Ergo, the greater the modularity, the better the structure of the identified communities) [174,175,177].

2.4.3.1. Clauset, Newman and Moore Algorithm (CNM)

This algorithm, also known as Fast Greedy, finds the modularity variation (ΔQ) that combines the different pairs of resulting communities and chooses the most significant modularity gain [170,177]. In other words, at the beginning, it will consider that each node is a community and the network is a multigraph, where the community given by a node and the values of the adjacency matrix will equal the number of edges between the communities. Subsequently, it calculates the change in modularity to find the pairs with the highest gain in the community and group them [177]. Accordingly, this algorithm focuses on feeding and refreshing the values of the matrix. ΔQ_{ij}, containing the most

prominent element of each row based on its max-heap it, obtaining; as a result, the running time $O(m\,d\log n)$ [181]; for a network of nodes and edges (n and m, respectively), and $d\sim\log n$ (dendrogram depth). Whilst, for sparse networks ($m\sim n$), the running time $O(n\log^2 n)$; which indicates that it is essentially linear [170,177]. For graph G, the community algorithm of any node i,j and community k can be calculated with the following formulas.

This algorithm has the following steps for each i: 1) Estimate the initial values of ΔQ_{ij} and a_i using equations 8a and 8b. Next, the most significant element of each row ΔQ_{ij} will be completed with the support of the max-heap of the Hamiltonian of the network (H), grouping the labels i,j of each corresponding pair of communities, 2) Choose the largest ΔQ_{ij} of H, then combine the related communities, update the matrix $\Delta Q, H,$ and a_i and increase Q by ΔQ_{ij}, and 3) Repeat the step 2 until you have a single community.

$$\Delta Q = \begin{cases} \frac{1}{2m} - \frac{k_i k_j}{(2m)^2} & if\ i,j\ are\ connected, \quad (8a) \\ 0 & otherwise, \end{cases}$$

and

$$a_i = \frac{\mathrm{k}_i}{2m}\ (8b)$$

The data structure of the adjacency matrix allows the updates in step 2 to be performed quickly, with only the adjustment of a few elements of ΔQ. Likewise, when joining communities i and j, this combined community will be labeled as j so that only the j^{th} row and column need to be updated, and the i^{th} row and column were removed entirely. Due to the above, the updated rules are now the following:

1) If k is connected to nodes i,j, then equation 8c is satisfied

$$\Delta Q'_{jk} = \Delta Q_{ik} + \Delta Q_{jk}\ (8c)$$

2) If k is connected to a i, but not to j, then equation 8d is satisfied

$$\Delta Q'_{jk} = \Delta Q_{ik} - 2a_j a_k\ (8d)$$

3) If k is connected to a j, but not to i, then equation 8e is satisfied

$$\Delta Q'_{jk} = \Delta Q_{jk} - 2a_i a_k \ (8e)$$

Update $a'_j = a_j + a_i$ and $a_i = 0$, it is trivial and will be done constantly [170].

2.4.3.2. Louvain Algorithm

This algorithm detects large network communities by maximizing modularity; when compared to other methods, such as the Girvan-Newman algorithm [167,168] and Spinglass algorithm [182], it has a longer calculation time; the limiting factor is the Random Access Memory (RAM) of the computer to use since it consumes a lot of this [169,177].

This algorithm is divided into two phases, which will be executed iteratively. During the first phase, a different community is assigned to each node in the network. Therefore, there are as many community nodes as possible at the beginning. The modularity gain will then be calculated to place the nodes in the appropriate communities. That is to say, the modularity gains when moving node i to community C can be calculated with the following formula. This process will be repeated for all nodes until no further improvements can be made to community detection. [169,177]. For graph G, community algorithm of any node i, j and community k can be calculated with the following formula.

$$\Delta Q = \left[\frac{\Sigma_{in} + \mathrm{k}_{i,in}}{2m} - \left(\frac{\Sigma_{tot} + \mathrm{k}_i}{2m}\right)^2\right] - \left[\frac{\Sigma_{in}}{2m} - \left(\frac{\Sigma_{tot}}{2m}\right)^2 - \left(\frac{\mathrm{k}_i}{2m}\right)^2\right] \qquad (9)$$

Where, Σ_{in} is the sum of the weights of the edges within C, Σ_{tot} is the sum of the weights of the edges within$C$, k_i is the sum of the weights of the edges incident to the nodes of i, k_{in} is the sum of the weights of the edges from i to the nodes in C, and m is the sum of the weights of all edges in the network [169].

The other part of the process is building a network, in which the nodes will be the communities detected in the first part of the process. The weights of the edges of the new nodes will be given by the sum of the edges that the nodes of the two communities have, respectively, making the edges between these nodes of the same community result in "automatic loops" in the new network community. Once this stage is completed, the first phase of the algorithm will be applied to the network again; this combination of processes will be called "pass." These can be iterated until the maximum modularity is

reached, making the algorithm relatively fast (running time: $O\ (n \log n)$) [169,177].

2.5. Functional Enrichment Analysis of the PPI of the VPH System

Functional enrichment was performed to identify the most functional pathways and processes for the PPI of the PVH system using gene set enrichment analysis using the Enrichrpy library [183–186] together with the GO_Biological_Process_2021 gen set (Gene Ontology: https://www.informatics.jax.org/vocab/gene_ontology/GO:0008150) (latest update from the database December 26, 2023) [187–189] and the GSEApy library [190,191] along with the genes sets MSigDB_Hallmark_2020 (Molecular Signatures Database: https://www.gsea-msigdb.org/gsea/msigdb/collections.jsp) (released in February, 2023) [192] and the KEGG_2021_Human (KEGG PATHWAY database: https://www.genome.jp/kegg/pathway.html) (Release 109.0, January 1, 2024) [193], as mentioned by [194], with some modifications [137]. This process allowed the recognition of statistically significant gene ontologies and enriched pathways. String v12.0 [135] rich profiles also can be calculated using the FunRich v3.1.4 tool [137,195,196].

3. RESULTS

Methods for detecting protein complexes have been analyzed and tested mainly in yeasts such as *Saccharomyces cerevisiae* because these organisms have been widely studied, and their data are curated. In this case, when evaluating data from *H. sapiens*, there are certain drawbacks since the data for generating Human PPIs mainly have noise or the protein complex interaction data size is too large compared to yeast [159]. In this work, 101 602 genes were analyzed, and after being filtered and purified, 985 genes were obtained: 501 nodes and 5 796 interactions (Figure 3) (S1 Table).

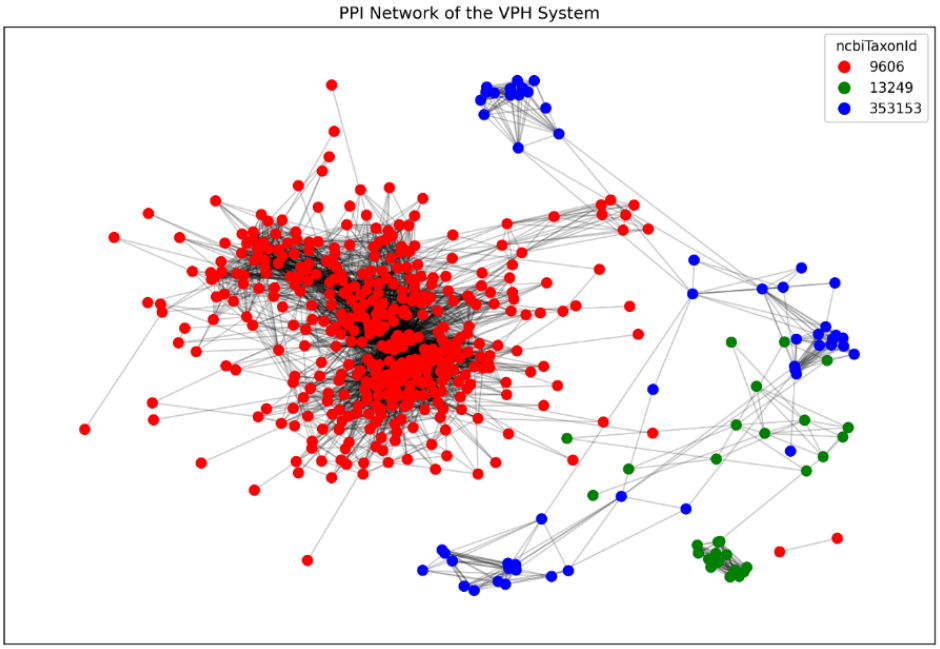

Figure 3. Graphic representation of the PPI Network of the PVH System. Colors are based on the NCBI-specific taxon identifiers for better schematization, 419 nodes in red (taxon identifier: 9606): Human/*H. sapiens*, 53 nodes in green (taxon identifier: 13249): Reduviidae/*R. prolixus* and 29 nodes in blue taxon identifier: 353153): *T. cruzi* CL Brener/*T. cruzi*.
Made by: Francisco Javier Mendoza Proaño.

3.1. Connectivity Analysis of the PPI Network of the PVH System

In the connectivity analysis, it is crucial to understand the structure of the PPI network in the PVH System of Chagas Disease. This is supported by analyzing the value of the degree of connectivity, the diameter of the network, the shortest path between nodes, and its clustering coefficients.

The analysis of the network diameter of the PPI Network of the PVH System indicates that it is partially connected since it has four sets of connected elements (S2 Table). Likewise, each of these has its diameter and number of components, which suggests that within the leading network, there are independent subnetworks within this set of interactions.

The component with the most significant number of elements is Connected Component 1 (S1_1 Fig), which has a diameter of 5; i.e., the maximum distance between any pair of nodes within it is five links. Likewise, this component is significantly enormous, indicating a network with good interconnection since it has 417 out of 501 possible.

However, Connected Components 2, 3, and 4 (S1_1, S1_2, S1_3 and S1_4 Fig) have smaller diameters but with more compact and less extensive subnetworks. The component with the most minor diameter and number of elements, 1 and 2, respectively, is Connected Component 2. This indicates that this component has a more isolated network than the others and that its interactions will be particular since they are only between two proteins.

On the other hand, the other Connected Components, 3 and 4, have a diameter of 5 and 6, respectively, along with sizes of 29 and 53 elements. These data suggest subnetworks of moderate size with indisputable complexity during their interactions with proteins. These data highlight the importance of understanding the network's structure at the level of its connected components, acquiring a more detailed perspective of the topology and functional association of protein interactions in Chagas Disease.

It is essential to remember that there are nodes with high and low connectivity degree values. The one with the highest degree of connectivity is the ALB node, which has a value of 189 and indicates its central role within the network. This is followed by other essential proteins, such as TNF (Tumor Necrosis Factor), IL6 (Interleukin 6), GAPDG (Glyceraldehyde-3-Phosphate Dehydrogenase), and IFNG (Interferon Gamma), with a connectivity degree of 188, 186, 139, and 137, respectively. These grade values are significantly high, highlighting their level of participation during protein interactions related to Chagas Disease and its pathogenesis.

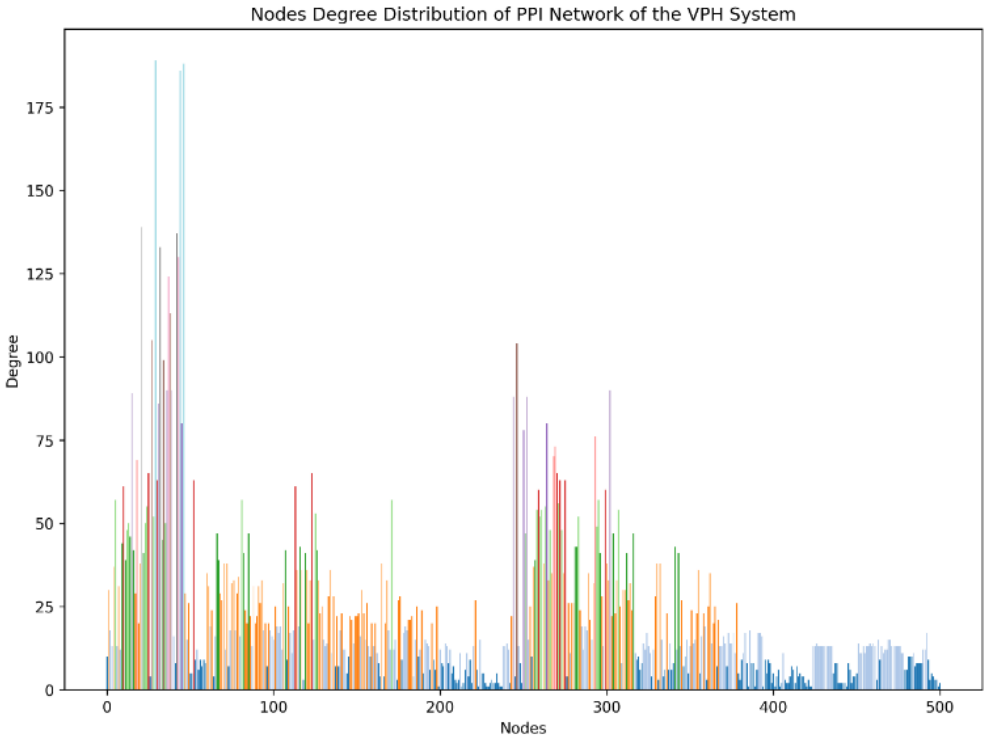

Figure 4. The degree distribution of the nodes in the graph indicates the existence of different interactions with high and low values. Nodes with high values will act as “huns” since they connect with multiple proteins. The pattern in the image indicates a non-uniform degree distribution with a particular upward bias in nodes with a high number of connections, e.g., the nodes ALB, TNF, IL6, GAPDH, or IFNG.
Made by: Francisco Javier Mendoza Proaño.

On the other hand, it is also vital to point out that there were nodes with low connectivity degree values. Among them, we can mention ZBTB43 (Zinc Finger and BTB Domain Containing 43), REG1A (Regenerating Family Member 1 Alpha), LGALS7B (Galectin 7B), PDZD4 (PDZ Domain Containing 4), and KEL (Kell Metallo-Endopeptidase), all with a value of degree 1. This highlights that the pre-theoretical diversity of the network is high, in addition to the fact that these peripheral non-nodes and their roles are more specific during the infection and cellular regularization processes of Chagas Disease (S3 Table) (Figure 4 and 5).

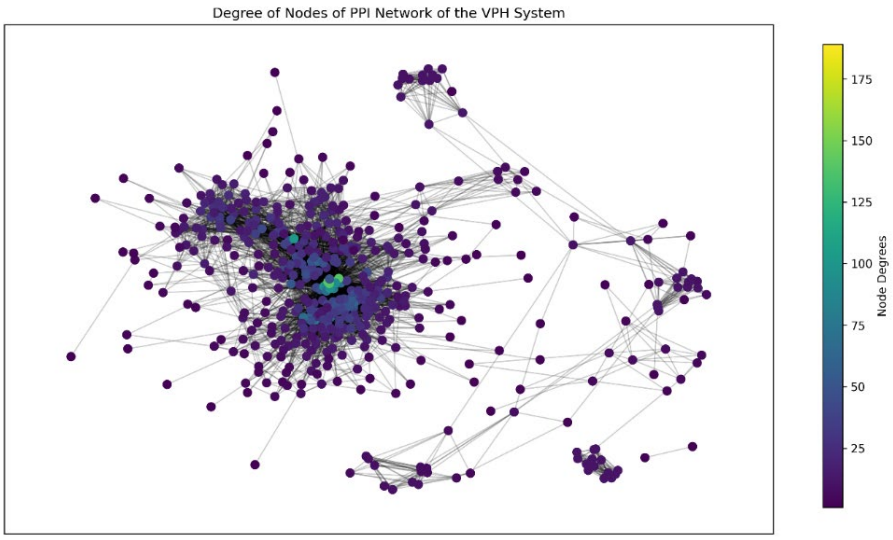

Figure 5. Graphic representation of the Degree Value of Nodes of PPI Network of the PVH System.
Made by: Francisco Javier Mendoza Proaño.

The analysis of the shortest paths is closely related to the connected components (S4 Table) and the network's diameter since only the elements that make up the different sets or communities of the network can be related. For this reason, the maximum number of paths will equal the network's diameter or subnetwork, depending on the case to be analyzed.

In the case of the nodes coming from *T. cruzi*, as an example, the path between Pyr4 (Orotidine 5'-phosphate decarboxylase) and Q4E4A8_TRYCC (Mismatch repair protein MSH6, putative) (pyr4 -> Q4D6I5_TRYCC -> Q4CQQ9_TRYCC -> Q4DPB9_TRYCC -> Q4D365_TRYCC -> Q4DBN2_TRYCC -> Q4E4A8_TRYCC), has a network diameter of 6, or the path of the nodes T1I0P5_RHOPR (Uncharacterized protein) and T1HPM3_RHOPR (Potassium channel tetramerisation-type BTB domain-containing protein) of the organism *R. prolixus*, whose path (T1I0P5_RHOPR -> R4G5E5_RHOPR -> T1HFL6_RHOPR -> R4G5K0_RHOPR -> R4FPL4_RHOPR -> T1HPM3_RHOPR) has a diameter of 5. Both indicate a chain of connections of various proteins. Therefore, they may be closely related to relevant metabolic interactions in the context of Chagas Disease.

Likewise, the shortest paths, such as that of the KRT15 (Keratin, type I cytoskeletal 15) and MAST2 (Microtubule-associated serine/threonine-protein kinase 2) nodes (KRT15 -> KRT14 -> VIM -> CFTR -> SLC9A3R1 -> MAST2) both from *H. sapiens*, reveal more specific protein associations since they may have connections of functional or signaling relevance of both nodes with their pathway proteins. Therefore, they would provide

relevant information on cellular pathways that could be involved in the disease.

Finally, it is worth mentioning the previously mentioned protein pathway, whose value degree was considerably high: MYH14 (Myosin-14) and IL6; MYH14 -> MYL2 -> ALB -> IL6, MSH4 (MutS protein homolog 4), and TNF pathway; path MSH4 -> MSH2 -> CCNB1 -> TNF and KRT2 (Keratin, type II cuticular Hb2) and ALB; path KRT2 -> KRT9 -> KRT14 -> ALB. As can be seen, all the mentioned roads have a diameter of 3, indicating their highly significant participation in dense communities.

Nodes with significant clustering coefficients were identified, among which T1HQF3_RHOPR (Ribosome production factor 2 homolog) stands out; T1HZK0_RHOPR (Uncharacterized protein); T1I0P5_RHOPR (Uncharacterized protein) and T1I2X8_RHOPR (Uncharacterized protein) (*R. prolixus*-associated proteins), Q4E4A8_TRYCC; Q4DIZ8_TRYCC (SH3 domain protein, putative); Q4DBA9_TRYCC (ATP-dependent RNA helicase) and Q4CUH4_TRYCC (SH3 domain-containing protein) (proteins associated with *T. cruzi*) and MTCP1 (Protein p13 MTCP-1); MUC16 (Mucin-16), MYH14, KIAA1217 (Sickle tail protein homolog), and HBA2 (Hemoglobin subunit alpha-2) (*H. sapiens*-associated proteins).

All the nodes above exhibit maximum clustering coefficients, i.e., 1, indicating that they are fully integrated into dense groups of neighboring nodes (Figure 6). Therefore, this agglomeration alludes to the fact that these can form or be related to protein complexes specific to certain metabolic pathways. However, there are also other nodes, such as GMDS (GDP-mannose 4,6-dehydratase), GSTA4 (glutathione S-transferase alpha 4), and KEL, whose clustering coefficient is 0. Therefore, this suggests a need for more dense local connections. With a minimum clustering coefficient, these nodes have a more peripheral role, acting more on more specific or isolated cellular functions (S5 Table).

On the contrary, the nodes with good value degree ALB, TNF, GAPDG, IL6, and IFNG obtained the following clustering coefficients: 0,1665; 0,1792; 0, 1781; 0, 1834 and 0,2653; correspondingly. This indicates that nodes like IFNG participate more significantly in dense communities than IL6 and GAPDH. On the other hand, the TNF node has a clustering coefficient similar to IL6 and GAPDH, indicating a similar network structure, although with slightly less cohesion than the others mentioned. Finally, ALB, having the lowest coefficient, tends to form less dense connections with an external role during the molecular interaction.

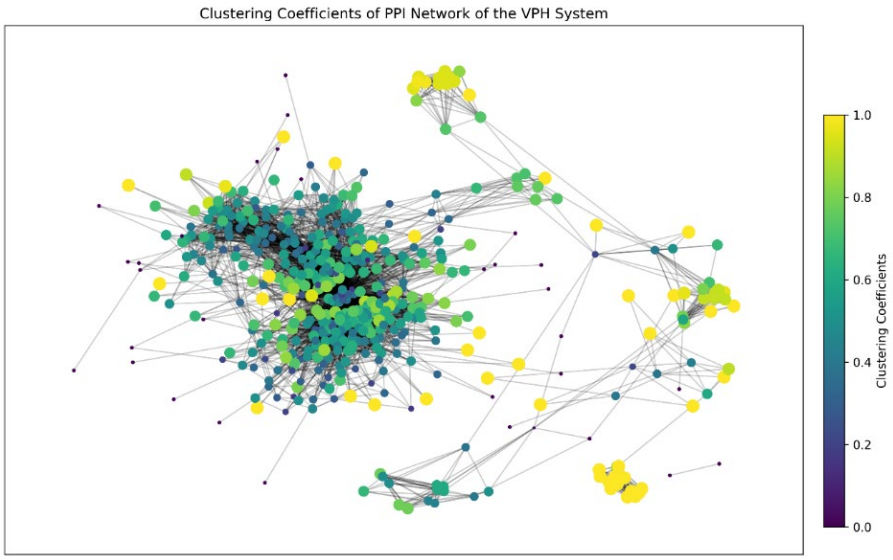

Figure 6. Graphic representation of the Clustering Coefficients of PPI Network of the PVH System.
Made by: Francisco Javier Mendoza Proaño.

3.2. Centrality Measures of the PPI Network of the PVH System

The centrality measures serve as indicators of relevant nodes in a network with the aim of better understanding its topology and dynamics. In other words, it shows specific nodes' roles during protein interaction. The most common centralities measures are degree centrality, closeness centrality, and betweenness centrality (S6 Table).

The first aspect analyzed in the study's exploration of network centrality measures is Degree Centrality (Figure 7). This measure identifies the importance of nodes in a network based on their connectivity. Nodes with notably high degree centrality values, such as ALB, TNF, IL6, GAPDH, and IFNG, play significant roles in the disease's pathogenesis. Conversely, MAN1B1 (mannosidase, alpha, class 1B, member 1), SLIRP (SRA Stem-Loop Interacting RNA Binding Protein), RNF115 (Ring Finger Protein 115), R4G5D9_RHOPR (putative 14-3-3-like protein), and Q4DI29_TRYCC (superoxide dismutase) have the lowest degree centrality values, indicating their biased intervention in Protein-Protein Interactions (PPIs). The variation in degree centrality values reflects the diverse functional roles of these proteins within the PPI network, highlighting the existence of subnetworks or available modules.

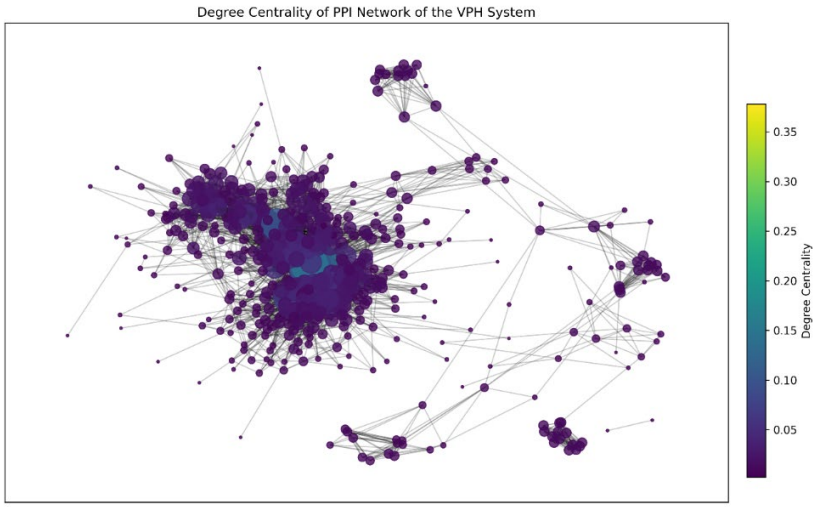

Figure 7. Graphic representation of the Degree Centrality of PPI Network of the PVH System.
Made by: Francisco Javier Mendoza Proaño.

Moving on to Closeness Centrality (Figure 8), this measure emphasizes nodes with significant proximity to others in the network. Nodes like ALB, TNF, IL6, GAPDH, TGFB1 (Transforming growth factor beta-1 proprotein), and IFNG are centrally located, facilitating shorter paths for communication with other proteins. In contrast, nodes like KEL and PDZD4 are more distant, requiring longer paths for information exchange. Notably, central nodes for the parasite and vector are identified, underscoring their crucial roles in data transmission within the PPIs.

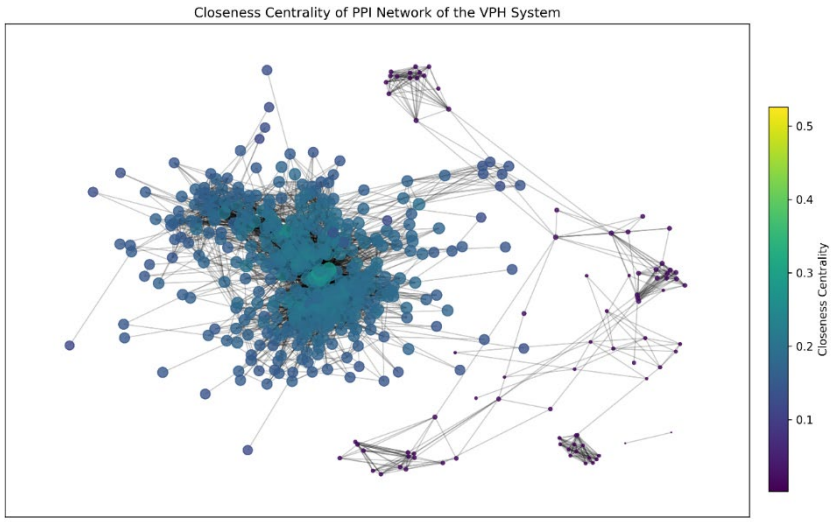

Figure 8. Graphic representation of the Closeness Centrality of PPI Network of the PVH System.
Made by: Francisco Javier Mendoza Proaño.

Betweenness Centrality analysis unveils proteins like ALB, GAPDH, and TNF as pivotal in information propagation between nodes (Figure 9). Their importance in metabolic pathways and potential roles as links between network modules are highlighted.

Conversely, KEL and PDZD4 nodes play a marginal role in communication routes. Specific proteins, such as Q4DNM5_TRYCC, Q4CL66_TRYCC, Q4E0X8_TRYCC, T1HFL6_RHOPR, R4G5E5_RHOPR, and T1HS28_RHOPR, are exemplified for their betweenness centrality values, providing insights into their biological processes in organisms.

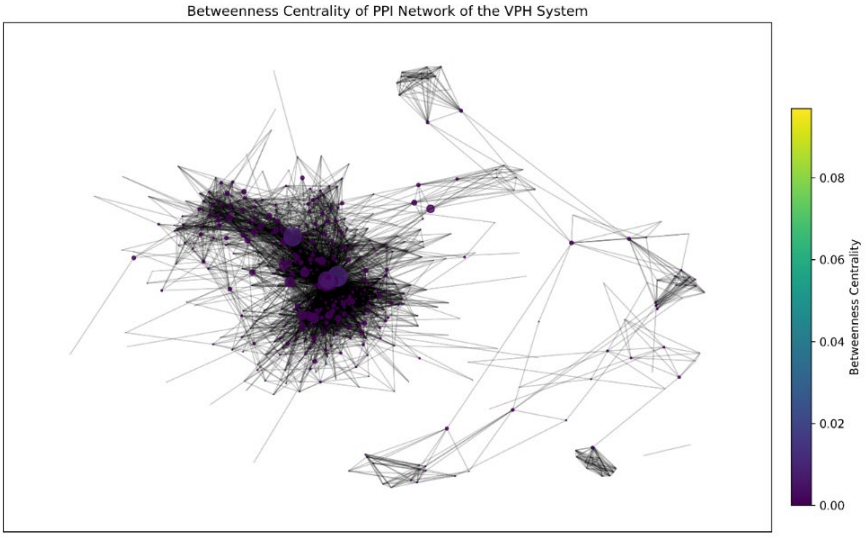

Figure 9. Graphic representation of the Betweenness Centrality of PPI Network of the PVH System.
Made by: Francisco Javier Mendoza Proaño.

3.3. Communities of the PPI Network of the PVH System

The Clauset-Newman-Moore (CNM) algorithm was employed to examine network structures. Seven communities emerged, each comprising various nodes, shedding light on the intricate organization, interconnections, and subnetworks formed by multiple nodes within biological contexts (Figure 10a). Community 1, housing 176 nodes, prominently features IL6, TNF, and ALB. This association suggests a significant link between biological processes related to the immune response and homeostasis in Chagas Disease, marking its importance for multifunctionality and central positioning in the PPI network. Conversely, communities 2 and 3, with 127 and 102 nodes, respectively, indicate subsets with greater specialization in protein interactions. Communities 4 and 5, associated with the parasite (53 nodes) and vector (29 nodes), provide crucial insights into these organisms' specific interactions, highlighting the parasite-vector-human interaction's biological complexity. Community 7, comprising only two nodes (KEL and PDZF4), notably signifies a more detailed and limited interaction within the PVH System (S7 Table).

The Louvain algorithm was also applied, revealing a more complex modular network structure with ten communities (Figure 10b). Communities 2 and 8, each with 114 nodes, and community 7, with 113 nodes, showcase greater complexity, protein diversity, and metabolic signal exchange. Community 2, akin to CNM community 1, features IL6 and TNF, emphasizing its relevance in the immune response and inflammation regulation in Chagas Disease. Community 4 (45 nodes) includes ALB, a node with high betweenness centrality, suggesting its role as a mediator in fundamental metabolic interactions due to its central positioning in the PPI network. Communities 11 and 5, with 2 and 14 nodes, respectively, represent more specialized subsets focused on identifying and understanding cellular regulatory signals within the PPI network. Finally, communities related to the vector and parasites (1, 3, 6, 9, and 10) offer insights into specific interactions concerning Chagas Disease, underscoring the network's modularity and highlighting the PVH System's high variability and biological complexity (S8 Table).

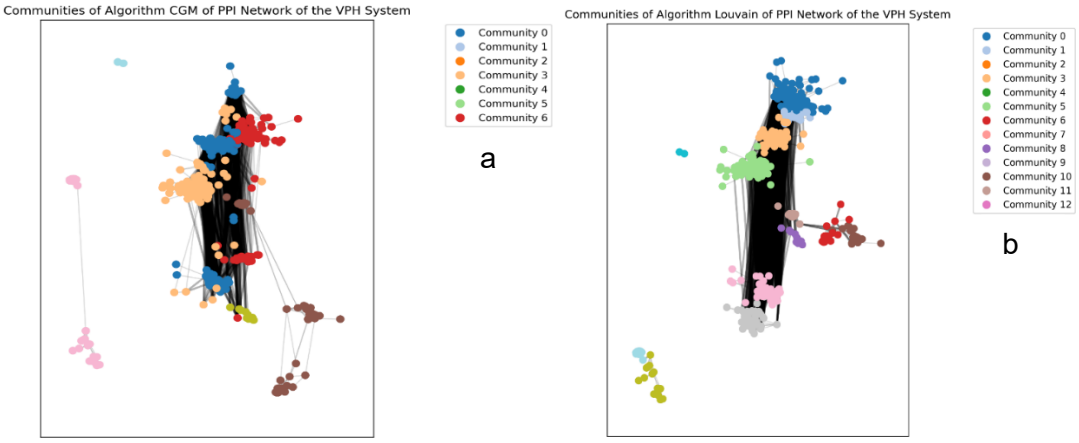

Figure 10. Graphic representation of the Communities of PPI Network of the PVH System., a) communities of CGM algorithm and b) Louvain algorithm.
Made by: Francisco Javier Mendoza Proaño.

3.4. Edges weight present in the Nodes of the PPI Network of the PVH System

The interaction levels and total amount of weight between the links give a quantifying perspective of the cohesion and strength of the PPI network (S9 Table). There were 5 796 interactions in this context, and the total weight was 3 770,39 (Figure 11). For the same reason, quantifying the energy of the edges (edge weights) indicates a robust, complex, and high-density network.

By dividing the network into its connected components, it was obtained that component 1 has the highest number of interactions and total weight between edges, 5 405 and 3 427,86, respectively (S2_1 Fig). These values indicate its importance in the global

network and that this component houses the vast majority of interactions and bond energy when compared to components 2, 3, and 4 (S2_1, S2_2, S2_3 and S2_4 Fig). Component 2 comprises a single interaction with an energy of 0,428, which is not insignificant and indicates a strong interaction despite its lack of connectivity. While components 3 and 4 have moderate interaction and energy levels proportional to the amount of components they have. These values indicate the specialty of these subnets.

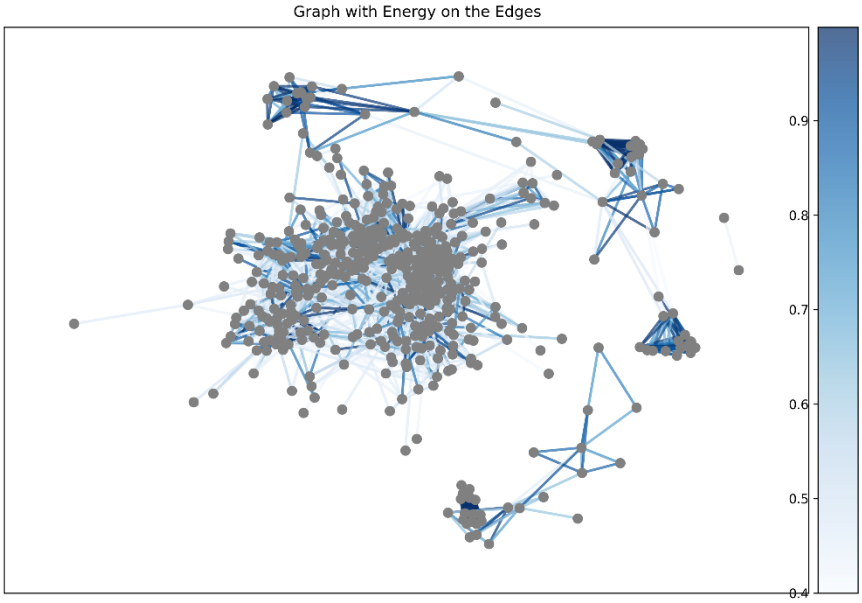

Figure 11. Graphic representation of the edge weight of the PPI Network of the PVH System.
Made by: Francisco Javier Mendoza Proaño.

On the other hand, the links with the highest level of interaction weight between the pairs of nodes were TNFRSF1A-TNF (TNF receptor superfamily member 1A), TNFRSF1A-TNFRSF1B (TNF receptor type 1-associated DEATH domain protein and TNF receptor-associated factor 2), HSPA5-PDIA3 (Transmembrane protein 132A and Protein disulfide-isomerase A3), SERPINE1-PLAT (Plasminogen activator inhibitor 1 and Phospholipase A and acyltransferase), and GAPDH-ALDOA (Fructose-bisphosphate aldolase A) all with 0,999; this value indicates that there is little interaction stable and weak between proteins, thermodynamically unfavorable. In contrast, the following node pairs have the worst interaction weight levels of the PPI network: HSP90AA1-CCL2 (Heat shock protein HSP 90-alpha and C-C motif chemokine 2), SERPINE1-DHFR (Dihydrofolate reductase), PARP1-TGFB1 (Poly [ADP-ribose] polymerase 1 and Transforming growth factor beta-1 proprotein), CCL4-IL12A (C-C motif chemokine 4 and I nterleukin-12 subunit alpha), and IL1RN-ADIPOQ (Interleukin-1 receptor antagonist protein and Adiponectin receptor protein 2) all with 0,4, indicating a strong affinity and stability, thermodynamically favorable (S9 Table).

3.5. Enrichment analysis of the PPI Network of the PVH System

In functional enrichment analysis, Enrichrpy provided insights into biologically relevant terms associated with the central nodes of the Protein-Protein Interaction (PPI) network in the PVH System. The cytokine-mediated signaling pathway (GO:0019221) emerged as a critical process (p-value: 1,78E-39, combined score: 655,39), involving genes like *CSF3*, *IL1RN*, *CD40* (Tumor necrosis factor receptor superfamily member 5), and TNF. These genes are linked to immune signaling processes and cytokine responses, aligning with the inflammation nature observed during human parasite infection. The cellular response to cytokine stimulus (GO:0071345) also played a significant role (p-value: 8,43E-31, combined score: 491,60), involving the same central genes and indicating their interconnected parts in immune response signaling pathways (S10 Table).

Moreover, positive regulation of intracellular signal transduction (GO:1902533, p-value: 1,11E-23, combined score: 293,26) highlighted genes *TNF*, *PIK3CG* (Phosphatidylinositol 4,5-bisphosphate 3-kinase catalytic subunit gamma isoform), and *AKT2* (RAC-beta serine/threonine-protein kinase), regulating intracellular signaling processes. Positive regulation of the multicellular organismal function (GO:0051240, p-value: 8,73E-20, combined score: 277,15) involved genes *TNF*, *FGF2* (Fibroblast growth factor 2), *IL10* (Interleukin-10), and *TGFB1* (Transforming growth factor beta-1 proprotein), suggesting their role in regulating broader biological processes. Lastly, platelet degranulation (GO:0002576, p-value: 5,53E-18, combined score: 448,09) emerged as a potential signaling mechanism during inflammatory responses and blood coagulation.

Moving to GSEApy, complement and coagulation cascade pathways stood out with significant gene involvement (adjusted p-value: 4,87E-23 and 6,03E-23, respectively) (S11 Table). Notable genes in these pathways include *SERPINA1* (Alpha-1-antitrypsin), *ITGAM* (Integrin subunit alpha M), *SERPINE1*, *C1R* (Complement C1r subcomponent), and *PLAT*. The allograft rejection pathway (adjusted p-value: 1,71E-21) involves genes *CD40*, *TNF*, and *IL12B*, indicating their role in immune signaling during graft rejection. Apoptosis (adjusted p-value: 6,53E-15) is associated with *FASLG* (Fas ligand), *TNF*, and *CASP1* (Caspase-1), suggesting activation in the presence of infection processes, leading to cellular senescence. The myogenesis pathway (adjusted p-value: 2,54E-14) involves *FGF2* and *AKT2*, indicating their role in regulating cellular functions and muscle regeneration.

4. DISCUSSION

4.1. Connectivity Nodes of the PPI Network of the PVH System

The identified connected components vary in diameter, suggesting the presence of independent subnetworks with different levels of complexity and extension [197]. Proteins within a PPI network often form complexes, and the concept of subnetworks aligns with the idea of protein complexes [198]. In the study, the analysis of connected components highlights the existence of subnetworks with distinct diameters and sizes. This modular organization reflects the potential presence of functionally related proteins within these subnetworks, resembling the idea of protein complexes [197].

The network's diameter, shortest paths, and clustering coefficients provide valuable insights into the interconnectedness of proteins. The study underscores the importance of understanding both highly connected proteins, such as ALB, TNF, IL6, GAPDG, and IFNG, and those with low connectivity, like ZBTB43, REG1A, LGALS7B, PDZD4, and KEL. This diversity emphasizes the network's complexity, with highly connected nodes playing central roles in the interaction network. In contrast, peripheral nodes contribute to specific functions during Chagas Disease infection and cellular regulation processes [158,159,197].

The concept of subnetworks aligns with the idea that proteins within complexes tend to interact with each other [198]. The study's findings resonate with previous research that utilizes network analysis to identify protein complexes based on topological features. However, challenges in predicting protein-protein interactions (PPI) directly from detected complexes are acknowledged, given potential issues with multiple subnetworks within complex and flexible interactions within independent subnetworks [164,198].

In studies such as [199], dynamic PPI networks were built, which used time series as the primary approach, which revealed the existence of subnetworks with different diameters. Therefore, these networks have temporal aspects, indicating a dynamic nature of the interactions between proteins at other times. These approaches are crucial in understanding the temporal dynamics of protein interactions during a disease.

The analysis of connected components and network properties presented in the Chagas Disease study provides valuable information for understanding the structure and

organization of the PPI network [113,166,197]. However, the discussion also highlights the complexities and challenges associated with predicting PPI solely based on detected protein complexes, emphasizing the need for a nuanced approach considering subnetworks and dynamic aspects in future studies [198].

4.2. Centrality Measures of the PPI Network of the PVH System

The analysis of centrality measures in the Protein-Protein Interaction (PPI) network of the PVH System sheds light on the crucial nodes and their roles in Chagas Disease. Degree Centrality identifies proteins with high connectivity, such as ALB, TNF, IL6, GAPDH, and IFNG, emphasizing their significance in the disease's pathogenesis. On the other hand, proteins like MAN1B1, SLIRP, RNF115, R4G5D9_RHOPR, and Q4DI29_TRYCC with low degree centrality values suggest a more specific or limited role in PPIs.

Closeness Centrality highlights the proximity of nodes within the network, with central nodes like ALB, TNF, IL6, GAPDH, TGFB1, and IFNG playing crucial roles in facilitating shorter paths for communication. The graphical representation in Figure 7 visualizes this centrality measure, emphasizing the central positioning of critical nodes.

The analysis of Betweenness Centrality identifies proteins like ALB, GAPDH, and TNF as pivotal in information propagation between nodes, potentially acting as links between different network modules. The discussion on the biological processes of specific proteins, such as Q4DNM5_TRYCC, Q4CL66_TRYCC, Q4E0X8_TRYCC, T1HFL6_RHOPR, R4G5E5_RHOPR, and T1HS28_RHOPR, provides insights into their roles in organisms.

The connectivity and protein interaction centrality values play a vital role in the flow of intraspecies interactions in the network, such as central proteins to many paths in the network (bottlenecks) or highly connected proteins (hubs). Likewise, central proteins have managed to attract the attention of researchers in recent years for their role in mediating cellular processes and their therapeutic potential [200].

Linking this information to the discussion on comorbid migraine and epilepsy, the centrality measures offer a perspective on the importance of specific proteins in complex diseases [197]. The central nodes identified, particularly ALB, play diverse roles, including potential connections to neurological processes [201].

ALB, a serum albumin, is highlighted for its essential role in maintaining colloid osmotic pressure and as a carrier for various compounds, including drugs. Its unique properties make it an attractive carrier for therapeutically active compounds. The discussion on the pharmacokinetics of serum albumins across species emphasizes the importance of understanding these differences for effective drug development [202].

Furthermore, the association between disease severity, albumin nadir, and rates of recurrent Clostridioides difficile infection (RCDC) is explored, suggesting that albumin nadir reflects disease severity. This information is valuable in understanding the clinical implications of albumin levels in disease progression [202,203].

The study on glicodendrimers and their therapeutic potential in chronic lymphocytic leukemia (CLL) illustrates the promising results of these compounds in inducing apoptosis in CLL cells without adverse effects on other blood components. This offers a glimpse into potential therapeutic strategies for CLL treatment [204].

4.3. Communities of the PPI Network of the PVH System

The CNM algorithm identified seven communities, revealing the complex subnetworks formed by various nodes within biological contexts. Community 1, characterized by nodes like IL6, TNF, and ALB, highlighted the significance of immune response and homeostasis in Chagas Disease. Communities 2 and 3 indicated subsets with greater specialization, while communities 4 and 5 focused on the parasite and vector, offering insights into their specific interactions. Community 7, with only two nodes (KEL and PDZF4), signified a more detailed and limited exchange within the PVH System.

Similarly, the Louvain algorithm revealed a more complex modular network structure with ten communities. Communities 2 and 8, along with community 7, showcased greater complexity, protein diversity, and metabolic signal exchange. Community 2, akin to CNM community 1, emphasized the relevance of IL6 and TNF in the immune response and inflammation regulation. Community 4, featuring ALB with high betweenness centrality, suggested its role as a mediator in fundamental metabolic interactions.

Linking this information to the discussion, the study presented a comprehensive exploration of community detection methods based on modularization, including the CNM algorithm [170]. The research addressed computational challenges, providing a high-performance toolkit for community detection in terms of runtime and memory usage.

Applying these methods to large-scale real networks demonstrated results consistent with the existing literature, emphasizing their appropriateness in contexts requiring immediate responses [197,205].

The study's limitations were acknowledged, pointing to potential future research directions [197]. The qualitative evaluation based on scholarly behaviors in the information retrieval (IR) field was noted, and the need for quantitative assessment across different domains and datasets was highlighted. The dynamic nature of communities and topics in social networks was discussed, emphasizing the importance of considering temporal aspects in community detection approaches [206].

Furthermore, the discussion compared the Louvain algorithm with the newly introduced Leiden algorithm. The Leiden algorithm addressed the shortcomings of the Louvain algorithm by providing guarantees of connected communities and local optimality in iterative applications. The experimental analysis convincingly demonstrated the superior scalability and accuracy of the Leiden algorithm over the Louvain algorithm [206–208].

4.4. Edges weight present in the Nodes of the PPI Network of the PVH System

The protein-protein interaction (PPI) network analysis results in the PVH System reveal a robust and complex network with 5 796 interactions, indicating high density and cohesion. The division into connected components highlights the centrality of component 1, with 5 405 interactions and 3 427,86 weight, emphasizing its global importance. Specific interactions, such as TNFRSF1A-TNF, exhibit weak stability, while HSP90AA1-CCL2 shows strong affinity. Figure 11 visualizes the edge weight in the subnetwork, indicating the proteins with the most impact during a perturbation in the system. In this case, the protein score range is from 0.1 to 0.999 (S9 Table) [209].

For example, some studies use gene expression data from the PPI network to assign a weight to each interaction in the network, which allows the construction of a weighted network, considering the importance of the centrality proteins and the weight of each node or edge [200,210]. Others focus on identifying protein complexes through methods such as edge weight and core binding structure (EWCA), demonstrating its effectiveness and accuracy compared to other complex identification methods in PPI networks [211]. On the other hand, some algorithms, such as Markov clustering, in conjunction with gene expression analysis, identify protein interaction complexes in networks and demonstrate improvements in the precision and F-measure of the model [212]. These studies

contribute to advancing the understanding of the identification of protein complexes using the edge weight of interactions.

4.5. Enrichment Analysis of the PPI Network of the PVH System

Enrichrpy and GSEApy analyses revealed significant pathways and biological terms linked to immune signaling, cytokine responses, and inflammation. The identified central genes such as *CSF3*, *IL1RN*, *CD40*, *TNF*, *PIK3CG*, *AKT2*, *FGF2*, *IL10*, and *TGFB1* play crucial roles in these processes. Likewise, the introduction of MTGO as a novel method for identifying functional modules in PPI networks provided a robust approach. The method's reliance on GO term attribution and topological measures ensured comprehensive coverage and overlap in the specified modules. MTGO's capability to detect significant GO terms and its superior performance in GO-related processes compared to other algorithms emphasized its effectiveness in accelerating PPI network analysis [112,213].

Likewise, protein-protein interaction studies indicate that biological functions with significant enrichments harbor critical processes for better cellular functioning, such as the cell surface receptor signaling pathway, stimulus-response, cell death, immune response defense, gene expression, the apoptotic process, T cell co-stimulation, cell-leukocyte cell adhesion, among others. Furthermore, the centrality-corrected minimal dominant set (CC-MDS) proteins with the highest significance around GO terms are aging genes, disease-associated genes, virus-targeting genes, transcription factors, and protein kinases [214].

On the other hand, studies identifying essential genes in neglected tropical diseases caused by *L. major*, such as that of [215] , analyzed the enrichment of the PPI network using the Database for Annotation, Visualization, and Integrated Discovery (DAVID) in set with GO annotations and KEGG paths and finding that the most significant GO terms were DNA replication (GO:0006260) and proton transport coupled to ATP hydrolysis (GO:0015991).

Likewise, other studies using StringDB functional analysis identified that the KEGG pathways most enriched in the case of interactions between *Plasmodium* and human proteins (malaria pathway) were the endocytosis pathway, ubiquitin-mediated proteolysis, the focal adhesion pathway, the regulation of the actin cytoskeleton pathway, and the bacterial invasion pathway of epithelial cells and the spliceosome. Furthermore,

they demonstrated that PFA0310c and FKBP35 are vital in linking proteins involved in malaria pathways [216]. Another study identified relevant patterns of interactions between dengue virus (DENV) and its two hosts, *H. sapiens* and the insect vector *Aedes aegypti*, finding that many proteins have functions related to GO terms, such as RNA processing, transcription, or regulation of the stress response, blood coagulation and hemostasis [217].

In the context of Chagas disease, exploring KIR genes, cytokines, and genetic factors has revealed their significant roles in infection susceptibility, disease progression, and the immune response. The study underscored the importance of genetic polymorphisms in cytokine genes in influencing disease outcomes. The comprehensive analysis of cytokines such as IL-10, IFN-γ, IL-6, IL-4, and TGF-β in the context of Chagas disease provided valuable insights into their roles in modulating immune responses and contributing to disease pathogenesis [218].

The analysis of genetic factors and cytokine interactions in Chagas disease demonstrated the complexity of the host-parasite interaction and its impact on disease outcomes. The potential of using PPI networks to explore the genetics of complex diseases, providing a framework for understanding the interplay between multiple genes and their contributions to disease complexity [178,218].

Studies such as that of [178] indicate the integration of T. cruzi minicircle DNA in the genome of patients with Chagas disease. This shows a unique perspective on its pathogenesis, relating it to genetic alterations with autoimmune reactions against cardiac tissues. For this reason, the construction of PPI networks and the analysis of topological centrality in different diseases caused by parasites highlighted the interactions between differentially expressed genes. [166].

5. CONCLUSIONS

In conclusion, exploring the Chagas Disease parasite-vector-human system through automated PPI analysis has yielded insightful perspectives. Our investigation focused on elucidating the connectivity, centrality measures, communities, weight considerations, and enrichment aspects within the intricate web of protein interactions.

Identifying pivotal genes, including *TNF*, *IL6*, and *GADPH*, underscores the significance of specific nodes in the network. With their pronounced roles and robust interactions, these genes stand out as potential targets for further research and therapeutic interventions. Their involvement suggests critical contributions to the molecular landscape of Chagas disease, offering valuable insights for understanding disease mechanisms.

Analyzing connectivity, centralities, and communities has revealed the complex interplay within the PPI. Shortest path analyses complemented other centrality measures, providing a holistic view of protein interactions in the context of Chagas disease. Furthermore, the variability in degree centralities offers detailed insights into the functional diversity of proteins, which is essential for understanding underlying biological complexity.

The combination of centrality measures, including degree, clustering coefficient, and closeness, facilitates a comprehensive understanding of the protein network's topology and dynamics associated with Chagas Disease. Our quantitative analysis of betweenness centrality delves into the network's topology, identifying key nodes and providing valuable information for understanding protein-protein interaction dynamics.

6. FUTURE PERSPECTIVE

The analysis of the PPI network of the PVH system in Chagas disease faces significant challenges around obtaining data on the interactome of *T. cruzi* and *R. prolixus*, which manifest during the disease. This and the need to update the organizations' databases above make generating a robust and solid network challenging. Despite this, more data is expected to be obtained to develop more effective and promising PPI networks.

For this, it is expected that advances in bioinformatics and molecular biology techniques will allow more protein interactions to be identified. In addition, they are developing methodologies with higher sensitivity and specific indices, where data from different databases are better related, and the generation of universal taxon and protein identifiers (PDB) would improve the quality and coverage of the interactome of these organisms.

In this sense, the analysis of PPI networks could provide indispensable information on the molecular mechanisms underlying the pathogenesis of Chagas disease. Likewise, by investigating the topology of the interaction networks, it would be possible to identify critical nodes (interest) and key routes that regulate different processes, such as signaling and response to cytokine stimulus, positive regulation of intracellular signal transduction, positive regulation of the multicellular organismal function, platelet degranulation, role in immune signaling during graft rejection, suggesting activation in the presence of infection processes, leading to cellular senescence or regulating cellular functions and muscle regeneration of the vector in humans.

For this reason, studying the structural properties and dynamics of the proteins with the most significant participation in the network could indicate potential molecular targets for treating Chagas disease. Due to those above, essential genes such as *IL6*, *TNF*, and *ALB* stand out in node connectivity analyses, centrality measures, community analysis, edge weight, or enrichment analysis. They may be targets when generating pharmacological strategies to combat *T. cruzi* infection and reduce its harmful effects in humans.

REFERENCES

1. Florin-Christensen M, Schnittger L. Parasitic protozoa of farm animals and pets. Parasitic Protozoa of Farm Animals and Pets. 2018. 1–438 p.
2. Omarova A, Tussupova K, Berndtsson R, Kalishev M, Sharapatova K. Protozoan parasites in drinking water: A system approach for improved water, sanitation and hygiene in developing countries. Int J Environ Res Public Health. 2018;15(3):1–18.
3. Adl SM, Simpson AGB, Farmer MA, Andersen RA, Anderson OR, Barta JR, et al. The new higher level classification of eukaryotes with emphasis on the taxonomy of protists. J Eukaryot Microbiol. 2005;52(5):399–451.
4. VICKERMAN K. THE EVOLUTIONARY EXPANSION OF THE TRYPANOSOMATID FLAGELLATES. Int J Parasitol. 1994;24(8):1179–201.
5. Yaeger RG. Medical Microbiology. 4th edition. Chapter 77 Protozoa: Structure, Classification, Growth, and Development General Concepts. 1996;1–4.
6. Tibayrenc M, Kjellberg F, Ayala FJ. A clonal theory of parasitic protozoa: The population structures of Entamoeba, Giardia, Leishmania, Naegleria, Plasmodium, Trichomonas, and Trypanosoma and their medical and taxonomical consequences. Proc Natl Acad Sci U S A. 1990;87(7):2414–8.
7. Stevens JR, Noyes HA, Schofield CJ, Gibson W. The molecular evolution of trypanosomatidae. Adv Parasitol. 2001;48:1–53.
8. Correa JP, Bacigalupo A, Yefi-Quinteros E, Rojo G, Solari A, Cattan PE, et al. Trypanosomatid infections among vertebrates of chile: A systematic review. Pathogens. 2020;9(8):1–21.
9. Ahmed A, Ijaz M, Ayyub RM, Ghaffar A, Ghauri HN, Aziz MU, et al. Balantidium coli in domestic animals: An emerging protozoan pathogen of zoonotic significance. Acta Trop [Internet]. 2020;203:105298. Available from: https://doi.org/10.1016/j.actatropica.2019.105298
10. Singh S. Isolation of free living amoebas as a highly sensitive index of water contamination. Water Int. 2000;25(3):403–9.
11. Shi Y, Queller DC, Tian Y, Zhang S, Yan Q, He Z, et al. The ecology and evolution of amoeba-bacterium interactions. Appl Environ Microbiol. 2021;87(2):1–16.
12. Guillén N. Pathogenicity and virulence of Entamoeba histolytica, the agent of amoebiasis. Virulence [Internet]. 2023;14(1):1–22. Available from: https://doi.org/10.1080/21505594.2022.2158656
13. Escalante AA, Ayala FJ. Evolutionary origin of Plasmodium and other Apicomplexa based on rRNA genes. Proc Natl Acad Sci U S A. 1995;92(13):5793–7.
14. Real E, Howick VM, Dahalan FA, Witmer K, Cudini J, Andradi-Brown C, et al. A single-cell atlas of Plasmodium falciparum transmission through the mosquito. Nat Commun [Internet]. 2021;12(1):1–13. Available from: http://dx.doi.org/10.1038/s41467-021-23434-z
15. Izidoro LFM, Ribeiro MC, Souza GRL, Sant'Ana CD, Hamaguchi A, Homsi-Brandeburgo MI, et al. Biochemical and functional characterization of an l-amino acid oxidase isolated from Bothrops pirajai snake venom. Bioorganic Med Chem. 2006;14(20):7034–43.
16. Stevens JR, Noyes HA, Dover GA, Gibson WC. The ancient and divergent origins of the human pathogenic trypanosomes, Trypanosoma brucei and T. cruzi. Parasitology. 1999;118(1):107–16.
17. Peck M. Molecular parasitology funding. Vol. 10, Scientist. 1996.
18. Barbosa RL, da Cunha JPC, Menezes AT, Melo R de FP, Elias MC, Silber AM, et al. Proteomic

analysis of Trypanosoma cruzi spliceosome complex. J Proteomics [Internet]. 2020;223(November 2019):103822. Available from: https://doi.org/10.1016/j.jprot.2020.103822

19. Kaufer A, Ellis J, Stark D, Barratt J. The evolution of trypanosomatid taxonomy. Parasites and Vectors. 2017;10(1):1–17.
20. Overath P, Haag J, Lischke A, O'hUigin C. The surface structure of trypanosomes in relation to their molecular phylogeny. Int J Parasitol. 2001;31(5–6):468–71.
21. Rassi A, Rassi A, Marcondes de Rezende J. American Trypanosomiasis (Chagas Disease). Infect Dis Clin North Am. 2012;26(2):275–91.
22. Fermino BR, Paiva F, Soares P, Tavares LER, Viola LB, Ferreira RC, et al. Field and experimental evidence of a new caiman trypanosome species closely phylogenetically related to fish trypanosomes and transmitted by leeches. Int J Parasitol Parasites Wildl. 2015;4(3):368–78.
23. World Health Organization. Research priorities for Chagas disease, human African trypanosomiasis and leishmaniasis. World Health Organ Tech Rep Ser [Internet]. 2012;(975):v–100. Available from: http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=med9&NEWS=N&AN=23484340
24. Richards CS, Williams KA, Villet MH. Predicting geographic distribution of seven forensically significant blowfly species (diptera: Calliphoridae) in South Africa. African Entomol. 2009;17(2):170–82.
25. Velásquez Y, Martínez-Sánchez AI, Thomas A, Rojo S. Checklist and distribution maps of the blow flies of Venezuela (Diptera, Calliphoridae, Mesembrinellidae). Zookeys. 2017;2017(645):103–32.
26. Badenhorst R, Villet MH. The uses of Chrysomya megacephala (Fabricius , 1794) (Diptera : Calliphoridae) in forensic entomology. 2018;1790. Available from: https://doi.org/10.1080/20961790.2018.1426136
27. Owings CG, Picard CJ. New distribution record for Lucilia cuprina (Diptera: Calliphoridae) in Indiana, United States. J Insect Sci. 2018;18(4):0–5.
28. Martins E, Neves JA, Moretti TC, Godoy WAC, Thyssen PJ. Breeding of ornidia obesa (Diptera: Syrphidae: Eristalinae) on pig carcasses in Brazil. J Med Entomol. 2010;47(4):690–4.
29. Zhang G, Hart ER, Weirauch C. A taxonomic monograph of the assassin bug genus Zelus Fabricius (Hemiptera: Reduviidae): 71 species based on 10,000 specimens. Biodivers Data J. 2016;4(1).
30. Olson DM, Prescott KK, Zeilinger AR, Hou S, Coffin AW, Smith CM, et al. Landscape Effects on Reproduction of Euschistus servus (Hemiptera: Pentatomidae), a Mobile, Polyphagous, Multivoltine Arthropod Herbivore. Environ Entomol. 2018;47(3):660–8.
31. Tripet F, Christe P, Møller AP. x. J Anim Ecol. 2002;71(5):735–48.
32. Medvedev S. Specific features of the distribution and host associations of fleas (Siphonaptera). 2002;(January 2002).
33. Keskin A. Description of a new flea species of the genus Ctenophthalmus Kolenati (Siphonaptera: Ctenophthalmidae) from Turkey. Parasitol Res. 2021;120(4):1203–7.
34. Clark NJ, Seddon JM, Šlapeta J, Wells K. Parasite spread at the domestic animal - wildlife interface: anthropogenic habitat use, phylogeny and body mass drive risk of cat and dog flea (Ctenocephalides spp.) infestation in wild mammals. Parasit Vectors. 2018;11(1):8.
35. Gálvez R, Musella V, Descalzo MA, Montoya A, Checa R, Marino V, et al. Modelling the current distribution and predicted spread of the flea species Ctenocephalides felis infesting outdoor dogs in Spain. Parasites and Vectors. 2017;10(1):4–11.
36. Crkvencic N, Šlapeta J. Climate change models predict southerly shift of the cat flea

(Ctenocephalides felis) distribution in Australia. Parasites and Vectors [Internet]. 2019;12(1):1–13. Available from: https://doi.org/10.1186/s13071-019-3399-6

37. Kiefer D, Stubbe M, Stubbe A, Gardner SL, Tserenorov D, Samyia R, et al. Siphonaptera of Mongolia and Tuva: results of the Mongolian- German Biological Expeditions since 1962 – years 1999-2003. Erforsch Biol Ressourcen der Mongolischen Volksrepublik. 2012;12:153–167.
38. Krasnov BR, Poulin R, Shenbrot GI, Mouillot D, Khokhlova IS. Host specificity and geographic range in haematophagous ectoparasites. Oikos. 2005;108(3):449–56.
39. Guo L, Wang S. A new species and a new record of Anthracoidea (Ustilaginales) from China. Mycotaxon. 2005;93:159–62.
40. Abramov AV, Medvedev SG. Notes on zoogeography and taxonomy of the badgers (Carnivora: Mustelidae: Meles) and some of their fleas (Siphonaptera: Ceratophyllidae: Paraceras). Zoosystematica Ross. 2003;11(2):397–402.
41. Cross CL, Martin E, Brendel L. Human Flea Pulex irritans Linnaeus, 1758 (Insecta: Siphonaptera: Pulicidae). Edis. 2022;2022(6):1–6.
42. Dellicour S, Michez D, Mardulyn P. Comparative phylogeography of five bumblebees: Impact of range fragmentation, range size and diet specialization. Biol J Linn Soc. 2015;116(4):926–39.
43. Gorris ME, Bartlow AW, Temple SD, Romero-Alvarez D, Shutt DP, Fair JM, et al. Updated distribution maps of predominant Culex mosquitoes across the Americas. Parasites and Vectors [Internet]. 2021;14(1):1–13. Available from: https://doi.org/10.1186/s13071-021-05051-3
44. Hardie RC, Franceschini N, Ribi W, Kirschfeld K. Distribution and properties of sex-specific photoreceptors in the fly Musca domestica. J Comp Physiol □ A. 1981;145(2):139–52.
45. Al-Khalaf AA, Nasser MG, Hosni EM. Global Potential Distribution of Sarcophaga dux and Sarcophaga haemorrhoidalis under Climate Change. Diversity. 2023;15(8).
46. Kozminsky E, Kraeva N, Ishemgulova A, Dobáková E, Lukeš J, Kment P, et al. Host-specificity of Monoxenous Trypanosomatids: Statistical Analysis of the Distribution and Transmission Patterns of the Parasites from Neotropical Heteroptera. Protist. 2015;166(5):551–68.
47. Leavengood JM. Forty New Country Records for Twenty-Two New World Coreidae and Two Alydidae (Hemiptera: Heteroptera). https://doi.org/104289/0013-87971222326 [Internet]. 2020 May 11 [cited 2024 Feb 7];122(2):326–34. Available from: https://bioone.org/journals/proceedings-of-the-entomological-society-of-washington/volume-122/issue-2/0013-8797.122.2.326/Forty-New-Country-Records-for-Twenty-Two-New-World-Coreidae/10.4289/0013-8797.122.2.326.full
48. AUAD AM, PIMENTA DS, SILVA DM, MONTEIRO PH, RESENDE TT. x. Rev Colomb Entomol. 2011;37(2):244–8.
49. Sahayaraj K, Francis Borgio J. Distribution of Metarhizium anisopliae (Metsch.) Sorokin (Deuteromycotina: Hyphomycetes) in Tamil Nadu, India, its biocontrol potential on Dysdercus cingulatus (Fab.) (Hemiptera: Pyrrhocoridae). Arch Phytopathol Plant Prot. 2009;42(5):424–35.
50. OLIVEIRA PS. On the mimetic association between nymphs of Hyalymenus spp. (Hemiptera: Alydidae) and ants. Zool J Linn Soc. 1985;83(4):371–84.
51. Carroll SP, Loye JE. Specialization of Jadera Species (Hemiptera: Rhopalidae) on the Seeds of Sapindaceae (Sapindales), and (Revolutionary Responses of Defense and Attack. Ann Entomol Soc Am. 1987;80(3):373–8.
52. Coelho A, Santos D, Andrés D, Ferreira F, Tropical RDB, Rica UDC, et al. Biogeography and areas of endemism of Prepops Reuter (Heteroptera : Miridae) The Miridae (Hemiptera : Heteroptera)

are. 2016;
53. Al-ghamdi AA, Nuru A, Khanbash MS, Smith DR. Apis mellifera jemenitica Ruttner Distribuci6n geografica y la variaci6n poblacional de Apis mellifera jemenitica Ruttner. 2013;52(3):124–33.
54. Melo MC, Montemayor SI. Biodiversity of the scentless plant bugs (Hemiptera: Rhopalidae) in southern South America. J Nat Hist. 2016;50(3–4):163–200.
55. Belkharchouche M, Berchi S, Mathieu B, Rakotoarivony I, Duhayon M, Baldet T, et al. Update of the Culicoides (Diptera: Ceratopogonidae) species checklist from Algeria with 10 new records. Parasites and Vectors [Internet]. 2020;13(1):1–15. Available from: https://doi.org/10.1186/s13071-020-04335-4
56. Hadj-Henni L, Djerada Z, Millot C, Augot D. Comprehensive characterisation of Culicoides clastrieri and C. festivipennis (Diptera: Ceratopogonidae) according to morphological and morphometric characters using a multivariate approach and DNA barcode. Sci Rep [Internet]. 2021;11(1):1–12. Available from: https://doi.org/10.1038/s41598-020-78053-3
57. Outammassine A, Zouhair S, Loqman S. Global potential distribution of three underappreciated arboviruses vectors (Aedes japonicus, Aedes vexans and Aedes vittatus) under current and future climate conditions. Transbound Emerg Dis. 2022;69(4):e1160–71.
58. Faúndez EI, Rocca JR. A new species of Oncopeltus Stål, 1868 (Heteroptera: Lygaeidae) in the nominate subgenus from Ecuador. Zootaxa. 2017;4238(2):249–52.
59. Podlipaev SA, Frolov AO, Kolesnikov AA. Proteomonas inconstans gen. nov., sp. nov. (Kinetoplastida: Trypanosomatidae), a parasite of Calocoris sexguttatus (Hemiptera: Miridae). Parazitologiia. 1990;24(4):339–345.
60. Taszakowski A, Nowińska A, Brożek J. Morphological study of the labial sensilla in Nabidae (Hemiptera: Heteroptera: Cimicomorpha). Zoomorphology [Internet]. 2019;138(4):483–92. Available from: https://doi.org/10.1007/s00435-019-00455-3
61. Kostygov AY, Malysheva MN, Ganyukova AI, Razygraev A V., Drachko DO, Yurchenko V, et al. The Roles of Mosquitoes in the Circulation of Monoxenous Trypanosomatids in Temperate Climates. Pathogens. 2022;11(11):1–16.
62. Takaoka Hi. Morphotaxonomic revision of Simulium (Gomphostilbia) (Diptera: Simuliidae) in the Oriental Region. Zootaxa [Internet]. 2012 Dec 10;3577(1):1. Available from: https://biotaxa.org/Zootaxa/article/view/zootaxa.3577.1.1
63. Barratt J, Kaufer A, Peters B, Craig D, Lawrence A, Roberts T, et al. Isolation of Novel Trypanosomatid, Zelonia australiensis sp. nov. (Kinetoplastida: Trypanosomatidae) Provides Support for a Gondwanan Origin of Dixenous Parasitism in the Leishmaniinae. PLoS Negl Trop Dis. 2017;11(1):1–26.
64. Orshan L, Elbaz S, Ben-Ari Y, Akad F, Afik O, Ben-Avi I, et al. Distribution and Dispersal of Phlebotomus papatasi (Diptera: Psychodidae) in a Zoonotic Cutaneous Leishmaniasis Focus, the Northern Negev, Israel. PLoS Negl Trop Dis. 2016;10(7):1–21.
65. Ronderos MM, Spinelli GR. On the subgenus Forcipomyia (Lasiohelea) in the neotropical region (Diptera: Ceratopogonidae). Trans Am Entomol Soc. 1999;125(1–2):151–61.
66. Heppner JB, Richman DB, Naranjo SE, Habeck D, Asaro C, Boevé JL, et al. Southern Green Stink Bug, Nezara viridula (L.) (Hemiptera: Heteroptera: Pentatomidae). In: Encyclopedia of Entomology [Internet]. Dordrecht: Springer Netherlands; 2008. p. 3471–2. Available from: http://link.springer.com/10.1007/978-1-4020-6359-6_4283

67. David Baldwin J, Dingle H. Geographic variation in the effects of temperature on life-history traits in the large milkweed bug Oncopeltus fasciatus. Oecologia. 1986;69(1):64–71.
68. Roy AS. Phthia picta (Hemiptera : Coreidae) [Internet]. 2016. Available from: https://gd.eppo.int/download/doc/453_minids_PHTIPI.pdf
69. Moloo SK. The distribution of Glossina species in Africa and their natural hosts. Int J Trop Insect Sci. 1993;14(04):511–27.
70. Muyobela J, Pirk CWW, Yusuf AA, Sole CL. Spatial distribution of Glossina morsitans (Diptera: Glossinidae) in Zambia: A vehiclemounted sticky trap survey and Maxent species distribution model. PLoS Negl Trop Dis [Internet]. 2023;17(7):1–20. Available from: http://dx.doi.org/10.1371/journal.pntd.0011512
71. Abad-Franch F, Monteiro FA, Pavan MG, Patterson JS, Bargues MD, Zuriaga MÁ, et al. Under pressure: phenotypic divergence and convergence associated with microhabitat adaptations in Triatominae. Parasites and Vectors [Internet]. 2021;14(1):1–21. Available from: https://doi.org/10.1186/s13071-021-04647-z
72. Bustamante Gomez M, Gonçalves Diotaiuti L, Gorla DE. Distribution of Pyrethroid Resistant Populations of Triatoma infestans in the Southern Cone of South America. PLoS Negl Trop Dis. 2016;10(3):1–15.
73. Hamilton PB, Stevens JR. Classification and Phylogeny of Trypanosoma cruzi. In: American Trypanosomiasis. First Edit. Elsevier; 2010. p. 321–38. Available from: http://dx.doi.org/10.1016/B978-0-12-384876-5.00013-7
74. Simpson AGB, Stevens JR, Lukeš J. The evolution and diversity of kinetoplastid flagellates. Trends Parasitol. 2006;22(4):168–74.
75. Hemmige V, Tanowitz H, Sethi A. Trypanosoma cruzi infection: A review with emphasis on cutaneous manifestations. Int J Dermatol. 2012;51(5):501–8.
76. Rassi A, Rassi A, Marin-Neto JA. Chagas disease. Lancet [Internet]. 2010;375(9723):1388–402. Available from: http://dx.doi.org/10.1016/S0140-6736(10)60061-X
77. Hotez PJ, Molyneux DH, Fenwick A, Kumaresan J, Sachs SE, Sachs JD, et al. Control of Neglected Tropical Diseases. N Engl J Med. 2007;357(10):1018–27.
78. Bern C, Kjos S, Yabsley MJ, Montgomery SP. Trypanosoma cruzi and chagas' disease in the united states. Clin Microbiol Rev. 2011;24(4):655–81.
79. Mathers CD, Ezzati M, Lopez AD. Measuring the burden of neglected tropical diseases: The global burden of disease framework. PLoS Negl Trop Dis. 2007;1(2).
80. Moncayo, Silveira AC. Current epidemiological trends of Chagas disease in Latin America and future challenges: Epidemiology, surveillance, and health policies. Am Trypanos Chagas Dis One Hundred Years Res Second Ed. 2017;104(May):59–88.
81. World Health Organization. Chagas disease : control and elimination. SIxty-Third World Healhy Assembly. 2010.
82. Viotti R, Vigliano C, Lococo B, Bertocchi G. Annals of Internal Medicine Article Long-Term Cardiac Outcomes of Treating Chronic Chagas Disease with Benznidazole versus No Treatment. Ann Intern Med. 2006;
83. Fabbro DL, Streiger ML, Arias ED, Bizai ML, Del Barco M, Amicone NA. Trypanocide treatment among adults with chronic Chagas disease living in Santa Fe City (Argentina), over a mean follow-up of 21 years: Parasitological, serological and clinical evolution. Rev Soc Bras Med Trop.

2007;40(1):1–10.

84. Viteri M, Omar D, Vector N. Chagas disease in Ecuador : a systematic review of epidemiological and entomological aspects. Rev Ecuatoriana Ciencia, Tecnol e Innovación en Salud Pública. 2021;(1):1–11.
85. Mita-Mendoza NK, McMahon E, Kenneson A, Barbachano-Guerrero A, Beltran-Ayala E, Cueva C, et al. Chagas Disease in Southern Coastal Ecuador: Coinfections with Arboviruses and a Comparison of Serological Assays for Chagas Disease Diagnosis. Am J Trop Med Hyg [Internet]. 2018 [cited 2023 Mar 31];99(6):1530. Available from: /pmc/articles/PMC6283517/
86. Aguilar HM, Abad-Franch F, Dias JCP, Junqueira ACV, Coura JR. Chagas disease in the Amazon Region. Mem Inst Oswaldo Cruz [Internet]. 2007 Aug 31;102(suppl 1):47–56. Available from: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0074-02762007000900010&lng=en&tlng=en
87. Organización Panamericana de la Salud. Enfermedad de Chagas [Internet]. [cited 2023 Apr 6]. Available from: https://www.paho.org/es/temas/enfermedad-chagas
88. Ministerio de Salud Pública. SUBSISTEMA DE VIGILANCIA SIVE-ALERTA ENFERMEDADES TRANSMITIDASx||| POR VECTORES ECUADOR, SE 14-2021 [Internet]. 2021 [cited 2023 Apr 6]. Available from: https://www.who.int/es/news-room/fact-sheets/detail/dengue-and-severe-dengue
89. Ministerio de SP. SUBSECRETARIA DE VIGILANCIA, PREVENCIÓN Y CONTROL DE LA SALUD DIRECCIÓN NACIONAL DE VIGILANCIA EPIDEMIOLÓGICA ENFERMEDADES TRANSMITIDAS POR VECTORES Ecuador 2023 SE 50 [Internet]. Ministerio De Salud Pública. 2024. Available from: https://www.salud.gob.ec/vacuna-de-astrazeneca-contra-la-covid-19-se-aplicara-en-ecuador/#
90. Diaz JH. Recognizing and reducing the risks of Chagas disease (American Trypanosomiasis) in travelers. J Travel Med. 2008;15(3):184–95.
91. Tyler KM, Engman DM. The life cycle of Trypanosoma cruzi revisited. Int J Parasitol. 2001;31(5–6):472–81.
92. Tyler KM, Olson CL, Engman DM. The Life Cycle Of Trypanosoma Cruzi. 2003;1–11.
93. Cohen JE. Infectious Diseases of Humans: Dynamics and Control. JAMA J Am Med Assoc [Internet]. 1992 Dec 16;268(23):3381. Available from: http://jama.jamanetwork.com/article.aspx?doi=10.1001/jama.1992.03490230111047
94. Moncayo A, Ortiz Yanine MI. An update on Chagas disease (human American trypanosomiasis). Ann Trop Med Parasitol. 2006;100(8):663–77.
95. Tanowitz H, Machado F, Jelicks L, Shirani J, Campos de Carvalho A, Spray D, et al. Perspectives on Trypanosoma cruzi-induced heart disease (Chagas disease). NIH Public Access. 2009;51(6):524–39.
96. Bern C, Montgomery SP, Katz L, Caglioti S, Stramer SL. Chagas disease and the US blood supply. Curr Opin Infect Dis. 2008;21(5):476–82.
97. Prata A. Clinical and epidemiological aspects of Chagas disease. Clin Epidemiol Asp Chagas Dis. 2001;1(2):92–100.
98. Flores-Chávez M, Faez Y, Olalla JM, Cruz I, Gárate T, Rodríguez M, et al. Fatal congenital Chagas' disease in a non-endemic area: a case report. Cases J. 2008;1(1):1–5.
99. Freilij H, Altcheh J. Congenital chagas' disease: Diagnostic and clinical aspects. Clin Infect Dis. 1995;21(3):551–5.
100. Hofflin JM, Sadler RH, Araujo FG, Page WE, Remington JS. Laboratory-acquired Chagas disease.

Trans R Soc Trop Med Hyg. 1987;81(3):437–40.

101. Centers for Disease Control and Prevention (CDC). Chagas Disease After Organ Transplantation — United States , 2001. Morb Mortal Wkly Rep. 2002;287(14):1795–8.
102. De Noya BA, Díaz-Bello Z, Colmenares C, Ruiz-Guevara R, Mauriello L, Zavala-Jaspe R, et al. Large urban outbreak of orally acquired acute chagas disease at a school in Caracas, Venezuela. J Infect Dis. 2010;201(9):1308–15.
103. Rodriguez-Morales AJ. Chagas disease: an emerging food-borne entity? J Infect Dev Ctries. 2008;2(2):149–50.
104. DIAS E, LARANJA FS, MIRANDA A, NOBREGA G. Chagas' disease; a clinical, epidemiologic, and pathologic study. Circulation. 1956;14(6):1035–60.
105. Dias JP, Bastos C, Araújo E, Mascarenhas AV, Netto EM, Grassi F, et al. Acute Chagas disease outbreak associated with oral transmission. Rev Soc Bras Med Trop. 2008;41(3):296–300.
106. Angheben A, Boix L, Buonfrate D, Gobbi F, Bisoffi Z, Pupella S, et al. Chagas disease and transfusion medicine: A perspective from non-endemic countries. Blood Transfus. 2015;13(4):540–50.
107. Das Neves Pinto AY, Ferreira AG, Da Costa Valente V, Harada GS, Da Silva Valente SA. Urban outbreak of acute chagas disease in amazon region of Brazil: Four-year follow-up after treatment with benznidazole. Rev Panam Salud Publica/Pan Am J Public Heal. 2009;25(1):77–83.
108. Peña-Callejas G, González J, Jiménez-Cortés JG, De Fuentes-Vicente JA, Salazar-Schettino PM, Bucio-Torres MI, et al. Enfermedad de Chagas: biología y transmisión de Trypanosoma cruzi. TIP Rev Espec en Ciencias Químico-Biológicas. 2022;25:1–19.
109. Duque TLA, Siqueira MS, Travassos LH, Moreira OC, Bozza PT, Melo RCN, et al. The induction of host cell autophagy triggers defense mechanisms against Trypanosoma cruzi infection in vitro. Eur J Cell Biol [Internet]. 2020 Jan;99(1):151060. Available from: https://linkinghub.elsevier.com/retrieve/pii/S017193351930144X
110. Mwangi VI, Martinez EG, Leda RL, Catunda MESLA, Dias A de S, Padron Antonio Y, et al. Resisting an invasion: A review of the triatomine vector (Kissing bug) defense strategies against a Trypanosoma sp infection. Acta Trop [Internet]. 2023;238(November 2022):106745. Available from: https://doi.org/10.1016/j.actatropica.2022.106745
111. Morrot A, Villar SR, González FB, Pérez AR. Evasion and immuno-endocrine regulation in parasite infection: Two sides of the same coin in chagas disease? Front Microbiol. 2016;7(MAY):1–10.
112. Vella D, Marini S, Vitali F, Di Silvestre D, Mauri G, Bellazzi R. MTGO: PPI Network Analysis Via Topological and Functional Module Identification. Sci Rep [Internet]. 2018;8(1):1–13. Available from: http://dx.doi.org/10.1038/s41598-018-23672-0
113. Rodriguez-Soca Y, Munteanu CR, Dorado J, Pazos A, Prado-Prado FJ, González-Díaz H. Trypano-PPI: A web server for prediction of unique targets in trypanosome proteome by using electrostatic parameters of protein-protein interactions. J Proteome Res. 2010;9(2):1182–90.
114. Snider J, Kotlyar M, Saraon P, Yao Z, Jurisica I, Stagljar I. Fundamentals of protein interaction network mapping. Mol Syst Biol. 2015;11(12):1–20.
115. Hao T, Peng W, Wang Q, Wang B, Sun J. Reconstruction and application of protein–protein interaction network. Int J Mol Sci. 2016;17(6).
116. Kitano H. Biological robustness. Nat Rev Genet 2004 511 [Internet]. 2004 Nov [cited 2023 Dec 31];5(11):826–37. Available from: https://www.nature.com/articles/nrg1471

117. Chatr-Aryamontri A, Breitkreutz BJ, Oughtred R, Boucher L, Heinicke S, Chen D, et al. The BioGRID interaction database: 2015 update. Nucleic Acids Res. 2015;43(D1):D470–8.
118. Xenarios I, Fernandez E, Salwinski L, Duan XJ, Thompson MJ, Marcotte EM, et al. DIP: The Databases of Interacting Protein: 2001 update. Nucleic Acids Res. 2001;29(1):239–41.
119. Montojo J, Zuberi K, Rodriguez H, Bader GD, Morris Q. GeneMANIA: Fast gene network construction and function prediction for Cytoscape. F1000Research. 2014;3:1–7.
120. Ma W, Zhang L, Zeng P, Huang C, Li J, Geng B, et al. An analysis of human microbe-disease associations. Brief Bioinform. 2017;18(1):85–97.
121. Ammari MG, Gresham CR, McCarthy FM, Nanduri B. HPIDB 2.0: a curated database for host-pathogen interactions. Database (Oxford). 2016;2016:1–9.
122. Cariello NF. human hprt gene. 1994;22(17):3547–8.
123. Cariello NF, Craft TR, Vrieling H, Van Zeeland AA, Adams T, Skopek TR. Human HPRT mutant database: Software for data entry and retrieval. Environ Mol Mutagen. 1992;20(2):81–3.
124. Orchard S, Kerrien S, Abbani S, Aranda B, Bhate J, Bridge A, et al. Exchange Consortium (IMEx). Nat Methods. 2012;9(4):345–50.
125. 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(D1):1228–33.
126. Orchard S, Ammari M, Aranda B, Breuza L, Briganti L, Broackes-Carter F, et al. The MIntAct project - IntAct as a common curation platform for 11 molecular interaction databases. Nucleic Acids Res. 2014;42(D1):358–63.
127. Launay G, Salza R, Multedo D, Thierry-Mieg N, Ricard-Blum S. MatrixDB, the extracellular matrix interaction database: Updated content, a new navigator and expanded functionalities. Nucleic Acids Res. 2015;43(D1):D321–7.
128. Chatr-aryamontri A, Ceol A, Palazzi LM, Nardelli G, Schneider MV, Castagnoli L, et al. MINT: The Molecular INTeraction database. Nucleic Acids Res. 2007;35(SUPPL. 1):2006–8.
129. Goll J, Rajagopala S V., Shiau SC, Wu H, Lamb BT, Uetz P. MPIDB: The microbial protein interaction database. Bioinformatics. 2008;24(15):1743–4.
130. Cuesta-Astroz Y, Santos A, Oliveira G, Jensen LJ. An integrative method to unravel the host-parasite interactome: An orthology-based approach. bioRxiv. 2017;
131. Durmuş Tekir S, Çakir T, Ardiç E, Sayilirbaş AS, Konuk G, Konuk M, et al. PHISTO: Pathogen-host interaction search tool. Bioinformatics. 2013;29(10):1357–8.
132. Urban M, Pant R, Raghunath A, Irvine AG, Pedro H, Hammond-Kosack KE. The Pathogen-Host Interactions database (PHI-base): Additions and future developments. Nucleic Acids Res. 2015;43(D1):D645–55.
133. Cowley MJ, Pinese M, Kassahn KS, Waddell N, Pearson J V., Grimmond SM, et al. PINA v2.0: Mining interactome modules. Nucleic Acids Res. 2012;40(D1):862–5.
134. 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. 2011;39(SUPPL. 1):691–7.
135. Szklarczyk D, Kirsch R, Koutrouli M, Nastou K, Mehryary F, Hachilif R, et al. The STRING database in 2023: protein-protein association networks and functional enrichment analyses for any sequenced genome of interest. Nucleic Acids Res. 2023;51(1 D):D638–46.
136. Miryala SK, Anbarasu A, Ramaiah S. Discerning molecular interactions: A comprehensive review

on biomolecular interaction databases and network analysis tools. Gene [Internet]. 2018;642:84–94. Available from: http://dx.doi.org/10.1016/j.gene.2017.11.028

137. Singh N, Rai S, Bhatnagar R, Bhatnagar S. Network analysis of host-pathogen protein interactions in microbe induced cardiovascular diseases. In Silico Biol. 2020;14(3–4):115–33.
138. Watanabe Costa R, da Silveira JF, Bahia D. Interactions between Trypanosoma cruzi secreted proteins and host cell signaling pathways. Front Microbiol. 2016;7(MAR):1–9.
139. Pech-Canul ÁDLC, Monteón V, Solís-Oviedo RL. A Brief View of the Surface Membrane Proteins from Trypanosoma cruzi. J Parasitol Res. 2017;2017.
140. Paba J, Ricart CAO, Fontes W, Santana JM, Teixeira ARL, Marchese J, et al. Proteomic analysis of Trypanosoma cruzi developmental stages using isotope-coded affinity tag reagents. J Proteome Res. 2004;3(3):517–24.
141. Liu Y zhi, Wang H, Zhu J hao, Han D min, Kudinha T, Kong F rong, et al. Isobaric tag for relative and absolute quantitation-based comparative proteomic analysis of human pathogenic Prototheca zopfii genotype 2 and environmental genotype 1 strains. J Microbiol Immunol Infect [Internet]. 2018;51(3):302–11. Available from: https://doi.org/10.1016/j.jmii.2016.07.003
142. Goh K Il, Cusick ME, Valle D, Childs B, Vidal M, Barabási AL. The human disease network. Proc Natl Acad Sci U S A. 2007;104(21):8685–90.
143. Thorne T, Stumpf MPH. Generating confidence intervals on biological networks. BMC Bioinformatics. 2007;8:1–10.
144. Keilhauer EC, Hein MY, Mann M. Accurate protein complex retrieval by affinity enrichment mass spectrometry (AE-MS) rather than affinity purification mass spectrometry (AP-MS). Mol Cell Proteomics. 2015;14(1):120–35.
145. Shanmugasundram A, Starns D, Böhme U, Amos B, Wilkinson PA, Harb OS, et al. TriTrypDB: An integrated functional genomics resource for kinetoplastida. PLoS Negl Trop Dis. 2023;17(1):1–20.
146. Amos B, Aurrecoechea C, Barba M, Barreto A, Basenko EY, Bażant W, et al. VEuPathDB: The eukaryotic pathogen, vector and host bioinformatics resource center. Nucleic Acids Res. 2022;50(D1):D898–911.
147. Bateman A, Martin MJ, Orchard S, Magrane M, Ahmad S, Alpi E, et al. UniProt: the Universal Protein Knowledgebase in 2023. Nucleic Acids Res. 2023;51(D1):D523–31.
148. Chapman C, Stolee KT. Exploring regular expression usage and context in python. In: ISSTA 2016 - Proceedings of the 25th International Symposium on Software Testing and Analysis [Internet]. Association for Computing Machinery, Inc; 2016 [cited 2023 Dec 31]. p. 282–93. Available from: https://dl.acm.org/doi/10.1145/2931037.2931073
149. Salunke S. Selenium Webdriver in Python: Learn with Examples. 1st ed. 2016 [cited 2023 Dec 31];282–293. Available from: https://doi.org/10.1145/2931037.2931073
150. Gilpin W. PyPDB: A Python API for the Protein Data Bank. Bioinformatics. 2016;32(1):159–60.
151. Pandey B. Analysis of Protein-Protein Interaction Networks Using High Performance Scalable Tools. 2018;113. Available from: https://scholarworks.uno.edu/honors_theses/113
152. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, et al. The Protein Data Bank. Nucleic Acids Res [Internet]. 2000 Jan 1 [cited 2023 Dec 31];28(1):235–42. Available from: https://dx.doi.org/10.1093/nar/28.1.235
153. Hunt J. Python DB-API. 2023 [cited 2024 Jan 1];369–75. Available from: https://link.springer.com/chapter/10.1007/978-3-031-40336-1_33

154. Xu Y, Luo XC. PyPathway: Python Package for Biological Network Analysis and Visualization. J Comput Biol. 2018;25(5):499–504.
155. Moreews F, Simon H, Siegel A, Gondret F, Becker E. PAX2GRAPHML: A python library for large-scale regulation network analysis using BioPAX. Bioinformatics. 2021;37(24):4889–91.
156. Hasan M, Kumar N, Majeed A, Ahmad A, Mukhtar S. Protein–Protein Interaction Network Analysis Using NetworkX. In: Methods in Molecular Biology [Internet]. Humana Press Inc.; 2023 [cited 2023 Dec 31]. p. 457–67. Available from: https://link.springer.com/protocol/10.1007/978-1-0716-3327-4_35
157. Li H, Zhou Y, Zhang Z. Network analysis reveals a common host–pathogen interaction pattern in Arabidopsis immune responses. Front Plant Sci. 2017;8(May):1–13.
158. Sohaee N, Forst C V. Identification of functional modules in a PPI network by bounded diameter clustering. J Bioinform Comput Biol. 2010;8(6):929–43.
159. Keretsu S, Sarmah R. Weighted edge based clustering to identify protein complexes in protein–protein interaction networks incorporating gene expression profile. Comput Biol Chem. 2016 Dec 1;65:69–79.
160. Kitsak M, Ganin A, Elmokashfi A, Cui H, Eisenberg DA, Alderson DL, et al. Finding shortest and nearly shortest path nodes in large substantially incomplete networks by hyperbolic mapping. ArXiv. 2022;1–26.
161. Watts D, Strogatz S. Collectivedynamics of 'small-world' networks. Nature. 1998;393(4):6684.
162. Freeman LC. Centrality in social networks. Soc Networks. 1979;1(3):215–39.
163. Mistry D, Wise RP, Dickerson JA. DiffSLC : A graph centrality method to detect essential proteins of a protein-protein interaction network. 2017;1–25.
164. Kumar N, Mukhtar MS. Ranking Plant Network Nodes Based on Their Centrality Measures. Entropy. 2023;25(4).
165. Singh A, Singh RR, Iyengar SRS. Node-weighted centrality: a new way of centrality hybridization. Comput Soc Networks [Internet]. 2020;7(1). Available from: https://doi.org/10.1186/s40649-020-00081-w
166. Pfeifer B, Saranti A, Holzinger A. GNN-SubNet: Disease subnetwork detection with explainable graph neural networks. Bioinformatics. 2022;38:II120–6.
167. Girvan M, Newman MEJ. Community structure in social and biological networks. Proc Natl Acad Sci U S A. 2002;99(12):7821–6.
168. Newman MEJ, Girvan M. Finding and evaluating community structure in networks. Phys Rev E - Stat Nonlinear, Soft Matter Phys. 2004;69(2 2):1–15.
169. Blondel VD, Guillaume JL, Lambiotte R, Lefebvre E. Fast unfolding of communities in large networks. J Stat Mech Theory Exp. 2008;2008(10).
170. Clauset A, Newman MEJ, Moore C. Finding community structure in very large networks. Phys Rev E - Stat Physics, Plasmas, Fluids, Relat Interdiscip Top. 2004;70(6):6.
171. Yadav SS, Chouhan U. Analysis of protein-protein interaction network of laminopathy based on topological properties. Biomed Pharmacol J. 2018;11(2):1091–103.
172. Hagberg A, Swart P, Schult D. Exploring network structure, dynamics, and function using NetworkX. 2008;836. Available from: https://www.osti.gov/servlets/purl/960616
173. Newman MEJ. Networks of networks - An introduction. 1 st. Oxford University Press; 2010. 772 p.
174. Newman M. Networks [Internet]. 2nd ed. Vol. 1. Oxford University Press; 2018. 793 p. Available

from: https://academic.oup.com/book/27884
175. Barabasi AL. NETWORK SCIENCE PERSONAL INTRODUCTION. In: Network Science. 2nd ed. 2016. p. 912.
176. Cormen TH, Leiserson CE, Rivest RL, Stein C. Introduction to algorithms. 4th ed. The MIT Press Cambridge, Massachusetts London, England. 2022. 1291 p.
177. Rahiminejad S, Maurya MR, Subramaniam S. Topological and functional comparison of community detection algorithms in biological networks. BMC Bioinformatics. 2019;20(1):1–25.
178. Safari-Alighiarloo N, Taghizadeh M, Rezaei-Tavirani M, Goliaei B, Peyvandi AA. Protein-protein interaction networks (PPI) and complex diseases. Gastroenterol Hepatol from Bed to Bench. 2014;7(1):17–31.
179. Asur S, Ucar D, Parthasarathy S. An ensemble framework for clustering protein-protein interaction networks. Bioinformatics. 2007;23(13):29–40.
180. Mishra B, Kumar N, Mukhtar MS. Systems biology and machine learning in plant–pathogen interactions. Mol Plant-Microbe Interact. 2019;32(1):45–55.
181. Deshmukh HV. Community Detection in Cyber Networks [Internet]. Purdue University; 2018. Available from: https://docs.lib.purdue.edu/open_access_theses
182. Reichardt J, Bornholdt S. Statistical mechanics of community detection. Phys Rev E - Stat Nonlinear, Soft Matter Phys. 2006;74(1):1–14.
183. Chen EY, Tan CM, Kou Y, Duan Q, Wang Z, Meirelles GV, et al. Enrichr: Interactive and collaborative HTML5 gene list enrichment analysis tool. BMC Bioinformatics. 2013;14.
184. Kuleshov M V, Jones MR, Rouillard AD, Fernandez NF, Duan Q, Wang Z, et al. Enrichr : a comprehensive gene set enrichment analysis web server 2016 update. 2016;44(May):90–7.
185. Evangelista JE, Xie Z, Marino GB, Nguyen N, Clarke DJB, Ma A. Enrichr-KG : bridging enrichment analysis across multiple libraries. 2023;51(May):168–79.
186. Xie Z, Bailey A, Kuleshov M V., Clarke DJB, Evangelista JE, Jenkins SL, et al. Gene Set Knowledge Discovery with Enrichr. Curr Protoc. 2021;1(3):1–51.
187. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene Ontology: tool for the unification of biology. Nat Genet [Internet]. 2000 May;25(1):25–9. Available from: https://www.nature.com/articles/ng0500_25
188. Aleksander SA, Balhoff J, Carbon S, Cherry JM, Drabkin HJ, Ebert D, et al. The Gene Ontology knowledgebase in 2023. Genetics [Internet]. 2023;224(1):1–14. Available from: https://doi.org/10.1093/genetics/iyad031
189. Thomas PD, Ebert D, Muruganujan A, Mushayahama T, Albou LP, Mi H. PANTHER: Making genome-scale phylogenetics accessible to all. Protein Sci. 2022;31(1):8–22.
190. Peltz G. Data and text mining GSEApy : a comprehensive package for performing gene set enrichment analysis in Python. 2023;39(November 2022):1–3.
191. Fang Z, Liu X, Peltz G. GSEApy: a comprehensive package for performing gene set enrichment analysis in Python. Bioinformatics. 2023;39(1):1–3.
192. Liberzon A, Birger C, Thorvaldsdóttir H, Ghandi M, Mesirov JP, Tamayo P. The Molecular Signatures Database Hallmark Gene Set Collection. Cell Syst. 2015;1(6):417–25.
193. Jin Z, Sato Y, Kawashima M, Kanehisa M. KEGG tools for classification and analysis of viral proteins . Protein Sci. 2023;(July).
194. Xie C, Mao X, Huang J, Ding Y, Wu J, Dong S, et al. KOBAS 2.0: A web server for annotation and

identification of enriched pathways and diseases. Nucleic Acids Res. 2011;39(SUPPL. 2):316–22.
195. Pathan M, Keerthikumar S, Ang CS, Gangoda L, Quek CYJ, Williamson NA, et al. FunRich: An open access standalone functional enrichment and interaction network analysis tool. Proteomics. 2015;15(15):2597–601.
196. Bateman A, Martin MJ, O'Donovan C, Magrane M, Apweiler R, Alpi E, et al. UniProt: A hub for protein information. Nucleic Acids Res. 2015;43(D1):D204–12.
197. Liu P, Yang L, Shi D, Tang X. Prediction of protein-protein interactions related to protein complexes based on protein interaction networks. Biomed Res Int. 2015;2015.
198. Shao M, Yang Y, Guan J, Zhou S. Choosing appropriate models for protein-protein interaction networks: A comparison study. Brief Bioinform. 2013;15(5):823–38.
199. Zhao J, Lei X, Wu FX. Predicting protein complexes in weighted dynamic PPI networks based on ICSC. Complexity. 2017;2017(Mcl):1–11.
200. Saha S, Sengupta K, Chatterjee P, Basu S, Nasipuri M. Analysis of protein targets in pathogen-host interaction in infectious diseases: A case study on Plasmodium falciparum and Homo sapiens interaction network. Brief Funct Genomics. 2018;17(6):441–50.
201. Shen Z, Pu S, Cao X, Tang M, Wang S, Bai D, et al. Bioinformatics and network pharmacology analysis of drug targets and mechanisms related to the comorbidity of epilepsy and migraine. Epilepsy Res. 2023 Jan 1;189:107066.
202. Roopenian DC, Low BE, Christianson GJ, Proetzel G, Sproule TJ, Wiles M V. Albumin-deficient mouse models for studying metabolism of human albumin and pharmacokinetics of albumin-based drugs. MAbs. 2015;7(2):344–51.
203. Rotramel A, Poritz LS, Messaris E, Berg A, Stewart DB. PPI Therapy and Albumin are Better Predictors of Recurrent Clostridium difficile Colitis than Choice of Antibiotics. J Gastrointest Surg. 2012;16(12):2267–73.
204. Franiak-Pietryga I, Ostrowska K, Maciejewski H, Appelhans D, Misiewicz M, Ziemba B, et al. PPI-G4 Glycodendrimers Upregulate TRAIL-Induced Apoptosis in Chronic Lymphocytic Leukemia Cells. Macromol Biosci. 2017;17(5):1–8.
205. Vieira VDF, Xavier CR, Ebecken NFF, Evsukoff AG. Performance evaluation of modularity based community detection algorithms in large scale networks. Math Probl Eng. 2014;2014.
206. Ding Y. Community detection: Topological vs. topical. J Informetr. 2011;5(4):498–514.
207. Zeng J, Yu H. A Scalable Distributed Louvain Algorithm for Large-Scale Graph Community Detection. Proc - IEEE Int Conf Clust Comput ICCC. 2018;2018-Septe:268–78.
208. Traag VA, Waltman L, van Eck NJ. From Louvain to Leiden: guaranteeing well-connected communities. Sci Rep. 2019;9(1):1–12.
209. Bhattacharyya M, Chakrabarti S. Identification of important interacting proteins (IIPs) in Plasmodium falciparum using large-scale interaction network analysis and in-silico knock-out studies. Malar J. 2015;14(1):1–17.
210. Yu Y, Kong D. Protein complexes detection based on node local properties and gene expression in PPI weighted networks. BMC Bioinformatics [Internet]. 2022;23(1):1–15. Available from: https://doi.org/10.1186/s12859-021-04543-4
211. Wang R, Liu G, Wang C. Identifying protein complexes based on an edge weight algorithm and core-attachment structure. BMC Bioinformatics. 2019;20(1):1–20.
212. Zhang Z, Song J, Tang J, Xu X, Guo F. Detecting complexes from edge-weighted PPI networks via

genes expression analysis. BMC Syst Biol [Internet]. 2018;12(Suppl 4). Available from: http://dx.doi.org/10.1186/s12918-018-0565-y

213. Wahab Khattak F, Salamah Alhwaiti Y, Ali A, Faisal M, Siddiqi MH. Protein-Protein Interaction Analysis through Network Topology (Oral Cancer). J Healthc Eng. 2021;2021.
214. Zhang XF, Ou-Yang L, Zhu Y, Wu MY, Dai DQ. Determining minimum set of driver nodes in protein-protein interaction networks. BMC Bioinformatics. 2015;16(1):1–13.
215. Lari N, Jalal R, Minuchehr Z, Rajabian Noghondar M. Identifying miltefosine-resistant key genes in protein–protein interactions network and experimental verification in Iranian Leishmania major. Mol Biol Rep [Internet]. 2019;46(5):5371–88. Available from: https://doi.org/10.1007/s11033-019-04992-4
216. Ghedira K, Hamdi Y, El Béji A, Othman H. An Integrative Computational Approach for the Prediction of Human- Plasmodium Protein-Protein Interactions. Biomed Res Int. 2020;2020.
217. Doolittle JM, Gomez SM. Mapping protein interactions between dengue virus and its human and insect hosts. PLoS Negl Trop Dis. 2011;5(2).
218. Ayo CM, Dalalio MMDO, Visentainer JEL, Reis PG, Sippert EÂ, Jarduli LR, et al. Genetic Susceptibility to Chagas disease: An overview about the infection and about the association between disease and the immune response genes. Biomed Res Int. 2013;2013.

SUPPORTING INFORMATION

In the following link, there is a repository with the main tables and figures of the research: https://github.com/JaviMendoza9/PPI_Network_of_the_PVH_Systeam/tree/JaviMendoza9-patch-1/PPI_PVH

S1 Fig. Graphic representation of the PPI Subnetwork of the PVH System.

- S1_1 Fig. Diameter and Nodes of Connected Component 1: 5 and 417 (https://github.com/JaviMendoza9/PPI_Network_of_the_PVH_Systeam/blob/JaviMendoza9-patch-1/PPI_PVH/Results/Connectivity_Analysis/S3_1_Figure.png)
- S1_2 Fig. Diameter and Nodes of Connected Component 2: 1 and 2 (https://github.com/JaviMendoza9/PPI_Network_of_the_PVH_Systeam/blob/JaviMendoza9-patch-1/PPI_PVH/Results/Connectivity_Analysis/S3_2_Figure.png)
- S1_3 Fig. Diameter and Nodes of Connected Component 3: 5 and 29 (https://github.com/JaviMendoza9/PPI_Network_of_the_PVH_Systeam/blob/JaviMendoza9-patch-1/PPI_PVH/Results/Connectivity_Analysis/S3_3_Figure.png)
- S1_4 Fig. Diameter and Nodes of Connected Component 4: 6 and 53 (https://github.com/JaviMendoza9/PPI_Network_of_the_PVH_Systeam/blob/JaviMendoza9-patch-1/PPI_PVH/Results/Connectivity_Analysis/S3_4_Figure.png)

S2 Fig. Graphic representation of the edge weight of the PPI Subnetwork of the PVH System.

- S2_1 Fig. Interaction level in component 1: 5405; Total amount of weight in component 1: 3427.8640000000028 (https://github.com/JaviMendoza9/PPI_Network_of_the_PVH_Systeam/blob/JaviMendoza9-patch-1/PPI_PVH/Results/Proteins/S12_1_Figure.png)
- S2_2 Fig. Interaction level in component 2: 1; Total amount of weight in component 2: 0.428 (https://github.com/JaviMendoza9/PPI_Network_of_the_PVH_Systeam/blob/JaviMendoza9-patch-1/PPI_PVH/Results/Proteins/S12_2_Figure.png)

- S2_3 Fig. Interaction level in component 3: 122; Total amount of weight in component 3: 113.70700000000004 (https://github.com/JaviMendoza9/PPI_Network_of_the_PVH_Systeam/blob/JaviMendoza9-patch-1/PPI_PVH/Results/Proteins/S12_3_Figure.png)
- S2_4 Fig. Interaction level in component 4: 268; Total amount of weight in component 4: 228.39200000000017 (https://github.com/JaviMendoza9/PPI_Network_of_the_PVH_Systeam/blob/JaviMendoza9-patch-1/PPI_PVH/Results/Proteins/S12_4_Figure.png)

S1 Table. PVH System PPI Network Interactions (https://github.com/JaviMendoza9/PPI_Network_of_the_PVH_Systeam/blob/JaviMendoza9-patch-1/PPI_PVH/Results/S3_Table.xlsx)

S2 Table. Components connected to the PPI network of the PVH System (https://github.com/JaviMendoza9/PPI_Network_of_the_PVH_Systeam/blob/JaviMendoza9-patch-1/PPI_PVH/Results/Connectivity_Analysis/S3_Table.tsv)

S3 Table. Degree Value top to bottom of the PPI network of the PVH System (https://github.com/JaviMendoza9/PPI_Network_of_the_PVH_Systeam/blob/JaviMendoza9-patch-1/PPI_PVH/Results/Connectivity_Analysis/S4_Table.xlsx)

S4 Table. Shortest path of the PPI network of the PVH System (https://github.com/JaviMendoza9/PPI_Network_of_the_PVH_Systeam/blob/JaviMendoza9-patch-1/PPI_PVH/Results/Connectivity_Analysis/S5_Table.xlsx)

S5 Table. Clustering coefficients of the PPI network of the PVH System (https://github.com/JaviMendoza9/PPI_Network_of_the_PVH_Systeam/blob/JaviMendoza9-patch-1/PPI_PVH/Results/Connectivity_Analysis/S6_Table.xlsx)

S6 Table. Centrality measures of the PPI network of the PVH System (https://github.com/JaviMendoza9/PPI_Network_of_the_PVH_Systeam/blob/JaviMendoza9-patch-1/PPI_PVH/Results/Centrality_Measures/S7_Table.xlsx)

S7 Table. CNM communities of the PPI network of the PVH System (https://github.com/JaviMendoza9/PPI_Network_of_the_PVH_Systeam/blob/JaviMend

oza9-patch-1/PPI_PVH/Results/Community_Detection/S8_Table.xlsx)

S8 Table. Louvain communities of the PPI network of the PVH System (https://github.com/JaviMendoza9/PPI_Network_of_the_PVH_Systeam/blob/JaviMendoza9-patch-1/PPI_PVH/Results/Community_Detection/S9_Table.xlsx)

S9 Table. Edges weight of the PPI network of the PVH System (https://github.com/JaviMendoza9/PPI_Network_of_the_PVH_Systeam/blob/JaviMendoza9-patch-1/PPI_PVH/Results/Proteins/S10_Table.xlsx)

S10 Table. Enrichrpy of the PPI network of the PVH System (https://github.com/JaviMendoza9/PPI_Network_of_the_PVH_Systeam/blob/JaviMendoza9-patch-1/PPI_PVH/Results/Functional_Enrichment/S10_Table.xlsx)

S11 Table. GSEApy of the PPI network of the PVH System (https://github.com/JaviMendoza9/PPI_Network_of_the_PVH_Systeam/blob/JaviMendoza9-patch-1/PPI_PVH/Results/Functional_Enrichment/S11_Table_human.xlsx)