

Integrative omics reveals candidate drug-repurposing targets in *Tritrichomonas foetus*

Joanna Dąbrowska · Maciej Kochanowski · Jacek Sroka

Received: 1 July 2026 / Accepted: 1 September 2026

© The Author(s) 2026

Abstract

Background Feline tritrichomonosis is a globally distributed chronic large-bowel diarrhoea of cats caused by the flagellated protozoan *Tritrichomonas foetus*. The only drug with demonstrated clinical efficacy is the off-label 5-nitroimidazole ronidazole, which has a narrow therapeutic window and against which clinical resistance has now been documented across three continents. The last comprehensive in silico drug-target prioritisation for *T. foetus* was published a decade ago and proposed 113 candidates based on homology to ChEMBL alone; the genomic and pharmacological landscape has since changed substantially, but no integrative reanalysis has been undertaken.

Results Starting from the predicted proteome of strain K (25030 proteins), an integrative subtractive pipeline combining filtering against the feline host proteome, three pharmacological reference databases, essential-gene resources from three model eukaryotes, and the PHI-base virulence reference returned 433 priority candidates. Application of the same pipeline to the chromosome-scale strain KV-1 assembly yielded 436 candidates, with 66% cross-strain recovery as reciprocal best hits. Gene-level expression analysis showed that 82.7% of priority candidates were actively transcribed in the feline isolate G10/1, and a proof-of-concept docking screen against three top-ranked candidates recovered internally consistent, structure-based affinity rankings that await experimental validation. Nine of the ten most highly expressed candidates correspond to a protein family with clinically advanced small-molecule inhibitors; the tenth, a Rab4B GTPase, belongs to the broader, pharmacologically tractable Ras/GTPase superfamily but has no Rab4B-specific inhibitor characterised to date.

Conclusions This work provides a substantially expanded and cross-validated drug-target landscape for *T. foetus*. Nine of the ten most highly expressed priority candidates fall into protein families for which clinically advanced small-molecule inhibitors are already available, namely Hsp70 and Hsp90 chaperones, the CDC48/p97 AAA+ ATPase, the ubiquitin-proteasome pathway, and serine/threonine kinases; the tenth, a Rab4B GTPase, belongs to the broader Ras/GTPase superfamily but has no Rab4B-specific inhibitor characterised to date. Together they offer an immediate opportunity for drug-repurposing studies in feline tritrichomonosis, a disease that lacks any approved therapy and against which off-label ronidazole resistance is increasingly documented.

Keywords *Tritrichomonas foetus* · Drug-target prioritisation · Subtractive proteomics · Comparative genomics · Molecular docking · Drug repurposing · Feline tritrichomonosis

This Article in Press is shared early to give you faster access to new research. It is citable and carries a permanent DOI. The final edited version will replace it automatically.



Background

Tritrichomonas foetus is a flagellated protozoan of the phylum Parabasalia (order Trichomonadida, family Tritrichomonadidae) that causes tritrichomonosis in cattle and cats, whereas in pigs it occurs as an intestinal commensal [1, 2, 3, 4, 5, 6, 7, 8]. The trophozoite is pear-shaped, with three anterior flagella and a fourth recurrent flagellum forming an undulating membrane, and a posteriorly protruding axostyle. As an obligate anaerobe, the parasite lacks canonical mitochondria and generates ATP via substrate-level phosphorylation in hydrogenosomes, which oxidatively decarboxylate pyruvate and release molecular hydrogen. In *T. foetus*, as in *Trichomonas vaginalis*, these organelles are also the site of reductive activation of 5-nitroimidazoles, the only drug class currently used to treat feline tritrichomonosis.

In cats, *T. foetus* colonises the distal ileum and colon, causing chronic large-bowel diarrhoea with intermittent passage of malodorous mucoid faeces, occasional haematochezia, and tenesmus. Infection occurs predominantly in young animals from high-density environments such as catteries, shelters and multi-cat households, where faecal–oral transmission and asymptomatic carriage sustain pathogen circulation [5, 6]. Feline tritrichomonosis has a global distribution, with reported prevalence in at-risk populations ranging from approximately 10% to 59%, depending on geographical region, diagnostic methodology and study design. Across continents, infection has been documented in North America, Europe, Asia, South America and Australasia, again with substantial variability between studies; reported estimates include 31% in show cats in the United States, 14.4% in diarrhoeic cats in the United Kingdom, and 10.6% in a nationwide survey in Poland, increasing to 14.9% in cats under one year of age [9–13].

Treatment options for feline tritrichomonosis are limited [15]. Ronidazole, a 5-nitroimidazole structurally related to metronidazole, is currently the only drug with demonstrated efficacy against feline *T. foetus* in randomised controlled trials [14, 18, 19]. It is not licensed for use in cats in any country and is prescribed off-label, usually as a compounded capsule formulation, with the recommended dosage of 30 mg/kg orally once daily for 14 days representing the upper limit of tolerability [16]. Clinical use is constrained by a narrow therapeutic window: reversible neurotoxicity, including ataxia, hyperaesthesia, tremor and seizure-like episodes, is reported in a substantial proportion of treated cats and often necessitates dose reduction or discontinuation [17], while teratogenic and mutagenic effects documented in laboratory species further restrict use in breeding queens and kittens. Clinical and in vitro resistance of feline *T. foetus* to ronidazole has now been documented in the United States, Australia and, most recently, Hungary, with minimum lethal concentrations of resistant isolates ranging from several-fold to more than thirty-fold above those of susceptible reference strains [20–22]. Other agents offer no satisfactory alternative: tinidazole is less effective than ronidazole, metronidazole is largely ineffective against feline isolates, and paromomycin is contraindicated because oral administration to cats has been associated with severe nephro- and ototoxicity, including acute renal failure, deafness and cataract formation [23–25]. In line with these limitations, both the European Advisory Board on Cat Diseases and the Companion Animal Parasite Council explicitly state that no approved treatment for feline tritrichomonosis is currently available.

At the biochemical level, ronidazole and other 5-nitroimidazoles act as prodrugs whose antiprotozoal activity depends on intracellular reduction of the nitro group within the hydrogenosome. The reduction is catalysed by pyruvate: ferredoxin oxidoreductase (PFOR), which generates reduced ferredoxin during the oxidative decarboxylation of pyruvate; subsequent electron transfer from reduced ferredoxin to the nitro group produces a cytotoxic nitro radical anion that damages DNA and essential proteins [26, 27]. Stable anaerobic resistance, including in clinically relapsing feline isolates, develops through progressive loss of PFOR activity and decreased ferredoxin expression, which collectively abolish drug activation; in *T. vaginalis* an alternative thioredoxin reductase-mediated activation route also requires inactivation before full resistance is acquired [28, 29]. The reliance of current chemotherapy on a single hydrogenosomal activation node therefore creates a clear therapeutic vulnerability and motivates the search for drug targets that do not depend on this pathway.

The combination of underdiagnosis, chronic morbidity, growing clinical resistance and the absence of any licensed therapy underscores the need for novel anti-trichomonal targets in *T. foetus*. Computational, comparative-genomics approaches to drug-target prioritisation, in which a pathogen proteome is screened against the host proteome, validated drug-target collections, essential-gene resources of model eukaryotes and curated virulence-factor databases, have been widely adopted in antiparasitic drug discovery and are particularly well suited to neglected veterinary pathogens for which experimental screening is unlikely to be pursued. The first such in silico analysis for *T. foetus* was published by Morin-Adeline and colleagues, who assembled *de novo* transcriptomes from bovine and feline isolates and proposed 113 candidate targets based on homology to ChEMBL. Since then, the genomic and pharmacological landscape has changed substantially: a curated reference proteome of strain K is now available through Ensembl Protists, a chromosome-scale assembly of strain KV-1 has been released by the Friedrich-Loeffler-Institut, and the pharmacological databases and essential-gene resources routinely used for target prioritisation have expanded markedly. Despite these advances, no comprehensive in silico drug-target prioritisation for *T. foetus* has been undertaken since the original 2014 study.

The aim of the present study was therefore to perform an updated and comprehensive in silico identification and prioritisation of drug-target candidates in *T. foetus*. Starting from the available genomic resources for strain K, we applied an integrative comparative pipeline combining exclusion of host homologues, screening against an expanded drug-target reference, prioritisation by orthology to essential genes in three evolutionarily distant model eukaryotes and integration with PHI-base host–pathogen interaction data. Priority candidates were then cross-validated against the integrated three-isolate transcriptome of Alonso and colleagues and against quantitative gene-level expression in the feline isolate G10/1, functionally annotated with InterProScan, and subjected to AlphaFold-based molecular docking of representative candidates against clinically relevant inhibitors. To the best of our knowledge, this work represents the most comprehensive in silico drug-target prioritisation for *T. foetus* reported to date.

Methods

Reference proteomes and ancillary databases

The complete predicted proteome of *T. foetus* strain K (assembly ASM183968v1, NCBI accession GCA_001839685.1; 25030 protein sequences) was retrieved from Ensembl Protists (Ensembl Genomes release 62). The reference proteome of the parasite host *F. catus* (assembly *F. catus*_Fca126_mat1.0, NCBI accession GCF_018350175.1; 71463 protein sequences including all annotated isoforms) was retrieved from NCBI RefSeq in April 2026. Three ancillary protein collections were assembled for the homology-based filtering steps. The first was a unified drug-target reference comprising 17,460 non-redundant sequences from ChEMBL release 36 [31], the IUPHAR/BPS Guide to Pharmacology release 2026.1 [32], and SwissProt entries cross-referenced to DrugBank [33] in UniProtKB release 2026_01, with redundancy removed by collapsing entries sharing identical UniProtKB accessions. The second was an essential-gene reference of 4876 sequences combining experimentally validated essential genes from three evolutionarily distinct model eukaryotes: *S. cerevisiae* (1376 sequences from the Saccharomyces Genome Database phenotype dataset, classified as “inviable” or null-deletion lethal), *S. pombe* (1446 sequences from PomBase, FYPO:0002061), and *P. falciparum* (2054 sequences from PlasmoDB, defined as genes with a mutagenesis index score ≤ 0.2 in the piggyBac saturating mutagenesis dataset of Zhang et al. [34]). Each sequence header was prefixed with its source organism to permit per-organism tallying during downstream filtering. The third was the PHI-base virulence reference comprising 9807 sequences retrieved from the PHI-base GitHub master snapshot in April 2026 [35]. *P. falciparum* was included pragmatically rather than for phylogenetic proximity: to our knowledge, no genome-wide, experimentally derived essentiality dataset (e.g. systematic knockout or knockdown screen) is currently available for *T. vaginalis* or other closely related anaerobic parabasalids, whereas PlasmoDB provides a genome-wide piggyBac transposon-insertion essentiality

screen with direct experimental essentiality calls. We revisit the resulting phylogenetic distance as a limitation in the Discussion.

Subtractive proteomics pipeline

All-versus-all pairwise sequence comparisons were performed with DIAMOND v2.1.8 in blastp mode [36], using a BLOSUM62 substitution matrix and an E-value cut-off of 1×10^{-5} . Four sequential filters were applied to the *T. foetus* strain K proteome. First, pathogen proteins sharing $\geq 70\%$ amino-acid identity over $\geq 50\%$ of the query length with any *F. catus* protein were removed, following established practice in subtractive proteomics for parasitic protozoa [37]. The remaining proteins were subsequently searched against the unified drug-target reference, and matches with $\geq 40\%$ identity and $\geq 50\%$ query coverage were retained as putatively druggable. Druggable candidates were further screened against the essential-gene reference, and proteins with homologues ($\geq 30\%$ identity, $\geq 50\%$ query coverage) in all three reference organisms (*S. cerevisiae*, *S. pombe*, *P. falciparum*) were classified as having strong essentiality support, following the three-organism criterion of Doyle and colleagues [49]. In parallel, the same druggable set was searched against PHI-base, and proteins with significant similarity to virulence factors ($\geq 30\%$ identity, $\geq 50\%$ query coverage) were flagged as candidate effectors of host colonisation. The intersection of druggable, essential, and PHI-base-related proteins defined the priority candidate set used for all downstream analyses. The 70% host-exclusion and 30% essentiality/PHI-base thresholds follow conventions established in previous subtractive-genomics pipelines for neglected pathogens [37, 49], and the 40% druggability threshold was set conservatively above the essentiality/PHI-base thresholds to require closer homology to a validated drug target than to a generic essential or virulence-associated gene. We did not perform a dedicated sensitivity analysis varying each threshold independently; this is noted as a limitation in the Discussion. We also note that the PHI-base intersection is intended as an additional plausibility-weighting step, on the reasoning that a protein already documented as relevant to pathogen\2013host interaction or virulence in a related organism is more likely to have a phenotypic consequence when targeted, rather than as an independent confirmation of essentiality; PHI-base and the essential-gene reference are not mutually exclusive categories, and their overlap may partly reflect shared representation of core cellular-maintenance genes.

De novo annotation of strain KV-1

The chromosome-scale assembly of *T. foetus* strain KV-1 was annotated with BRAKER3 v3.0.8 [38] in protein-and-RNA-Seq mode. The OrthoDB v11 protists partition was supplied as the external protein evidence set, and paired-end RNA-Seq from a feline *T. foetus* isolate (SRR1688628, see "Expression validation in a feline isolate" section) was provided as transcript evidence. GeneMark-ETP and AUGUSTUS were used as integrated ab initio predictors within the BRAKER3 workflow. Predicted proteins were extracted from the resulting GFF3 file with gffread v0.12.7. The same four-step subtractive pipeline ("Subtractive proteomics pipeline" section) was then applied independently to the KV-1 protein set. BUSCO assessment (eukaryota_odb10, protein mode) returned near-identical profiles for the KV-1 and strain K proteomes (KV-1: C 12.9%, D 4.3%; strain K: C 11.8%, D 3.9%); the comparable, low duplication indicates that the higher KV-1 gene count does not reflect redundant gene models, and the uniformly low completeness reflects the poor fit of the eukaryotic marker set to Parabasalia.

Cross-strain reciprocal best hits

To assess the conservation of priority candidates between strain K (reference annotation) and the chromosome-scale strain KV-1 assembly, reciprocal best hits (RBH) were computed with DIAMOND blastp in both directions (K $\rightarrow$ KV-1 and KV-1 $\rightarrow$ K) using identical parameters (E-value $\leq 1 \times 10^{-5}$, $\geq 30\%$ identity, $\geq 50\%$ query coverage). A protein pair was retained as an RBH only when each member was the top-scoring hit of the other.

Expression validation in a feline isolate

Gene-level expression of priority candidates was quantified using a previously published RNA-Seq dataset from feline isolate G10/1 (NCBI Sequence Read Archive accession SRR1688628; Morin-Adeline et al. 2014). Paired-end reads (2×100 nt, $\sim 32 \times 10^6$ read pairs) were aligned to the strain K reference genome with HISAT2 v2.2.2 [39] using the default spliced-alignment parameters and the `--no-mixed --no-discordant` flags. The reference FASTA was pre-processed to remove whitespace from sequence headers prior to index construction. Read counts per gene were obtained with featureCounts v2.0.6 from the Subread package [40], operating on the strain K GTF annotation with the following parameters: paired-end mode, fragment counting (`-p --countReadPairs`), exon-level features (`-t exon`), gene-level summarisation (`-g gene_id`), unstranded mode (`-s 0`), and the requirement that both mates of a fragment be aligned to the same chromosome (`-B -C`). Effective gene lengths were computed as the sum of non-overlapping exon spans per gene, after merging overlapping intervals. Transcripts Per Million (TPM) values were computed and a candidate was considered expressed when $\text{TPM} > 1$.

Functional annotation

Priority candidates were annotated with InterProScan v5.59-91.0 [41] in the Galaxy Europe instance, querying the Pfam-A v36 member database. Gene Ontology terms, InterPro family identifiers, and pathway annotations (Reactome, MetaCyc) were retrieved via the `--goterms`, `--iprlookup`, and `--pathways` options. Pfam family distribution was assessed by tallying the number of priority candidates per Pfam accession.

Structural validation by molecular docking

A proof-of-concept structural screen was performed on three top-ranked priority candidates. Targets were selected on the basis of expression rank in the feline isolate, availability of an AlphaFold model, and the existence of one or more clinically or pre-clinically relevant inhibitors that could serve as positive controls. AlphaFold v2 monomer predictions [42] were downloaded from the AlphaFold Protein Structure Database for *T. foetus* Hsp90-2 (UniProtKB A0A1J4KTM2), a CIPK8-like serine/threonine kinase (A0A1J4JU77), and the CDC48/p97 ATPase (A0A1J4K6T2); mean per-residue confidence (pLDDT) ranged from 74 to 88. Ramachandran analysis (MolProbity criteria) placed 91–96% of residues in favoured regions across the three models, and the few outliers (1–2%) were confined to low-confidence disordered loops and termini rather than the folded binding pockets. Candidate binding pockets were predicted with fpocket v4.0 [43] and the pocket with the highest druggability score was selected for docking in each case. Receptors and grid boxes (22–24 Å edges, centered on pocket centroids) were prepared for AutoDock with Meeko v0.6.1. Clinically or pre-clinically relevant ligands and a natural-ligand control (geldanamycin, tanespimycin, radicicol, staurosporine, GSK650394, bortezomib, CB-5083, NMS-873, and ATP) were retrieved as canonical SMILES from PubChem (CIDs in Additional file 2: Table S2). Three-dimensional conformers were generated with RDKit v2024.03 using the ETKDGv3 embedding algorithm followed by MMFF94 energy minimisation, and converted to PDBQT format with Meeko. The boronic-acid boron atoms of bortezomib were substituted with sp^3 carbon prior to docking, since the AutoDock Vina scoring function does not parameterise boron; this substitution preserves relative ranking across bortezomib dockings [45]. Docking was performed with AutoDock Vina v1.2.5 [44, 45] using `exhaustiveness=4` and `num_modes=3`. For each receptor–ligand combination the lowest-energy pose was retained and the corresponding binding affinity (kcal mol^{-1}) was recorded as the docking score. Pharmacologically validated inhibitors served as positive controls (geldanamycin for Hsp90-2, staurosporine for CIPK8, CB-5083 for CDC48), and ATP served as a natural-ligand binding control for all three receptors. Repeating the docking at exhaustiveness 8 and 16 (the AutoDock Vina default is 8) left the binding affinities (within approximately 0.4 kcal/mol) and the relative ranking of the three targets unchanged, confirming robustness to the search-exhaustiveness parameter.

Software environment

Custom analyses (sequence parsing, TPM computation, RBH analysis) were implemented in Python 3.12 using Biopython v1.83 and pandas v2.2.1. Pipeline orchestration was performed on the Galaxy Europe platform and on a local Linux workstation.

Results

Identification of 433 priority drug-target candidates in *T. foetus* strain K

The complete predicted proteome of *T. foetus* strain K (25030 protein sequences, median length 392 amino acids; Ensembl Protists release 62) was used as the starting point for the four-step subtractive proteomics pipeline (Fig. 1). Functional descriptions were available for 15,909 of 25,030 entries (63.6%); the remainder corresponded to uncharacterised proteins. Of the initial proteome, 24,942 sequences (99.6%) survived the host-homology filter against the *F. catus* reference proteome at the $\geq 70\%$ identity threshold. Screening of the host-filtered proteome against the unified drug-target reference ($\geq 40\%$ identity, $\geq 50\%$ query coverage) returned 1444 proteins classified as putatively druggable.

Essentiality screening of the druggable set against three evolutionarily distinct model eukaryotes ($\geq 30\%$ identity, $\geq 50\%$ query coverage) returned 835 proteins with an essential homologue in *S. cerevisiae*, 884 in *S. pombe*, and 808 in *P. falciparum*. The conservative intersection of all three organisms yielded 524 candidates, designated as having very strong essentiality support. In parallel, screening of the druggable set against the PHI-base virulence reference at the same identity and coverage thresholds returned 2178 *T. foetus* proteins with significant similarity to known virulence factors. The intersection of these two layers defined the final priority set of 433 candidates, hereafter referred to as priority drug-target candidates. The complete annotated list of 433 priority candidates with Pfam domain assignments, GO terms, and TPM values is provided in Additional file 1: Table S1. Of the 524 essential candidates, 91 were excluded solely at the PHI-base intersection step, yielding the 433 priority set; an independent reconstruction of the pipeline reproduced this reduction exactly.

Numbers in each box correspond to protein sequences retained after the respective filtering step. The pipeline starts from the predicted proteome of *T. foetus* strain K (Ensembl Protists release 62) and applies four sequential filters. The drug-target reference combines ChEMBL release 36, IUPHAR/BPS Guide to Pharmacology, and DrugBank-derived SwissProt entries. The essential-gene reference combines validated essential genes from *S. cerevisiae* (SGD), *S. pombe* (PomBase) and *P. falciparum* (PlasmoDB). PHI-base v4 serves as the virulence reference. The priority candidate set ($n=433$) represents the intersection of the druggable, essential and PHI-base subsets and is the focus of all downstream characterisation analyses (cross-strain RBH, InterProScan annotation, expression validation in the feline isolate G10/1, and molecular docking). Filtering thresholds were $\geq 70\%$ identity for the host-homology filter, $\geq 40\%$ identity with $\geq 50\%$ query coverage for the drug-target screen, and $\geq 30\%$ identity with $\geq 50\%$ query coverage for the essentiality and PHI-base filters.

Numbers refer to protein sequences retained after each filtering step. The druggable subset comprises proteins of the host-filtered proteome with homology to validated drug targets curated in the unified ChEMBL+IUPHAR/BPS+DrugBank reference. The essentiality subset required orthology to experimentally validated essential genes in all three model eukaryotes (*S. cerevisiae*, *S. pombe* and *P. falciparum*) and is therefore designated 3 of 3. The PHI-base subset comprises proteins with homology to curated virulence-associated factors. The priority candidate set (final row, in bold) represents the intersection of the druggable, essential and PHI-base subsets and constitutes the focus of all downstream analyses. Percentages refer to the input proteome of 25,030 sequences.

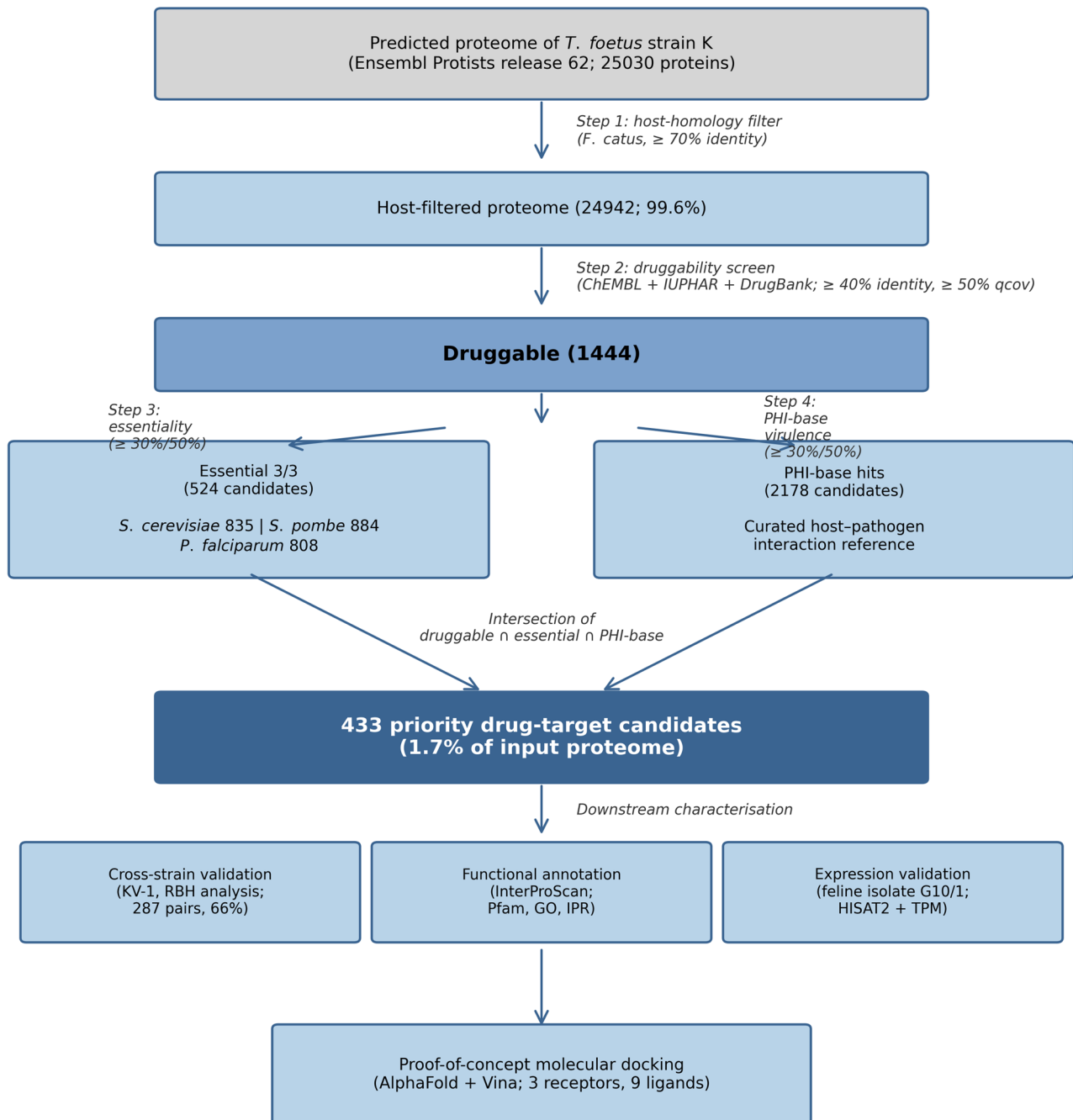


Fig. 1 Subtractive proteomics workflow for *T. foetus* drug-target prioritisation

De novo annotation of strain KV-1

The analysis was repeated on the chromosome-scale assembly of strain KV-1 (GCA_041260205.1; 148.7 Mb across five chromosomes), for which gene annotation was not publicly available. BRAKER3 v3.0.8 was run in protein-and-RNA-Seq mode using the OrthoDB v11 protists partition as protein evidence and feline-isolate RNA-Seq SRR1688628 as transcript evidence, yielding 43,965 predicted protein-coding genes. This figure exceeds the curated Ensembl annotation of strain K (25030 proteins) by approximately 1.8-fold. Application of the same

four-step subtractive pipeline to the KV-1 predicted proteome, under identical parameters and thresholds as those applied to strain K, produced 436 priority candidates.

Cross-strain reciprocal best hits

Reciprocal best hits (RBH) were computed between the strain K and KV-1 priority candidate sets to assess the extent of cross-strain conservation (Fig. 2). A total of 287 RBH pairs were recovered, corresponding to 65.8% of the strain KV-1 priority candidates and 66.3% of the strain K candidates. No priority candidate identified in KV-1 was genuinely novel with respect to strain K. Every KV-1 candidate without an RBH could be traced to either a KV-1-specific gene fragmentation event or to a closely related paralogue in strain K. The curated strain K annotation was retained as the primary reference for all downstream analyses; the KV-1 assembly served as a chromosome-scale cross-validation. The full list of 287 RBH pairs is provided in Additional file 3: Table S3.

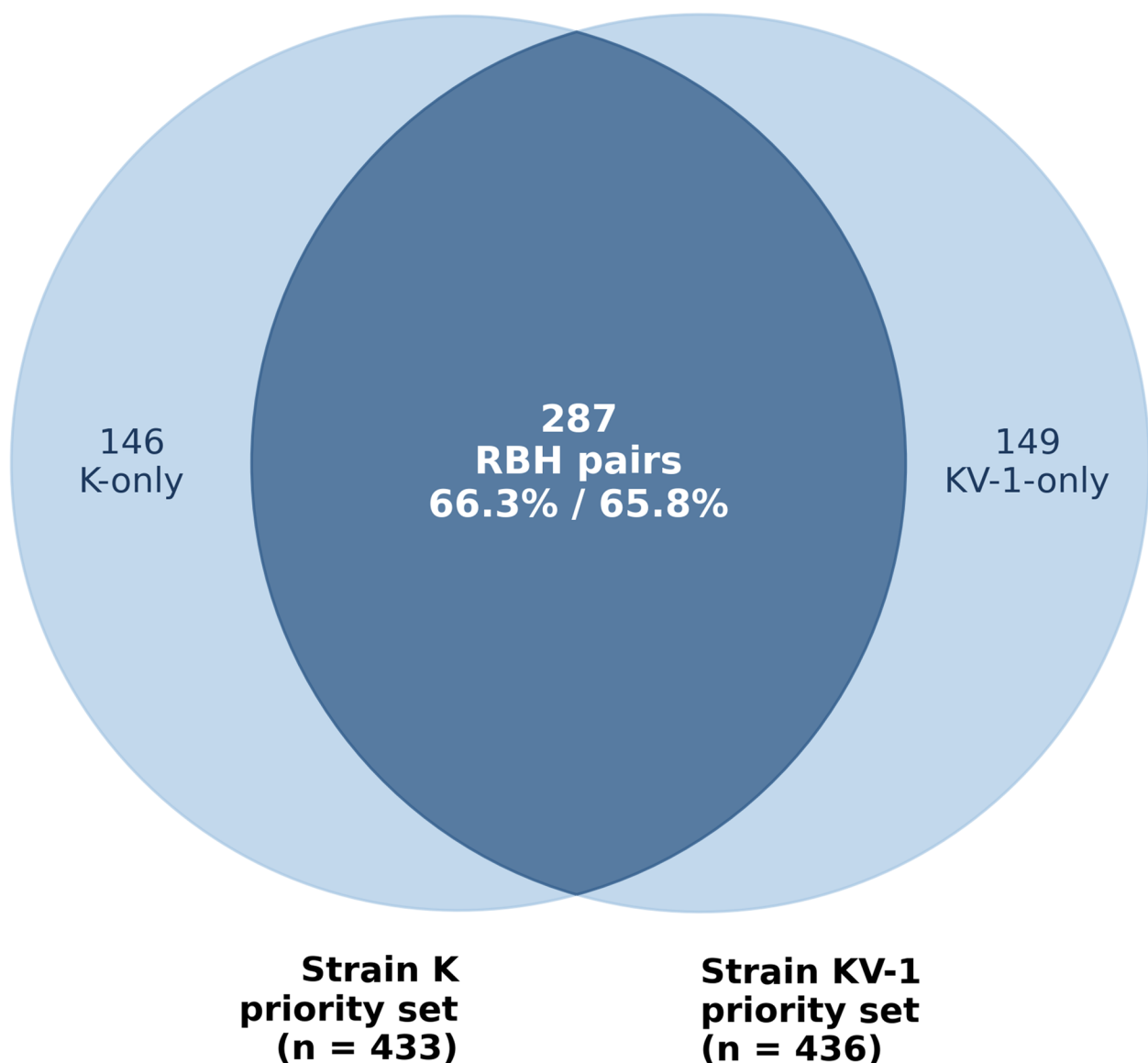


Fig. 2 Cross-strain conservation of priority candidates between *T. foetus* strain K and strain KV-1

Reciprocal best hits were computed with DIAMOND blastp in both directions ($K \rightarrow KV-1$ and $KV-1 \rightarrow K$) using identical parameters ($E\text{-value} \leq 1 \times 10^{-5}$, $\geq 30\%$ identity, $\geq 50\%$ query coverage). The Venn diagram shows the size of each priority candidate set (433 in strain K, 436 in strain KV-1) and the intersection of 287 RBH pairs, corresponding to 66.3% of strain K candidates and 65.8% of strain KV-1 candidates. Strain-specific candidates (146 in K, 149 in KV-1) correspond either to KV-1 ab initio gene fragmentation events or to closely related paralogues in the alternative annotation. The full list of 287 RBH pairs with strain K annotations and TPM values is provided in Additional file 3: Table S3.

Functional annotation of priority candidates

InterProScan annotation of the 433 priority candidates against the Pfam-A v36 database returned at least one Pfam domain assignment for every candidate (433/433; 100%), with Gene Ontology terms available for 428 (98.8%) and InterPro family identifiers for all 433. A total of 1688 Pfam hits were recorded, and 96 of the 433 candidates (22%) carried two or more annotated domains. The five most frequent Pfam accessions corresponded to the protein kinase domain (PF00069; $n=123$ candidates), the Ras family of small GTPases (PF00071; $n=121$), the calcineurin-like phosphoesterase superfamily (PF00149; $n=109$), and the DEAD/DEAH box helicase domains (PF00270 and PF00271; $n=18$ each), followed by the Hsp70 family (PF00012; $n=9$). Protein kinases, Ras-family GTPases, and phosphoesterases, three biologically related categories central to eukaryotic signal transduction, together accounted for 352 of the 433 priority candidates (81.3%) (Fig. 3). Over-representation of these families among the priority candidates was statistically significant by Fisher's exact test against the strain K proteome (Benjamini-Hochberg corrected): Ras-family GTPases (odds ratio 28.9, $q=8e-117$), protein kinases

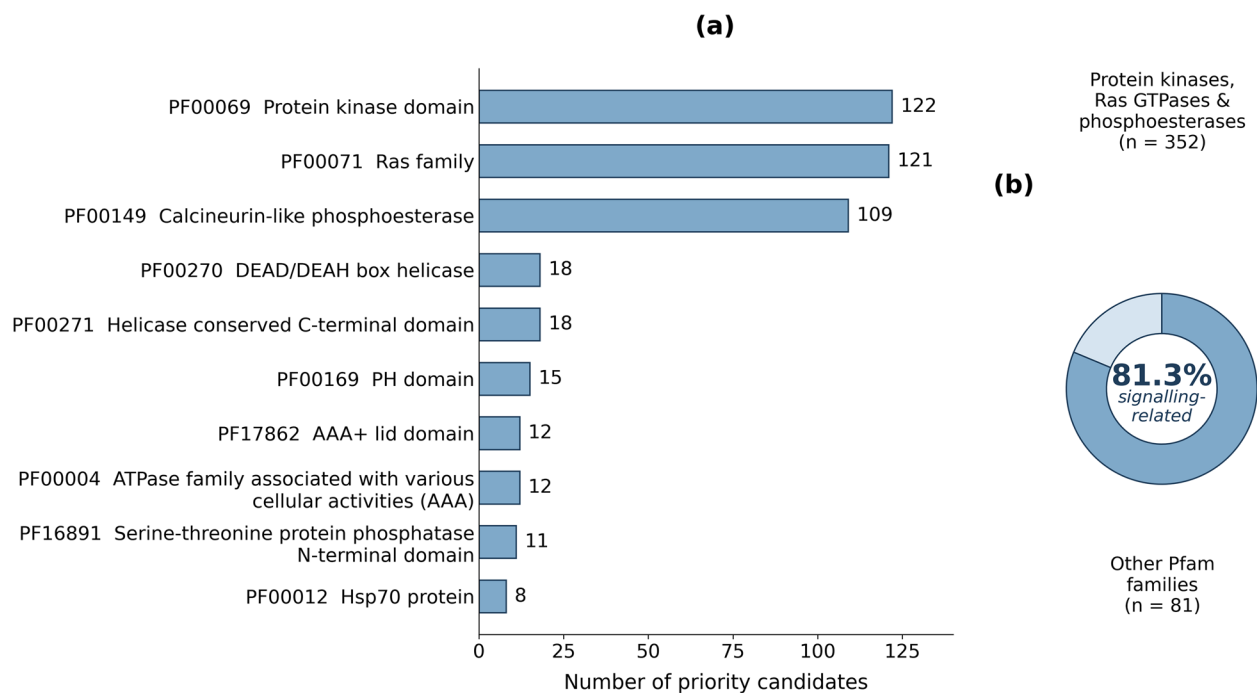


Fig. 3 Pfam family distribution among the 433 priority drug-target candidates. **(a)** Horizontal bar chart of the ten most frequent Pfam-A accessions among the priority candidate set, ranked by the number of candidates carrying at least one domain instance. Pfam family assignments were obtained with InterProScan v5.59-91.0 querying Pfam-A v36. **(b)** Donut chart showing the combined representation of protein kinases (PF00069), Ras-family small GTPases (PF00071), and calcineurin-like phosphoesterases (PF00149); these three biologically related signal-transduction categories together account for 352 of 433 unique priority candidates (81.3%). The remaining 81 candidates carry diverse Pfam families including helicases, AAA+ ATPases and chaperones.

(OR 5.7, $q=9e-43$), Hsp70 (OR 10.5), AAA+ ATPases (OR 8.4) and Hsp90 (OR 171.6), whereas ubiquitin was not enriched.

Gene-level expression of priority candidates in the feline isolate

Gene-level expression of the 433 priority candidates was quantified in the feline *T. foetus* isolate G10/1 using the publicly available RNA-Seq dataset SRR1688628 (Morin-Adeline et al. 2014). Paired-end reads were aligned to the strain K reference genome with an overall alignment rate of 91.18%. Of the 32×10^6 read pairs, 72.4% were assigned to annotated genes by featureCounts; the remainder consisted of multi-mapping reads (8.5%), reads aligned outside annotated features (10.9%), and unmapped or low-quality pairs. The median TPM across the strain K transcriptome was 2.20, with 15,849 genes (62.6%) reaching TPM > 1.

Within the priority candidate set, 358 of 433 candidates (82.7%) showed TPM > 1 in the feline isolate, 278 (64.2%) exceeded TPM 5, 231 (53.3%) exceeded TPM 10, and 61 (14.1%) exceeded TPM 100 (Fig. 4). The ten most highly expressed priority candidates comprised canonical chaperones, including Hsp70 (OHS93735; TPM 4289), two Hsp90-2 paralogues (OHS98422 and OHT14649; TPM 3128 and 2231), the hydrogenosomal Hsp60 (OHT00431; TPM 782), and the endoplasmic-reticulum BiP/GRP78 chaperone (OHS92902 and OHT11192; TPM 1155 and 999); together with the AAA+ ATPase CDC48/p97 (OHT07071; TPM 1971), the ubiquitin-S27a fusion (OHT13114; TPM 1835), a CBL-interacting serine/threonine kinase (OHT02689; TPM 776), and a Ras-related Rab-4B GTPase (OHT15241; TPM 759). Strikingly, nine of these ten candidates correspond to a protein family with an extensively characterised, clinically advanced small-molecule inhibitor; the exception is the Rab4B GTPase, for which no Rab4B-specific clinically advanced inhibitor currently exists \u2014 its inclusion here reflects membership of the broader, pharmacologically tractable Ras/GTPase superfamily rather than a Rab4B-specific precedent. This near-complete overlap still defines a translational shortlist for downstream drug-repurposing studies.

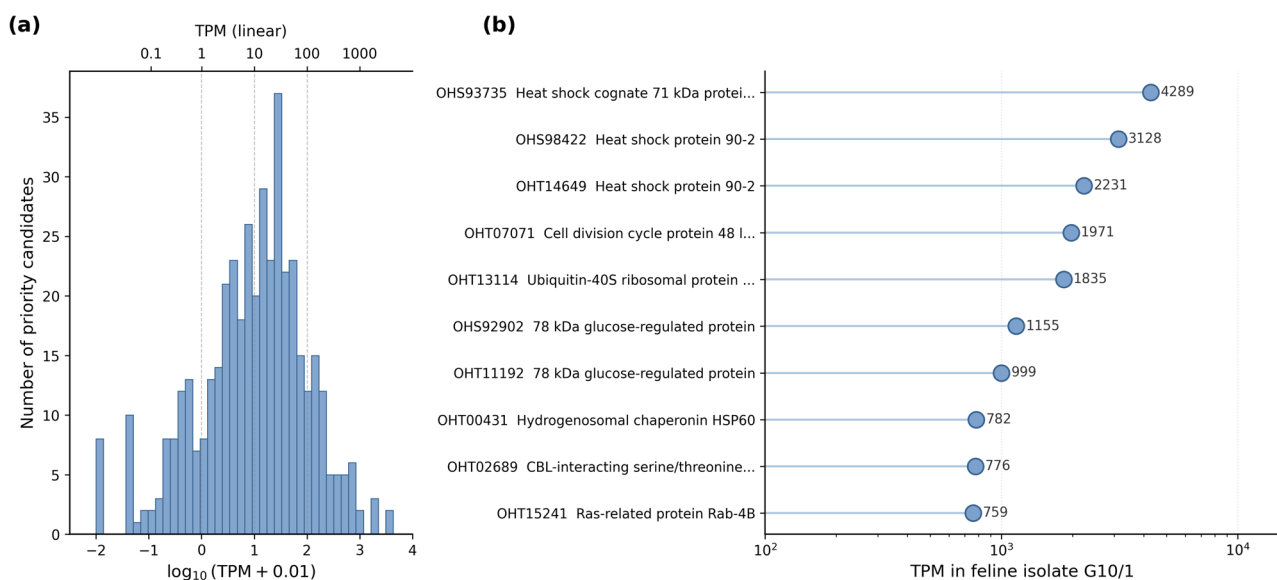


Fig. 4 Gene-level expression of priority candidates in the feline *T. foetus* isolate G10/1. Gene-level expression was quantified by HISAT2 alignment of RNA-Seq dataset SRR1688628 [30] followed by featureCounts and TPM normalisation against the strain K reference genome. **(a)** Histogram of $\log_{10}(\text{TPM} + 0.01)$ values for the 433 priority candidates; vertical dashed lines indicate TPM thresholds of 1, 10 and 100, and the secondary top axis shows TPM values on a linear scale. Of the 433 priority candidates, 358 (82.7%) were expressed above TPM 1, 278 (64.2%) above TPM 5, 231 (53.3%) above TPM 10, and 61 (14.1%) above TPM 100. **(b)** Ranked dot plot of the ten most highly expressed priority candidates, labelled with their UniProt accession, protein description, and TPM value in the feline isolate

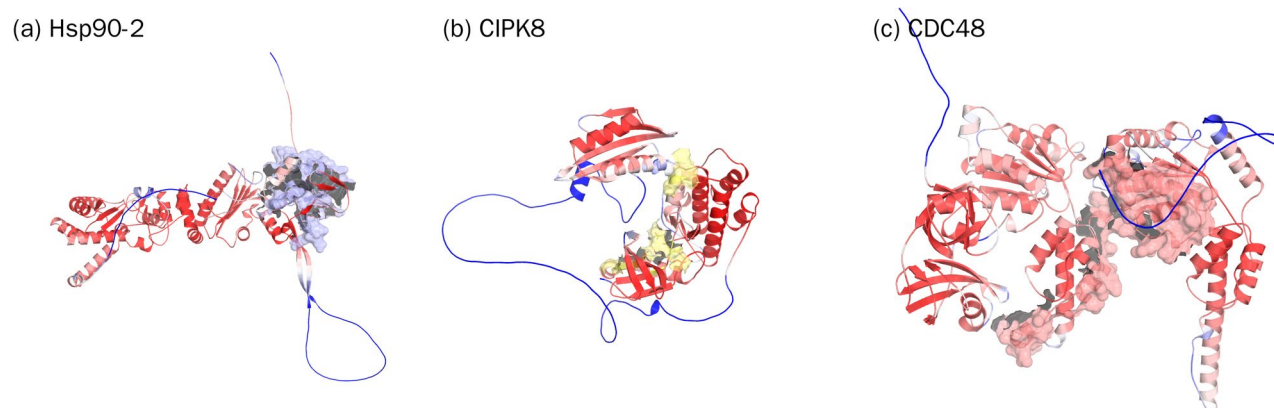


Fig. 5 AlphaFold structural models and predicted binding pockets of three top-ranked priority candidates. AlphaFold v2 monomer predictions were retrieved from the AlphaFold Protein Structure Database for *T. foetus* Hsp90-2 (A0A1J4KTM2; mean pLDDT 84.8), the CIPK8-like serine/threonine kinase (A0A1J4JU77; mean pLDDT 74.1), and the CDC48/p97 AAA+ ATPase (A0A1J4K6T2; mean pLDDT 80.4). Cartoon backbones are coloured by per-residue pLDDT confidence (blue, low; red, high). The highest-druggability binding pocket predicted by fpocket v4.0 is shown as a transparent surface for each receptor. In panel (b), CIPK8 catalytic residues K37, D87 (HRD motif) and D150–F151 (DFG motif) are shown as sticks. These predicted pockets were used as the target sites for molecular docking (Fig. 6)

Table 1 AutoDock Vina docking affinities for representative priority candidates against clinically relevant inhibitors

Receptor	UniProt	Ligand	Class	Affinity (kcal mol ⁻¹)
Hsp90	A0A1J4KTM2	bortezomib	negative control	-6.10
Hsp90	A0A1J4KTM2	ATP	natural ligand	-5.75
Hsp90	A0A1J4KTM2	geldanamycin	positive control	-5.27
Hsp90	A0A1J4KTM2	17-AAG	analogue	-4.04
Hsp90	A0A1J4KTM2	radicicol	analogue	-4.03
CIPK8	A0A1J4JU77	staurosporine	positive control	-10.37
CIPK8	A0A1J4JU77	GSK650394	selective inhibitor	-7.42
CIPK8	A0A1J4JU77	ATP	natural ligand	-7.11
CDC48	A0A1J4K6T2	CB-5083	positive control	-7.63
CDC48	A0A1J4K6T2	bortezomib	negative control	-7.46
CDC48	A0A1J4K6T2	NMS-873	selective inhibitor	-6.89
CDC48	A0A1J4K6T2	ATP	natural ligand	-6.71

Docking was performed with AutoDock Vina v1.2.5 using exhaustiveness=4 and num_modes=3. For each receptor–ligand combination the lowest-energy pose is reported. Receptors are *T. foetus* Hsp90-2 (UniProt A0A1J4KTM2), a CIPK8-like serine/threonine kinase (A0A1J4JU77) and the CDC48/p97 AAA+ ATPase (A0A1J4K6T2). Ligand structures were retrieved from PubChem (CIDs and SMILES in Additional file 2: Table S2). Pharmacologically validated inhibitors served as positive controls (geldanamycin for Hsp90-2, staurosporine for CIPK8, CB-5083 for CDC48); ATP served as a natural-ligand binding control for all three receptors. Affinity values are sorted from strongest (most negative) to weakest within each receptor block

Molecular docking of representative priority candidates

Three top-ranked priority candidates were subjected to a proof-of-concept docking screen against pharmacologically validated inhibitors (Fig. 5; Table 1): Hsp90-2 (OHT14649), the CIPK8-like serine/threonine kinase (OHT02689), and the CDC48/p97 AAA+ ATPase (OHT07071). Candidate binding pockets predicted on each AlphaFold model yielded druggability scores of 0.65 for Hsp90-2 (an inter-domain ATP-binding region), 0.51 for CIPK8 (a canonical kinase ATP pocket lined by the K37 catalytic lysine, the D87 HRD motif, and the D150–F151 DFG motif), and 0.99 for CDC48 (the D1/D2 inter-domain region).

Docking against pharmacologically relevant ligands recovered a coherent affinity hierarchy across all three receptors (Fig. 6). The CIPK8 kinase displayed strong binding for the broad-spectrum kinase inhibitor

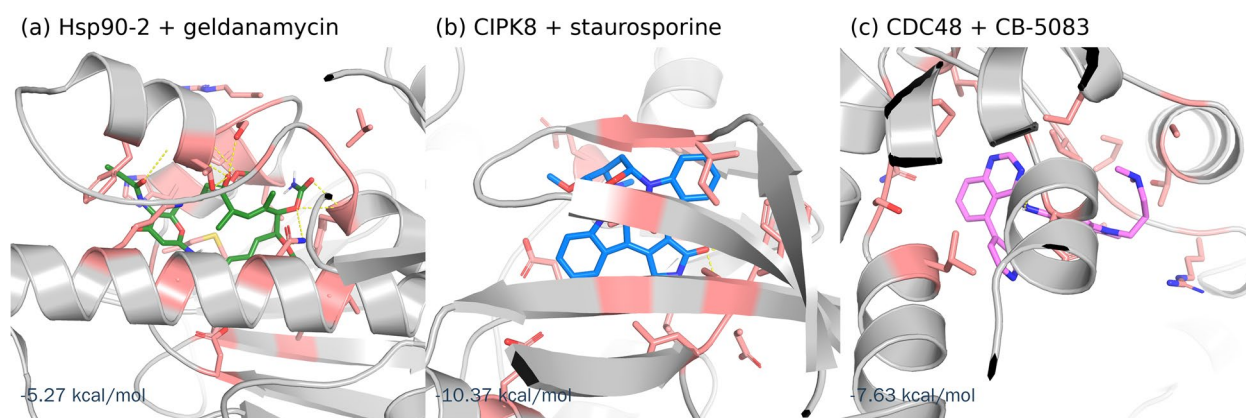


Fig. 6 Molecular docking poses of clinically validated inhibitors against the three top-ranked priority candidates. Lowest-energy docked poses from AutoDock Vina (exhaustiveness=4, num_modes=3) for the strongest pharmacologically meaningful ligand of each receptor. **a** Hsp90-2 with geldanamycin (positive control, -5.27 kcal mol $^{-1}$). **b** CIPK8-like kinase with staurosporine (positive control, -10.37 kcal mol $^{-1}$). **c** CDC48/p97 with CB-5083 (positive control, -7.63 kcal mol $^{-1}$). Receptors are shown as cartoon (grey); ligands are shown as sticks coloured by element (carbon: panel-specific colour); pocket residues within 4 Å of the ligand are shown as thin sticks; predicted hydrogen bonds are shown as dashed lines. For each receptor, the docking box was centred on the highest-druggability pocket predicted by fpocket and shown for the same receptor in Fig. 5

staurosporine (-10.37 kcal mol $^{-1}$), substantially exceeding affinities recovered for the Sgk-selective inhibitor GSK650394 (-7.42 kcal mol $^{-1}$) and for ATP (-7.11 kcal mol $^{-1}$). CDC48 docked favourably with the clinical AAA+ ATPase inhibitor CB-5083 (-7.63 kcal mol $^{-1}$), followed by the allosteric VCP inhibitor NMS-873 (-6.89 kcal mol $^{-1}$) and ATP (-6.71 kcal mol $^{-1}$). Docking of Hsp90-2 against geldanamycin, tanespimycin, and radicicol recovered weaker affinities ranging from -4.0 to -5.3 kcal mol $^{-1}$.

Discussion

The present study reports the first comprehensive in silico drug-target prioritisation for *Tritrichomonas foetus* since the work of Morin-Adeline and colleagues a decade ago [30]. Starting from the predicted proteome of strain K, we identified 433 priority drug-target candidates through sequential subtractive filtering against the feline host proteome, three pharmacological reference databases, essential-gene resources from three model eukaryotes and the PHI-base virulence reference, representing almost a four-fold increase over the 113 candidates proposed in 2014. Application of the same pipeline to a *de novo* annotation of the chromosome-scale strain KV-1 assembly produced 436 candidates, and 66% of candidates from each strain were recovered as reciprocal best hits in the other. Gene-level expression analysis showed that 82.7% of priority candidates are actively transcribed in the feline isolate G10/1, and a proof-of-concept docking screen against three top-ranked candidates recovered biologically meaningful affinity hierarchies for pharmacologically validated inhibitors. These findings provide a rational starting point for downstream experimental validation of antiprotozoal drug targets in *T. foetus*.

A direct comparison with the work of Morin-Adeline and colleagues [30], which remains the only prior in silico drug-target study for *T. foetus*, highlights both the continuity and the substantial expansion of the present analysis. The 2014 study assembled *de novo* transcriptomes from the bovine BP-4 and feline G10/1 isolates, recovered 7,547 orthologous transcripts, and proposed 113 candidate targets on the basis of homology to ChEMBL alone. The present study extends this earlier framework in four respects. First, it uses the full predicted proteome of strain K (25,030 proteins), which is approximately three times larger than the 2014 transcriptome and includes proteins not necessarily expressed under the in vitro culture conditions sampled by Morin-Adeline and colleagues. Second, it incorporates two pharmacological references additional to ChEMBL, namely

the IUPHAR/BPS Guide to Pharmacology and DrugBank-derived SwissProt entries, which together provide a more complete view of the validated drug-target space. Third, it explicitly prioritises candidates by orthology to experimentally validated essential genes from three model eukaryotes, an approach that was infeasible in 2014 because the underlying resources (notably the saturating mutagenesis dataset of *Plasmodium falciparum* [34]) had not yet been generated. Fourth, it cross-validates candidates against quantitative gene-level expression in the feline isolate context, a step that is absent from the 2014 analysis. The integrative three-isolate transcriptome of Alonso and colleagues [46], which expanded the transcriptional resource by adding the porcine PIG30/1 isolate, contributes complementary host-specific information but does not, on its own, perform target prioritisation.

Set against the broader in silico drug-target literature for protozoan parasites, the present work occupies a methodological niche between large-scale shared-target studies and species-specific screens. Galindo-Hernández and colleagues [47] applied an ontology-based orthology approach to five protozoan proteomes (*Entamoeba histolytica*, *Giardia lamblia*, *Trichomonas vaginalis*, *Trypanosoma cruzi*, and *Leishmania mexicana*) and identified twelve drug targets shared across all five species. Their pan-protozoan shortlist is intentionally restrictive and emphasises broad-spectrum targets; the present study, in contrast, recovers a *T. foetus*-focused set that includes both pan-eukaryotic essentials (kinases, GTPases, chaperones) and species-relevant signatures derived from PHI-base homology. The 40-candidate antimalarial shortlist of Sturm and colleagues [48], generated through a comparable subtractive pipeline applied to *P. falciparum*, is closer in spirit to the present analysis but operates at a different scale of prioritisation: their pipeline uses a stricter 15% global identity threshold to enrich for essential proteins and yields a smaller final set, whereas our four-criterion intersection yields a larger but more heterogeneous candidate space. The identity distribution we observed between *T. foetus* candidates and their best drug-target hits had a median of 35.5%, reflecting the substantial phylogenetic distance separating Parabasalia from the model organisms most heavily represented in pharmacological databases. Application of the same 70% identity threshold used at the host-exclusion step would have retained only two candidates (ubiquitin and the 26 S proteasome subunit) and would have eliminated essentially all biological signal. This is consistent with the more stringent identity cut-offs typically applied in studies of more divergent eukaryotic pathogens. The three-organism essentiality intersection [49] provides an approximately five-fold enrichment for true essentiality predictions relative to single-organism calls and is particularly important for Parabasalia, which possess a divergent translation apparatus and several non-canonical genome features that make orthology-based prioritisation more challenging than in bacterial pathogens.

The cross-strain reciprocal-best-hits analysis between strain K and strain KV-1 indicates that the priority set is robust to the choice of reference assembly. No priority candidate identified in KV-1 was genuinely novel with respect to strain K; every KV-1 candidate without an RBH could be traced to either a KV-1-specific gene fragmentation event in the ab initio annotation or to a closely related paralogue in strain K. This robustness contrasts with the situation reported for *T. vaginalis*, the closest sequenced relative of *T. foetus* within the Parabasalia, where the highly repetitive nature of the genome has historically complicated short-read assembly and annotation [50]. The recent chromosome-level reannotation of a *T. vaginalis* clinical isolate [51] has substantially improved the situation, but inter-strain comparisons of drug-target candidates in *T. vaginalis* are still scarce. The convergence of our two independent annotations of *T. foetus*, namely a curated draft genome (strain K, Ensembl Protists) and a *de novo* annotation of a chromosome-scale assembly (strain KV-1, BRAKER3), therefore provides an important quality-control checkpoint for downstream studies and validates the use of the strain K annotation as the primary reference for the foreseeable future.

The Pfam family distribution of the 433 priority candidates is strongly biased towards protein classes that are considered classically druggable, with protein kinases (PF00069; $n=123$), Ras-family GTPases (PF00071; $n=121$), and calcineurin-like phosphoesterases (PF00149; $n=109$) together accounting for 352 of the 433 priority candidates (81.3%). Although 5-nitroimidazoles activate within *T. foetus* hydrogenosomes and remain the only drug class used in feline tritrichomonosis, this predominance of kinases, GTPases and proteostasis factors points towards therapeutic strategies that complement or move beyond classical nitroimidazole activation pathways. This enrichment is consistent with the broader antiprotozoal drug-discovery literature. In *P. falciparum*,

the parasite kinome of approximately 90 kinases [52] is one of the most extensively investigated target classes in malaria research, with PfPI4K β -targeting compounds reaching clinical trials and several CDPK inhibitors progressing through preclinical evaluation [53, 54]. In *G. lamblia*, bumped kinase inhibitors targeting kinases with small gatekeeper residues have been validated in vitro against both wild-type and metronidazole-resistant strains [55], and Rho-family GTPases have recently been validated as targets for repurposed drugs including dasatinib and imatinib [56]. In *T. vaginalis*, the 20 S proteasome has been validated as a drug target through peptide inhibitor screening, with selected inhibitors showing low-nanomolar potency and up to 200-fold selectivity over mammalian cells [57]. The strong representation of ubiquitin and proteasome-pathway components among our top-expressed priority candidates is consistent with these findings and suggests that proteasome-targeting compounds developed against *T. vaginalis* may serve as a starting point for *T. foetus* drug discovery.

The cross-validation of priority candidates against quantitative gene-level expression in the feline isolate G10/1 confirms that the prioritisation is not an artefact of comparative genomics and identifies the subset of candidates most likely to be relevant to feline tritrichomonosis specifically. The 82.7% rate of actively expressed candidates (TPM > 1) is substantially higher than the 52% rate recovered in our earlier first-pass analysis based on Pfam-family matching against the FPKM-thresholded transcriptome of Alonso and colleagues [46], confirming that proper gene-level alignment with HISAT2 and featureCounts is the appropriate validation strategy for this question.

Notably, nine of the ten most highly expressed priority candidates corresponds to a protein family with an extensively characterised small-molecule inhibitor in the broader drug-discovery literature: Hsp70 and Hsp90 (geldanamycin, tanespimycin and several derivatives in clinical or preclinical evaluation against cancer and parasitic diseases [58, 59]), the AAA+ ATPase CDC48/p97 (CB-5083, advanced through Phase I clinical trials [60]), the ubiquitin-proteasome pathway (bortezomib, a clinically approved oncology drug [57, 61]), and serine/threonine kinases (staurosporine and its many derivatives [62, 63]); the tenth, the Rab4B GTPase, belongs to the broader Ras/GTPase superfamily but has no Rab4B-specific clinically advanced inhibitor characterised to date. Mapping these established pharmacology onto the most actively transcribed *T. foetus* targets defines a drug-repurposing shortlist for downstream experimental evaluation.

The three targets selected for docking occupy central, druggable roles in parasite proteostasis and signalling: Hsp90-2 is an ATP-dependent molecular chaperone that stabilises a broad client repertoire, CDC48/p97 is an essential AAA+ ATPase that extracts ubiquitinated substrates for proteasomal degradation, and the CIPK8-like enzyme is a serine/threonine protein kinase whose specific role in the parasite remains uncharacterised; inhibition of the two ATP-dependent chaperone and segregase functions is expected to compromise proteostasis and parasite viability. The proof-of-concept docking screen provides preliminary structural support for the druggability of the prioritised target set and illustrates both the strengths and the limitations of static structure-based validation. The CIPK8-like serine/threonine kinase recovered a binding affinity of -10.37 kcal mol $^{-1}$ for staurosporine, comparable to literature-reported affinities for staurosporine against canonical eukaryotic kinases [62, 63], and substantially exceeded affinities for the Sgk-selective inhibitor GSK650394 and for ATP itself. CDC48 docked favourably with the clinical AAA+ ATPase inhibitor CB-5083 and showed a coherent affinity hierarchy that is consistent with the cross-species applicability of VCP/p97 pharmacology between *T. foetus* and the human ortholog. In contrast, Hsp90-2 docked weakly with geldanamycin, tanespimycin and radicicol (all affinities in the range of -4.0 to -5.3 kcal mol $^{-1}$), most likely reflecting the limitations of docking a closed apo N-terminal conformation in which the geldanamycin-binding pocket is partially occluded. This single-state limitation is well documented for AlphaFold predictions of chaperones with conformationally dynamic ATP-binding domains [64, 65]. Multi-state structure prediction or molecular-dynamics relaxation would be required to validate Hsp90 as a docking target in *T. foetus*. Structural models were available in the AlphaFold Database for seven of the ten prioritised candidates, and the three docked targets were those additionally possessing a pharmacologically validated inhibitor suitable as a positive control; de novo modelling with AlphaFold 3 for the remaining candidates would enable systematic docking across the full priority set. The inhibitors tested range from an approved drug (bortezomib) and clinically trialled agents (tanespimycin, CB-5083) to pre-clinical tool compounds, so the repurposing

rationale rests on the clinical validation of the target classes rather than of each individual probe, several of which would require medicinal-chemistry optimisation before veterinary application.

None of the nine ligands used in this docking screen has documented safety or pharmacokinetic (ADMET) data in cats. Of particular concern, geldanamycin shows pronounced hepatotoxicity in animal models, which has limited its own clinical development despite strong Hsp90 affinity [71]. Tanespimycin and CB-5083 have undergone clinical safety evaluation in humans but not in cats, and the remaining compounds (radicol, staurosporine, GSK650394, NMS-873) are pre-clinical tool inhibitors with no systematic toxicology profile. Bortezomib is the only agent with an established human safety record, reflecting its regulatory approval for multiple myeloma and mantle cell lymphoma. We have therefore added a clinical/therapeutic status field for each ligand to Additional file 2 (Table S2). We emphasise that the repurposing rationale of this study rests on validating druggable target classes rather than on proposing any of the tested ligands as ready-to-use feline therapeutics: progression toward veterinary application would require feline-specific ADMET profiling and, in most cases, structural optimisation or replacement with safer analogues within the same pharmacological class, for example later-generation, lower-toxicity Hsp90 inhibitors in place of geldanamycin.

An important consideration when interpreting the present results is the asymmetry between genomic and transcriptomic conservation across host-adapted *T. foetus* lineages. Whole-genome sequencing of bovine, porcine and feline *T. foetus* isolates has revealed substantial genomic divergence between the feline lineage and the closely related bovine/porcine clade, consistent with the proposal that feline *T. foetus* may represent a distinct species [69]. In contrast, comparative transcriptomics has demonstrated striking conservation of expressed gene content between bovine and feline isolates [30, 46]. This apparent paradox is resolved by noting that the most pronounced genomic differences fall within non-coding and rapidly evolving regions. The protein-coding repertoire, particularly the deeply conserved drug-target families of chaperones, AAA+ ATPases and protein kinases targeted by our pipeline, remains highly conserved across host-adapted lineages. The chromosome-scale feline *T. foetus* genome required for a fully feline-referenced analysis is not yet publicly available; this represents a general limitation of the field rather than a defect of the present study, and has recently been flagged as a priority for the *T. foetus* research community [70]. Our integrative design addresses this limitation pragmatically by validating priority candidate expression directly in a feline isolate (G10/1; 82.7% transcribed at TPM > 1), bridging the bovine reference to the feline disease context.

The 70% host-exclusion threshold is permissive by design, intended to retain broadly conserved, classically druggable protein families rather than to enforce strict host selectivity at the sequence-identity stage; applying a 70% threshold at the essentiality stage instead would leave only two candidates, illustrating how sensitive the candidate set is to this choice. This trades sensitivity for selectivity: several top-ranked candidates, including the Hsp90-2 and Hsp70 chaperones prioritised here, belong to protein families that are highly conserved with their mammalian/feline orthologs (e.g. Hsp70/Hsp90, ubiquitin-related proteins, protein kinases). For these families, therapeutic selectivity cannot be established at the homology-filtering stage and will instead depend on structure-based exploitation of species-specific binding-pocket features during subsequent medicinal-chemistry work.

Several limitations of the present study should be acknowledged. The essentiality predictions rely on orthology to validated essential genes in three model eukaryotes rather than on direct genetic interrogation in *T. foetus*, which remains poorly tractable to transfection and gene knockout [66]. This orthology-based essentiality filter has a specific blind spot: seven PFOR-family sequences in the strain K proteome passed the host-homology and druggability filters but were removed at the essentiality step, because orthology to three aerobic model eukaryotes cannot recognise an anaerobic, hydrogenosome-restricted enzyme with no aerobic counterpart. PFOR is the enzyme responsible for reductive activation of ronidazole, so its exclusion from the priority set illustrates a general limitation of orthology-based essentiality calling rather than evidence against its biological importance. To address this, we compiled a complementary, orthology-independent “anaerobic essential” gene set comprising PFOR, ferredoxins and other hydrogenosomal components, which recovers PFOR and related enzymes outside the standard essentiality filter. This distinction also clarifies the therapeutic rationale of the present study relative to ronidazole: because ronidazole requires PFOR-mediated bioactivation, resistance can arise from a single point

of failure (loss of PFOR activity), whereas the chaperone, ATPase and kinase targets prioritised here (Hsp90-2, CIPK8-like kinase, CDC48/p97) act through direct inhibition of constitutively required proteins and do not depend on parasite-mediated pro-drug activation. The docking screen was restricted to three candidates and to pharmacologically validated inhibitors, and broader virtual screening together with molecular-dynamics refinement will be required to convert structural hits into actionable leads. We did not explore the relationship between the priority set and the ronidazole-resistance phenotype of feline *T. foetus* [67, 68]. To our knowledge, no publicly available RNA-Seq dataset directly compares ronidazole-resistant and ronidazole-susceptible *T. foetus* isolates: published resistance studies report minimum lethal concentrations and phenotypic characterisation rather than transcriptomic data, and the only related transcriptomic resource is a comparative bovine-versus-feline genotype study that does not address resistance status. Generating matched resistant/susceptible transcriptomes would be a natural next step for directly testing the clinical relevance of the priority candidates identified here, and selective toxicity beyond the host-exclusion filter remains an open question, as several top-ranked candidates belong to chaperone families with mammalian orthologs and will require structure-based exploitation of species-specific binding-site features. We did not perform a dedicated sensitivity analysis varying the identity/coverage thresholds independently at each filtering step; such an analysis would be a valuable direction for follow-up work. The expression-validation step used a single publicly available RNA-Seq run because, to our knowledge, no biologically replicated RNA-Seq dataset for a feline *T. foetus* isolate is currently public; the resulting TPM values were therefore treated as a qualitative filter (transcribed versus not detectably transcribed at a given threshold) rather than as a quantitative expression estimate with an associated measure of biological variability.

The integrative pipeline presented here provides a tractable starting point for downstream experimental validation of antiprotozoal drug targets in *T. foetus*. The subset of ten highly expressed candidates for which clinically advanced inhibitors already exist defines an immediate drug-repurposing opportunity for feline tritrichomonosis, a disease for which no approved therapy is currently available and against which clinical resistance to the only off-label drug of choice has been documented across three continents.

Conclusions

We report the most comprehensive in silico drug-target prioritisation conducted for *Tritrichomonas foetus* since the seminal work of Morin-Adeline and colleagues a decade ago. Through sequential subtractive filtering of the strain K proteome and cross-validation against an independently annotated chromosome-scale assembly of strain KV-1, we identified 433 priority drug-target candidates that are robust to the choice of reference annotation, biased towards classically druggable protein families, and predominantly transcribed in the feline isolate G10/1.

The single most actionable outcome of this analysis is the identification of a drug-repurposing shortlist. Nine of the ten most highly expressed priority candidates belong to a protein family for which clinically advanced or approved small-molecule inhibitors already exist; the tenth, a Rab4B GTPase, belongs to the broader Ras/GTPase superfamily but has no Rab4B-specific inhibitor characterised to date. These candidates include Hsp70 and Hsp90 chaperones, the CDC48/p97 AAA+ ATPase, the ubiquitin-proteasome pathway, and serine/threonine kinases. They provide an immediate translational opportunity for feline tritrichomonosis, a disease for which no approved therapy is currently available and against which clinical resistance to the only off-label drug of choice is increasingly documented.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1186/s12864-026-13330-5>.

Acknowledgements We thank the Galaxy Europe community and the European Bioinformatics Institute (EMBL-EBI) for maintaining the public bioinformatic resources used in this study, including Ensembl Protists, InterProScan and the Alpha-Fold Protein Structure Database. We also acknowledge the developers of all open-source software cited in the Methods section.

Authors' contributions JD conceived the project, designed the bioinformatic strategy and supervised the work. MK implemented the subtractive proteomics pipeline, performed the comparative genomics, expression validation and structural docking analyses. JS contributed parasitological expertise and clinical interpretation. JD wrote the manuscript with input from MK and JS. All authors read and approved the final manuscript.

Funding This work was supported by the Internal Research Fund of the National Veterinary Research Institute (PIWet-PIB), Puławy, Poland, under project F/123. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Data availability All data generated or analysed during this study are included in this published article and its additional files. The 433 priority drug-target candidates with annotations and expression values are provided in Additional file 1 (Table S1), the ligand library with PubChem CIDs and canonical SMILES in Additional file 2 (Table S2), and the 287 cross-strain reciprocal best hit pairs in Additional file 3 (Table S3). The *T. foetus* strain K predicted proteome is available from Ensembl Protists release 62. RNA-Seq reads were obtained from NCBI SRA accession SRR1688628. AlphaFold v2 monomer predictions were retrieved from the AlphaFold Protein Structure Database. The custom pipeline scripts used for the subtractive proteomics, cross-strain reciprocal best-hit, expression, and docking analyses, together with the BRAKER3 gene annotation of strain KV-1 generated in this study, will be deposited in a public repository (GitHub) and archived under a permanent DOI (Zenodo) prior to publication. These materials are available to the editors and reviewers on request during peer review.

Declarations

Ethics approval and consent to participate Not applicable.

Consent for publication Not applicable.

Competing interests The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

References

1. Gookin JL, Breitschwerdt EB, Levy MG, Lappin MR, Stebbins ME, Stebbins KE. Diarrhea associated with trichomonosis in cats. *J Am Vet Med Assoc*. 1999;215(10):1450–4.
2. Levy MG, Gookin JL, Poore M, Birkenheuer AJ, Dykstra MJ, Litaker RW. *Tritrichomonas foetus* and not *Pentatrichomonas hominis* is the etiologic agent of feline trichomonal diarrhea. *J Parasitol*. 2003;89(1):99–104.
3. Tolbert MK, Gookin J. *Tritrichomonas foetus*: a new agent of feline diarrhea. *Compend Contin Educ Vet*. 2009;31(8):374–81.
4. Yao C, Köster LS. *Tritrichomonas foetus* infection, a cause of chronic diarrhea in the domestic cat. *Vet Res*. 2015;46:35.
5. Stockdale HD, Givens MD, Dykstra CC, Blagburn BL. *Tritrichomonas foetus* infections in surveyed pet cats. *Vet Parasitol*. 2009;160(1–2):13–7.
6. Bastos BF, Brener B, de Figueiredo MA, Leles D, Mendes-de-Almeida F. What is known about *Tritrichomonas foetus* infection in cats? *Rev Bras Parasitol Vet*. 2019;28(1):1–11.
7. Frey CF, Müller N. *Tritrichomonas* — systematics of an enigmatic genus. *Mol Cell Probes*. 2012;26(3):132–6.
8. Walden HS, Dykstra C, Dillon A, Rodning S, Givens D, Bird R, et al. A new species of *Tritrichomonas* (*Sarcomastigophora: Trichomonida*) from the domestic cat. *Parasitol Res*. 2013;112(6):2227–35.
9. Gookin JL, Stebbins ME, Hunt E, Burlone K, Fulton M, Hochel R, et al. Prevalence of and risk factors for feline *Tritrichomonas foetus* and *Giardia* infection. *J Am Vet Med Assoc*. 2004;224(8):1300–6.

10. Gunn-Moore DA, McCann TM, Reed N, Simpson KE, Tennant B. Prevalence of *Tritrichomonas foetus* infection in cats with diarrhoea in the UK. *J Feline Med Surg*. 2007;9(3):214–8.
11. Dąbrowska J, Karamon J, Kochanowski M, Sroka J, Skrzypek K, Zdybel J, Cencek T. *Tritrichomonas foetus*: a study of prevalence in animal hosts in Poland. *Pathogens*. 2020;9(3):203.
12. Dąbrowska J, Karamon J, Kochanowski M, Sroka J, Zdybel J, Różycki M, Cencek T. *Tritrichomonas foetus* as a causative agent of tritrichomonosis in different animal hosts. *J Vet Res*. 2019;63(4):533–41.
13. Dąbrowska J, Karamon J, Kochanowski M, Sroka J, Zdybel J, Różycki M, Cencek T. *Tritrichomonas foetus* infection in cat — first detection in Poland. *Acta Parasitol*. 2015;60(4):605–8.
14. Gookin JL, Copple CN, Papich MG, Poore MF, Stauffer SH, Birkenheuer AJ, Twedt DC, Levy MG. Efficacy of ronidazole for treatment of feline *Tritrichomonas foetus* infection. *J Vet Intern Med*. 2006;20(3):536–43.
15. Gookin JL, Hanrahan K, Levy MG. The conundrum of feline trichomonosis: the more we learn the ‘trickier’ it gets. *J Feline Med Surg*. 2017;19(3):261–74.
16. Papich MG, LeVine DN, Gookin JL, Davidson GS, Stagner WC, Hayes RB. Ronidazole pharmacokinetics in cats following delivery of a delayed-release guar gum formulation. *J Vet Pharmacol Ther*. 2013;36(4):399–407.
17. Rosado TW, Specht A, Marks SL. Neurotoxicosis in 4 cats receiving ronidazole. *J Vet Intern Med*. 2007;21(2):328–31.
18. Lim S, Park SI, Ahn KS, Oh DS, Shin SS. Efficacy of ronidazole for treatment of cats experimentally infected with a Korean isolate of *Tritrichomonas foetus*. *Korean J Parasitol*. 2012;50(2):161–4.
19. Grellet A, Makhoulouf SE, Desquilbet L, Hovhannessian F, Boogaerts C, Dore V, et al. Efficacy of guar gum-based ronidazole capsules as a treatment for *Tritrichomonas foetus* infection in cats. *J Feline Med Surg*. 2017;19(2):177–84.
20. Gookin JL, Stauffer SH, Dybas D, Cannon DH. Documentation of *in vivo* and *in vitro* aerobic resistance of feline *Tritrichomonas foetus* isolates to ronidazole. *J Vet Intern Med*. 2010;24(4):1003–7.
21. Rush GM, Šlapeta J. Evidence of self-resolution of feline trichomonosis in a pair of single household cats due to ronidazole-resistant *Tritrichomonas foetus*. *Vet Parasitol*. 2021;300:109609.
22. Kerek Á, Yurt A, Szabó Á, Tuska-Szalay B, Jerzsele Á. Antimicrobial resistance and natural alternatives: *in vitro* efficacy of Hungarian propolis against feline and bovine *Tritrichomonas foetus*. *Front Vet Sci*. 2025;12:1635358.
23. Gookin JL, Stauffer SH, Coccaro MR, Poore MF, Levy MG, Papich MG. Efficacy of tinidazole for treatment of cats experimentally infected with *Tritrichomonas foetus*. *Am J Vet Res*. 2007;68(10):1085–8.
24. LeVine DN, Papich MG, Gookin JL, Davidson GS, Davis JL, Hayes RG. Ronidazole pharmacokinetics after intravenous and oral immediate-release capsule administration in healthy cats. *J Feline Med Surg*. 2011;13(4):244–50.
25. Gookin JL, Riviere JE, Gilger BC, Papich MG. Acute renal failure in four cats treated with paromomycin. *J Am Vet Med Assoc*. 1999;215(12):1821–3.
26. Lindmark DG, Müller M. Antitrichomonad action, mutagenicity, and reduction of metronidazole and other nitroimidazoles. *Antimicrob Agents Chemother*. 1976;10(3):476–82.
27. Kulda J. Trichomonads, hydrogenosomes and drug resistance. *Int J Parasitol*. 1999;29(2):199–212.
28. Land KM, Delgadillo-Correa MG, Tachezy J, Vanacova S, Hsieh CL, Sutak R, Johnson PJ. Targeted gene replacement of a ferredoxin gene in *Trichomonas vaginalis* does not lead to metronidazole resistance. *Mol Microbiol*. 2004;51(1):115–22.
29. Leitsch D, Kolarich D, Binder M, Stadlmann J, Altmann F, Duchêne M. *Trichomonas vaginalis*: metronidazole and other nitroimidazole drugs are reduced by the flavin enzyme thioredoxin reductase and disrupt the cellular redox system. Implications for nitroimidazole toxicity and resistance. *Mol Microbiol*. 2009;72(2):518–36.
30. Morin-Adeline V, Lomas R, O’Meally D, Stack C, Conesa A, Šlapeta J. Comparative transcriptomics reveals striking similarities between the bovine and feline isolates of *Tritrichomonas foetus*: consequences for *in silico* drug-target identification. *BMC Genomics*. 2014;15:955.
31. Zdrazil B, Felix E, Hunter F, Manners EJ, Blackshaw J, Corbett S, et al. The ChEMBL Database in 2023: a drug discovery platform spanning multiple bioactivity data types and time periods. *Nucleic Acids Res*. 2024;52(D1):D1180–92.
32. Harding SD, Armstrong JF, Faccenda E, Southan C, Alexander SPH, Davenport AP, et al. The IUPHAR/BPS Guide to Pharmacology in 2024. *Nucleic Acids Res*. 2024;52(D1):D1438–49.
33. Knox C, Wilson M, Klinger CM, Franklin M, Oler E, Wilson A, et al. DrugBank 6.0: the DrugBank Knowledgebase for 2024. *Nucleic Acids Res*. 2024;52(D1):D1265–75.
34. Zhang M, Wang C, Otto TD, Oberstaller J, Liao X, Adapa SR, et al. Uncovering the essential genes of the human malaria parasite *Plasmodium falciparum* by saturation mutagenesis. *Science*. 2018;360(6388):eaap7847.
35. Urban M, Cuzick A, Seager J, Wood V, Rutherford K, Venkatesh SY, et al. PHI-base in 2022: a multi-species phenotype database for Pathogen-Host Interactions. *Nucleic Acids Res*. 2022;50(D1):D837–47.
36. Buchfink B, Reuter K, Drost HG. Sensitive protein alignments at tree-of-life scale using DIAMOND. *Nat Methods*. 2021;18(4):366–8.
37. Hosen MI, Tanmoy AM, Mahbuba DA, Salma U, Nazim M, Islam MT, Akhter S. Application of a subtractive genomics approach for *in silico* identification and characterization of novel drug targets in *Mycobacterium tuberculosis* F11. *Interdiscip Sci*. 2014;6(1):48–56.

38. Gabriel L, Brúna T, Hoff KJ, Ebel M, Lomsadze A, Borodovsky M, Stanke M. BRAKER3: Fully automated genome annotation using RNA-Seq and protein evidence with GeneMark-ETP, AUGUSTUS, and TSEBRA. *Genome Res.* 2024;34(5):769–77.
39. Kim D, Paggi JM, Park C, Bennett C, Salzberg SL. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat Biotechnol.* 2019;37(8):907–15.
40. Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics.* 2014;30(7):923–30.
41. Blum M, Chang HY, Chuguransky S, Grego T, Kandasaamy S, Mitchell A, et al. The InterPro protein families and domains database: 20 years on. *Nucleic Acids Res.* 2021;49(D1):D344–54.
42. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, et al. Highly accurate protein structure prediction with AlphaFold. *Nature.* 2021;596(7873):583–9.
43. Le Guilloux V, Schmidtke P, Tuffery P. Fpocket: an open source platform for ligand pocket detection. *BMC Bioinformatics.* 2009;10:168.
44. Trott O, Olson AJ. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem.* 2010;31(2):455–61.
45. Eberhardt J, Santos-Martins D, Tillack AF, Forli S. AutoDock Vina 1.2.0: New docking methods, expanded force field, and Python bindings. *J Chem Inf Model.* 2021;61(8):3891–8.
46. Alonso N, Sturm A, Pereira LM, Bhumkar A, Yates JR 3rd, Garner B, et al. Comparative transcriptomic analysis of *Tritrichomonas foetus* isolates from different host species reveals strain- and host-specific gene expression patterns. *BMC Genomics.* 2022;23(1):394.
47. Galindo-Hernández O, Reyes-García MR, Téllez-Sosa J, Martínez-Barnette J, Saldaña-Balmori Y, Hernández-Hernández JM. An ontology-based bioinformatics approach identifies twelve shared drug targets in five protozoan parasites of public health interest. *Front Cell Infect Microbiol.* 2022;12:826112.
48. Sturm A, Cohen S, Dimitrov D, Tassinari E, Marini G, Bartolommei G, et al. Subtractive proteomics-based *in silico* identification of essential proteins as potential drug targets in *Plasmodium falciparum*. *Asian Pac J Trop Med.* 2014;7(Suppl 1):S480–9.
49. Doyle MA, Gasser RB, Woodcroft BJ, Hall RS, Ralph SA. Drug target prediction and prioritization: using orthology to predict essentiality in parasite genomes. *BMC Genomics.* 2010;11:222.
50. Bradic M, Warring SD, Tooley GE, Scheid P, Secor WE, Land KM, et al. Genetic indicators of drug resistance in the highly repetitive genome of *Trichomonas vaginalis*. *Genome Biol Evol.* 2017;9(6):1658–72.
51. Tao Z, Wang Y, Ke H, Zhang J, Zhao H, Wang H, et al. Chromosome-level genome assembly and annotation of a *Trichomonas vaginalis* clinical isolate. *Sci Data.* 2024;11(1):524.
52. Talevich E, Mirza A, Kannan N. Structural and evolutionary divergence of eukaryotic protein kinases in Apicomplexa. *BMC Evol Biol.* 2011;11:321.
53. Cabrera DG, Horatscheck A, Wilson CR, Basarab G, Eyermann CJ, Chibale K. Plasmodial kinase inhibitors targeting malaria: recent developments. *J Med Chem.* 2024;67(3):1761–98.
54. Marapana DS, Cowman AF. Uncovering *Plasmodium falciparum* kinase function. *Trends Parasitol.* 2021;37(7):604–6.
55. Hennessey KM, Rogiers IC, Shih HW, Hulverson MA, Choi R, McCloskey MC, et al. Identification and validation of small-gatekeeper kinases as drug targets in *Giardia lamblia*. *PLoS Negl Trop Dis.* 2016;10(11):e0005107.
56. Hortua-Triana MA, Vella SA, Márquez-Nogueras KM, Watanabe A, Etheridge RD, Moreno SNJ. Rho GTPase signaling in *Toxoplasma gondii* and related apicomplexan parasites as a therapeutic target. *Front Cell Infect Microbiol.* 2025;15:1567890.
57. Bibo-Verdugo B, Wang SC, Almaliti J, Ta AP, Jiang Z, Wong DA, et al. The proteasome as a drug target in the metazoan pathogen, *Schistosoma mansoni*. *ACS Infect Dis.* 2023;9(2):209–19.
58. Roy SK, Mishra DR, Patra T, Sahu R, Mohanty C, Mohapatra D, et al. Hsp90 inhibitors: a target-based strategy in drug discovery. *Drug Discov Today.* 2012;17(11–12):582–9.
59. Wright C, Iyer AKV, Yakisich JS, Azad N. Hsp70 inhibitors in cancer treatment: a review. *Eur J Pharmacol.* 2018;819:174–84.
60. Anderson DJ, Le Moigne R, Djakovic S, Kumar B, Rice J, Wong S, et al. Targeting the AAA ATPase p97 as an approach to treat cancer through disruption of protein homeostasis. *Cancer Cell.* 2015;28(5):653–65.
61. Adams J. The proteasome: a suitable antineoplastic target. *Nat Rev Cancer.* 2004;4(5):349–60.
62. Rüegg UT, Burgess GM, Staurosporine. K-252 and UCN-01: potent but nonspecific inhibitors of protein kinases. *Trends Pharmacol Sci.* 1989;10(6):218–20.
63. Tanramluk D, Schreyer A, Pitt WR, Blundell TL. On the origins of enzyme inhibitor selectivity and promiscuity: a case study of protein kinase binding to staurosporine. *Chem Biol Drug Des.* 2009;74(1):16–24.
64. del Alamo D, Sala D, Mchaourab HS, Meiler J. Sampling alternative conformational states of transporters and receptors with AlphaFold2. *eLife.* 2022;11:e75751.
65. Saldaña T, Escobedo N, Marchetti J, Zea DJ, Mac Donagh J, Velez Rueda AJ, et al. Impact of protein conformational diversity on AlphaFold predictions. *Bioinformatics.* 2022;38(10):2742–8.

66. Hirt RP. *Trichomonas vaginalis* virulence factors: an integrative overview. Sex Transm Infect. 2013;89(6):439–43.
67. Gookin JL, Stauffer SH, Dybas D, Cannon DH. Documentation of *in vivo* and *in vitro* aerobic resistance of feline *Tritrichomonas foetus* isolates to ronidazole. J Vet Intern Med. 2010;24(4):1003–7.
68. Foster DM, Gookin JL, Poore MF, Stebbins ME, Levy MG. Outcome of cats with diarrhea and *Tritrichomonas foetus* infection. J Am Vet Med Assoc. 2021;225(6):888–92.
69. Dąbrowska J, Keller I, Karamon J, Kochanowski M, Gottstein B, Cencek T, Frey CF, Müller N. Whole genome sequencing of a feline strain of *Tritrichomonas foetus* reveals massive genetic differences to bovine and porcine isolates. Int J Parasitol. 2020;50(3):227–33.
70. Dąbrowska J, Sroka J. A multi-omics perspective on *Tritrichomonas foetus*: from genomics to future directions. Int J Mol Sci. 2025;26(17):8343.
71. Abdullah O, Omran Z. Geldanamycins: potent Hsp90 inhibitors with significant potential in cancer therapy. Int J Mol Sci. 2024;25(20):11293.

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

Authors and Affiliations

Joanna Dąbrowska¹ · Maciej Kochanowski² · Jacek Sroka³

✉ Joanna Dąbrowska
joanna.dabrowska@piwet.pulawy.pl

¹ Department of Virology and Viral Diseases of Animals, National Veterinary Research Institute, Partyzantów 57 Avenue, Puławy 24-100, Poland

² Department of Bacteriology and Bacterial Animal Diseases, National Veterinary Research Institute, Partyzantów 57 Avenue, Puławy 24-100, Poland

³ Department of Parasitology and Invasive Diseases, Bee Diseases and Aquatic Animal Diseases, National Veterinary Research Institute, Partyzantów 57 Avenue, Puławy 24-100, Poland