

RESEARCH ARTICLE

Targeting tRNA-synthetase interactions towards novel therapeutic discovery against eukaryotic pathogens

Paul Kelly^{1,2}, Fatemeh Hadi-Nezhad³, Dennis Y. Liu⁴, Travis J. Lawrence^{3,5}, Roger G. Linington⁴, Michael Ibba^{1,2,6*}, David H. Ardell^{3,7*}

1 The Ohio State University Molecular, Cellular and Developmental Biology Program, The Ohio State University, Columbus, Ohio, United States of America, **2** Center for RNA Biology, The Ohio State University, Ohio, United States of America, **3** Quantitative and Systems Biology Program, University of California, Merced, California, United States of America, **4** Department of Chemistry, Simon Fraser University, Burnaby, British Columbia, Canada, **5** Biosciences Division, Oak Ridge National Lab, Oak Ridge, Tennessee, United States of America, **6** Department of Microbiology, The Ohio State University, Columbus, Ohio, United States of America, **7** Department of Molecular & Cell Biology, University of California, Merced, California, United States of America

* ibba.1@osu.edu (MI); dardell@ucmerced.edu (DHA)



OPEN ACCESS

Citation: Kelly P, Hadi-Nezhad F, Liu DY, Lawrence TJ, Linington RG, Ibba M, et al. (2020) Targeting tRNA-synthetase interactions towards novel therapeutic discovery against eukaryotic pathogens. *PLoS Negl Trop Dis* 14(2): e0007983. <https://doi.org/10.1371/journal.pntd.0007983>

Editor: Rhoel Ramos Dinglasan, University of Florida, UNITED STATES

Received: July 22, 2019

Accepted: December 10, 2019

Published: February 27, 2020

Copyright: © 2020 Kelly et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the manuscript and its Supporting Information files.

Funding: This work was supported by the National Institutes of Health grant AI127582 to MI, DA and RL (<https://www.nih.gov>). Oak Ridge National Laboratory is managed by UT-Battelle, LLC, for the U.S. Department of Energy under contract DE-AC05-00OR22725. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Abstract

The development of chemotherapies against eukaryotic pathogens is especially challenging because of both the evolutionary conservation of drug targets between host and parasite, and the evolution of strain-dependent drug resistance. There is a strong need for new non-toxic drugs with broad-spectrum activity against trypanosome parasites such as *Leishmania* and *Trypanosoma*. A relatively untested approach is to target macromolecular interactions in parasites rather than small molecular interactions, under the hypothesis that the features specifying macromolecular interactions diverge more rapidly through coevolution. We computed tRNA Class-Informative Features in humans and independently in eight distinct clades of trypanosomes, identifying parasite-specific informative features, including base pairs and base mis-pairs, that are broadly conserved over approximately 250 million years of trypanosome evolution. Validating these observations, we demonstrated biochemically that tRNA:aminoacyl-tRNA synthetase (aaRS) interactions are a promising target for anti-trypanosomal drug discovery. From a marine natural products extract library, we identified several fractions with inhibitory activity toward *Leishmania major* alanyl-tRNA synthetase (AlaRS) but no activity against the human homolog. These marine natural products extracts showed cross-reactivity towards *Trypanosoma cruzi* AlaRS indicating the broad-spectrum potential of our network predictions. We also identified *Leishmania major* threonyl-tRNA synthetase (ThrRS) inhibitors from the same library. We discuss why chemotherapies targeting multiple aaRSs should be less prone to the evolution of resistance than monotherapeutic or synergistic combination chemotherapies targeting only one aaRS.

Competing interests: The authors have declared that no competing interests exist.

Author summary

Trypanosome parasites pose a significant health risk worldwide. Conventional drug development strategies have proven challenging given the high conservation between humans and pathogens, with off-target toxicity being a common problem. Protein synthesis inhibitors have historically been an attractive target for antimicrobial discovery against bacteria, and more recently for eukaryotic pathogens. Here we propose that exploiting pathogen-specific tRNA-synthetase interactions offers the potential for highly targeted drug discovery. To this end, we improved tRNA gene annotations in trypanosome genomes, identified functionally informative trypanosome-specific tRNA features, and showed that these features are highly conserved over approximately 250 million years of trypanosome evolution. Highlighting the species-specific and broad-spectrum potential of our approach, we identified natural product inhibitors against the parasite translational machinery that have no effect on the homologous human enzyme.

Introduction

Developing therapies against eukaryotic pathogens has proven challenging due to high conservation between the infectious agent drug target and their host counterpart [1]. Of particular concern is the trypanosome parasite *Leishmania* that infects upwards of 2 million individuals every year and accounts for more than 50,000 deaths annually [2]. While current treatments of amphotericin B and miltefosine are commonly prescribed to patients with leishmanial infections, they have undesired off-target cytotoxicity, leading to poor patient compliance and low-dose administration [3], and ultimately contributing to the rise of strain-dependent drug resistance [4,5]. There is a strong need for new nontoxic drugs with broad-spectrum activity against different species of *Leishmania* and other trypanosomes [6,7].

Given their essential role in protein synthesis, aminoacyl-tRNA synthetases (aaRSs) have been an attractive target for antimicrobial therapeutics [8]. AaRSs are essential enzymes found in all domains of life that are responsible for the correct pairing of free amino acids in the cell to their cognate tRNA [9]. AaRSs perform their activity in two steps: first, a free amino acid is activated by the enzyme through the hydrolysis of ATP, forming an aminoacyl-adenylate. Second, the amino acid is transferred to its corresponding tRNA before being released into the aminoacyl-tRNA pool [9]. Given the complex pool of free amino acids and uncharged tRNAs in the cell, aaRSs have co-evolved discrete mechanisms to ensure mutually exclusive amino acid activation and cognate tRNA recognition [10]. The sequence/structural determinants (or anti-determinants) that lead to accurate aaRS-tRNA recognition are also known as the tRNA identity elements. The primary tRNA identity elements that aid in cognate aminoacylation have been extensively studied for several decades [11,12]. For example, across all three domains of life, all tRNA^{Ala} isoacceptors contain a conserved G:U base pair in the acceptor stem that is recognized by alanyl-tRNA synthetase (AlaRS), leading to accurate Ala-tRNA^{Ala} synthesis in the cell [13–15].

While some aaRS inhibitors have successfully made it to the clinic, including the IleRS-targeting mupirocin [16], ProRS inhibitor halofuginone [17], and the LeuRS inhibitor tavaborole [18,19], there are likely many potential aaRS drugs still to be identified. Target-based approaches relying on structural data and sequence identity have previously been used to try and predict novel trypanosome aaRS drug targets [20–22] with some recent success [23]. While structure-based approaches have their utility, exploiting tRNA-aaRS interactions has been under-explored for its therapeutic potential. In particular, while interactions with small

molecules are expected to be quite conserved across species, the evolutionary diversification of tRNA identity element interactions through coevolution with aaRSs opens the possibility of greater species-specific inhibition.

While major identity elements have been experimentally characterized for many aaRS-tRNA pairs in various model systems, much less is known about how tRNA identity elements evolve and diverge over the Tree of Life. Recent theoretical advances explain how tRNA identity elements can evolve and diverge in a phylogenetically informative way, even while under strong selective constraints [24]. In earlier work, we developed a bioinformatic method to predict tRNA identity elements [25]. Our bioinformatic predictions are called Class-Informative Features (CIFs), based on the statistic of structure-conditioned Shannon Information [26], and visualized through graphs called Function Logos [25]. In later work, we applied two other statistics, Information Difference and Kullback-Leibler Divergence, to facilitate pairwise comparisons of CIFs between two taxa, in two new visualizations called Information Difference (ID) logos and Kullback-Leibler Divergence (KLD) logos, respectively [27]. ID logos visualize gains and losses of CIFs, which in this work we call gains and losses of information, while KLD logos visualize the functional conversion of CIFs from one functional type of tRNA to another, which in this work we call change of functional information. In the present work, we integrate together all three statistics (structure-conditioned information about function, ID, and KLD) and apply it to the problem of identifying parasite-specific tRNA identity elements. Our approach visualizes functionally informative features in parasite tRNAs that have either gained or retained functional information relative to humans, altered functional associations, or both, since divergence from their common ancestor with humans.

Our modeling approach integrates genomic tRNA sequence variation across multiple tRNA gene families of different functions, revealing potentially useful information about the specification of substrate identity for all aaRSs simultaneously. The multiplicity of aaRSs in cells provides multiple potential targets for inhibition of essential parasite enzymes, opening the door to improved combination chemotherapies. Advances in systems biology and chemogenomics have fueled interest in combination chemotherapies to benefit from synergistic drug interactions [28–31] and combat the evolution of resistance [32]. Combination chemotherapies are naturally effective, for example, in the pathogenic defenses of arthropods [33] and have yielded exciting antifungal [34] and antihelminthic [35] therapies. Additionally, artemisinin-based combination therapies (ACTs) are the primary treatment plan for *Plasmodium falciparum* malaria infections [36,37].

Here, we report our improved annotation of TriTrypDB genomes and new methodologies for predicting conserved identity elements across biological domains. As proof of principle, we screened for identity element divergence between trypanosomes and humans to search for new therapeutic targets for these eukaryotic pathogens. Validating our computational approaches, we found several natural product fractions that inhibit *Leishmania major* AlaRS activity but have no effect on the homologous human enzyme. The fractions we identified also have inhibitory activity against *Trypanosoma cruzi* AlaRS, showing that our approach holds promise towards identifying new broad-spectrum anti-trypanosomal therapies.

Methods

Annotation, clustering and filtering of tRNA genes in TriTrypDB genomes

We downloaded data for 46 genome assemblies from TriTrypDB version 41 released December 5th, 2018. We ran tRNAscan-SE v.2.0.0 installed via BioConda in February, 2019 [38] and Aragorn v.1.2.38 [39] using option “-i116” (implying a maximum intron length in search targets of 116 base-pairs) on this data. We unified gene records from the two finders if they

overlapped by at least one base-pair, had consistent strand-orientations and end-displacements less than or equal to 4 bp. To independently identify initiator tRNA genes, we computed edit distances [40] of CAT-anticodon-containing genes implemented in the function `stringdist()` from its R package v. 0.9.5.5 and clustered them agglomeratively using Ward's minimum variance method [41] implemented in the function `hclust()` with method `ward.D2` from the base R stats package, examining clusters for the initiator-distinguishing features described in [42]. All statistical analyses and sequence processing for annotation and clustering were carried out in R v.3.6 [43].

To further investigate these gene records, we examined their genetic clustering in TriTrypDB genomes as defined by co-occurrence within a distance of 1000 bp on either strand. We computed gene function content distances of tRNA gene-clusters as pairwise Jaccard distances considering gene clusters as sets of functions using `stringdist()` and clustered them with Ward's method using function `hclust()` with method `ward.D2`. We finalized our annotation union gene-set by retaining 3616 genes that had an Aragorn score above 106 bits or a tRNAscan-SE2 score above 49 bits, and reannotating sequences as described in the Results.

Prediction of divergent tRNA Class-Informative Features (CIFs) in humans and parasites

To compare CIFs between TriTrypDB genomes and humans and to have sufficient data to estimate trypanosome CIFs, we defined eight phylogenetic clades for 39 of the 46 trypanosome genomes as shown in Table 1. These clades were based on a composite of phylogenetic results in the literature [44–47]. CIFs were subsequently estimated for each clade independently, by pooling tRNA genes within clades. We removed two incomplete genomes from analysis, *T. rangeli* SC58 and *T. cruzi* CL-Brenner that had fewer than half the number of tRNA genes identified in any other genome and missed more than two functional classes. We filtered the gene annotation union gene set of 3616 genes, removing selenocysteine genes, pseudogenes, truncated genes, and genes of ambiguous function, leaving 3488 high-confidence functionally annotated gene records from 44 genomes in TriTrypDB v.41 for alignment. To this set we added 431 high-confidence human tRNA gene records downloaded from GtRNADB [48] on May 15, 2019 (in the file “hg38-tRNAs.fas”), excluding two human selenocysteine tRNA genes, to yield a grand total of 3919 tRNA genes from 45 genomes for structural alignment. We aligned this alignment gene-set of 3919 genes using COVEA v.2.4.4 [49] to the eukaryotic tRNA covariance model supplied with tRNAscan-SE v.1 [50]. The output alignment was manually edited in SEAVIEW [51] to correct the misalignment of 595 human and trypanosome tRNA genes (almost exclusively of type tRNA^{Leu} and tRNA^{Ser}) at Sprinzl coordinates 45 and 47 and exclude majority-gap/insertion and variable-arm-containing sites (Sprinzl coordinates are a standardized coordinate system that encodes both the consensus universal secondary structure of tRNAs, and conserved, more functionally-specific structures like the long variable arms of tRNA^{Leu} and tRNA^{Ser} [52]). Sequences were further processed with the FAST toolkit to partition genes into clades [53]. After excluding an additional 464 genes from five genomes not included in our defined clades, 3455 aligned trypanosome and human genes remained. More statistics on the CIF estimation gene sets by clade are shown in Table 1 and additional notes and code to reproduce the data workflow are provided in Supplementary Online Materials (Code and Data).

For each clade independently, we computed function logos [25], Information Difference logos and Kullback-Leibler Divergence logos [27] with a newly written Python 3 program tSFM (tRNA Structure-Function Mapper) v0.9.14 available on github (<https://github.com/tlawrence3/tsfm>), which we describe briefly here, and more fully in a forthcoming publication.

Table 1. Clades and genomes analyzed, with statistics on CIF estimation gene sets.

Clade	Genome Assemblies	Total Genes	Mean Genes/Genome (Std. Dev.)	Grand Mean %G	Grand Mean %C	Grand Mean %T
Humans	Hg38	431	431	32.5	25.7	23.0
<i>L. major</i> Clade (n = 8)	1. <i>L. major</i> Friedlin 2. <i>L. major</i> LV39c5 3. <i>L. major</i> SD75 4. <i>L. tropica</i> L590 5. <i>L. aethiopic</i> a L147 6. <i>L. gerbilli</i> LEM452 7. <i>L. turanica</i> LEM423 8. <i>L. arabica</i> LEM1108	664	83 (0.6)	31.9	26.1	23.4
<i>L. infantum</i> Clade (n = 3)	1. <i>L. donovani</i> BHU1220, 2. <i>L. donovani</i> BPK282A1 3. <i>L. infantum</i> JPCM5	250	83.33 (0.2)	31.9	26.1	23.3
<i>L. mexicana</i> Clade (n = 2)	1. <i>L. amazonensis</i> MHOMBR71973M2269 2. <i>L. mexicana</i> MHOMGT2001U1103	148	74 (4.7)	31.8	25.9	23.4
Viannia Subclade (n = 4)	1. <i>L. braziliensis</i> MHOMBR75M2904 2. <i>L. braziliensis</i> MHOMBR75M2903 3. <i>L. panamensis</i> MHOMPA94PSC1 4. <i>L. panamensis</i> MHOMCOL81L13	327	81.75 (2.5)	31.8	26.1	23.4
<i>L. enriettii</i> Clade (n = 2)	1. <i>Leishmania</i> sp MARLEM2494 2. <i>Leishmania enriettii</i> LEM3045	160	80 (0.5)	31.9	25.9	23.4
<i>Leptomonas/ Crithidia</i> Clade (n = 3)	1. <i>Crithidia fasciculata</i> CfCl 2. <i>Leptomonas seymouri</i> ATCC30220 3. <i>Leptomonas pyrrocoris</i> H10	300	100 (2.4)	31.8	26.0	23.3
American <i>Trypanosoma</i> (n = 11)	1. <i>T. grayi</i> ANR4 2. <i>T. cruzi</i> CL Brener Esmeraldo-like 3. <i>T. cruzi</i> CL Brener Non-Esmeraldo-like 4. <i>T. cruzi cruzi</i> Dm28c 5. <i>T. cruzi</i> Dm28c 6. <i>T. cruzi</i> Esmeraldo 7. <i>T. cruzi</i> JRcl4 8. <i>T. cruzi</i> marinkelleiB7 9. <i>T. cruzi</i> SylvioX10-1 10. <i>T. cruzi</i> SylvioX10-1-2012 11. <i>T. cruzi</i> Tulac2	773	70.27 (5.7)	32.0	26.3	23.0
African <i>Trypanosoma</i> (n = 6)	1. <i>T. bruceigambiense</i> DAL972 2. <i>T. brucei</i> Lister427 3. <i>T. brucei</i> TREU927 4. <i>T. evansi</i> STIB805 5. <i>T. congolense</i> IL3000, 6. <i>T. vivax</i> Y486	402	67 (2.1)	32.2	26.1	23.1

<https://doi.org/10.1371/journal.pntd.0007983.t001>

tSFM provides a command-line user interface for estimating function, ID, and KLD logos using our published methods. tSFM additionally calculates tRNA CIFs for secondary-structure feature pairs, in addition to single-site features. Class-Informative Feature Pairs are elements of the Cartesian product set $C = f \times f \times BP$, where $f = \{A, C, G, U, -\}$ is the set of single-site features we consider and BP is the set of structurally-paired Sprinzl Coordinates involved in potential base-pairing interactions along the four arms of the planar clover-leaf consensus secondary structure of tRNAs [52]. We ran tSFM with option “-x 1” corresponding to computing exact expected entropies for samples of size one by the method of [54] or by the Bayesian Nemenman–Shafee–Bialek (NSB) entropy estimator [55] otherwise.

Briefly, we computed the gain-of-information of a CIF in a particular functional class and trypanosome clade as its information difference in bits, with that clade as foreground and humans as background, multiplied by the normalized ratio of posterior-to-prior odds of the

CIF in that functional class in trypanosomes and humans, corresponding to letter heights in ID logos, and measured in bits. We computed change-of-function of a CIF in a particular functional class and trypanosome clade as its Kullback-Liebler Divergence in bits, with that clade as foreground and humans as background, multiplied by the normalized ratio of posterior-to-prior odds of the CIF in that functional class, corresponding to letter heights in KLD logos and measured in bits. To avoid division by zero when calculating KLD, we added pseudocounts to either the background or the foreground posterior distributions when one or more of the 21 functional classes was not observed. When calculating the normalized ratio of posterior-to-prior odds for a specific functional class, we only added pseudocounts to the background posterior distribution. Furthermore, to avoid inaccuracies, we defined the KLD of a feature to be zero when its frequency in the background is less than or equal to five.

We wrote a custom script in R 3.6 to visualize CIFs within each cluster for each functional class of tRNA in a structural context, and color the parasite CIFs according to whether those CIFs have gained information or changed functional information relative to human since divergence from their common ancestor. All data and scripts are provided as supplementary data.

AaRS cloning and protein purification

Leishmania major (*Lm*) AlaRS and *Lm* ThrRS-encoding genes were codon optimized, synthesized, and sub-cloned into pUC57 (GenScript). Engineered flanking NdeI and SmaI restriction sites were used to clone the aaRS genes into pTYB2, creating in-frame C-terminal intein fusions. The resulting expression vectors were transformed into the *E. coli* expression strain BL21 (DE3). The gene encoding *Trypanosoma cruzi* (*Tc*) AlaRS was codon optimized, synthesized, and directly cloned into NdeI and XhoI cut sites in the pET21b expression vector (GenScript). The resulting plasmid expressed *Tc* AlaRS under T7 control and was in-frame with a C-terminal 6x-His tag. The pET21b-*Tc* AlaRS vector was transformed into the *E. coli* expression strain XJb (DE3) (Zymo Research).

Both *Lm* AlaRS and ThrRS were purified by growing cells to an OD₆₀₀ of ~0.5 and cooling on ice for 30 minutes. Protein induction was initiated by the addition of IPTG to a final concentration of 500 μM and cells continued to grow at 16°C for 16 hours. Cells were harvested by centrifugation and lysed by sonication in Buffer A (25 mM HEPES pH 7.2, 500 mM NaCl, 3 mM DTT) with cOmplete mini protease inhibitor (Sigma) added. Clarified lysate was added to a chitin resin column (NEB) and washed with Buffer A. The intein tag was cleaved by the incubation of Buffer B (25 mM HEPES pH 7.2, 100 mM NaCl, and 100 mM DTT) on the resin bed overnight at 4°C. Protein was dialyzed in two stages in Buffer C (25 mM HEPES pH 7.2, 30 mM NaCl, 6 mM BME, and 10% - 50% glycerol).

Trypanosoma cruzi (*Tc*) AlaRS-expressing cells were grown to an OD₆₀₀ ~0.3 and then cooled to 18°C and induced with 500 μM IPTG. Cells were grown for an additional 16 hours at 18°C before harvesting by centrifugation. Cell pellets were re-suspended in lysis buffer [Buffer I (500 mM Tris-HCl pH 8.0, 300 mM NaCl, and 10 mM imidazole) with cOmplete mini protease inhibitor (Sigma)], sonicated, clarified, and cell lysate passed over a TALON metal affinity column (Takara). After washing the column with Buffer I, protein was eluted with Buffer II (Buffer I with 250 mM imidazole). Protein was dialyzed in two stages to remove imidazole and to store the enzyme in 50% glycerol.

Human AlaRS was expressed in *E. coli* Rosetta (DE3) (Novagen) from pET21a which encodes the human AlaRS gene in-frame with a C-terminal 6x-His tag (expression plasmid provided by Karin Musier-Forsyth, Ohio State University). Cells were grown to an OD₆₀₀ of ~0.5 and cooled on ice for 30' before inducing expression with 500 μM IPTG. Upon induction,

cells grew for an additional 16 hours at 20°C before harvesting. Human AlaRS was purified as described above with the addition of 5 mM β-mercaptoethanol to both Buffer I and Buffer II. All enzyme concentrations were determined by active site titration [56,57] using [¹⁴C]-alanine (Perkin Elmer) and [¹⁴C]-threonine (American Radiochemicals).

Preparation of *in vitro* transcribed tRNA

Lm tRNA^{Ala} (chr11. trna1-AlaCGC), *Lm* tRNA^{Thr} (chr23. trna6-ThrTGT), and *Tc* tRNA^{Ala} (TctRNA-Ala.03) DNA sequences were cloned into EcoRI and XbaI restriction sites in pUC18 by slow cooling complementary synthetic DNA oligos and ligation as previously described [58]. PCR was used to amplify 50 μg DNA template from the pUC18-tRNA plasmids to be used for T7 runoff transcription. *In vitro* transcription was performed with 40 mM Tris-HCl pH 8, 2 mM spermidine, 22 mM MgCl₂, 5 mM DTT, 50 μg/mL BSA, 4 mM NTPs, 20 mM 5'GMP, 20 U Protector RNase Inhibitor, 2 U pyrophosphatase, DNA template, and T7 RNAP at 42°C for 16 hours. Transcription products were purified on a Diethylaminoethyl (DEAE) Sephacel (GE Healthcare) column in 20 mM Tris-HCl pH 8.0, 5 mM MgCl₂, and 250 mM NaCl. tRNA was eluted from the resin with 1 M NaCl. The RNA was precipitated overnight at -20°C in 1/10th volume sodium acetate and 3x volume ethanol and re-suspended in RNase-free H₂O.

Marine natural product library

The marine natural products screening library comprises 5,304 fractions from organic extracts of marine-derived Actinobacterial fermentations (1 litre culture, following our standard protocol [59]). All fractions are stored as concentrated stock solutions in DMSO in standard 96-well format. The library is comprised of extracts of marine sediment-derived bacterial strains, containing a cross section of gram-positive genera and enriched in Actinobacterial strains, hand-collected from over 70 discrete dive sites on the West coast of the United States from the Channel Islands of Southern California to the San Juan Islands in Northern Washington.

Crude extracts were fractionated in to six sub-fractions on Seppak C₁₈ cartridges using a stepwise elution profile (20, 40, 60, 80, 100% MeOH/ H₂O, 100% EtOAc). The resulting fractions were solubilized in DMSO (1 mL per fraction), 4 μL aliquots diluted 1:5 in DMSO, and arrayed in 384 well format (17 x 384 well plates). The MNP library screened in this assay consisted of a focused group of bacterial extract pre-fractions that had already demonstrated activity against *Leishmania* in a prior whole cell assay against *L. donovani* amastigotes. The MNP library was also counter-screened in a mammalian system against HeLa cells [59]. Fractions with acute cell cytotoxicity were removed from the screening library, resulting in a set of test extracts with demonstrated activity against *L. donovani* and low/ no cytotoxicity against HeLa cells. Following primary screening against *L. donovani* amastigotes, 120 active fractions were arrayed as serial dilutions (8 x 2-fold dilutions; 50–0.4 μM) in 96 well format for aaRS screening.

Screen for aminoacylation inhibitors

Serial dilutions from the marine natural product (MNP) library were screened using the following protocol. Aminoacylation reactions were performed at room temperature using 10 mM DTT, 8 mM ATP, 5 μM tRNA, 60–80 μM [¹⁴C]-Ala or [¹⁴C]-Thr, 100–500 nM aaRS, and DMSO or MNP samples. After incubating the reaction for either 15 or 20 minutes, 1 μL of the reaction was spotted on 5% pre-soaked TCA 3 MM Whatman filter paper. The precipitated tRNA-bound filter paper was washed 3x with 5% TCA, washed once with ethanol, and dried. The dried filter paper was exposed overnight on a phosphor imager screen and imaged the

following day. Qualitatively, the phosphor image screen was examined for a change in signal intensity relative to the DMSO control; a decrease in phosphor image intensity indicates partial or full inhibition of the reaction in the presence of the inhibitor. While active concentrations were unknown for each of the MNP mixes, the serial dilution helped prevent false-positive identification. All lead candidates from the preliminary screen were confirmed using similar reaction conditions; the reactions were monitored over a time course and placed at 37°C. Samples were quantified using a scintillation counter.

Pyrophosphate exchange

Amino acid activation was monitored using ATP/PP_i exchange as previously described [60]. Reactions were performed at 37°C in 100 mM HEPES pH 7.2, 30 mM KCl, 10 mM MgCl₂, 2 mM NaF, 2 mM ATP, 2 mM [³²P]-PP_i (Perkin Elmer), 90 μM alanine, 160 nM AlaRS, and DMSO or aaRS inhibitor. At increasing time points, aliquots of the reaction mixture were quenched in a charcoal solution containing 1% activated charcoal, 5.6% HClO₄, and 75 mM PP_i. Quenched reactions were vacuum filtered on to 3MM Whatman filter discs, washed three times with 5 mL of water and once with 5 mL of ethanol. After drying the filter discs, charcoal-bound radiolabeled ATP was quantified on a scintillation counter. Relative endpoint amino acid activation was determined by comparing the inhibitor-treated enzymes to their respective DMSO control samples.

Results

Custom annotation of tRNA genes and gene clusters in TriTrypDB genomes

We obtained 4381 unified gene records from the raw output of two tRNA gene-finders, Aragorn and tRNAscan-SE v.2.0, to TriTrypDB v.41. Of these, 3597 were found by both gene-finders, 750 were found by Aragorn only, and 34 were found by tRNAscan-SE 2.0.0 only. We identified the same 76 genes as initiator tRNA genes, using either tRNAscan-SE 2.0.0's profile-based predictions or our own edit-distance-based clustering approach, by finding the unique set of genes carrying conserved initiator tRNA features as described in [42].

To further refine the final annotated gene set, we identified tRNA gene clusters in TriTrypDB genomes using a maximum intergenic distance criterion of 1000 bp on either strand. Doubling this distance criterion did not substantially increase cluster number or size. After filtering 4381 gene records by their gene-finder scores as described below, 3616 high-confidence gene records remained, of which 77% occur in clusters of size two or greater (Fig 1). The largest tRNA gene clusters were of size ten, accounting for 9% of total genes. We used Jaccard distance as a gene functional content distance to hierarchically cluster tRNA gene clusters with similar gene functional contents, and found that distance cutoffs between 0.680 and 0.692 defined intuitively reasonable similar, distinct, and putatively homologous tRNA gene cluster variant groups that we found to be conserved either within each of the *Leishmania* and *Trypanosoma* genera, or across both genera, with substantial evidence of evolution in gene organization and content of gene clusters within groups through gene duplication, divergence, inversion and other changes (S1–S3 Tables). Some evidence for whole tRNA gene cluster duplication and paralogy within our gene cluster variant groups (labeling variants by the concatenation of their tRNA gene function / aminoacylation identities in gene order, using IUPAC one-letter-codes for amino acids to stand for identities) include one gene cluster variant DSA, with a frequency of three in the genome of *T. cruzi* DM28c and a frequency of two in the genome of *T. cruzi* DM28c, and nine other cluster variants with frequencies of two in one

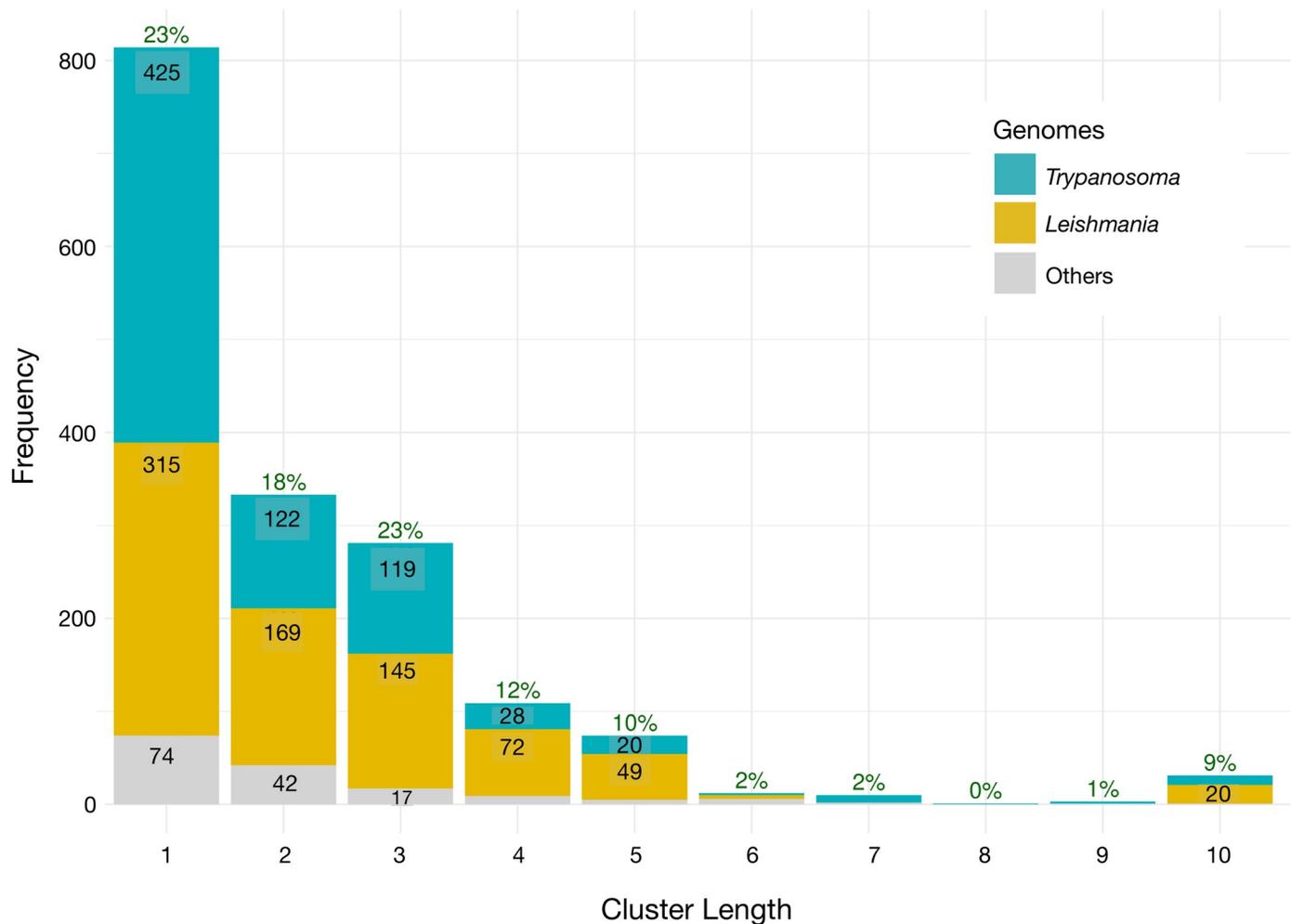


Fig 1. tRNA gene cluster size distribution for *Trypanosoma*, *Leishmania*, and other TriTrypDB version 41 genomes. Green labels at tops of stacks show percentages of total tRNA genes in clusters of given length. Numbers within each bar show frequencies of gene clusters of that length.

<https://doi.org/10.1371/journal.pntd.0007983.g001>

or more genomes, including tRNA gene cluster variants ASD, FEV, NARK, and VYMEMSL occurring twice in the genome of *T. cruzi* Tulacl2, variants PTN and VEF occurring twice in the genome of *T. cruzi* DM28c, variant EVRH occurring twice in *L. arabica* LEM1108, variant FEV occurring twice in *T. cruzi* SylvioX10-1, variant GAL in *L. tropica* L590, and variant HEF in *T. congolense* IL3000. The conservation of tRNA gene cluster variants and groups spanning TriTrypDB genome assemblies of different genera is indisputable, but further statistical and phylogenetic characterization of them may best be undertaken via long-read genome resequencing, as tRNA gene clusters can be difficult to assemble reliably from short-read sequencing data.

Using the similarities of tRNA gene clusters across genomes, we found putative homologs for some of the 45 functionally ambiguous but high-scoring genes marked as pseudogenes or truncated by tRNAscan-SE 2.0.0 as well as some genes detected only by Aragorn. With these results in mind, we plotted the densities of gene-finder scores according to whether they were found by both or only one gene-finder, showing clear evidence of a small fraction of Aragorn-only genes with high scores, making up about 1% of our total finalized gene set (Fig 2 and S4 Table).



Fig 2. Density plots of gene-finder scores according to source of detection.

<https://doi.org/10.1371/journal.pntd.0007983.g002>

Based on this evidence, we retained 3616 genes from 46 TriTrypDB genome assemblies that had an Aragorn bit-score of at least 107 or a tRNAscan-SE 2.0 bit-score of at least 50, including 36 genes found by Aragorn only and 1 gene found by tRNAscan-SE 2.0 only. These score cut-offs separated Aragorn-only genes within conserved gene clusters from singletons, which had lower scores (S1 Fig). At time of publication, a more recent version of tRNAscan-SE, 2.0.5, could find 24 of the 36 high-scoring ara-only genes in our union set. The 24 additionally found genes were identical or had only 1 or 2 base differences in sequence to other high-scoring genes in our defined intersection set. However, we observed them to lie either near the ends or close to ambiguous segments (strings of Ns) in the genome assembly sequences.

The median number of genes per genome in our raw union annotation gene set was 82. Among these were 45 functionally ambiguous but high scoring genes, including 2 with identity unassigned by both gene-finders, 6 marked as pseudogenes or truncated by tRNAscan-SE 2.0.0, 4 containing sequence ambiguities, and 33 with conflicting structural and anticodon annotations, including ten intron-containing genes predicted as tRNA^{Tyr} genes by tRNAscan-SE and tRNA^{Asn} genes by Aragorn, all from genomes in the American *Trypanosoma* clade. We annotated these ten as tRNA^{Tyr} genes following tRNAscan-SE 2.0.0, which are known to contain introns in that clade [61], as this helped complete the sets of functional types for 8 genomes in that clade (S2 Fig). After score filtering, three genome assemblies that were not excluded from further analysis remained incompletely annotated (S2 Fig): *T. Cruzi* dm28c (missing a gene for tRNA^{Phe}), *T. congolense* IL3000 (missing genes for tRNA^{ASP} and tRNA^{Ser}) and *T. vivax* Y486 (missing a gene for tRNA^{Tyr}). Candidate genes were available in our raw

union annotation gene-set to complement some but not all of the 46 genomic gene sets that were missing classes (S2 Fig), including one Aragorn-only candidate gene of 100 bits for tRNA^{Tyr} in *T. cruzi* MarinkelleiB7, between 5–8 low- or marginal-scoring genes for tRNA^{Ser} in *T. congolense* IL3000 (3 from tRNAscan-SE2 with bit scores between 22 and 25 and 5 potentially overlapping genes predicted from Aragorn with scores between 100 and 106), a tRNAscan-SE2-only candidate gene for tRNA^{Tyr} scoring 36 bits in *T. vivax* Y486, and an Aragorn-only candidate gene for tRNA^{Asp} scoring 100 bits in *P. confusum*. S4 Table shows structural and functional statistics on our score-filtered annotation gene sets. Our score-filtered union annotation gene-set was further filtered and pooled into defined clade gene-sets as described in the Methods section. Table 1 shows mean and standard deviations of tRNA gene number and pooled composition statistics by clade. Every clade contained genes for all 21 functional types excluding tRNA^{Sec} (Table 1 and S2 Fig). S5 Table gives gene numbers and compositions of gene sets by individual genome. Table 1 and S5 Table show that the overall divergence in gene compositions is not great, and follows phylogenetic expectations based on phylogeny. The gene compositions of the *Leishmania* clades are quite similar, and different from those of both humans and *Trypanosoma*. Human gene composition is most divergent, with African *Trypanosoma* gene composition second most divergent from those of *Leishmania*. American and African *Trypanosoma* compositions are less divergent from each other. However, *Trypanosoma* and human tRNA gene compositions are different from those of *Leishmania* in different ways. Human tRNA genes are richer in purines while *Trypanosoma* tRNA genes are richer in G and C. American *Trypanosoma* are also GC-rich, but less so than African *Trypanosoma*. S5 Table shows that there is little heterogeneity of gene set compositions by genome assembly within clades, with the greatest variation appearing in *Trypanosoma*, particularly African *Trypanosoma*.

Divergent class-informative features between humans and TriTrypDB genomes

We developed a bioinformatic workflow that combines information from tRNA function logos estimated from a parasite clade and Information Difference (ID) logos [25] and Kullback-Leibler Divergence (KLD) logos between the parasite clade and humans [27]. The workflow quantitates tRNA features that are functionally informative in the parasite clade and have either gained or retained functional information or altered functional association since divergence of the parasite clade and humans from their common ancestor. We found many examples of highly informative trypanosome CIFs that have been gained, retained or changed functional information since divergence from their common ancestor with humans, and most of these divergent CIFs have been strongly conserved in trypanosomes over 231–283 million years of evolutionary divergence between *Leishmania* and *Trypanosoma* [62], for example among alanine tRNAs (Fig 3) and threonine tRNAs (Fig 4). Structural bubble-plot visualizations at single-site resolution of these CIF divergence measures are provided for all functional classes in supplementary materials, showing that some classes have diverged much more than others. Even though they are calculated at single-site resolution, CIF divergences are correlated across structurally paired sites. Inspection of single-site function logos across taxa confirms the conservation of parasite-specific CIFs and reveals A- and U-containing features underlying the signals shown in Figs 3 and 4, including some sharing of divergent features between tRNA^{Ala} and tRNA^{Thr} functional classes, for example at Sprinzl coordinate 39 (Figs 5 and S24–S27). Figs 6–8 show base-pair function logos for Humans, the *L. major* clade and the American *Trypanosoma* respectively, showing that both Class Informative Base-Pairs and Class-Informative Mis-Pairs can be relatively conserved, and that recurring hot-spots of CIF evolution

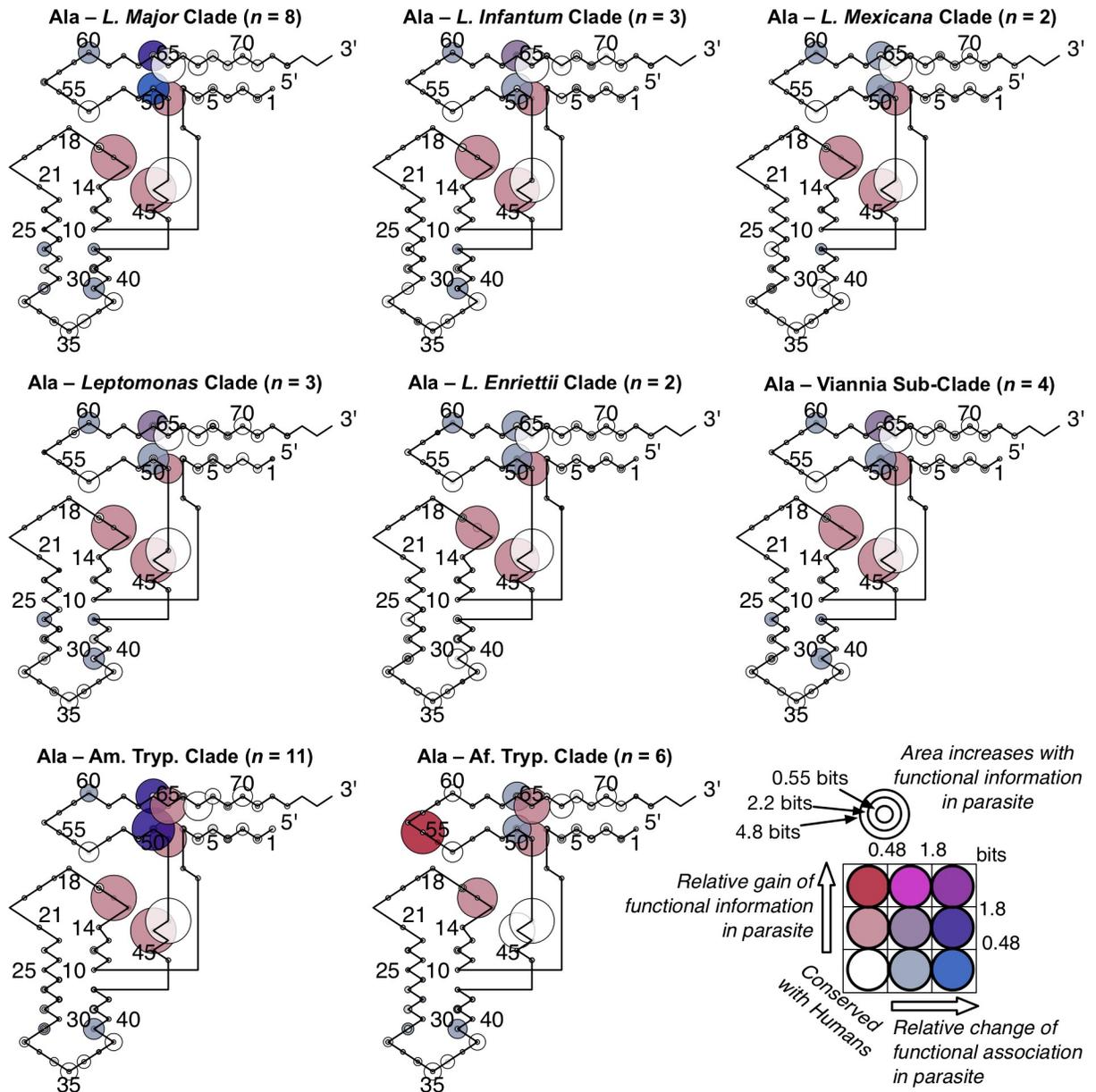


Fig 3. Conserved divergence of parasite tRNA^{Ala} CIFs across eight phylogenetic clades of *Leishmania* and *Trypanosoma*. Evolutionary divergence of trypanosomes relative to *Leishmania major* increases clockwise from *Leishmania major*.

<https://doi.org/10.1371/journal.pntd.0007983.g003>

appear in the data, yielding insight to mechanisms of CIF evolution. Inspection of Class-Informative Base-Pairs and Mis-Pairs shows that a U:A informative base-pair diverged in tRNA^{Thr} to an adjacent site-pair, from 31:39 to 30:40, and that a U-U informative mismatch was gained in tRNA^{Ala} at site-pair 6:66 in trypanosomes relative to humans (Figs 6–8). Full function logo results for all clades are provided in S24–S40 Figs.

Our computational screen for tRNA CIF divergence as shown in S3–S23 Figs, show that tRNA^{Ala} and tRNA^{Thr}, are among those tRNA functional types that have the greatest number of sites and site-pairs with the largest CIF divergence relative to humans and would be good

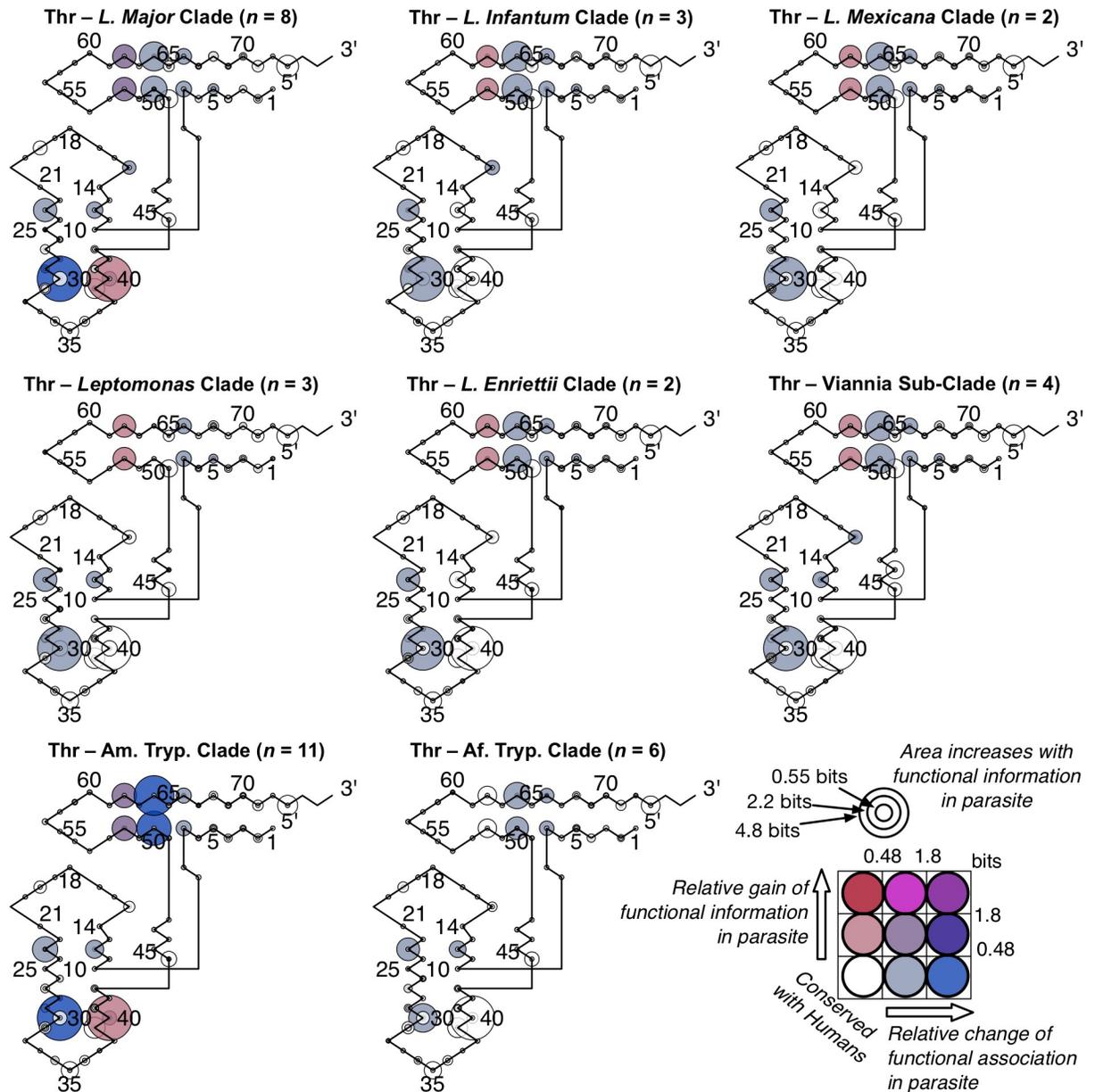


Fig 4. Conserved divergence of parasite tRNA^{Thr} CIFs across eight phylogenetic clades of *Leishmania* and *Trypanosoma*. Evolutionary divergence of trypanosomes relative to *Leishmania major* increases clockwise from *Leishmania major*.

<https://doi.org/10.1371/journal.pntd.0007983.g004>

potential candidates for therapeutic targeting. Contrast, for example, our results for Trypanosome tRNA^{Tyr} or tRNA^{Trp}, which show trypanosomal tRNA CIFs that are strongly conserved with humans, as shown in S9 or S10 Figs. These observations led us to follow-up and investigate tRNA^{Ala} and tRNA^{Thr} because both these tRNA types, and their accompanying cognate synthetases, are readily reconstituted *in vitro*. Previous experience with *in vitro* reconstitution of AlaRS and ThrRS enzymes in the Ibba lab aided us in troubleshooting problems and avoiding false-positive identification of inhibitors, as artifactual changes in enzyme activity are common if technical care is not taken.

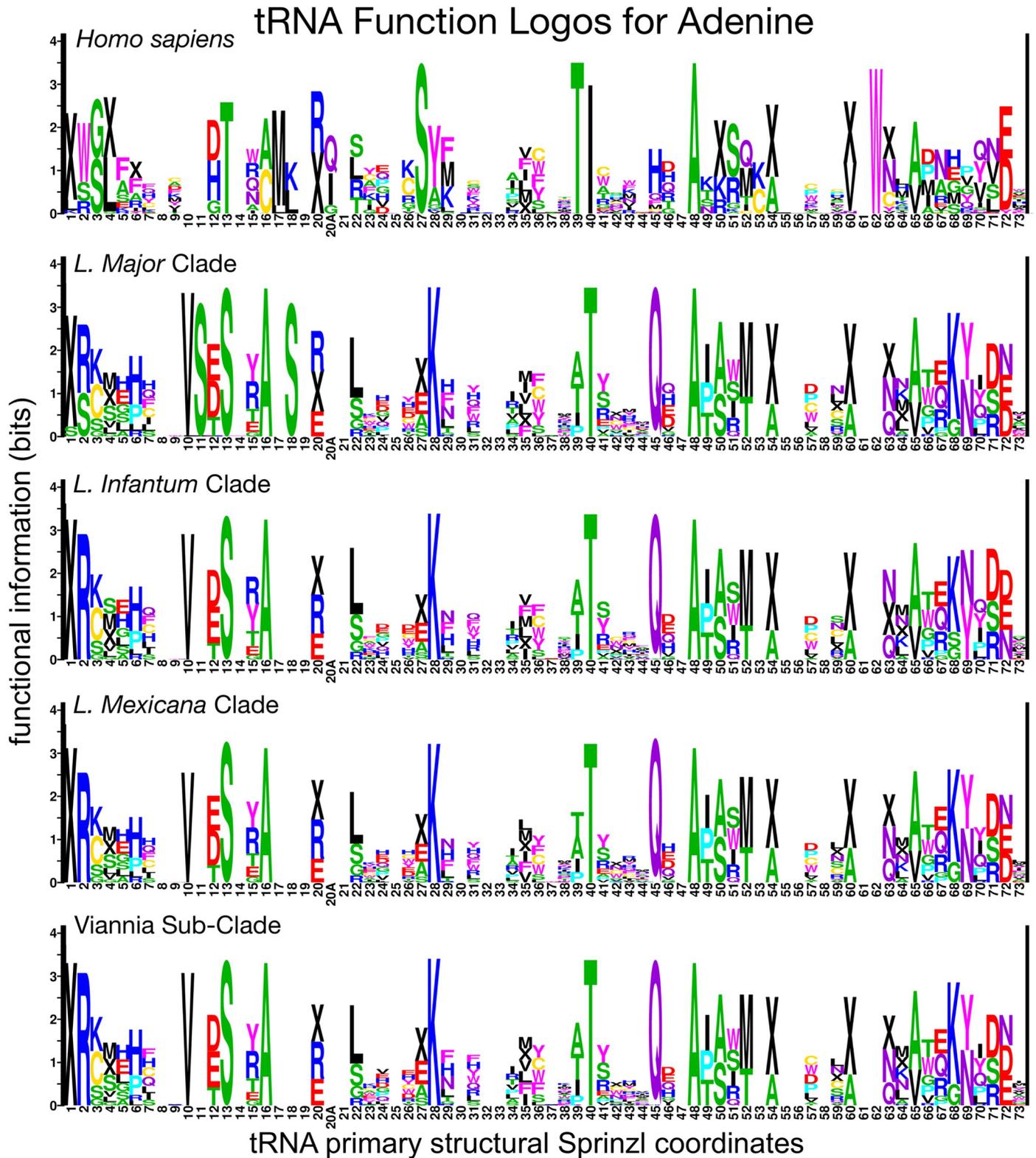


Fig 5. Adenine function logos for humans and four clades of *Leishmania*. Complete single-site function logo results are shown in S24–S31 Figs. The total height of a stack of letters at any site quantifies the information potentially gained about the functional type of a tRNA by a tRNA-binding protein if it recognizes the specific feature corresponding to that site and logo, for example Adenine at Sprinzl coordinate 16 (or some modification that biosynthetically depends on A16) in the case of the left-most boxed site. The letters within each stack symbolize functional types of tRNAs, wherein IUPAC one-letter amino acid codes represent elongator tRNA

aminoacylation identities and "X" symbolizes initiator tRNAs. The relative heights of letters within each stack quantifies the over-representation of tRNA functional types carrying that feature relative to the background frequency determined by gene frequencies of functional types (as calculated through the normalized log-odds).

<https://doi.org/10.1371/journal.pntd.0007983.g005>

AaRS screen identified *Leishmania major* AlaRS inhibitors

Using the pre-validated MNP library, we developed a medium-throughput phosphorimaging-based aminoacylation screen to identify aaRS inhibitors *in vitro* (Fig 9A and 9B). From the one hundred and twenty complex inhibitory mixes tested in the MNP library, we qualitatively identified four potential *Lm* AlaRS inhibitors as determined by a decrease in the overall tRNA-

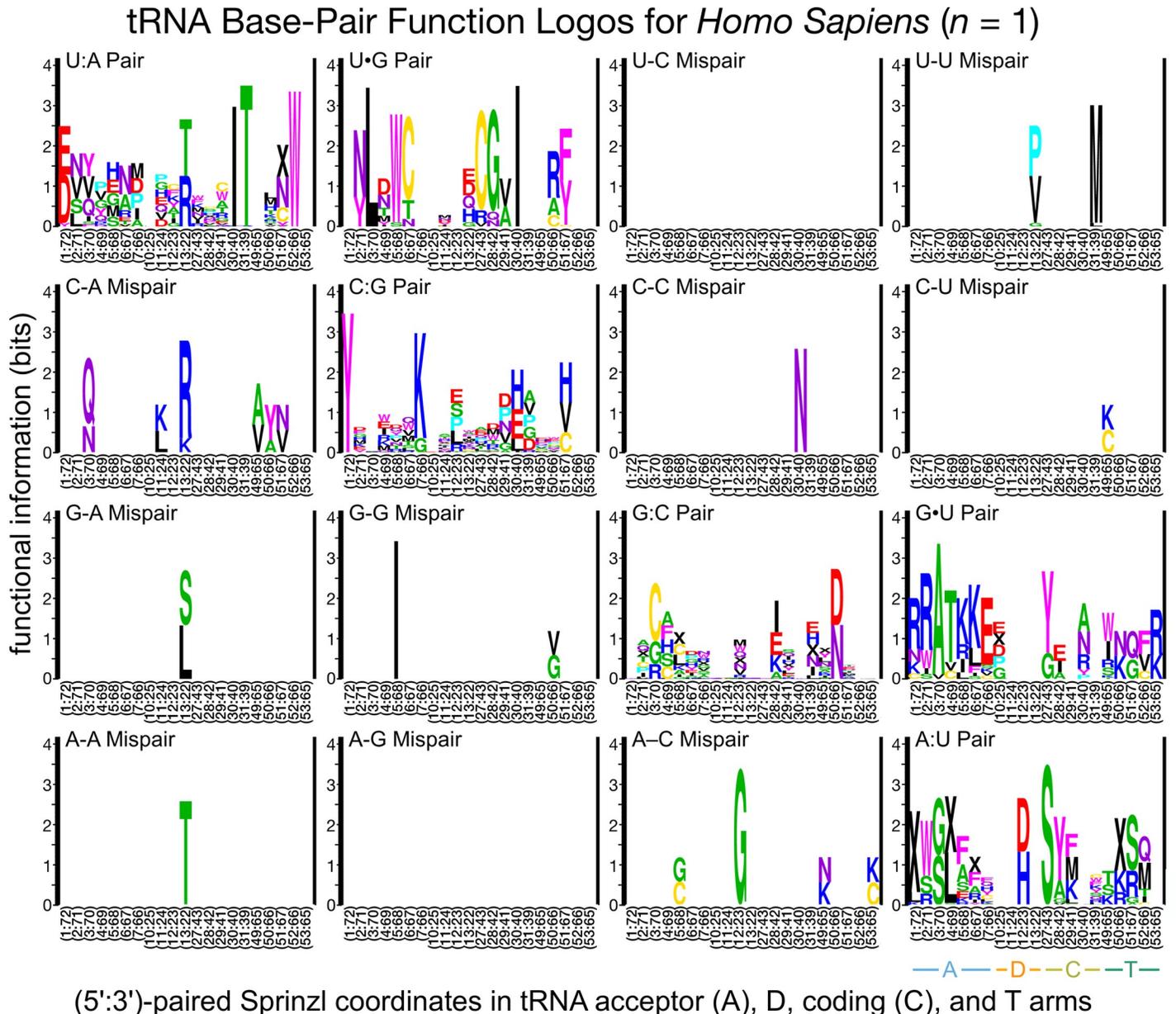
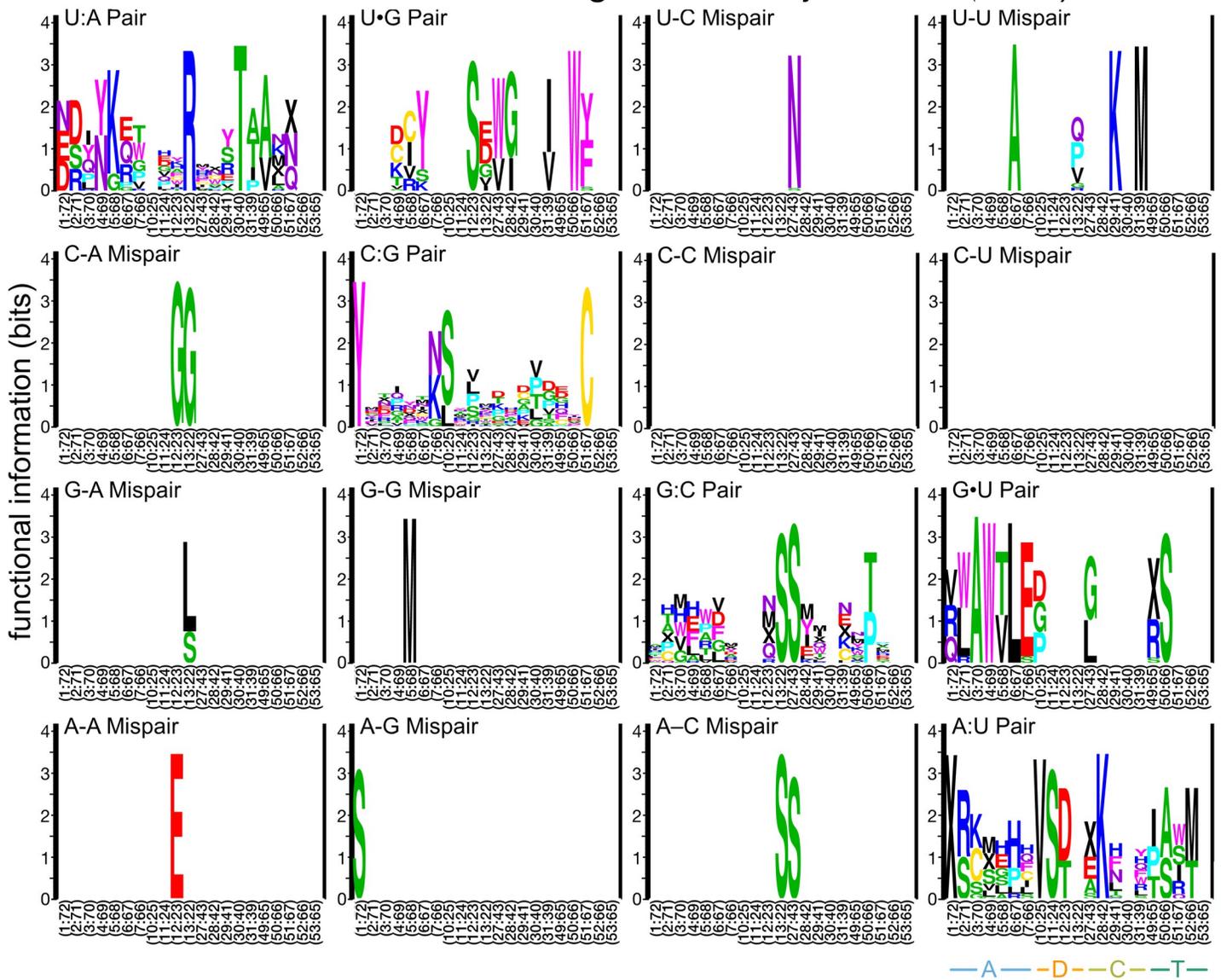


Fig 6. Function logos for tRNA Class-Informative Base-Pairs and Class-Informative Mis-Pairs in humans. The meanings of letters, stack heights and letter heights are all the same as in Fig 5.

<https://doi.org/10.1371/journal.pntd.0007983.g006>

tRNA Base-Pair Function Logos for *L. Major* Clade ($n = 8$)



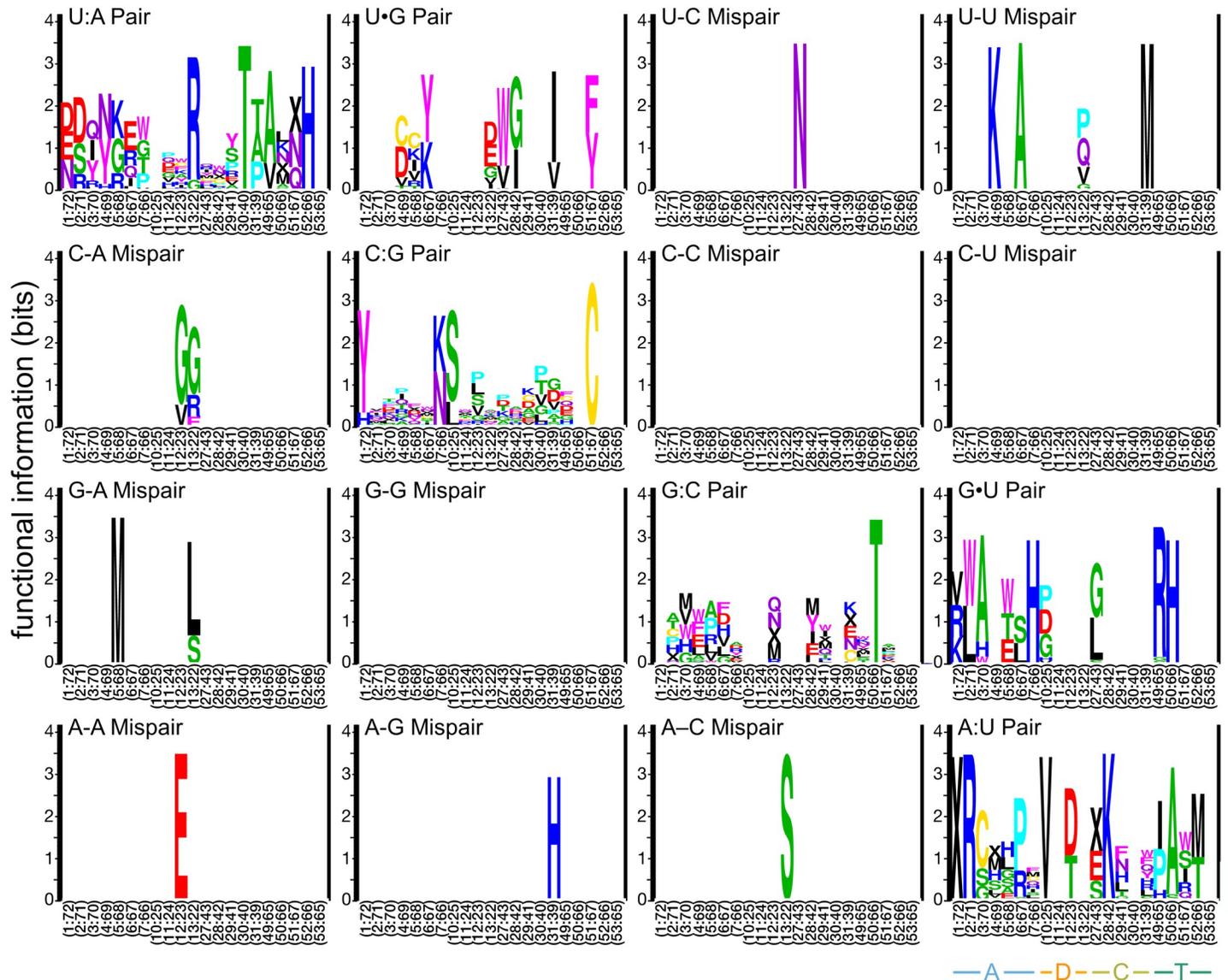
(5':3')-paired Sprinzl coordinates in tRNA acceptor (A), D, coding (C), and T arms

Fig 7. Function logos for tRNA Class-Informative Base-Pairs and Class-Informative Mis-Pairs in the *L. major* clade. The meanings of letters, stack heights and letter heights are all the same as in Fig 5.

<https://doi.org/10.1371/journal.pntd.0007983.g007>

aminoacylation signal. These four candidates were then re-screened using time-dependent quantitative approaches and we concluded that three of the four mixes, 1881C, 2059D, and 2096B were altering aminoacylation, with inhibitory activities ranging between 80% and 99% (Fig 9C). Using means and standard deviations of four replicate scintillation count-per-minute endpoints under the DMSO control condition or without added enzyme, we calculated an acceptable Z-factor of 0.67 for this follow-up aminoacylation time-course-based assay (S6 Table).

tRNA Base-Pair Function Logos for American *Trypanosoma* ($n = 11$)



(5':3')-paired Sprinzl coordinates in tRNA acceptor (A), D, coding (C), and T arms

Fig 8. Function logos for tRNA Class-Informative Base-Pairs and Class-Informative Mis-Pairs in the American *Trypanosoma* clade. The meanings of letters, stack heights and letter heights are all the same as in Fig 5.

<https://doi.org/10.1371/journal.pntd.0007983.g008>

Since the aminoacylation screen discerns total net changes to the aaRS activity, we attempted to identify which step of the aaRS catalyzed reaction is being affected by the MNPs. To observe any tRNA-independent effects on aaRS function, we used pyrophosphate exchange to monitor ATP-dependent amino acid activation. From this experiment, we were able to conclude that our inhibitors were perturbing amino acid activation, with lead compounds ranging in inhibitory activity between 45% and 95%. The differences in MNP activity between amino acid activation and tRNA-dependent aminoacylation highlight the multiple aaRS activities that can be targeted in our network predictions. To validate the predictive tool for identifying anti-trypanosomal drugs, we counter-screened the newly identified *Lm*. AlaRS inhibitors

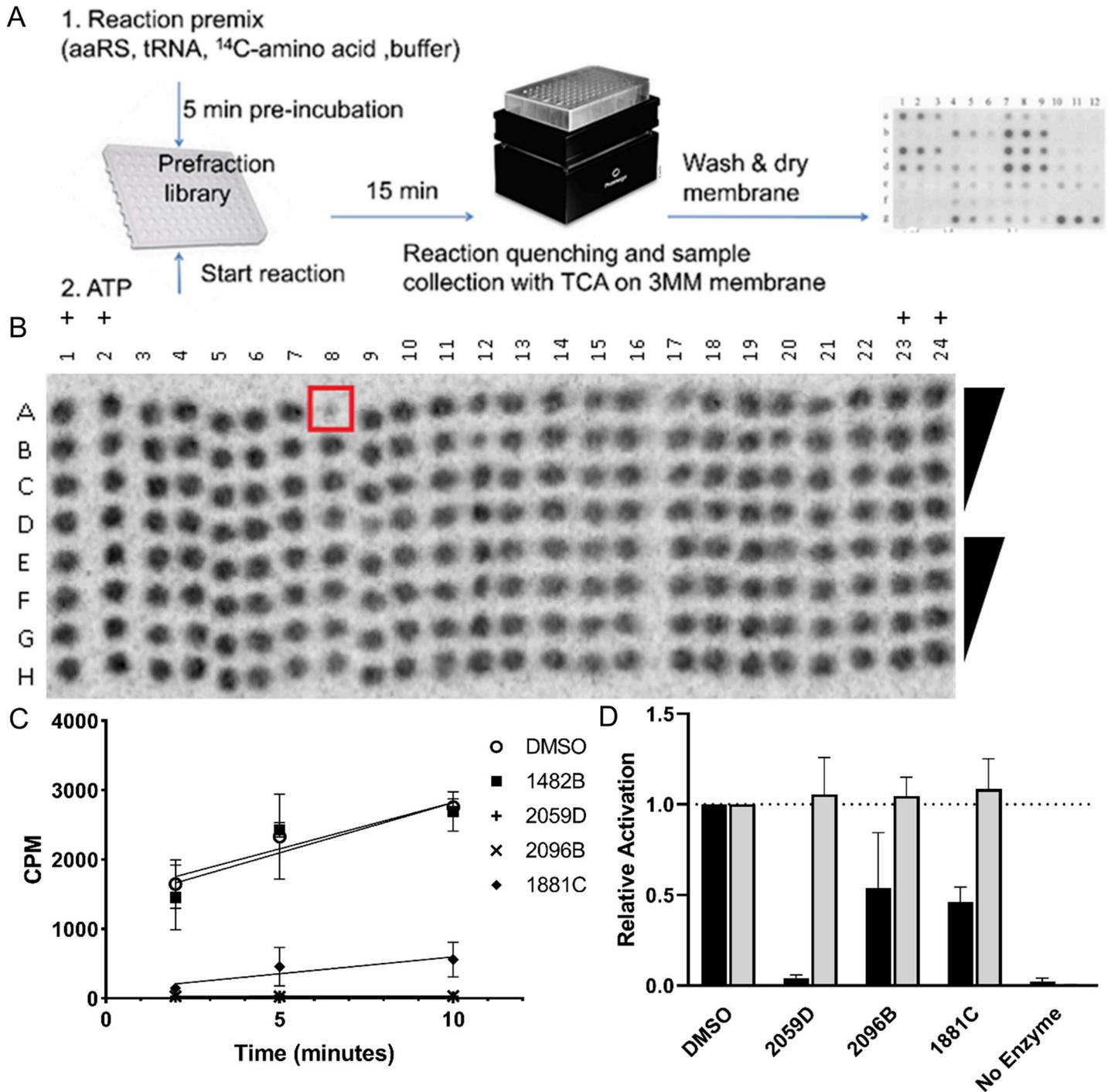


Fig 9. Identification of *Leishmania major* AlaRS inhibitors. A) Workflow to identify aminoacylation inhibitors (details described in Methods). B) Representative image of the MNP chemical screen. The spot boxed in red is an example of a predicted inhibitor depicted by the decrease in signal intensity. DMSO positive control (+). C) Three of the four identified inhibitors prevented the accumulation of Ala-tRNA^{Ala} formation, 1428B was a false-positive result from our preliminary screen. D) The three identified inhibitors perturbed *L. major* AlaRS activation (black) but had no effect on human AlaRS (gray). The relative amino acid activation is plotted relative to the DMSO control. Error bars indicate the standard deviation of three replicates.

<https://doi.org/10.1371/journal.pntd.0007983.g009>

against the human AlaRS enzyme. Treatment of the human AlaRS enzyme with the MNP inhibitors had no effect on amino acid activation (Fig 9D). Combined with our original screening data, these results show the utility of our computational and biochemical workflow to identify new novel therapeutics that have minimal cross-reactivity with the human homolog of the parasite drug target.

Natural product library inhibitors of *Leishmania major* ThrRS

As our network predictions identified CIF divergence among many functional classes of tRNAs (as shown in the Supplementary materials), we also wanted to determine if our MNP library would find inhibitors against non-AlaRS aaRS. The most concentrated mixes from our MNP library were re-screened against *Lm* Threonyl-tRNA synthetase (ThrRS) and tRNA^{Thr} aminoacylation (Fig 10A). The preliminary screen led to the identification of eight extracts with inhibitory activity. Those fractions were re-analyzed using quantitative aminoacylation reactions and the results show that that two of the candidates did not inhibit aminoacylation, two inhibited the reaction at ~50%, and four had greater than 75% inhibition (Fig 10B). In

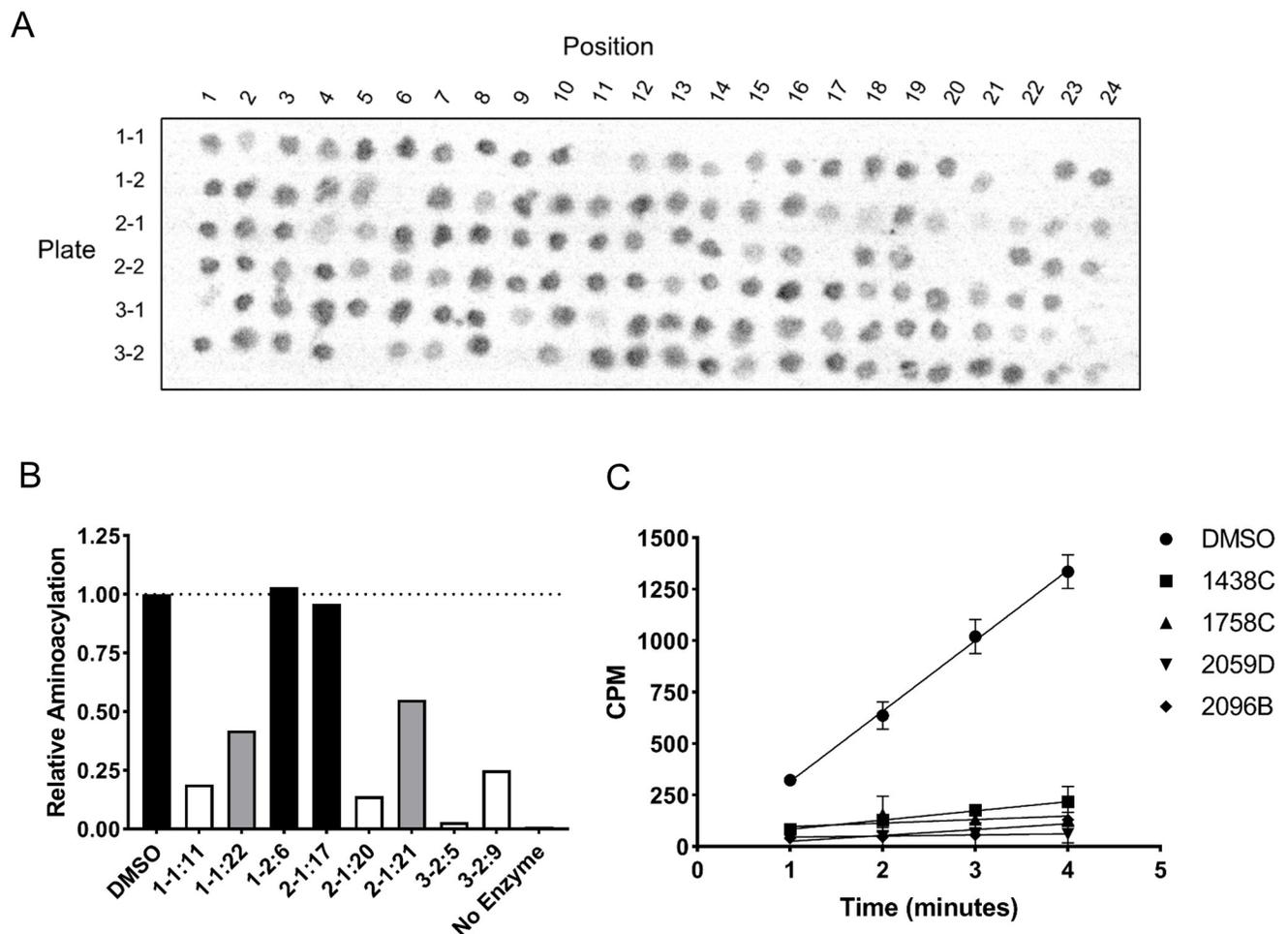


Fig 10. Identification of *Leishmania major* ThrRS inhibitors. A) The MNP library was re-screened at the highest concentrations to qualitatively identify *Lm* ThrRS aminoacylation inhibitors. Plate IDs reference the position within the original library and not library IDs. B) Eight inhibitors were qualitatively identified from the preliminary screen. Two of the candidates did not inhibit aminoacylation (black), two inhibited at ~50% activity (gray), and four inhibited at greater than 25% (white). C) All four active inhibitors continued to perturb aminoacylation over a time course experiment. Error bars indicate the standard deviation of three replicates.

<https://doi.org/10.1371/journal.pntd.0007983.g010>

addition, two of the four most active inhibitors (2059D and 2096B) also had activity against *Lm* AlaRS (Fig 10C). The cross-reactivity of these inhibitors may be a consequence of the extensively conserved aaRS architecture found between AlaRS and ThrRS [63,64].

Predictive network interactions identified broad-spectrum anti-trypanosomal targets

The tRNA-aaRS network analyses suggested that parasite-specific tRNA^{Ala} identity elements were highly conserved between the *Leishmania* and *Trypanosoma* genera (Figs 3,4,5 and 11A). To test this hypothesis, we purified *Tc* AlaRS and screened our three active *Lm* AlaRS inhibitors in an aminoacylation inhibition assay using *Tc* AlaRS and tRNA^{Ala}. Supporting our network prediction, all three *Lm* inhibitors also had activity against the *Tc* enzyme, with activities ranging between 40% and 95% total inhibition (Fig 11B). While these activities were slightly reduced compared to their effect of the *Lm* AlaRS enzyme (Fig 9C), these results highlight the additional potential utility of our computational methodologies as a means of identifying broad-spectrum antimicrobials for closely-related clades.

Separation of active components from natural products extracts

From the initial set of 120 extracts with activity against *L. donovani* parasites, four extracts showed corresponding activity in the initial aaRS assay. Of these, three (1881C, 2059D, and 2096B) were prioritized for chemical follow up, based on potent, dose-dependent biological activity. Initially, each sample was separated into 10 sub-fractions using HPLC (Phenomenex Synergy C18, 5 μ , 4.6 x 250 mm). Screening of these fractions identified one fraction (2096B F10) with potent activity. To generate additional material, the producing organism was cultured on large scale (1 L, GNZ medium with 20 g XAD-7 resin), filtered, and the resin/cell slurry extracted with organic solvents (2:1 CH₂Cl₂/ MeOH, 400 mL). The crude extract was fractionated using an automated Combiflash chromatography system (C₁₈ cartridge; 20, 40, 60, 80, 100% MeOH/ H₂O, 100% EtOAc) and the resulting fractions subjected to biological screening (S1 Text). Two fractions (C and D) showed strong activity and were subjected to subsequent separation to give 10 sub-fractions each (S41–S44 Figs). Of these, fraction 2096D F10 showed the strongest reproducible activity (S45 Fig). However, subsequent fractionation steps yielded sub-fractions with very low quantities of material. Review of these sub-fractions by UPLC-ESI-qTOF mass spectrometry did not identify any individual mass signatures consistent with a candidate bioactive molecule. Similarly, lack of material precluded the identification of diagnostic signals in the NMR spectra for these subfractions. Provisional information from these analyses, including NMR and MS signatures from earlier fractions and the non-polar nature of the active fractions, suggest that the active component is likely a bioactive lipid, although the precise nature of the structure of this metabolite remains unknown. The isolate producing the bioactive substance was collected on April 20th, 2012 from marine sediment off the coast of Kellet Bluff, Henry Island, WA US under the permit issuing authority of the Washington Department of Fish and Wildlife (permit # 12–034).

Discussion

Systems-biology driven identification of trypanosome-specific drug targets

tRNA CIFs apply an information criterion using function logos, rather than a conservation criterion using conventional sequence logos, to bioinformatically predict tRNA identity elements. Even though we did not apply a conservation criterion in our predictions, when we applied our information criterion independently over different trypanosome clades, we found that tRNA CIFs were

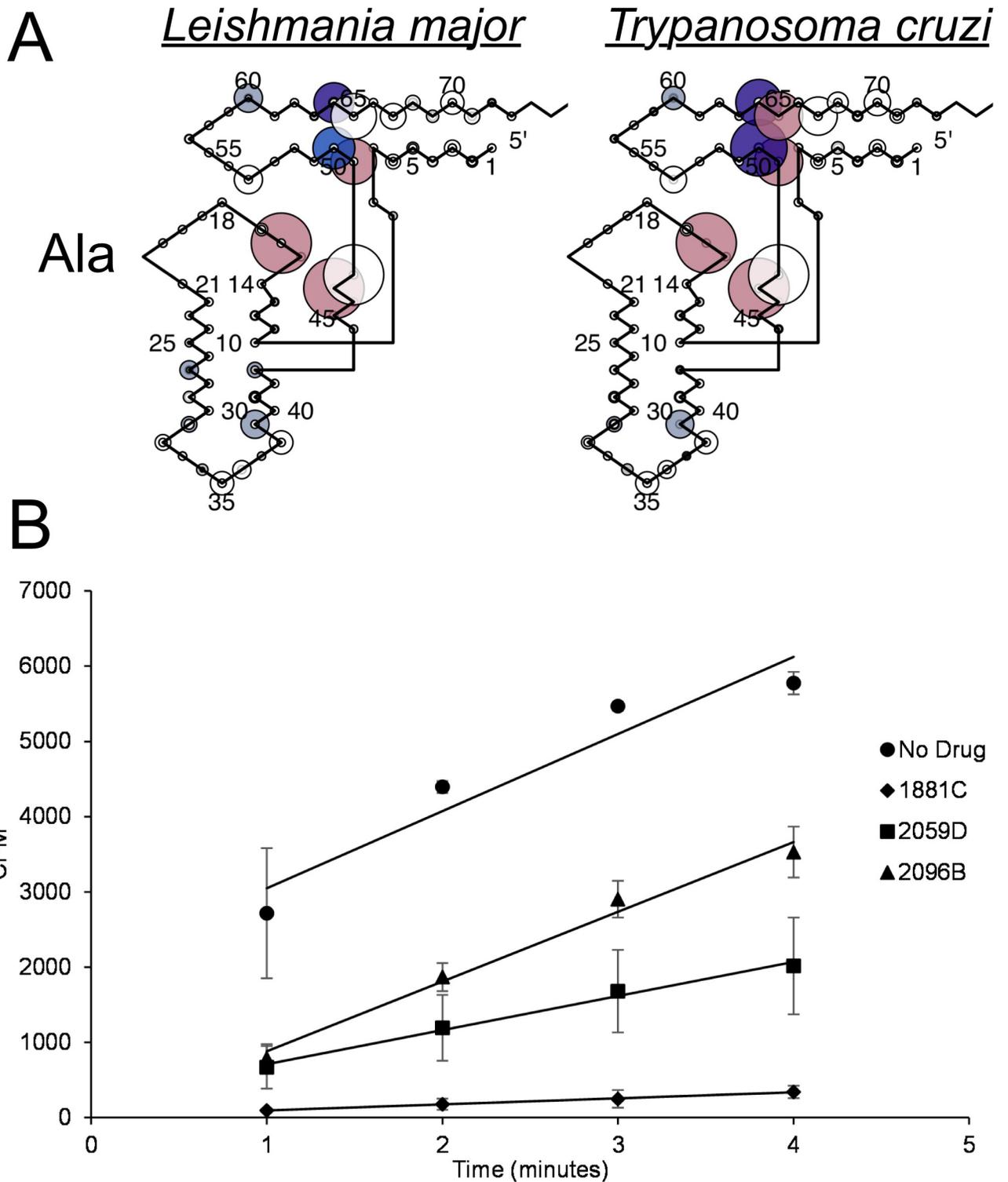


Fig 11. *Leishmania major* and *Trypanosoma cruzi* AlaRS have conserved tRNA identity elements. A) CIF Divergence Models for tRNA^{Ala} in *Leishmania major* and *Trypanosoma cruzi* B) The three identified *Lm* AlaRS inhibitors also have activity against the *Tc* AlaRS enzyme. Error bars indicate the standard deviation of three replicates.

<https://doi.org/10.1371/journal.pntd.0007983.g011>

highly conserved over 250 million years of trypanosome evolution. A biological interpretation of this result of tRNA CIF conservation within trypanosomes (and also between trypanosomes and humans) is that the information contained in tRNA CIFs is functional in specifying substrate identity to tRNA-binding proteins such as aaRSs. That is to say, tRNA-binding proteins themselves exploit the information contained in tRNA CIFs to identify their tRNA substrates against the background of all possible tRNAs, with which they must interact to varying degrees. We present a systems biological theory for the function and divergence of tRNA CIFs in [24].

Maintaining efficient and accurate translation is predicated on catalytically productive interactions between aaRSs and free tRNAs in the cell. While the major identity elements for a given aaRS-tRNA pair are generally conserved, here we have identified divergent features within tRNAs that apparently contribute to divergent RNA-protein interactions in trypanosomes. Much of the focus in this work was on the phylogenetic divergence of identity elements among alanine tRNAs. This class of tRNAs strongly support the utility of our computational analyses as the tRNA^{Ala} identity elements have been one of the most well characterized to date [13,14]. Interestingly, it was recently shown that the conserved G3:U70 base pair is recognized by AlaRS using three distinct mechanisms across all domains of life [15]. This observation highlights that even highly conserved identity elements may be recognized and discriminated against by distinct biophysical aaRS interactions, which may therefore be stronger potentially specific drug targets than previously anticipated. The dominant association of G3:U70 with tRNA^{Ala} is conserved among all trypanosome clades and humans in our data (Figs 6–8 and S32–S40 Figs).

The primary objective of this research was to develop a computational workflow to quantify divergence of functionally informative features of tRNAs across different evolutionary clades. The practical application of this work is to use the information gained from our computational analyses to identify novel therapeutic targets that may be of use in the clinic. As described above, tRNA^{Ala} and tRNA^{Thr} were specifically chosen because of their amenability for *in vitro* reconstitution, while the computational results shown in Supplementary Figures show that other leishmanial tRNA/aaRS pairs could serve as additional therapeutic targets either using our MNP library or other available libraries. While interesting, those discoveries are outside the scope of the present work and left to future investigations.

Inhibition of aminoacyl-tRNA synthetases

A goal of this work was to identify divergent tRNA identity elements in trypanosome parasites. We predicted that parasite-specific tRNA-aaRS interactions would be identified, sufficiently divergent from homologous human machinery to be strong candidates for drug discovery. Interestingly, our network divergence analysis led to the discovery of tRNA-independent, amino acid activation inhibitors that were specific to trypanosomes. We interpret this as consistent with our goal, because tRNAs and aaRSs must coevolve to accommodate changes to structure and mechanism that evolve on either side of their interactions. Presumably, divergence in the structural mechanism of amino acid activation in trypanosome AlaRSs has also changed how they interact with their tRNA substrates. By integrating information from many tRNA functional classes, we gain leverage to interpret divergence in structure and function of the much more structurally complex ensemble of aaRSs as a system. Our tRNA-based network approach identifies potential aaRS targets that may not have been initially predicted when analyzing aaRS functional sequences in isolation.

Chemotherapeutic inhibition of multiple aminoacyl-tRNA synthetases may be relatively resistance-proof

Two of the fractions we described were effective inhibitors of both AlaRS and ThrRS in parasites. Although monotherapeutic inhibitors of aaRSs are highly effective [65], combination

therapies involving multiple aaRSs have not been studied. Because aminoacylation pathways are integrated in parallel at the ribosome, the slowest aminoacylation pathway can be rate-limiting for protein synthesis and growth [24,66]. Thus, we expect the inhibition of multiple aaRSs to be antagonistic relative to Loewe Additivity expectations, in keeping with the Highest Single Agent (HSA) model [67,68]: single- or multiple-drug inhibition of multiple aaRSs should mask the potentially growth-restorative effects of resistance mutations arising in any one parasite aaRS gene. It is known that antagonistic combination chemotherapies are less prone to the evolution of resistance than single-drug or synergistic combination chemotherapies [32,69,70,71]. Therefore, chemotherapeutic inhibition of multiple aaRSs should be relatively less prone to the evolution of resistance than monotherapeutic or synergistic combination chemotherapeutic inhibition of single aaRSs. Further work is needed to test this hypothetical benefit.

Conclusion

Trypanosome parasites pose a significant health risk worldwide. While current therapies exist, they are often also accompanied by off-target cytotoxicity and can lead to the rise of antimicrobial resistance. Here we have demonstrated that targeting tRNA-synthetase interactions have been an underexplored avenue for drug discovery. Using a combination of predictive computational tRNA network analyses and biochemical validation, we showed that aminoacyl-tRNA synthetases are a promising target for broad-spectrum anti-trypanosomal discovery with no significant consequence to the human counterpart target.

Supporting information

S1 Fig. Aragorn Score Distributions of Singleton vs. Co-Clustered Ara-Only Genes.
(PDF)

S2 Fig. Annotated and Missing Functional Classes in Union Annotation Gene Set.
(PDF)

S3 Fig. Bubbleplots of CIF divergence between humans and eight clades of trypanosomes for tRNA^{Leu}.
(PDF)

S4 Fig. Bubbleplots of CIF divergence between humans and eight clades of trypanosomes for tRNA^{Ile}.
(PDF)

S5 Fig. Bubbleplots of CIF divergence between humans and eight clades of trypanosomes for tRNA^{Val}.
(PDF)

S6 Fig. Bubbleplots of CIF divergence between humans and eight clades of trypanosomes for tRNA^{Arg}.
(PDF)

S7 Fig. Bubbleplots of CIF divergence between humans and eight clades of trypanosomes for tRNA^{Cys}.
(PDF)

S8 Fig. Bubbleplots of CIF divergence between humans and eight clades of trypanosomes for tRNA^{Met}.
(PDF)

S9 Fig. Bubbleplots of CIF divergence between humans and eight clades of trypanosomes for tRNA^{Glu}.

(PDF)

S10 Fig. Bubbleplots of CIF divergence between humans and eight clades of trypanosomes for tRNA^{Gln}.

(PDF)

S11 Fig. Bubbleplots of CIF divergence between humans and eight clades of trypanosomes for tRNA^{Tyr}.

(PDF)

S12 Fig. Bubbleplots of CIF divergence between humans and eight clades of trypanosomes for tRNA^{Trp}.

(PDF)

S13 Fig. Bubbleplots of CIF divergence between humans and eight clades of trypanosomes for tRNA^{Ser}.

(PDF)

S14 Fig. Bubbleplots of CIF divergence between humans and eight clades of trypanosomes for tRNA^{Thr}.

(PDF)

S15 Fig. Bubbleplots of CIF divergence between humans and eight clades of trypanosomes for tRNA^{Pro}.

(PDF)

S16 Fig. Bubbleplots of CIF divergence between humans and eight clades of trypanosomes for tRNA^{His}.

(PDF)

S17 Fig. Bubbleplots of CIF divergence between humans and eight clades of trypanosomes for tRNA^{Gly}.

(PDF)

S18 Fig. Bubbleplots of CIF divergence between humans and eight clades of trypanosomes for tRNA^{Asp}.

(PDF)

S19 Fig. Bubbleplots of CIF divergence between humans and eight clades of trypanosomes for tRNA^{Asn}.

(PDF)

S20 Fig. Bubbleplots of CIF divergence between humans and eight clades of trypanosomes for tRNA^{Lys}.

(PDF)

S21 Fig. Bubbleplots of CIF divergence between humans and eight clades of trypanosomes for tRNA^{Phe}.

(PDF)

S22 Fig. Bubbleplots of CIF divergence between humans and eight clades of trypanosomes for tRNA^{Ala}.

(PDF)

S23 Fig. Bubbleplots of CIF divergence between humans and eight clades of trypanosomes for tRNA^{iMet}.

(PDF)

S24 Fig. Single-site function logos for Adenine in all clades, part I.

(PNG)

S25 Fig. Single-site function logos for Adenine in all clades, part II.

(PNG)

S26 Fig. Single-site function logos for Uracil in all clades, part I.

(PNG)

S27 Fig. Single-site function logos for Uracil in all clades, part II.

(PNG)

S28 Fig. Single-site function logos for Guanine in all clades, part I.

(PNG)

S29 Fig. Single-site function logos for Guanine in all clades, part II.

(PNG)

S30 Fig. Single-site function logos for Cytosine in all clades, part I.

(PNG)

S31 Fig. Single-site function logos for Cytosine in all clades, part II.

(PNG)

S32 Fig. Paired-site function logos for humans.

(PNG)

S33 Fig. Paired-site function logos for *L. major* clade.

(PNG)

S34 Fig. Paired-site function logos for *L. infantum* clade.

(PNG)

S35 Fig. Paired-site function logos for *L. mexicana* clade.

(PNG)

S36 Fig. Paired-site function logos for Viannia subclade.

(PNG)

S37 Fig. Paired-site function logos for *L. enriettii* clade.

(PNG)

S38 Fig. Paired-site function logos for *Leptomonas/Crithidia* clade.

(PNG)

S39 Fig. Paired-site function logos for American *Trypanosoma* clade.

(PNG)

S40 Fig. Paired-site function logos for African *Trypanosoma* clade.

(PNG)

S41 Fig. UPLC-qTOF base peak chromatogram of (A) RL12-182-HVF-D Sep-Pak Fraction C Subfraction 10–9 and (B) extracted ion chromatogram of peak 316.28 *m/z* in active region of

trace.

(PDF)

S42 Fig. UPLC-qTOF mass spectrum for peak at 316.28 m/z.

(PDF)

S43 Fig. Schematic diagram of the large-scale fermentation and extraction of RL12-182-HVF-D.

(PDF)

S44 Fig. HPLC-UV trace of the final preparative isolation step for the RL12-182-HVF-D active compound. Red box denotes active peak.

(PDF)

S45 Fig. Marine natural product extract 2096D F10 inhibits *Leishmania major* AlaRS aminoacylation.

(PDF)

S1 Table. Similar and putatively homologous tRNA gene cluster variant groups of length at least three occurring in at least two *Leishmania* genomes.

(PDF)

S2 Table. Similar and putatively homologous tRNA gene cluster groups of length at least three occurring in at least two *Trypanosoma* genomes.

(PDF)

S3 Table. Similar and putatively homologous tRNA gene cluster groups of length at least three spanning both *Leishmania* (L) and *Trypanosoma* (T) genera genomes. Only groups 1, 3 and 4 contain five gene cluster variants conserved in both genera, namely, ASD, DSA, LXP (two variants), and QLL, where “X” represents the initiator iMet tRNA gene.

(PDF)

S4 Table. Gene length, structure and functional type statistics on final annotation gene sets.

(PDF)

S5 Table. Nucleotide Composition of TriTryp tRNA Gene-Sets by Individual Genome Assembly, Organized by Clade (alternating background in the order of [Table 1](#), boldface) or Excluded (at end of table in roman font).

(PDF)

S6 Table. Values of four replicate count-per-minute endpoints (taken at time $t = 10$ min. in time-course measurements as in [Fig 9C](#)) of *Leishmania major* AlaRS under DMSO or without added enzyme, their means and standard deviations, and their corresponding Z-factor in the scintillation counter-based aminoacylation time-course assay.

(PDF)

S1 Code and Data. Code and data to reproduce computational results.

(TGZ)

S1 Text. Supplementary methods for bacterial fermentation, natural product extraction, and active compound identification using HPLC-UV-MS and UPLC-ESI-qTOF-MS.

(PDF)

Acknowledgments

The authors would like to thank Dr. Tammy Bullwinkle for assistance in generating expression constructs for this work and Amin Boroomand for initial work on annotation of TriTrypDB genomes.

Author Contributions

Conceptualization: Roger G. Linington, Michael Ibba, David H. Ardell.

Data curation: Paul Kelly, Fatemeh Hadi-Nezhad, David H. Ardell.

Formal analysis: Fatemeh Hadi-Nezhad, David H. Ardell.

Funding acquisition: Roger G. Linington, Michael Ibba, David H. Ardell.

Investigation: Paul Kelly, Fatemeh Hadi-Nezhad, Dennis Y. Liu, Roger G. Linington, David H. Ardell.

Methodology: Paul Kelly, Fatemeh Hadi-Nezhad, Dennis Y. Liu, Travis J. Lawrence, Roger G. Linington, Michael Ibba, David H. Ardell.

Project administration: Roger G. Linington, Michael Ibba, David H. Ardell.

Resources: Paul Kelly, Fatemeh Hadi-Nezhad, Dennis Y. Liu, Travis J. Lawrence, Roger G. Linington, Michael Ibba, David H. Ardell.

Software: Fatemeh Hadi-Nezhad, Travis J. Lawrence, David H. Ardell.

Supervision: Roger G. Linington, Michael Ibba, David H. Ardell.

Validation: Paul Kelly, Fatemeh Hadi-Nezhad, Dennis Y. Liu, Travis J. Lawrence, Michael Ibba, David H. Ardell.

Visualization: Paul Kelly, Fatemeh Hadi-Nezhad, Dennis Y. Liu, Travis J. Lawrence, Michael Ibba, David H. Ardell.

Writing – original draft: Paul Kelly, Fatemeh Hadi-Nezhad, Dennis Y. Liu, Travis J. Lawrence, Roger G. Linington, Michael Ibba, David H. Ardell.

Writing – review & editing: Paul Kelly, Fatemeh Hadi-Nezhad, Travis J. Lawrence, Roger G. Linington, Michael Ibba, David H. Ardell.

References

1. Fairlamb AH, Gow NA, Matthews KR, Waters AP (2016) Drug resistance in eukaryotic microorganisms. *Nat Microbiol* 1: 16092. <https://doi.org/10.1038/nmicrobiol.2016.92> PMID: 27572976
2. World Health O (2010) Control of the leishmaniasis. *World Health Organ Tech Rep Ser*: xii-xiii, 1–186, back cover.
3. Crowther GJ, Shanmugam D, Carmona SJ, Doyle MA, Hertz-Fowler C, et al. (2010) Identification of attractive drug targets in neglected-disease pathogens using an in silico approach. *PLoS Negl Trop Dis* 4: e804. <https://doi.org/10.1371/journal.pntd.0000804> PMID: 20808766
4. Polonio T, Efferth T (2008) Leishmaniasis: drug resistance and natural products (review). *Int J Mol Med* 22: 277–286. PMID: 18698485
5. Handman E, Kedzierski L, Uboldi AD, Goding JW (2008) Fishing for anti-leishmania drugs: principles and problems. *Adv Exp Med Biol* 625: 48–60. https://doi.org/10.1007/978-0-387-77570-8_5 PMID: 18365658
6. Singh S, Sivakumar R (2004) Challenges and new discoveries in the treatment of leishmaniasis. *J Infect Chemother* 10: 307–315. <https://doi.org/10.1007/s10156-004-0348-9> PMID: 15614453
7. Davis AJ, Kedzierski L (2005) Recent advances in antileishmanial drug development. *Curr Opin Investig Drugs* 6: 163–169. PMID: 15751739

8. Ho JM, Bakkalbasi E, Soll D, Miller CA (2018) Drugging tRNA aminoacylation. *RNA Biol* 15: 667–677. <https://doi.org/10.1080/15476286.2018.1429879> PMID: 29345185
9. Ibba M, Soll D (2000) Aminoacyl-tRNA synthesis. *Annu Rev Biochem* 69: 617–650. <https://doi.org/10.1146/annurev.biochem.69.1.617> PMID: 10966471
10. Salazar JC, Ahel I, Orellana O, Tumbula-Hansen D, Krieger R, et al. (2003) Coevolution of an aminoacyl-tRNA synthetase with its tRNA substrates. *Proc Natl Acad Sci U S A* 100: 13863–13868. <https://doi.org/10.1073/pnas.1936123100> PMID: 14615592
11. Giege R, Sissler M, Florentz C (1998) Universal rules and idiosyncratic features in tRNA identity. *Nucleic Acids Res* 26: 5017–5035. <https://doi.org/10.1093/nar/26.22.5017> PMID: 9801296
12. Ardell DH (2010) Computational analysis of tRNA identity. *FEBS Lett* 584: 325–333. <https://doi.org/10.1016/j.febslet.2009.11.084> PMID: 19944694
13. Hou YM, Schimmel P (1989) Evidence that a major determinant for the identity of a transfer RNA is conserved in evolution. *Biochemistry* 28: 6800–6804. <https://doi.org/10.1021/bi00443a003> PMID: 2684266
14. McClain WH, Foss K (1988) Changing the identity of a tRNA by introducing a G-U wobble pair near the 3' acceptor end. *Science* 240: 793–796. <https://doi.org/10.1126/science.2452483> PMID: 2452483
15. Chong YE, Guo M, Yang XL, Kuhle B, Naganuma M, et al. (2018) Distinct ways of G:U recognition by conserved tRNA binding motifs. *Proc Natl Acad Sci U S A* 115: 7527–7532. <https://doi.org/10.1073/pnas.1807109115> PMID: 29967150
16. Silvian LF, Wang J, Steitz TA (1999) Insights into editing from an ile-tRNA synthetase structure with tRNA^{ile} and mupirocin. *Science* 285: 1074–1077. PMID: 10446055
17. Zhou H, Sun L, Yang XL, Schimmel P (2013) ATP-directed capture of bioactive herbal-based medicine on human tRNA synthetase. *Nature* 494: 121–124. <https://doi.org/10.1038/nature11774> PMID: 23263184
18. Rock FL, Mao W, Yaremchuk A, Tukalo M, Crepin T, et al. (2007) An antifungal agent inhibits an aminoacyl-tRNA synthetase by trapping tRNA in the editing site. *Science* 316: 1759–1761. <https://doi.org/10.1126/science.1142189> PMID: 17588934
19. Seiradake E, Mao W, Hernandez V, Baker SJ, Plattner JJ, et al. (2009) Crystal structures of the human and fungal cytosolic Leucyl-tRNA synthetase editing domains: A structural basis for the rational design of antifungal benzoxaboroles. *J Mol Biol* 390: 196–207. <https://doi.org/10.1016/j.jmb.2009.04.073> PMID: 19426743
20. Merritt EA, Arakaki TL, Gillespie JR, Larson ET, Kelley A, et al. (2010) Crystal structures of trypanosomal histidyl-tRNA synthetase illuminate differences between eukaryotic and prokaryotic homologs. *J Mol Biol* 397: 481–494. <https://doi.org/10.1016/j.jmb.2010.01.051> PMID: 20132829
21. Larson ET, Kim JE, Zucker FH, Kelley A, Mueller N, et al. (2011) Structure of *Leishmania major* methionyl-tRNA synthetase in complex with intermediate products methionyladenylate and pyrophosphate. *Biochimie* 93: 570–582. <https://doi.org/10.1016/j.biochi.2010.11.015> PMID: 21144880
22. Larson ET, Kim JE, Castaneda LJ, Napuli AJ, Zhang Z, et al. (2011) The double-length tyrosyl-tRNA synthetase from the eukaryote *Leishmania major* forms an intrinsically asymmetric pseudo-dimer. *J Mol Biol* 409: 159–176. <https://doi.org/10.1016/j.jmb.2011.03.026> PMID: 21420975
23. Huang W, Zhang Z, Ranade RM, Gillespie JR, Barros-Alvarez X, et al. (2017) Optimization of a binding fragment targeting the "enlarged methionine pocket" leads to potent *Trypanosoma brucei* methionyl-tRNA synthetase inhibitors. *Bioorg Med Chem Lett* 27: 2702–2707. <https://doi.org/10.1016/j.bmcl.2017.04.048> PMID: 28465105
24. Collins-Hed AI, Ardell DH (2019) Match fitness landscapes for macromolecular interaction networks: Selection for translational accuracy and rate can displace tRNA-binding interfaces of non-cognate aminoacyl-tRNA synthetases. *Theor Popul Biol*.
25. Freyhult E, Moulton V, Ardell DH (2006) Visualizing bacterial tRNA identity determinants and antideterminants using function logos and inverse function logos. *Nucleic Acids Res* 34: 905–916. <https://doi.org/10.1093/nar/gkj478> PMID: 16473848
26. Schneider TD (2010) A brief review of molecular information theory. *Nano Commun Netw* 1: 173–180. <https://doi.org/10.1016/j.nancom.2010.09.002> PMID: 22110566
27. Freyhult E, Cui Y, Nilsson O, Ardell DH (2007) New computational methods reveal tRNA identity element divergence between Proteobacteria and Cyanobacteria. *Biochimie* 89: 1276–1288. <https://doi.org/10.1016/j.biochi.2007.07.013> PMID: 17889982
28. Fitzgerald JB, Schoeberl B, Nielsen UB, Sorger PK (2006) Systems biology and combination therapy in the quest for clinical efficacy. *Nat Chem Biol* 2: 458–466. <https://doi.org/10.1038/nchembio817> PMID: 16921358

29. Stumpf MP, Robertson BD, Duncan K, Young DB (2007) Systems biology and its impact on anti-infective drug development. *Prog Drug Res* 64: 1, 3–20. https://doi.org/10.1007/978-3-7643-7567-6_1 PMID: 17195469
30. Russell RB, Aloy P (2008) Targeting and tinkering with interaction networks. *Nat Chem Biol* 4: 666–673. <https://doi.org/10.1038/nchembio.119> PMID: 18936751
31. Berger SI, Iyengar R (2009) Network analyses in systems pharmacology. *Bioinformatics* 25: 2466–2472. <https://doi.org/10.1093/bioinformatics/btp465> PMID: 19648136
32. Anderson JB (2005) Evolution of antifungal-drug resistance: mechanisms and pathogen fitness. *Nat Rev Microbiol* 3: 547–556. <https://doi.org/10.1038/nrmicro1179> PMID: 15953931
33. Kroiss J, Kaltenpoth M, Schneider B, Schwinger MG, Hertweck C, et al. (2010) Symbiotic Streptomyces provide antibiotic combination prophylaxis for wasp offspring. *Nat Chem Biol* 6: 261–263. <https://doi.org/10.1038/nchembio.331> PMID: 20190763
34. Jansen G, Lee AY, Epp E, Fredette A, Surprenant J, et al. (2009) Chemogenomic profiling predicts antifungal synergies. *Mol Syst Biol* 5: 338. <https://doi.org/10.1038/msb.2009.95> PMID: 20029371
35. Hu Y, Platzer EG, Bellier A, Aroian RV (2010) Discovery of a highly synergistic anthelmintic combination that shows mutual hypersusceptibility. *Proc Natl Acad Sci U S A* 107: 5955–5960. <https://doi.org/10.1073/pnas.0912327107> PMID: 20231450
36. Organization WH (2015) Guidelines for the treatment of malaria: World Health Organization.
37. Organization WH (2018) Artemisinin resistance and artemisinin-based combination therapy efficacy: status report. World Health Organization.
38. Lowe TM, Chan PP (2016) tRNAscan-SE On-line: integrating search and context for analysis of transfer RNA genes. *Nucleic Acids Res* 44: W54–57. <https://doi.org/10.1093/nar/gkw413> PMID: 27174935
39. Laslett D, Canback B (2004) ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences. *Nucleic Acids Res* 32: 11–16. <https://doi.org/10.1093/nar/gkh152> PMID: 14704338
40. Levenshtein VI. Binary codes capable of correcting deletions, insertions, and reversals. *Sov. Phys. Dokl.* 10, 8, 707–710. Original in Russian in *Dokl. Akad. Nauk SSSR* 163, 4, 845–848, 1965.
41. Ward JH Jr (1963) Hierarchical grouping to optimize an objective function. *Journal of the American statistical association* 58: 236–244.
42. Marck C, Grosjean H (2002) tRNomics: analysis of tRNA genes from 50 genomes of Eukarya, Archaea, and Bacteria reveals anticodon-sparing strategies and domain-specific features. *RNA* 8: 1189–1232. <https://doi.org/10.1017/s1355838202022021> PMID: 12403461
43. Team RC (2018) R: A language and environment for statistical computing.
44. de Souza DAS, Pavoni DP, Krieger MA, Ludwig A (2018) Evolutionary analyses of myosin genes in trypanosomatids show a history of expansion, secondary losses and neofunctionalization. *Sci Rep* 8: 1376. <https://doi.org/10.1038/s41598-017-18865-y> PMID: 29358582
45. Hughes AL, Piontkivska H (2003) Molecular phylogenetics of Trypanosomatidae: contrasting results from 18S rRNA and protein phylogenies. *Kinetoplastid Biol Dis* 2: 15. <https://doi.org/10.1186/1475-9292-2-15> PMID: 14613495
46. Pothirat T, Tantiworawit A, Chaiwarith R, Jariyapan N, Wannasan A, et al. (2014) First isolation of Leishmania from Northern Thailand: case report, identification as *Leishmania martiniquensis* and phylogenetic position within the *Leishmania enriettii* complex. *PLoS Negl Trop Dis* 8: e3339. <https://doi.org/10.1371/journal.pntd.0003339> PMID: 25474647
47. Kelly S, Ivens A, Mott GA, O'Neill E, Emms D, et al. (2017) An Alternative Strategy for Trypanosome Survival in the Mammalian Bloodstream Revealed through Genome and Transcriptome Analysis of the Ubiquitous Bovine Parasite Trypanosoma (Megatrypanum) theileri. *Genome Biol Evol* 9: 2093–2109. <https://doi.org/10.1093/gbe/evx152> PMID: 28903536
48. Chan PP, Lowe TM (2016) GtRNAdb 2.0: an expanded database of transfer RNA genes identified in complete and draft genomes. *Nucleic Acids Res* 44: D184–189. <https://doi.org/10.1093/nar/gkv1309> PMID: 26673694
49. Eddy SR, Durbin R (1994) RNA sequence analysis using covariance models. *Nucleic Acids Res* 22: 2079–2088. <https://doi.org/10.1093/nar/22.11.2079> PMID: 8029015
50. Lowe TM, Eddy SR (1997) tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res* 25: 955–964. <https://doi.org/10.1093/nar/25.5.955> PMID: 9023104
51. Gouy M, Guindon S, Gascuel O (2010) SeaView version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Mol Biol Evol* 27: 221–224. <https://doi.org/10.1093/molbev/msp259> PMID: 19854763

52. Sprinzl M, Horn C, Brown M, Iudovitch A, Steinberg S (1998) Compilation of tRNA sequences and sequences of tRNA genes. *Nucleic Acids Res* 26: 148–153. <https://doi.org/10.1093/nar/26.1.148> PMID: 9399820
53. Lawrence TJ, Kauffman KT, Amrine KC, Carper DL, Lee RS, et al. (2015) FAST: FAST Analysis of Sequences Toolbox. *Front Genet* 6: 172. <https://doi.org/10.3389/fgene.2015.00172> PMID: 26042145
54. Schneider TD, Stormo GD, Gold L, Ehrenfeucht A (1986) Information content of binding sites on nucleotide sequences. *J Mol Biol* 188: 415–431. [https://doi.org/10.1016/0022-2836\(86\)90165-8](https://doi.org/10.1016/0022-2836(86)90165-8) PMID: 3525846
55. Nemenman I, Shafee F, Bialek W. Entropy and inference, revisited; 2002. pp. 471–478.
56. Wilkinson AJ, Fersht AR, Blow DM, Winter G (1983) Site-directed mutagenesis as a probe of enzyme structure and catalysis: tyrosyl-tRNA synthetase cysteine-35 to glycine-35 mutation. *Biochemistry* 22: 3581–3586. <https://doi.org/10.1021/bi00284a007> PMID: 6615786
57. Ibba M, Kast P, Hennecke H (1994) Substrate specificity is determined by amino acid binding pocket size in *Escherichia coli* phenylalanyl-tRNA synthetase. *Biochemistry* 33: 7107–7112. <https://doi.org/10.1021/bi00189a013> PMID: 8003476
58. Sampson JR, Uhlenbeck OC (1988) Biochemical and physical characterization of an unmodified yeast phenylalanine transfer RNA transcribed in vitro. *Proc Natl Acad Sci U S A* 85: 1033–1037. <https://doi.org/10.1073/pnas.85.4.1033> PMID: 3277187
59. Schulze CJ, Bray WM, Woerhmann MH, Stuart J, Lokey RS, et al. (2013) "Function-first" lead discovery: mode of action profiling of natural product libraries using image-based screening. *Chem Biol* 20: 285–295. <https://doi.org/10.1016/j.chembiol.2012.12.007> PMID: 23438757
60. Bullwinkle TJ, Reynolds NM, Raina M, Moghal A, Matsa E, et al. (2014) Oxidation of cellular amino acid pools leads to cytotoxic mistranslation of the genetic code. *Elife* 3.
61. Padilla-Mejia NE, Florencio-Martinez LE, Figueroa-Angulo EE, Manning-Cela RG, Hernandez-Rivas R, et al. (2009) Gene organization and sequence analyses of transfer RNA genes in Trypanosomatid parasites. *BMC Genomics* 10: 232. <https://doi.org/10.1186/1471-2164-10-232> PMID: 19450263
62. Lukes J, Skalicky T, Tyc J, Votypka J, Yurchenko V (2014) Evolution of parasitism in kinetoplastid flagellates. *Mol Biochem Parasitol* 195: 115–122. <https://doi.org/10.1016/j.molbiopara.2014.05.007> PMID: 24893339
63. Gowri VS, Ghosh I, Sharma A, Madhubala R (2012) Unusual domain architecture of aminoacyl tRNA synthetases and their paralogs from *Leishmania major*. *BMC Genomics* 13: 621. <https://doi.org/10.1186/1471-2164-13-621> PMID: 23151081
64. Beebe K, Ribas De Pouplana L, Schimmel P (2003) Elucidation of tRNA-dependent editing by a class II tRNA synthetase and significance for cell viability. *EMBO J* 22: 668–675. <https://doi.org/10.1093/emboj/cdg065> PMID: 12554667
65. Agarwal V, Nair SK (2012) Aminoacyl tRNA synthetases as targets for antibiotic development. *Med-ChemComm* 3: 887–898.
66. Elf J, Ehrenberg M (2005) Near-critical behavior of aminoacyl-tRNA pools in *E. coli* at rate-limiting supply of amino acids. *Biophys J* 88: 132–146. <https://doi.org/10.1529/biophysj.104.051383> PMID: 15501947
67. Lehar J, Zimmermann GR, Krueger AS, Molnar RA, Ledell JT, et al. (2007) Chemical combination effects predict connectivity in biological systems. *Mol Syst Biol* 3: 80. <https://doi.org/10.1038/msb4100116> PMID: 17332758
68. Cokol M (2013). Drugs and their interactions. *Current Drug Discovery Technologies* 10:106–113. <https://doi.org/10.2174/1570163811310020003> PMID: 23237673
69. Hegreness M, Shores N, Damian D, Hartl D, Kishony R (2008) Accelerated evolution of resistance in multidrug environments. *Proc Natl Acad Sci U S A* 105: 13977–13981. <https://doi.org/10.1073/pnas.0805965105> PMID: 18779569
70. Michel J, Yeh PJ, Chait R, Moellering RC Jr., Kishony R (2008) Drug interactions modulate the potential for evolution of resistance. *Proc Natl Acad Sci U S A* 105: 14918–14923. <https://doi.org/10.1073/pnas.0800944105> PMID: 18815368
71. Torella JP, Chait R, Kishony R (2010) Optimal drug synergy in antimicrobial treatments. *PLoS Comput Biol* 6: e1000796. <https://doi.org/10.1371/journal.pcbi.1000796> PMID: 20532210