Pathogenic Protist Transmembranome database (PPTdb): a web-based platform for searching and analysis of protist transmembrane proteins

Chi-Ching Lee1,2, Po-Jung Huang2,3, Yuan-Ming Yeh2, Sin-You Chen1, Cheng-Hsun Chiu2,4, Wei-Hung Cheng5* and Petrus Tang4,5*

From The 8th Annual Translational Bioinformatics Conference
Seoul, South Korea. 31 October - 2 November 2018

Abstract

Background: Pathogenic protist membrane transporter proteins play important roles not only in exchanging molecules into and out of cells but also in acquiring nutrients and biosynthetic compounds from their hosts. Currently, there is no centralized protist membrane transporter database published, which makes system-wide comparisons and studies of host-pathogen membranomes difficult to achieve.

Results: We analyzed over one million protein sequences from 139 protists with full or partial genome sequences. Putative transmembrane proteins were annotated by primary sequence alignments, conserved secondary structural elements, and functional domains. We have constructed the PPTdb (Pathogenic Protist Transmembranome database), a comprehensive membrane transporter protein portal for pathogenic protists and their human hosts. The PPTdb is a web-based database with a user-friendly searching and data querying interface, including hierarchical transporter classification (TC) numbers, protein sequences, functional annotations, conserved functional domains, batch sequence retrieving and downloads. The PPTdb also serves as an analytical platform to provide useful comparison/mining tools, including transmembrane ability evaluation, annotation of unknown proteins, informative visualization charts, and iterative functional mining of host-pathogen transporter proteins.

Conclusions: The PPTdb collected putative protist transporter proteins and offers a user-friendly data retrieving interface. Moreover, a pairwise functional comparison ability can provide useful information for identifying functional uniqueness of each protist. Finally, the host and non-host protein similarity search can fulfill the needs of comprehensive studies of protists and their hosts. The PPTdb is freely accessible at http://pptdb.cgu.edu.tw.

Background

Bioactive molecules that cross through the extracellular barrier mainly rely on channels, pores, or energy-consuming pumps composed of transporter proteins [1]. Transporters are specifically necessary for parasitic protists to salvage essential molecules for survival, such as nucleotides/ nucleosides, carbohydrates, and amino acids, from the human host [2–8]. In addition, membrane transporters also play a role in the drug resistance of protists [9]. These pivotal roles of transporters for parasitism strongly prompt parasitologists to characterize protist transporters in order to extend knowledge in protist pathogenesis and innovate new treatment strategies. Indeed, more than 50% of therapeutic drugs targeted to receptors or transporters on the membrane, emphasizing the importance of membrane proteins in disease control

© The Author(s). 2019 Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated.
However, it is difficult to characterize the biological functions of membrane proteins because of uncertain experimental procedures [11]. Thus, a functional prediction based on protein sequence would be helpful to concentrate on the interested membrane proteins. Although whole genome sequencing of important human pathogenic protists has been conducted during the last decade, information about protist transporters has been lacking. Therefore, the goal of this study was to build the human Pathogenic Protist Transmembranome database (PPTdb) to provide classification and system-wide comparison to accelerate the exploration of protist transporters.

Several databases collecting transporter information have been constructed as a resource of known and putative transporter proteins. The Transporter Classification Database (TCDB) is the only classification organization for transporters that documents at least 1000 transporter families based on function and phylogeny derived from published studies [12]. TransportDB 2.0 has recorded potential transporters from ~2700 sequenced genomes using the following criteria: (1) primary amino acid sequence identity, and (2) predicted structural homology and topology [13]. PPTdb classification was based on protein sequence alignment to known human transporters, coupled with transmembrane domain prediction and gene ontology (GO) categorization to increase the accuracy of transporter annotation in sequenced protist genomes.

Advanced applications such as expression profiles and polymorphisms of transporters are included in the Human Transporter Database (HTD) [14]. The Yeast Transporter Protein database (YTPdb) contains membrane topology, post-translational modifications, and a "wiki-like" freely updatable platform [15]. The PPTdb is not only a resource but also an interactive functional search engine for potential protist transporters. The query section of PPTdb is composed of sequence identity to known human transporters and conserved transporter characteristics. The summary of a PPTdb search consists of a pairwise functional comparison of protist to human transporters that distinguishes either protist-specific transporters or human homologs. Furthermore, a functional comparison between two protists can be examined to highlight the commonalities between or uniqueness of transporters in each organism. An iterative search panel is another option for multiple functional queries in order to obtain more specific targets.

As the first transporter collection for protists, the PPTdb has globally analyzed more than one million protein transporters collected from sub-databases of the EupathDB family and from the NCBI genome archive. The classification of potential protist transporters depends on the sequence similarity to human host transporters, putative transmembrane domains, and GO functional groupings. The user-friendly interface allows one to use different strengths to filter out the desired transporters and to conduct inter-species comparisons to humans and other protists. Additionally, the iterative functional mining allows for the entry of more than one keyword to find possible transporters within one round of searching. Comparing to previous databases, PPTdb offers an easy-to-use data querying interface, such as human homolog gene filter, iterative functional mining, and PPTdb also contains regularly updated gene and amino acid data by a back-end automatically data processing pipeline. The PPTdb thus provides a platform for parasitologists to understand the field of transporters and discover new therapeutic targets in important protists by comparing human homologous genes and protist uniqueness transporter genes.

Construction and content
Sequence characterization and annotation
The data processing pipeline was mainly built by service-side command-line PHP and a suite of in-house developed text-processing and data retrieval modules which were implemented by Bash scripting language.

To gather the general and taxonomic information of each organism, we developed an automatic data retrieving and processing pipeline. Customized shell and PHP scripts were used to extract general and taxonomic information from NCBI taxonomy data portal. The taxonomy data were used to build the hierarchical species selection module on the front page of the website. To ensure that all the annotations were annotated using the same software environment, we re-annotated all the protein sequences based on functional and secondary structure information, such as transmembrane prediction by TMHMM, and functional annotations by PROSITE, Pfam, and Gene Ontology.

Web-interface and database architecture
The PPTdb is a database providing real-time user interaction functionalities. Its user interface was implemented by HTML5, jQuery, and Ajax which can be opened by most modern web browsers. Functional annotations, species categories, sequence alignments and general information were stored in a relational database implemented by MySQL, which guarantees short latency and instant responses by online queries.

The PPTdb collected all protein sequence BLASTp results; traditional SQL queries such as ‘join’ and ‘nested-query’ require significant waiting time. To minimize the data querying time, we designed an SQL and PHP two-step conditional query methodology. The first step was called SQL query. Information for each species was separately stored in a MySQL data table, which can reduce SQL querying time, and multiple species queries could be executed in parallel. User-selected gene lists sent from the web interface were ignored in the first stage query and passed on to step 2, and all genes fitting the functional and sequence similarity
conditions were selected from the database. The second step was the server-side gene query. The gene lists accepted from step 1 were stored in memory by creating a memory-cached search. According to our tests, this strategy could be executed at least twice as fast as a nested SQL query.

Amino acid sequences of all protist proteins are stored as FASTA files in our file-based database. A set of sequence-retrieving programs is used to extract user-selected gene sequences from web-based queries.

**Database contents**

The PPTdb collected all known protist-annotated protein sequences from the EupathDB and NCBI genome archive [16–18]. There were 139 protist genomes composed of *Entamoeba* and *Acanthamoeba* (11 species), *Cryptosporidium* (12 species), *Giardia* (6 species), *Microsporidia* (26 species), *Piroplasma* (8 species), *Plasmodium* (23 species), *Toxoplasma* (26 species), *Trichomonas* (one species), and *Trypanosomatidae* (26 species). All the human transporter protein sequences were retrieved from the TCDB data portal which provides the comprehensive annotations of transporter sequences, including functional classifications. To construct human transporter sequences for pathogen and host studies, we selected all human transporter sequences based on the annotation records of TCDB. Table 1 shows the summary of database statistics.

The PPTdb serves as a knowledge portal, an online functional mining platform, and a cross-species comparison tool specific to protist transporter proteins and human hosts. It is freely accessible at [http://pptdb.cgu.edu.tw](http://pptdb.cgu.edu.tw). Figure 1 shows the analysis workflow and data mining processes of the PPTdb.

The database had collected 1,055,827 protein sequences which were annotated based on the same annotating software and parameters, including 465,391 transmembrane domains, 4429 Pfams, 1064 superfamilies, 1896 GO terms, 668 PROSITE patterns, and 695 PROSITE profiles. Figure 2 represents the interactive web interface of the PPTdb.

**Utility and discussion**

**Transmembrane domain filter**

The PPTdb annotated all proteins of all collected protists by TMHMM, which is the most widely used transmembrane domain (TM) prediction tool. Users could select transmembrane candidate proteins of interest with a specific number of TMs, such as a six-TM potassium channel, or a single TM domain for an alpha-helix. Using the same TMs could easily identify proteins with similar structures or biological functions, which was easier than searching an entire dataset.

**One-click potential transporter gene finder**

We collected all 16,478 transporter proteins downloaded from the TCDB and annotated them by Gene Ontology terms to construct a transporter protein functional ontology dataset. Proteins collected in the TCDB share approximately 1300 GO terms (Additional file 1). Functional ontology terms were used to identify putative transporter proteins including functionally clarified and hypothetical proteins. Comparing to existing protist resources and collecting the feedbacks of testing members, the one-click potential transporter finder may be the most straightforward biological function-based gene finder for protists, of which annotations and sequences had not been comprehensively completed.

**Sequence homology search for human transporters**

Primary sequence alignment is an effective method to identify proteins as human homologs. Protists and their human-host homolog genes could be used on studies of drug design and host-parasite interactions, because most functional elements share similar sequence identities, such as secondary structural domains and conserved functional elements. The PPTdb executed all-against-all amino acid sequence alignment by BLASTp on proteins of protist organisms and human transporter proteins adopted from TCDB. It also provided a sequence homology search interface allowing the user to preset the search criteria, including sequence identity, similarity, and a ratio of alignment length versus length of target or query proteins. A higher ratio indicates similarity to human proteins. The search result was visualized by a Venn diagram helping users to select or rule out human homology proteins. Users could click the checkbox on the Venn diagram to select protist-unique proteins, human transporter homology proteins, or both. Genes fitting the above criteria were summarized in functional component charts, and listed in Type-n-Search dynamic tables on the bottom side of the web page. Moreover, the table delivered transporter classification (TC) IDs for the human protein homologs which were identified by BLASTp alignment.

**Functional component charts**

The selected putative transporter genes using the above filters were summarized using functional component charts by delivering the top 10 domains/components, allowing users to identify the most dominant functional elements from the genomic point of view. There were five functional component charts: GO, Pfam, PROSITE pattern, PROSITE

---

**Table 1** Database statistics of PPTdb

| Category             | Number of entries |
|----------------------|-------------------|
| Protists             | 139               |
| Protein sequences    | 1,039,394         |
| GO terms             | 1896              |
| Transmembrane domains| 465,391           |
Fig. 1 Analysis workflow of PPTdb Protein sequences collected were assigned functional annotations by the same software packages with equal parameters, including PROSITE, Pfam, Gene Ontology, and Super Family. The all-against-all primary sequence alignment executed by BLASTp was performed as well, and they were then stored in our back-end database. The system was armed with a high-speed database query system that allows the user to obtain the data instantly from the database. Four filters (transmembrane domain (TM), transporter GOs, iterative functional mining, and sequence homology) were used to help users to narrow their search results.
All of them were highly associated with biological functions.

**Dynamic type-n-search table**

All the listed data including genes and GO terms were delivered by a Type-n-Search table, allowing users to narrow down the selection data. The data rows could be sorted by ascending or descending value (text data were sorted by alphabetical order, while the numeric data were sorted by number) by clicking the header in the data table. The keyword search box located in the top-right corner of the table could perform partial string matching for all columns in the table.

**Download page for sequence and annotation retrieval**

The PPTdb provided a data retrieval tool for protein sequences and annotations targeting genes selected by the functional filters and homology search tool mentioned above. Download by annotations delivers the gene ID and annotations. Download by BLAST results contained all the primary sequence similarity results of users’ selected genes against all protein sequences recorded in the TransportDB. Download by sequences provided all amino acid sequences formatted in a FASTA text file. All three download links were dynamically generated by the user modifying any search parameters above. Files were zipped in a tar.gz format which could be unzipped by conventional file compression tools such as 7-zip in Windows and MacOS or tar in Linux operating systems.

**Iterative functional mining**

Most protist resources provided basic search functionalities including a gene ID/name or keyword search. However, if a user wanted to search by specific biological function such...
Comparison to other protist transporter resources
Currently, there was no published transporter-system database specifically designed for protists. Two general-purpose transporter protein databases (TCDB and TransportDB) could be used for protist transporter system studies. However, both databases collected a limited number of protists, which could not provide sufficient resources for protist transporter protein studies. The TCDB contained data for more than ten thousand transporter proteins and provided a transporter classification (TC) system recognized by the International Union of Biochemistry and Molecular Biology (IUBMB), generating systematical functional categories. However, there were only 75 protists in this database and fewer than 300 protist transporters. TransportDB provided clean and comprehensive data resources for transporter systems targeting to sequenced genomes which contained ~2500 bacteria species; however, there were fewer than 40 eukaryotic species (and only 11 protists) in the latest published version.

Both TCDB and TransportDB provided interfaces for transporter protein studies; however, neither of them offered mining tools for protist versus human homolog gene searches, protist to protist comparison, or further protist-associated studies. Table 2 listed the major differences between PPTdb, TCDB, and TransportDB from the protist research point of view.

Iterative functional mining workflow: an example of use
Horizontal gene transfer events were known as evolutionary driving forces of eukaryotes [19]. For example, nucleotide transporter (NTT) gene acquisition was reported as a major evolutionary innovation of Microsporidia which were intracellular parasites of animals and human [20].

As a proof of our user-friendly interface, sequences of putative NTTs were identified from seven Microsporidia species using PPTdb’s iterative functional mining workflow. By using traditional one-way search interface that only allowed single query once a time, users must enter an exactly correct term, for example, “nucleotide transporter” returned only 1 GO descriptions in A. Algerae PRA339. This was becausrse words between “nucleotide” and “transport” could not be searched by the general query method. In iterative functional mining, one could first search word “nucleotide” which returned all GO descriptions which contained keyword “nucleotide”. Then, the type-n-search data table of PPTdb allowed a
secondary refining search by entering “transport” to narrow down search results. This two-way search interface offered a high flexibility especially on multiple keyword combinations which could not be searched by one-way search, for example, nucleotide-sugar transmembrane transporter activity (GO:0005338), guanine nucleotide transmembrane transporter activity (GO:0001409), and nucleotide transmembrane transporter activity (GO:0015215). Moreover, the type-n-search table instantaneously returned database query candidates. Users could obtain putative results by just entering a few characters instead of entire searching keywords. For example, entering “transport” would return all the candidate entries including “transport”, “transporter”, and “transmembrane transporter” which could be act as a search guidance to inform users that there was more than one “transport” associated keywords in the GO description database.

In this example, 22 putative transporters were identified by those nucleotide transport associated GO descriptions from seven Microsporidia species (Table 3). Interestingly,

| Table 2 | A comparison table between PPTdb and other protist transporter resources |
|---------|-------------------------------------------------|
| **Tool** | **PPTdb** | **TCDB** | **TransportDB** |
| Protist Genomes | 139 | 75* | 11 |
| Designed for | Protist-human transporter studies | Transporter classification | Transporter knowledgebase |
| Annotations | | | |
| Human homology search | | | |
| Functional keyword search | | | |
| Pairwise protist comparison | | | |
| Visualized comparison results | | | |
| Integrated Transport classification id | | | |

*full function; □ partial function; – NA
*cross reference to EupathDB by species name
*PPTdb provides iterative functional mining, other resources offer only one-way search

| Table 3 | Search entries of NTTs in seven Microsporidia species |
|---------|-------------------------------------------------|
| **Species** | **Homology search** | **NTT (putative transporters)** | **Transcript Id** | **Annotation** | **Protein Length** |
| A. algerae PRA109 | Portist uniqueness | 2 (60) | H311_01875 | hypothetical protein | 250 |
| | Human homolog | 2 (59) | H311_05190 | hypothetical protein | 120 |
| | | | H311_02925 | hypothetical protein | 562 |
| | | | H311_03867 | hypothetical protein | 327 |
| A. algerae PRA339 | Portist uniqueness | 1 (51) | H312_02827 | hypothetical protein | 539 |
| | Human homolog | 1 (49) | H312_02469 | hypothetical protein | 562 |
| | | | AEWD_081300 | ATP/ADP translocase | 559 |
| | | | AEWD_100320 | ATP/ADP translocase | 536 |
| | | | AEWD_100430 | ATP/ADP translocase | 543 |
| E. cuniculi EC1 | Portist uniqueness | 4 (41) | AEWD_100450 | ATP/ADP translocase | 553 |
| | Human homolog | 0 (41) | – | – | – |
| N. bombycis CQ1 | Portist uniqueness | 1 (80) | NBO_251gi001 | ADP/ATP carrier protein 1 | 206 |
| | Human homolog | 2 (59) | NBO_55g0018 | ADP,ATP carrier protein 1 | 484 |
| P. neurophilia strain MK1 | Portist uniqueness | 2 (37) | M153_7540001844 | ATP:ADP Antiporter (AAA) Family | 568 |
| | | | M153_1368000483 | ATP:ADP Antiporter (AAA) Family | 568 |
| | Human homolog | 1 (57) | M153_3260001719 | ATP:ADP Antiporter (AAA) Family | 560 |
| V. corneae ATCC 50505 | Portist uniqueness | 3 (56) | VCG_00013 | hypothetical protein | 240 |
| | | | VCG_00014 | hypothetical protein | 329 |
| | | | VCG_00919 | hypothetical protein | 609 |
| | Human homolog | 0 (46) | – | – | – |
| V. culicis subsp. floridensis | Portist uniqueness | 2 (57) | VCLG_00419 | hypothetical protein | 552 |
| | Human homolog | 1 (44) | VCLG_01129 | hypothetical protein | 566 |

The parameters used in this example: BLASTp identity: 30%; one-click-transporter filter: on; Number of transmembrane domains ≥ 1
more than half of the results are hypothetical proteins. The protist uniqueness genes and human transporter homologs could be easily separated from the search. Finally, we could download all the FASTA sequences by clicking the Download link of the web page. Tools such as MAFFT [21] and ClustalW [22] could do the multiple sequence alignment and deliver the phylogenetic tree of these sequences. Detailed information of steps would be found on the demonstration page of PPTdb (http://pptdb.cgu.edu.tw/demo.php).

Specific search strategies for putative transporter proteins
PPTdb offered the precisely functional search instead of general keyword search used by general purposed databases such as Entrez Gene [17, 18], UniProt [23], and EuPathDB [16]. In addition to keyword search, PPTdb also provided specialized pre-set search filters for putative transporter proteins, including one-click-potential transporter genes, number of transmembrane domains, and iterative functional search boxes. Table 4 illustrated a search term “nucleotide” for putative transporter proteins with 1 transmembrane domain of species Acanthamoeba castellanii str. Neff. It was complicated to mimic all possible search movements for every user of these databases that provided several advanced search tools. The most straightforward search strategy was the keyword search and filters provided by each database. The search would be set if the database offers pre-set filters, such as number of transmembrane domains. However, the keyword search was used without pre-set filters, such as putative transporter proteins. The result showed that the functional filter and iterative search functionalities could make the search more user-friendly than other general purposed databases. All the searched genes IDs from PPTdb could be downloaded in a text format and then be executed more detailed data retrieving processes, such as genetic information from Entrez Gene database, protein 3D structures and pathways from UniProt database.

Future work
PPTdb collected all the putative transporter protist proteins and provides a user-friendly data querying interface. The next goal is to collect the human validated transporter proteins by offering a system for putative and validated proteins comparison. The collection of 3D structures of protist transporters will be the goal as well. We believe these future works will make PPTdb more useful on potential protist associated treatment or drug development.

Conclusions
The PPTdb had been specifically designed for protist transporter system studies and provides a data query portal, an online comparison tool, and a flexible functional search interface. For all putative, hypothetical, or curated protist proteins, the PPTdb provided functional annotations. The PPTdb also offered a straightforward protist-human homology search interface for pathogen and host studies.

Table 4 Search comparisons PPTdb and several general purpose genomic databases

| Items                        | PPTdb                                      | Entrez Gene     | UniProt          | EuPathDB       |
|------------------------------|--------------------------------------------|-----------------|------------------|----------------|
| Search term                  | nucleotide                                 | nucleotide AND  | nucleotide transporter AND | nucleotide transporter |
|                              |                                            | Acanthamoeba castellanii str. Neff.     | Acanthamoeba castellanii str. Neff. | Acanthamoeba castellanii str. Neff. |
| Pre-search settings          | Click A. castellanii str. Neff              | –               | –                | Click AmoebaDB |
|                              | Set number of transmembrane domain: 1      | –               | –                | 1135 genes for keyword search |
|                              | One-click potential transporters: checked  | –               | –                | 76 genes with  |
|                              |                                            | –               | –                | 1 transmembrane domain |
| Number of searched items     | 12 GO terms                                 | 3 Prosite Patterns | No number of transmembrane domain filter | No number of transmembrane domain filter |
|                              |                                            | 5 Pfams         |                   |                 |
|                              |                                            | 105             | b1303            | b1290           |
|                              |                                            |                 |                  | 0/76            |

aThere are several search combinations of each database, it is reasonable that different search combinations may lead to different search results. This demo demonstrates one search strategy of each database.
bNo transmembrane domain filter.
cThere is no specific column for putative transporter genes of EuPathDB. Using “transporter” as keyword is the most possible search action used by most users.
Additional file

Additional file 1: GO terms used in the PPTdb (XLSX 45 kb)

Abbreviations
BLASTp: Basic local alignment search tool for protein; GO: Gene Ontology; HTD: Human Transporter Database; HTML5: HyperText markup language S; IUBMB: International Union of Biochemistry and Molecular Biology; MySQL: My structured query language; NCBI: National Center for Biotechnology Information; NTT: Nucleotide transporter; PHP: PHP hyperext preprocessor; PPTdb: Pathogenic Prost Transmembrane database; TC: Transporter classification; TCDB: Transporter classification database; TM: Transmembrane domain; TMHMM: Transmembrane Helices; Hidden Markov Model; YTPdb: Yeast Transporter Protein Database

Acknowledgement
Not applicable

Funding
This study was supported by grants from the Chang Gung Memorial Hospital Research Funding (CMRPD1G0541-1G0543) and Ministry of Science and Technology, Taiwan (MOST 106–2221-E-182-068). Publication of this article was sponsored by the grants MOST 107–2320-B-182-021-MY3 and BMRPO56.

Availability of data and materials
The web site of PPTdb is freely accessible at http://pptdb.cgu.edu.tw.

About this article
This article has been published as part of BMC Bioinformatics Volume 20 Supplement 13, 2019: Selected articles from the 8th Translational Bioinformatics Conference: Bioinformatics. The full contents of the supplement are available online at https://bmcbioinformatics.biomedcentral.com/articles/supplements/ volume-20-supplement-13.

Authors’ contributions
Software, CCL, P.H. and YMY; Validation, CCL, P.H. and YMY; Investigation, CCL, SYC. and WHC.; Resources, CCL; Data Curation, CCL. and SYC.; Writing-Original Draft Preparation, CCL. and WHC.; Writing-Review & Editing, CHC., WHC. and P.T.; Visualization, CCL, P.H. and YMY.; Supervision, P.T. All authors have read and approved the final manuscript.

Ethics approval and consent to participate
Not applicable

Consent for publication
Not applicable

Competing interests
The authors declare that they have no competing interests.

Publisher’s Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details
1Department and Graduate Institute of Computer Science and Information Engineering, Chang Gung University, Taoyuan, Taiwan. 2Genomic Medicine Core Laboratory, Chang Gung Memorial Hospital, Linkou, Taiwan. 3Department of Biomedical Sciences, Chang Gung University, Taoyuan, Taiwan. 4Molecular Infectious Disease Research Center, Chang Gung Memorial Hospital, Linkou, Taiwan. 5Department of Parasitology, College of Medicine, Chang Gung University, Taoyuan, Taiwan. Published: 24 July 2019

References
1. Dean P, Major P, Nakjang S, Hint RP, Embley TM. Transport proteins of parasitic protozoa and their role in nutrient salvage. Front Plant Sci. 2014;5:153.
2. Vasudevan G, Carter NS, Drew ME, Beverley SM, Sanchez MA, Seyfang A, Ullman B, Landfear SM. Cloning of Leishmania nucleoside transporter genes by rescue of a transport-deficient mutant. Proc Natl Acad Sci USA. 1998;95:9783–8.
3. Tsabuis AD, Kunji ERS, Goldberg AV, Luccooq JM, Hirt RP, Embley TM. A novel route for ATP acquisition by the remnant mitochondria of Encephalitozoon cuniculi. Nature. 2008;453:533–6.
4. Tjaden J, Haferkamp L, Boxma B, Tiellens AGM, Huynen M, Hackstein JHP. A divergent ADP/ATP carrier in the hydrogenosomes of Trichomonas gallinae argues for an independent origin of these organelles. Mol Microbiol. 2004;51:1439–46.
5. Landfear SM. Glucose transporters in parasitic protozoa. Methods Mol Biol. 2010;637:245–62.
6. Blume M, Hilics M, Rodriguez-Contreras D, Sanchez M, Landfear S, Lucius R, Matuschewski K, Gupta N. A constitutive pan-hexose permease for the Plasmodium life cycle and transfgenic models for screening of antimalarial sugar analogs. FASEB J. 2011;25:1218–29.
7. Inbar E, Schlisselberg D, Suter Gromeyer M, Renttsch D, Zilberstein D. A versatile proline/saline transporter in the unicellular pathogen Leishmania donovani regulates amino acid homeostasis and osmotic stress responses. Biochem. J. 2013;449:555–66.
8. Shaked-Mishan P, Suter Gromeyer M, Yoel-Almagor T, Holland N, Zilberstein D, Renttsch D. A novel high-affinity arginine transporter from the human parasitic protozoan Leishmania donovani. Mol Microbiol. 2006;60:30–8.
9. Fidock DA, Normura T, Talley AK, Cooper RA, Dzekunov SM, Ferdig MT, Usos LM, Sidhu BA, Naudé B, Detsch KM, Su XZ, Woottom JC, Roeppe PD, Wellens TE. Mutations in the P. falciparum digestive vacuole transmembrane protein PICT2 and evidence for their role in chloroquine resistance. Mol Cell. 2006;6:861–71.
10. Rask-Andersen M, Almén MS, Schloth H. Trends in the exploitation of novel drug targets. Nat Rev Drug Discov. 2011;10:579–90.
11. Carpenter EP, Beis K, Cameron AD, Iwata S. Overcoming the challenges of membrane protein crystallography. Curr Opin Struct Biol. 2008;18:581–6.
12. Saier MH, Reddy VS, Tsi BV, Ahmed MS, Li C, Moreno-Hagelsieb G. The transporter classification database (TCDB): recent advances. Nucleic Acids Res. 2018;46:D372–9.
13. Elbourne LDH, Tetu SG, Hassan KA, Paulsen IT. TransportDB 2.0. a database for exploring membrane transporters in sequenced genomes from all domains of life. Nucleic Acids Res. 2017;45:D320–4.
14. Ye A, Liu QR, Li CY, Zhao M, Qu H. Human transporter database: comprehensive knowledge and discovery tools in the human transporter genes. PLoS One. 2014;9:e88883.
15. Brohée S, Barriot R, Moreau Y, André B. YTPdb: a wiki database of yeast membrane transporters. Biochim Biophys Acta. 2010;1798:1008–12.
16. Aurrecoechea C, Barreto A, Basenky EY, Brestelli J, Brunk BP, Cade S, Crouch K, Doherty R, Fallek D, Fischer S, Gajria B, Harb OS, Heiges M, Hertz-Fowler C, Hu S, Iodice J, Kissingar JC, Lawrence C, Li W, Pinney DF, Pullman JA, Roos DS, Shankugusundram A, Silva-Franco F, Steinbiss S, Stoeckert CJ, Sprull D, Wang H, Warrenfeltz S, Zheng J. EuPathDB: the eukaryotic pathogen genomics database resource. Nucleic Acids Res. 2017;45:D581–91.
17. Sayers EW, Barrett T, Benson DA, Bryant SH, Lipman DJ, Ostell J, Pruitt KD, Schuler GD, Sequeira E, Sherry ST, Shumway M, Sirotkin K, Souvorov A, Starchenko G, Tatusova TA, Wagner L, Yanchenko E, Ye J. Database resources of the National Center for biotechnology information. Nucleic Acids Res. 2009;37:D15.–19.
18. Benson DA, Karch-Mizichaki I, Lipman DJ, Ostell J, Sayers EW. GenBank. Nucleic Acids Res. 2009;37:D26–31.
19. Schönknecht G, Weber APM, Lerner MJ. Horizontal gene acquisitions by eukaryotes as drivers of adaptive evolution. Bioessays. 2014;36:9–20.
20. Dean P, Sendra KM, Williams TA, Watson AK, Major P, Nakjang S, Kozhevnikova E, Goldberg AV, Kunji ERS, Hirt RP, Embley TM. Transporter gene acquisition and innovation in the evolution of microsporidia intracellular parasites. Nat Commun. 2018;9:1709.
21. Katoh K, Rozewicki J, Yamada KD. MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. Brief Bioinform. 2017;18:3059.
22. Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 1994;22:4673–80.

23. Bateman A. UniProt: a hub for protein information. Nucleic Acids Res. 2015;43:D204–12.