



# Cascade Ligand- and Structure-Based Virtual Screening to Identify New Trypanocidal Compounds Inhibiting Putrescine Uptake

Lucas N. Alberca<sup>1</sup>, María L. Sbaraglini<sup>1</sup>, Juan F. Morales<sup>1</sup>, Roque Dietrich<sup>1</sup>,  
María D. Ruiz<sup>2</sup>, Agustina M. Pino Martínez<sup>3</sup>, Cristian G. Miranda<sup>3</sup>, Laura Fraccaroli<sup>2</sup>,  
Catalina D. Alba Soto<sup>3</sup>, Carolina Carrillo<sup>2</sup>, Pablo H. Palestro<sup>1</sup> and Alan Talevi<sup>1\*</sup>

<sup>1</sup> Laboratory of Bioactive Compounds Research and Development (LIDeB), Medicinal Chemistry, Department of Biological Science, Exact Sciences College, National University of La Plata, Buenos Aires, Argentina, <sup>2</sup> Institute of Sciences and Technology Dr César Milstein (ICT Milstein), Argentinean National Council of Scientific and Technical Research (CONICET), Buenos Aires, Argentina, <sup>3</sup> Department of Microbiology, Parasitology and Immunology, School of Medicine, Institute of Microbiology and Medical Parasitology (CONICET), University of Buenos Aires, Buenos Aires, Argentina

## OPEN ACCESS

### Edited by:

Jorge Enrique Gómez Marín,  
University of Quindío, Colombia

### Reviewed by:

Claudia M. Calvet Alvarez,  
University of California, San Diego,  
United States  
Vishvanath Tiwari,  
Central University of Rajasthan, India

### \*Correspondence:

Alan Talevi  
alantalevi@gmail.com

Received: 21 December 2017

Accepted: 04 May 2018

Published: 25 May 2018

### Citation:

Alberca LN, Sbaraglini ML,  
Morales JF, Dietrich R, Ruiz MD, Pino  
Martínez AM, Miranda CG,  
Fraccaroli L, Alba Soto CD, Carrillo C,  
Palestro PH and Talevi A (2018)  
Cascade Ligand- and  
Structure-Based Virtual Screening to  
Identify New Trypanocidal  
Compounds Inhibiting Putrescine  
Uptake.  
*Front. Cell. Infect. Microbiol.* 8:173.  
doi: 10.3389/fcimb.2018.00173

Chagas disease is a neglected tropical disease endemic to Latin America, though migratory movements have recently spread it to other regions. Here, we have applied a cascade virtual screening campaign combining ligand- and structure-based methods. In order to find novel inhibitors of putrescine uptake in *Trypanosoma cruzi*, an ensemble of linear ligand-based classifiers obtained by has been applied as initial screening filter, followed by docking into a homology model of the putrescine permease TcPAT12. 1,000 individual linear classifiers were inferred from a balanced dataset. Subsequently, different schemes were tested to combine the individual classifiers: MIN operator, average ranking, average score, average voting, with MIN operator leading to the best performance. The homology model was based on the arginine/agmatine antiporter (AdiC) from *Escherichia coli* as template. It showed 64% coverage of the entire query sequence and it was selected based on the normalized Discrete Optimized Protein Energy parameter and the GA341 score. The modeled structure had 96% in the allowed area of Ramachandran's plot, and none of the residues located in non-allowed regions were involved in the active site of the transporter. Positivity Predictive Value surfaces were applied to optimize the score thresholds to be used in the ligand-based virtual screening step: for that purpose Positivity Predictive Value was charted as a function of putative yields of active in the range 0.001–0.010 and the Se/Sp ratio. With a focus on drug repositioning opportunities, DrugBank and Sweetlead databases were subjected to screening. Among 8 hits, cinnarizine, a drug frequently prescribed for motion sickness and balance disorder, was tested against *T. cruzi* epimastigotes and amastigotes, confirming its trypanocidal effects and its inhibitory effects on putrescine uptake. Furthermore, clofazimine, an antibiotic

with already proven trypanocidal effects, also displayed inhibitory effects on putrescine uptake. Two other hits, meclizine and butoconazole, also displayed trypanocidal effects (in the case of meclizine, against both epimastigotes and amastigotes), without inhibiting putrescine uptake.

**Keywords:** Chagas disease, *Trypanosoma cruzi*, putrescine uptake, drug repositioning, drug repurposing, cinnarizine, virtual screening, positive predictive value

## INTRODUCTION

World Health Organization (WHO) describes neglected tropical diseases (NTDs) as a group of tropical diseases which mainly affect people living in poverty, lacking adequate sanitary conditions and in close contact with the infectious vectors (World Health Organization, 2015). One of the most important NTDs—in numerical terms—is Chagas disease, a parasitic disease endemic to Latin-America, caused by the infection by the protozoan *Trypanosoma cruzi*. This parasite can be transmitted to humans and more than 150 domestic and wild mammals, making complete eradication of the parasite practically impossible. The main route of transmission of *T. cruzi* is through the feces of the insect vector, known as *vinchuca*, a bug of the subfamily Triatominae. There are also other transmission routes, as congenital transmission and blood transfusions (Rassi et al., 2010), which are becoming increasingly important in the last years. WHO estimates based on 2010 data indicate that more than 6 million people are infected with *T. cruzi* worldwide, mostly in Latin-American countries (World Health Organization, 2015). However, several reports suggest that the actual number of infected people could be quite higher, reaching 10 million people (Ventura-Garcia et al., 2013; Stanaway and Roth, 2015; Browne et al., 2017).

Chagas disease presents two clinical phases. The initial or acute phase, which lasts between 4 and 8 weeks, is in general asymptomatic or might present as a self-limiting febrile illness. After the acute phase, an indeterminate, latent phase follows, with absence of clinical symptoms. About 60–70% of these people will remain in the indeterminate phase, but the remaining 30–40% will develop the symptomatic chronic phase characterized by damage to specific organs—particularly heart, esophagus, or colon. The chronic phase remains throughout life drastically reducing life expectancy among these patients (Nunes et al., 2013).

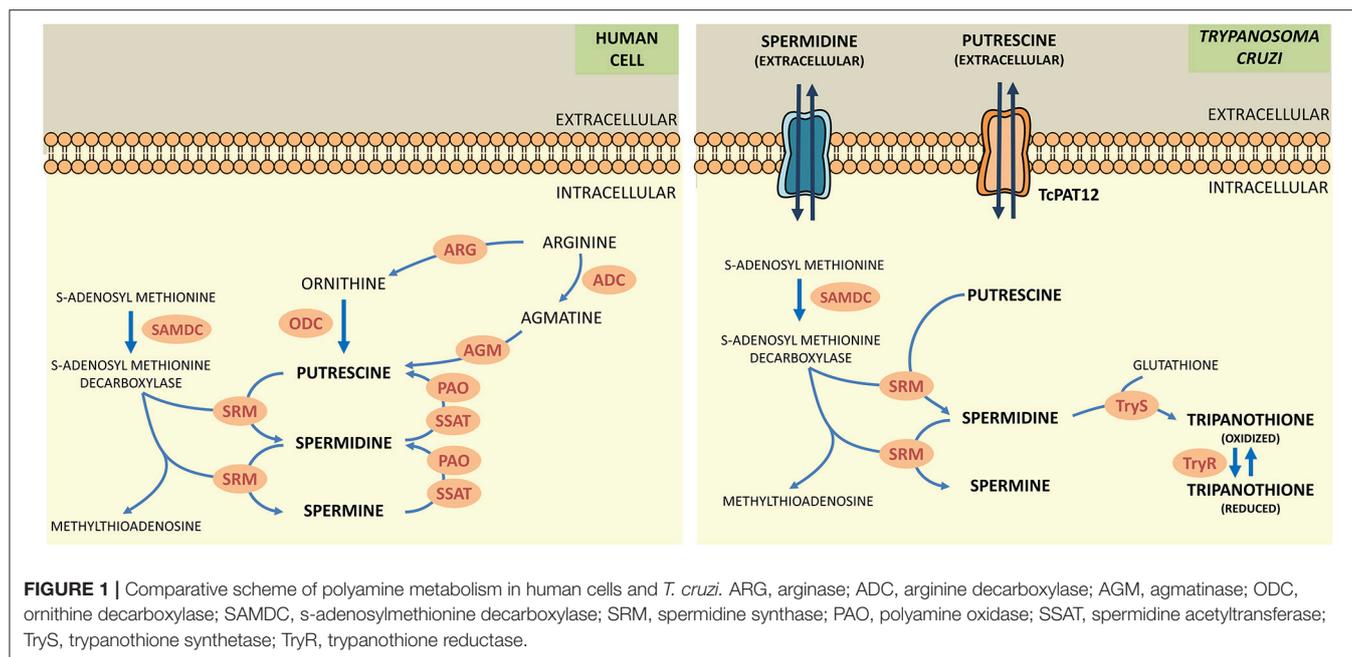
The only two approved drugs for the treatment of Chagas disease so far are Benznidazole and Nifurtimox, launched in the early 1970s. Both compounds are well-tolerated in children and effective during the acute phase. However, they present considerable side effects in adults, different susceptibility among *T. cruzi* strains and limited efficacy in adults in chronic phase (Morillo et al., 2015; Bermudez et al., 2016).

Drug repositioning (also known as drug repurposing, indication expansion and indication shift) represents an interesting strategy to approach the development of new medications for NTD (Ekins et al., 2011; Bellera et al., 2015; Ferreira and Andricopulo, 2016; Sbaraglini et al., 2016). It

consists in finding novel medical uses for existing drugs, including approved, experimental, discontinued and shelved drugs. Drug repurposing has several advantages over the search of *de novo* drugs. Since the new indication is built on already available pharmacokinetic and safety data, drug development time and costs can be considerably shortened. Possible manufacturing issues have also been solved. There are several successful cases of repositioned drugs in the field of NTDs: the anticancer drug eflornithine has been approved for the treatment of sleeping sickness and the antifungal drug amphotericin B has been repurposed for treatment of visceral leishmaniasis. To date, however, although there are several reports of drug candidates to be repositioned for the treatment of Chagas disease, none of these has yet been approved (Andrews et al., 2014; Klug et al., 2016; Sbaraglini et al., 2016). While initially drug repurposing stories arose from serendipitous observations, the drug discovery community has progressively adopted more systematic approximations to indication expansion (Ekins et al., 2011; Jin and Wong, 2014; Ferreira and Andricopulo, 2016), including genomic and structural biology tools, *in silico* screening and high-throughput screening platforms.

Polyamines (putrescine, spermidine, spermine) are low molecular weight polycations with crucial physiologic role in all the eukaryotic cells. They take part in fundamental cellular processes such as growth, differentiation, macromolecular biosynthesis and protection against oxidative damage. The polyamine metabolism in *T. cruzi* differs significantly from its human counterpart since the parasite lacks the enzymes arginine decarboxylase and ornithine decarboxylase, which are necessary for the biosynthesis of polyamines (Figure 1; Carrillo et al., 1999, 2003). Thus, *T. cruzi* depends on the incorporation of polyamines from the host cell. These functions are carried out by polyamine transporters such as the high-affinity putrescine permease TcPAT12 (or TcPOT1.1) which does not present homologous in the mammalian lineage (Carrillo et al., 2006). The importance of polyamines for parasites survival, the inability of the parasite to biosynthesize polyamines and the exclusivity of the putrescine transporter in *T. cruzi* makes putrescine uptake an attractive target for the search of new trypanocidal drugs (Hasne et al., 2016).

Back in 2016, we reported the first *in silico* drug repurposing campaign to discover novel inhibitors of polyamine uptake in *T. cruzi* (Alberca et al., 2016); such study applied an ensemble of ligand-based models to screen DrugBank 4.0 and Sweetlead databases and resulted in the identification of three candidates that impaired putrescine transport: paroxetine, triclabendazole



and sertaconazole. Here, we have improved our ligand-based computational models and complemented them with molecular docking based on a homology model of *TcPAT12*; the combined screening has been applied to identify two new *TcPAT12* potential inhibitors: clofazimine and cinnarizine. We also report, for the first time, the use of Positivity Predictive Value (PPV) surfaces analysis to select the score threshold that will be used in the virtual screening procedure. The two hits were assayed against *T. cruzi* epimastigotes and trypomastigotes, and the inhibitory effect on putrescine uptake was also determined.

## MATERIALS AND METHODS

### Ligand-Based Virtual Screening Dataset Collection

Polyamine analogs previously assayed against *T. cruzi* were compiled from literature. 256 polyamine analogs were found and conformed the dataset used here for model calibration and validation. Such dataset was curated using the standardization tool available in Instant JCHEM v. 17.2.6.0. We have labeled the 256 compounds as ACTIVE or INACTIVE according to their half-maximal effective concentrations ( $EC_{50}$ ) against *T. cruzi*. The ACTIVE category included compounds with  $EC_{50}$  below  $20 \mu\text{M}$ ; the remaining compounds were included in the INACTIVE category. Considering such cut-off, the dataset includes 116 actives and 140 inactives. The molecular diversity of the whole dataset and within each category can be appreciated in the heatmaps displayed in **Figure 2**, which show, for every compound pair in the database, the Tanimoto distance computed using ECFP<sub>4</sub> molecular fingerprints. The heatmap was built using Gitools v. 2.3.1 (Perez-Lamas and Lopez-Bigas, 2011) and Tanimoto distances were calculated using ScreenMD—Molecular Descriptor Screening

v. 5.5.0.1 (ChemAxon, 2002–2011). The dataset is included as Supplementary Information.

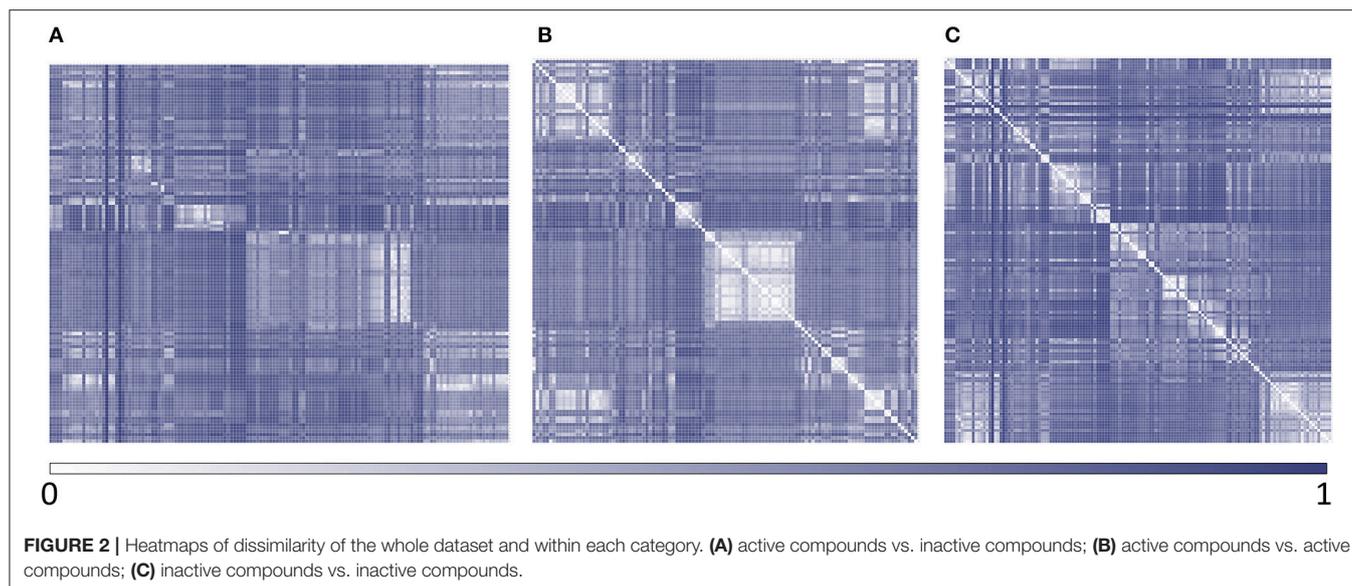
### Dataset Splitting

The resulting dataset of 256 polyamine analogs was divided into two groups using a representative sampling procedure: (a) training set, that was used to calibrate the models and; (b) test set, that was used to externally validate the models. The representative partition of the dataset resulted from a serial combination of clustering procedures: first, we have used the hierarchical clustering method included in LibraryMCS software (version 17.2.13.0–ChemAxon), which relies on the Maximum Common Substructure (MCS). From the resulting clusters, a compound from each cluster was randomly chosen and used as a seed to perform a non-hierarchical clustering using the k-means algorithm, as implemented in Statistica 10 Cluster Analysis module (Statsoft, 2011). Such procedure was performed in an independent manner for the ACTIVE and INACTIVE categories.

Seventy-five percent of the compounds in each cluster of the ACTIVE category were kept for the training set (totaling 87 compounds); an equal number of compounds were taken from the INACTIVE category clusters (62.14% of each INACTIVE cluster). The remaining 29 active and 53 inactive compounds were assigned to test set. Note that, to obtain the training set, we have under-sampled the INACTIVE category, so that a balanced calibration sample was obtained. In that way, model bias toward predicting a specific category was avoided.

### Molecular Descriptor Calculation and Modeling Procedure

Three thousand six hundred sixty-eight conformation-independent descriptors were computed with Dragon 6.0 software. A random subspace-based method was applied to



obtain 1,000 random descriptor subsets of no more than 200 potential independent variables each. In the random subspace approach, the features (independent variables, e.g., molecular descriptors) are randomly sampled and each model (learner) is trained on one subset of the feature space (El Habib Daho and Amine Chikh, 2015; Vyskovsky et al., 2016), which causes individual models not to over-focus on features that display high explanatory power in the training set.

A dummy, binary variable (class label) was used as dependent variable. Such variable was assigned observed values of 1 for compounds within the ACTIVE, and observed values of 0 for compounds in the INACTIVE category.

Using a Forward Stepwise approach and a semi-correlation approach (Toropova and Toropov, 2017), 1,000 linear classifiers were obtained, one from each of the random subsets of features. In order to avoid overfitting, only one molecular descriptor every 10 training instances was allowed into each model, with no more than 12 independent variables per model. No regression coefficient with *p*-value above 0.05 was allowed into the model.

R language and environment was used for all data analysis. The R package data table (<https://cran.r-project.org/package=data.table>) was used to handle datasets.

The robustness and predictive ability of the models were estimated through randomization and leave-group-out cross-validation tests. In the case of randomization, the class label was randomized across the compounds in the training set. The training set with the randomized dependent variable was then used to train new models from the descriptor selection step. Such procedure was repeated 10 times within each descriptor subset. It is expected that the randomized models will perform poorly compared to the real ones. Regarding the Leave-Group-Out cross-validation, 18-compound subsets were randomly removed from the training set in each cross-validation round and the model was regenerated. The resulting model was used to predict the class label for the 18 removed compounds. The procedure was repeated 10 times, removing each of the training set compounds once.

### Ensemble Learning

Classifier ensembles are known to handle complex data and to provide better generalization and accuracy than single model classifiers (El Habib Daho and Amine Chikh, 2015; Carbonneau et al., 2016; Min, 2016; Vyskovsky et al., 2016); they can be particularly useful to prevent overfitting when handling datasets that suffer from small sample size while their dimensionality is large (Vyskovsky et al., 2016), a quite frequent scenario in the drug design field.

The best individual classifiers were selected and combined (selective ensemble learning approach) using the area under the ROC curve metric (AUC ROC) as criterion of performance. Systematic combinations of the 2–100 best performing classifiers were analyzed. Four combination schemes were applied to obtain a combined score: MIN operator (which returns the minimum score among the individual scores of the combined models); Average Score; Average Ranking and Average Voting. Voting was computed according to the equation previously used by Zhang and Muegge (2006). AUC ROCs were obtained with the pROC package (Robin et al., 2011); the DeLong method was used to obtain 95% confidence intervals.

### Pilot (Retrospective) Screening Campaign

Through simulated ranking experiments, Truchon et al. demonstrated that the AUC ROC metric is dependent on the ratio of actives/inactives, and the standard deviation of the metric converges to a constant value when small yield of actives of the screened library ( $Y_a$ , also called ratio of actives or prevalence of activity) are used (ratios of actives below 0.05 seem to provide more robust results). Reasonably small  $Y_a$  also ensures that the saturation effect is constant or absent. A high number of decoys (around 1,000 or higher) contribute to a controlled statistical behavior, especially if poorly performing classifiers/methods are applied (Truchon and Bayly, 2007). Accordingly, we have performed pilot (retrospective) VS campaigns for the individual classifiers and classifier ensembles. For that purpose, we have dispersed the active compounds from each test set among a

large number of decoys obtained with the help of the enhanced Directory of Useful Decoys (DUD-E) (Mysinger et al., 2012). Each of the test set active compounds from each dataset was used as a query in the DUD-E website, thus generating paired putative inactive compounds (decoys) for each of those active compounds. As a result, the pilot database contained 29 known active compounds dispersed among 1302 DUD-E decoys and 53 known inactive compounds, adding up a total of 1384 compounds and displaying a  $Y_a$  of 0.021.

### Building Positivity Predictive Value Surfaces and Choosing an Adequate Score Threshold Value

A practical concern in virtual screening campaigns is to predict the actual probability that a predicted hit will prove truly active when submitted to experimental testing (the  $PPV$ ). Estimation of such probability is however obstructed due to its dependency on the  $Y_a$  of the screened library, which cannot be known *a priori*:

$$PPV = \frac{(Se)(Y_a)}{(Se)(Y_a) + (1 - Sp)(1 - Y_a)} \quad (1)$$

where  $Se$  represents the sensitivity associated to a given score cutoff value and  $Sp$  represents the specificity. Equation (1) was applied to build  $PPV$  surfaces. 3D plots illustrating the interplay between  $PPV$ , the  $Se/Sp$  ratio and  $Y_a$  were built for each individual model and for each model ensemble. Using the corresponding pilot database (as described in previous subsections),  $Se$  and  $Sp$  were computed in all the range of possible cutoff score values. Note that there is no guarantee that the  $Se$  and  $Sp$  associated to each score value (and thus, the ratio  $Se/Sp$ ) will be the same when applying the classifiers to other compound databases, e.g., in the real virtual screening campaign; nevertheless, since controlled statistical behavior is observed for database sizes of about 1000 compounds or more and  $Y_a$  below 0.05, we can reasonably assume that the ROC curve and derived metrics will be similar when applying the models to classify other large chemical databases with low  $Y_a$ . In order to build the  $PPV$  surfaces, and taking into consideration that in real VS applications  $Y_a$  is ignored *a priori* but invariably low,  $Y_a$  was varied between 0.001 and 0.010. The R package `plotly` (<https://cran.r-project.org/package=plotly>) was used to obtain all the  $PPV$  graphs. Visual analysis of the resulting  $PPV$  surfaces allowed us to select a score cutoff value with a desired range of  $PPV$ .

### Virtual Screening

Based on visual inspections of the resulting of  $PPV$  graphs, we have applied an 8-model ensemble using the MIN operator to combine individual models for virtual screening, choosing 0.354 as score threshold (above 0.354 compounds from the screened databases have labeled as predicted active compounds). Such threshold corresponds to a  $Se/Sp$  ratio of 0.666. It was checked that every hit belonged to the applicability domain of the model from the model ensemble that assigned the minimum score (which, was, according to our combination scheme, the one that ultimately decided if a compound was or was not labeled as an active). The leverage approach was used to assess if a hit belongs to the applicability domain, using  $3d/n$  as cutoff value, where  $d$  is

the number of descriptors in the correspondent model and  $n$  is the number of compounds in the training set.

We have used the 8-model ensemble to screen two databases: (a) DrugBank 4.0, an online database containing extensive information about the US Food and Drug Administration (FDA) approved, experimental, nutraceutical, illicit and investigational drugs (Law et al., 2014); (b) Sweetlead, a curated database of drugs approved by other international regulatory agencies, compounds isolated from traditional medicinal herbs, and regulated chemicals (Novick et al., 2013). These two databases were curated using Standardizer 16.10.10.0 (ChemAxon). We have applied the following actions to generate homogeneous representations of the molecular structure for the virtual screening: (1) Strip salts; (2) Remove Solvents; (3) Clear Stereo; (4) Remove Absolute Stereo; (5) Aromatize; (6) Neutralize; (7) Add Explicit Hydrogens; and (8) Clean 2D. Additionally, duplicate structures were removed using Instant JCHEM v. 17.2.6.0. Based in the results obtained and considering the most direct candidates for repositioning, two compounds were selected for experimental evaluation. Hits submitted to experimental testing were acquired from Sigma-Aldrich.

### Structure-Based Virtual Screening

The 24 best solutions from ligand based virtual screening were submitted to molecular docking calculations. To this end we employed as target our 3D model of the putrescine permease TcPAT12 previously constructed (Dietrich et al., 2018). This model was based on the arginine/agmatine antiporter (AdiC) from *E. coli* as template (Protein data bank accession code: 3L1L, Feng et al., 2000), which was identified through Blast sequence search and showed 23% of sequence identity and 64% coverage of the entire query sequence. The model of TcPAT12 architecture was achieved by Modeller Software (Webb and Sali, 2016). The best model has been selected based on the normalized Discrete Optimized Protein Energy parameter ( $z$ -DOPE) (Shen and Sali, 2006) and the GA341 score (which analyze the reliability of a model based on statistical potentials) (Melo and Sali, 2007). This structure had the highest number of residues in allowed area (96%) on Ramachandran plot, and none of the residues located in non-allowed regions were involved in the active site of the transporter.

The docking conditions were defined according to our previous studies, which pointed to Autodock4.2 (rigid mode) as the best software to discriminate known inhibitors from non-inhibitors of TcPAT12 through the docking score (Dietrich et al., 2018). The “docking active site” was set according to previous research of Soysa et al. (2013). They proposed the location of the putrescine-binding pocket in a region that includes Gly69, Cys66, Trp241, Ala244, Asn145, Cys396, Asn245, Tyr148, Tyr400 amino acids. Autodock4.2 calculations were computed in a grid with the default spacing (0.375 Å) between the  $44 \times 58 \times 40$  grid points in x, y, z directions respectively. Additionally, we performed the standard estimation for all the variables such as Marsilli-Gasteiger partial charges. We computed 100 docking runs for each compound, with a rigid target and flexible ligands (allowing the rotation of all non-ring torsion angles of the candidates).

## Experimental Assays

### Biological Activity Against Different *T. cruzi* Stages

For all assays, stock and working solutions of the candidate drugs were prepared using DMSO as solvent and test or control conditions were tested in triplicate.

Epimastigotes of the Y strain of *T. cruzi* were cultured at 28°C in BHT medium supplemented with 20 µg/ml haemin, 10% heat-inactivated fetal bovine serum (FBS), 100 µg/ml streptomycin and 100 U/ml penicillin. The antiproliferative activity of the candidates was tested at concentrations from 1 to 100 µM in cultures initiated at 10<sup>7</sup> cells/ml. After 4 and 8 days, the number of viable parasites were counted using a hemocytometer chamber under the light microscope (Bellera et al., 2013). Controls were performed under the same culture conditions with equal concentrations of DMSO as for candidate drugs. The EC<sub>50</sub> values were determined from dose response curves fitted to a sigmoidal equation (Boltzmann model) or extrapolated from linear fitting plots (Fernández et al., 2013).

*T. cruzi* trypomastigotes were purified at the parasitemia peak from peripheral blood of mice infected with the RA strain. Trypomastigotes (1 × 10<sup>4</sup> per well) were cultured in a 96 well-plate (final volume 200 µl) in RPMI medium supplemented with 10% FBS at 37°C in 5% CO<sub>2</sub> atmosphere. After 24 hours motile parasites were counted in hemocytometer chamber under the light microscope (Miranda et al., 2015). Results were expressed as % viability of trypomastigotes at 20 µM of the hits. Controls were performed under the same culture conditions with equal concentrations of DMSO as for candidate drugs. The negative control was cultured with PBS and the positive control was cultured with Benznidazole (20 µM).

*In vitro* evaluation of drug activity against intracellular amastigotes. J774 cells infected with bloodstream trypomastigotes of RA strain were cultured at 37°C in humidified incubator with 5% CO<sub>2</sub> in 24-well plates (1.5 × 10<sup>5</sup> cells per well; final volume of 500 µl in duplicate). After 24 h of incubation, increasing doses of freshly prepared dilutions of benznidazole, meclizine dihydrochloride, cinnarizine, or butoconazole were added at final concentrations of 5, 20, and 50 µM for meclizine, and cinnarizine and 1, 5, and 20 µM for butoconazole nitrate. Seventy-two hours later, medium was drained and cells were fixed for 10 min in ice-cold methanol and stained with 10% v/v of Giemsa solution for 15 min. The number of amastigotes per 100 host cells was recorded. Control cultures were incubated in medium alone or with equal DMSO concentrations. Cytotoxicity was analyzed in trypsinized cell suspensions, after addition of Propidium iodide (PI) (Sigma, St. Louis, USA) (5 µg/ml) 10 min prior to analysis by fluorescence and light microscopy of the number of viable and dead cells.

### Aminoacid/Polyamine Transport Assay

Aliquots of *T. cruzi* epimastigotes (3 × 10<sup>7</sup> parasites) starved for 3 hs in 2% glucose—phosphate-buffered saline (PBS) were collected, centrifuged at 1,500 g for 10 min and washed three times with PBS. Cells were then resuspended in 2 ml of PBS containing 5 µM (<sup>14</sup>C)-putrescine, or (<sup>14</sup>C) arginine, and candidates at a final concentration of 50 µM previously solubilized in DMSO. Aliquots were taken at different time

points, centrifuged and washed three times with 1 ml of ice cold 10 mM solution of unlabeled putrescine/arginine in PBS. Pellets were resuspended and radioactivity determined in UltimaGold XR (Díaz et al., 2014). All experiments were performed in triplicate.

The effect of the candidate on parasite viability under uptake assay conditions was evaluated through the tetrazolium salt (MTT) reduction assay (Mosmann, 1983). Briefly, 10 µl of 5 mg/ml MTT dye (3[4,5-dimethylthiazol-2-yl]-2,5-diphenyltetrazolium bromide) was added to the eppendorf tubes containing 10<sup>6</sup> parasites in 100 µl of BHT and the drug candidates at 50 µM solubilized in DMSO. After incubation for 3 h at 28°C, the tubes were spin-dried (3,000 rpm) and the pellet with the formazan crystals were dissolved with 100 µl of DMSO. The optical density (OD) was determined using a microplate reader (Labsystems Multiskan MS, Finland) at 570/695 nm. Under such conditions, the OD is proportional to the viable cell number in each well. All experiments were performed in triplicate.

## RESULTS

### Model Development and Validation and Virtual Screening

Two serial virtual screening methods were applied to find out putative *TcPAT12* inhibitors. First, we resorted to a computationally inexpensive ligand-based approach. For that purpose, 1,000 individual linear classifiers were obtained by applying a random subspace approximation. The individual models were externally validated by using an independent test set and, for a more challenging and realistic simulation, by retrospective screening of a simulated library where a small proportion of active compounds was dispersed among a high number of (mostly putative) decoys. 82, 57, and 25% of the individual classifiers displayed AUC ROCs above 0.8 for the training set, the test set and the simulated DUD-E database, in that order. This suggests some degree of overfitting and corroborates that the pilot (retrospective) screening campaign against the DUD-E database is the more challenging task for the classifiers. **Table 1** shows the eight individual classifiers that showed the best performance on the training and test sets, and also on the DUD-E database.

The best individual model included the following features:

Model 348: Class = 3.12777 + 0.03474 \* F06[C-C] + 0.20805 \* S-107 - 0.04291 \* F05[N-N] + 0.39611 \* C-039 - 0.34582 \* SM5\_B(s) + 6.25705 \* Eta\_epsi\_A + 0.53013 \* nSO2OH - 1.28338 \* SpMax\_H2 + 0.44827 \* Eig04\_AEA(ri) - 1.73390 \* ATSC1e + 0.05975 \* CATS2D\_09\_PL + 0.02805 \* SaaO  
Wilks' Lambda: 0.59  
 $F_{(12, 161)} = 9.23$   
 $p < 0.0000$

Dragon's nomenclature for the molecular descriptors has been kept in the previous expression. F06[C-C] refers to the count of C-C at a topological distance of 6; S-107 represents the count of R2S/RS-SR groups; F05[N-N] stands for the frequency of

**TABLE 1** | Values of the AUROC metric for the best eight individual models and the best 8-model ensemble.

Model	Training set	Test set	DUD-E database
8-MODEL ENSEMBLE (MIN)	0.851 (±0.0281)	0.885 (±0.0367)	0.976*** (±0.0085)
8-MODEL ENSEMBLE (RANKING)	0.886 (±0.0239)	0.878 (±0.0375)	0.976*** (±0.0082)
8-MODEL ENSEMBLE (AVERAGE)	0.891 (±0.0234)	0.887 (±0.0357)	0.970*** (±0.0096)
8-MODEL ENSEMBLE (VOTING)	0.833 (±0.0283)	0.810 (±0.0516)	0.959* (±0.0173)
348	<b>0.882 (±0.0250)</b>	<b>0.885 (±0.0360)</b>	<b>0.934 (±0.0123)</b>
706	0.809* (±0.0319)	0.736** (±0.0565)	0.924 (±0.0205)
981	0.843 (±0.0298)	0.837 (±0.0467)	0.922 (±0.0254)
557	0.778** (±0.0343)	0.818 (±0.0482)	0.919 (±0.0203)
123	0.850 (±0.0285)	0.882 (±0.0382)	0.918 (±0.0185)
693	0.860 (±0.0280)	0.828* (±0.0459)	0.913* (±0.0171)
560	0.775** (±0.0348)	0.779* (±0.0525)	0.911* (±0.0185)
746	0.844 (±0.0288)	0.820 (±0.0456)	0.910 (±0.0195)

AUROC values statistically different from the correspondent column for the best individual model (model 348). \* $p < 0.05$ , \*\* $p < 0.01$ , and \*\*\* $p < 0.001$ . The highest AUC for an individual model is indicated in bold.

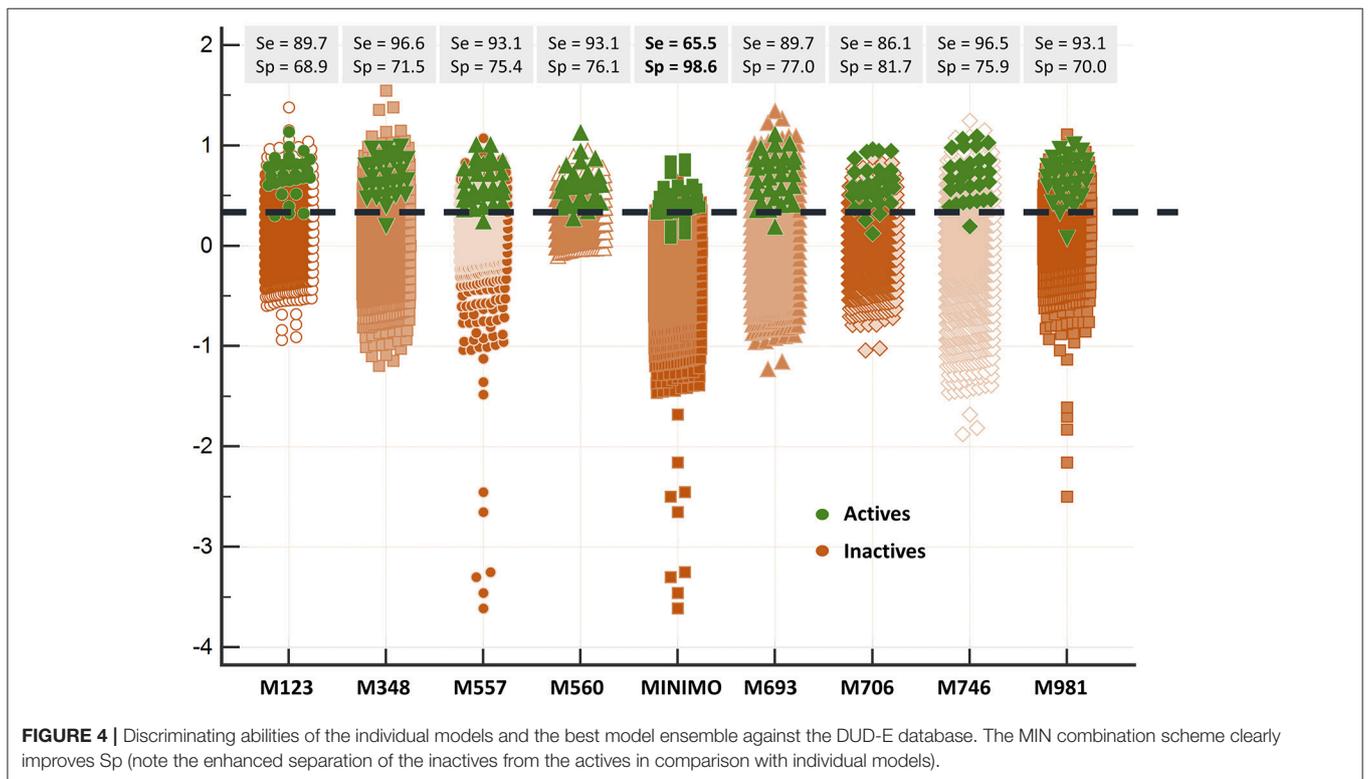
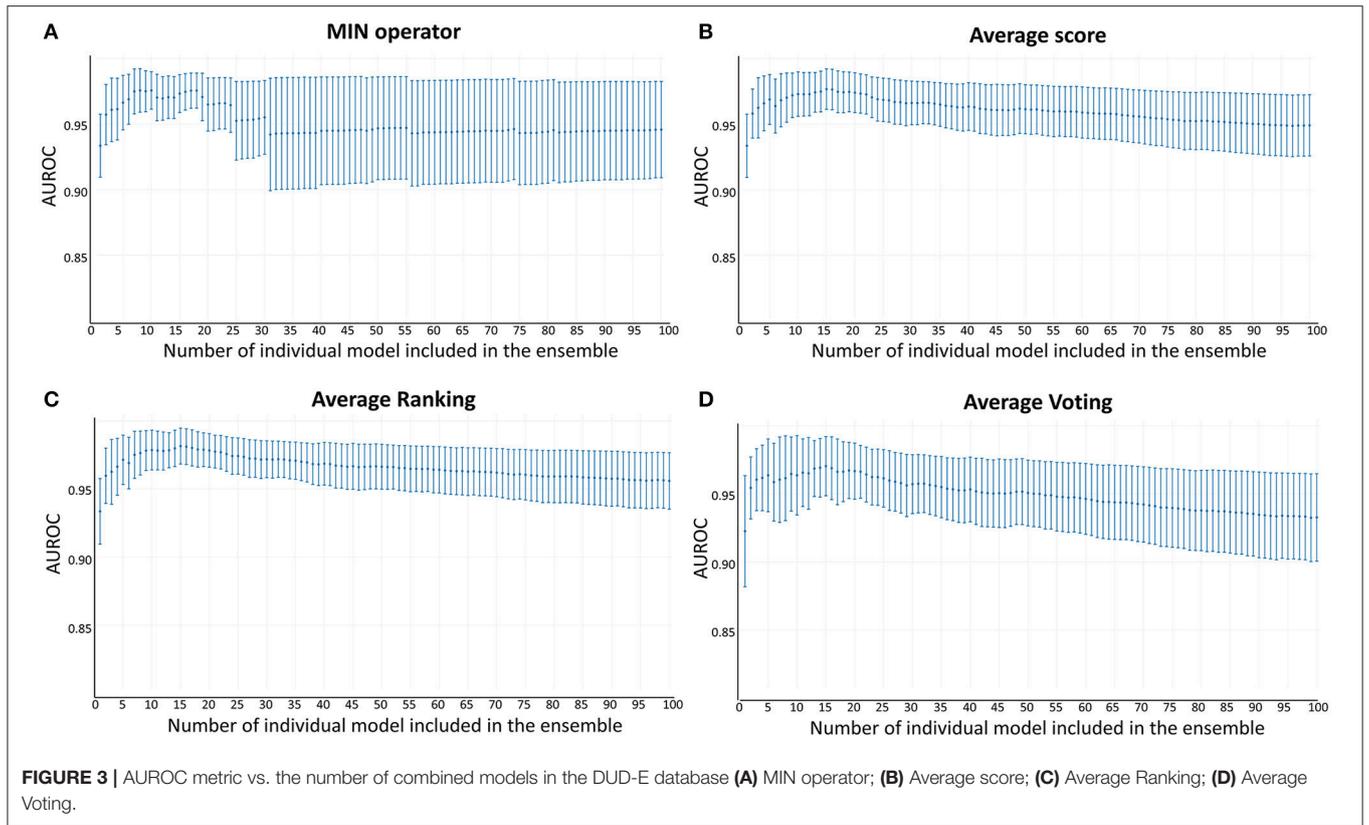
N - N at a topological distance of 5; C-039 refers to Ar-C(=X)-R groups; SM5\_B(s) corresponds to the spectral moment of order 5 from the Burden matrix weighted by I-State; Eta\_epsi\_A is the eta average electronegativity measure; nSO2OH stands for the number of sulfonic (thio-/dithio-) acids; SpMax\_H2 is the leading eigenvalue from reciprocal squared distance matrix; Eig04\_AEA(ri) denotes the fourth eigenvalue from augmented edge adjacency matrix weighted by resonance integral; ATSC1e corresponds to the Centred Broto-Moreau autocorrelation of lag 1 weighted by Sanderson electronegativity; CATS2D\_09\_PL is the CATS2D Positive-Positive at lag 09 and; SaaO corresponds to the sum of aaO E-states.

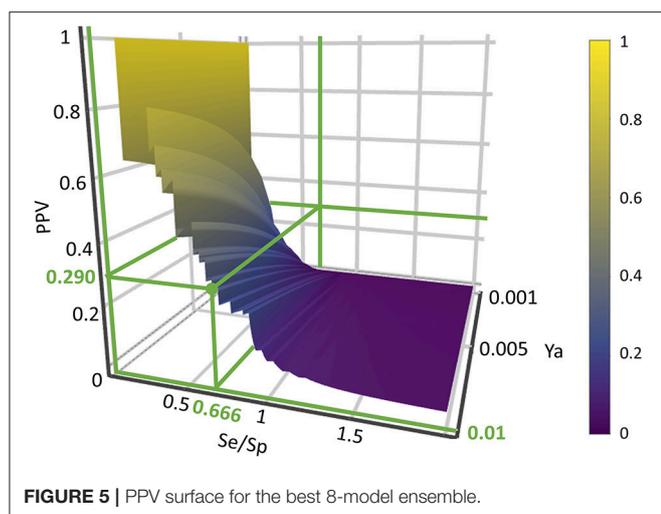
Whereas the performance of the best individual classifiers was quite satisfactory, we resorted to ensemble learning to obtain meta-classifiers with improved accuracy and a more robust behavior. The expectations on the model combination approach were confirmed statistically: no matter which combination scheme is applied, the model ensembles show a statistically similar behavior to the best individual model when classifying the training and test set compounds, while statistically outperforming the individual models when considering the DUD-E database ( $p < 0.05$  in all cases). The MIN, RANKING and AVERAGE combination schemes led to the best results (Table 1), with  $p$ -values below 0.001. When considering the influence of the number of combined models on the AUROC metric, it was observed that although the ensemble learning approach always seems to improve the results in comparison with the individual classifiers, combinations above 10 models tend to poorer behavior in terms of AUROC values but also regarding the standard deviation of the mean, which increases with the number of models included in the ensemble (Figure 3).

Since all the model combination approaches exhibited similar behavior against the test set and the DUD-E database for relatively low number (below 10) of combined classifiers, we chose to move to the real virtual screening campaign with the combination scheme based on the MIN. In our experience, this is the most conservative approach which tends to display the smallest false positive rate. This was confirmed when comparing

Sp and Se of the 8-model meta-classifier with those of the best performing individual models, where a substantial improvement in Sp was observed (Figure 4). This is a particularly good result in our context (small academic group from a mid-income country, with limited resources to invest in hit experimental validation); we often prefer to reduce the false positive rate even if this means losing sensitivity and sacrificing some active scaffolds. However, in this particular case we have chosen to refine the former criteria (prioritizing Sp) by resorting to what we have called PPV surfaces. With the help of these surfaces, the evolution of the most relevant metric for our purposes, the PPV value, can be visually (or, eventually, mathematically) optimized as a function of the Se/Sp ratio across a range of Ya values. For this analysis, we have used the association between Se/Sp and score values that have been observed in the pilot screening campaign. The strongest assumption of our approach is that the Se/Sp value observed for a given score during the pilot screening campaign against the DUD-E database will hold when screening other databases (e.g., the ones screened in real virtual screening applications). This is of course not necessarily true. However, since the AUROC values obtained for the DUD-E database are unequivocally high (always above 0.9 for the individual models and very close to the perfect value of 1 for the ensembles) while on the other hand the DUD-E database Ya ratio (quite below 0.05) and size (>1,000 compounds) speak of a controlled statistical behavior, we believe it is a reasonable assumption in the present setting.

When analyzing the PPV surface for the 8-model combination based on the MIN operator, it was observed that in the current scenario favoring Sp over Se has a positive impact on the PPV, thus resulting in higher probabilities of confirming *in silico* hits submitted to experimental validation (Figure 5). It should be emphasized that such behavior is not general: other situations linked to different PPV surface shapes might lead to an entirely different conclusion (that is, the need to prioritize Se over Sp to enhance PPV). The selection of the best cutoff value should be based in the specific PPV surface obtained in each particular modeling situation.





Using the PPV surface, we chose 0.354 as score threshold to be used in our real virtual screening campaign; such score is associated to a Se/Sp ratio of 0.666 for the 8-model ensemble based on the MIN operator. Whereas higher scores linked to lower Se/Sp ratio would, according to the PPV surface, result in improved PPVs, the resulting number of hits would also be substantially diminished, leaving very few choices for subsequent acquisition and experimental testing (the ligand-based screen using the 0.354 cutoff score resulted in only 24 hits, with just 15 of them being approved drugs, the most straightforward candidates to drug repurposing).

The 24 hits emerging from the ligand-based screening stage were submitted to the structure-based screening (docking into the *Tc*PAT12 homology model), with 17 of them surviving the docking stage and only 9 of them having achieved approved status. **Table 2** shows the hits selected through the combined ligand- and target-based approach, including the PPV range for the correspondent score value of the 8-model ensemble between  $Y_a$  values of 0.001 and 0.010. Note that 3 of the hits, namely clomifene, oxiconazole and clofazimine have previously been assayed against *T. cruzi*, with positive results (Sykes and Avery, 2013; Bellera et al., 2015; Kaiser et al., 2015). Most of the compounds have a docking score lower than the value found for the natural ligand putrescine ( $-6.0$ ), which was calculated previously in the same docking conditions. The results justify the selection of the structures as promising candidates. The exception was clofazimine, with a docking score of  $-1.58$ . In this particular case, the drug has been identified as trypanocidal agent within our research group, during an *in silico* screening to detect novel cruzipain inhibitors. The drug later confirmed its potential both in acute and chronic rodent models of Chagas disease (Bellera et al., 2015; Sbaraglini et al., 2016). In that occasion, though, it was observed that the potency of the drug against the parasite was higher than the inhibitory potency against cruzipain, suggesting multiple mechanisms of action besides cruzipain inhibition.

Based on access to the chosen compounds, we decided to test clofazimine, cinnarizine, meclizine dihydrochloride, and butoconazole nitrate (Sigma Aldrich) effects on putrescine uptake in *T. cruzi*. Cinnarizine, meclizine, and butoconazole

were also tested against *T. cruzi* epimastigotes, trypomastigotes, and amastigotes. The latter assays were omitted for clofazimine, since as already discussed, this drug had previously been tested against the different *T. cruzi* stages, with positive results. Before acquisition, it was verified if the compounds obeyed different druglikeness rules. All compounds satisfy Lipinski's rules (they satisfy three out of four of the Lipinski's proposition, only violating the proposition linked to  $\text{clogP}$ ), Veber's rule and the druglikeness criteria adopted by Monge et al. in previous studies (Lipinski et al., 2001; Veber et al., 2002; Monge et al., 2006). All in all, these analysis suggest that they have high probabilities of displaying good oral bioavailability (Lipinski's and Veber's rules) while also excluding some undesirable properties such as highly reactive chemical groups and possible toxicophores) (Monge's criteria). This is not a surprising result since all the hits submitted to experimental validation are approved drugs. A summary of the screening workflow and the number of hits surviving each step is presented in **Figure 6**.

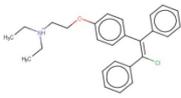
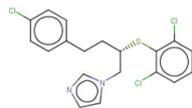
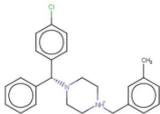
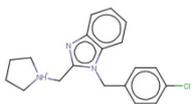
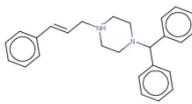
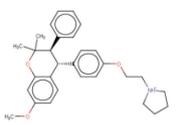
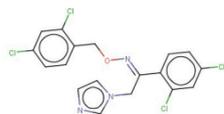
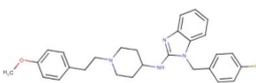
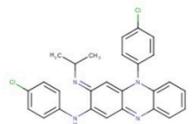
## Experimental Testing

The effect of different concentrations of cinnarizine, meclizine and butoconazole against *T. cruzi* epimastigotes was tested (**Figure 7**) and the correspondent  $EC_{50}$  was calculated. The three drugs showed a clear inhibition in a dose-dependent manner on the proliferation of epimastigotes, with a  $EC_{50}$  (at day 4) of  $6.05 \mu\text{M}$ ,  $8.39 \mu\text{M}$  and  $3.08 \mu\text{M}$ , respectively. The reference drug benznidazole displayed a  $EC_{50}$  of  $2.56 \mu\text{M}$  against epimastigotes. When tested against trypomastigotes at  $20 \mu\text{M}$ , cinnarizine displayed a slight inhibition in viability (30%). No inhibition was observed for the other two hits tested. When testing its effects against amastigotes, butoconazole showed a considerable cytotoxicity against J774 cells at all the assayed concentrations (100.0%;  $57.5 \pm 0.6\%$ ;  $48.3 \pm 6.3\%$  at  $50 \mu\text{M}$ ; 20 and  $5 \mu\text{M}$  respectively), and it was therefore disregarded for future investigations. Meclizine showed cytotoxicity against J774 cells at  $50 \mu\text{M}$  ( $13.9 \pm 2.0\%$ ), though no toxic effects were observed at lower concentrations. Cinnarizine showed no toxicity at any of the tested concentrations. Meclizine significantly reduced the number of amastigotes per 100 cells at the three assayed concentrations. Cinnarizine showed a weak effect against amastigotes, displaying almost a 50% inhibition at  $50 \mu\text{M}$  (**Figure 8**).

To determine if the mechanism of action of the candidates correlates with that predicted by our models, the inhibition of putrescine uptake by *T. cruzi* epimastigotes was determined (**Figure 9A**). A 10-fold molar excess of the candidate drugs were tested. Cinnarizine and clofazimine showed a clear effect on putrescine uptake with a significant initial velocity reduction to  $52.56 \pm 4.84$  and  $30.85 \pm 2.74\%$  respectively, compared with transport in control conditions. In contrast, meclizine and butoconazole did not display any inhibitory effect on putrescine uptake, which suggests that their trypanocidal effect is related to other mechanisms of action. Meclizine has previously been shown to inhibit cruzipain at low  $\mu\text{M}$  concentrations (Doak et al., 2010).

The specificity of the putrescine uptake inhibition by cinnarizine and clofazimine was checked by testing the effect of both active hits on other transporter of the same family

**TABLE 2** | Candidates selected through the combined ligand- and target-based approach.

Name	MIN Score	PPV% (Ya = 0.001)	PPV% (Ya = 0.01)	Structure	Score docking	Current therapeutic indication
Clomifene	0.4837	14.64	63.38		-9.37	Used mainly in female infertility due to anovulation
Butoconazole	0.4768	14.64	63.38		-9.87	Local treatment of vulvovaginal candidiasis
Meclizine	0.4546	12.30	58.61		-8.95, -6.64 <sup>a</sup>	Motion sickness and vertigo
Clemizole hydrochloride	0.4544	12.30	58.61		-8.91	Allergies
Cinnarizine	0.4273	9.20	50.56		-7.43	Motion sickness and vertigo
Centchroman	0.3798	5.67	37.76		-6.31	Primarily used as a contraceptive
Oxiconazole	0.3751	5.00	34.68		-10.44	Dermal fungal infections.
Astemizole	0.3666	4.25	30.96		-6.45	Allergies
Clofazimine	0.3605	4.06	29.92		-1.58	Leprosy

<sup>a</sup>The docking scores of both isomers were calculated.

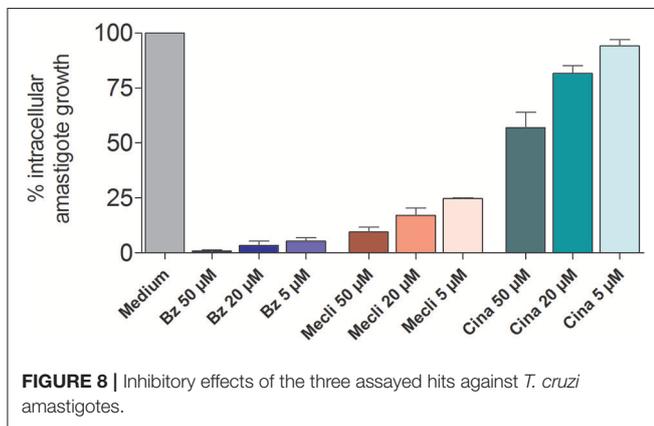
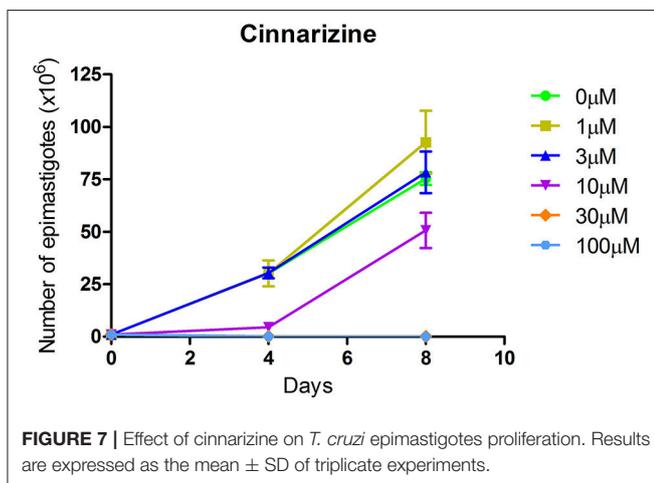
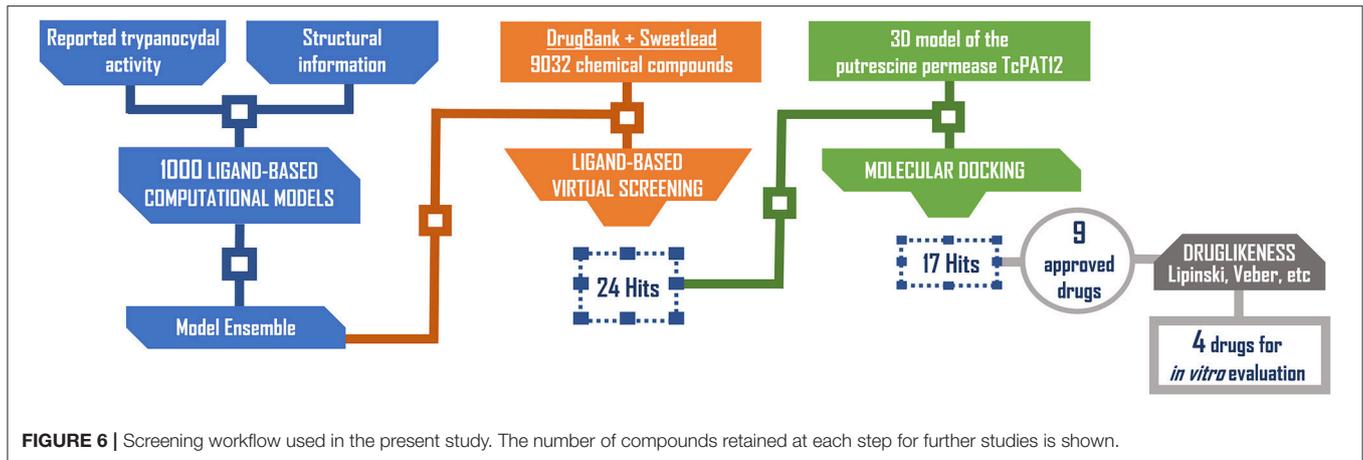
(Figure 9B). The MTT assay indicated that under the uptake conditions (with 50  $\mu$ M of the drug and 1% DMSO), neither cinnarizine nor clofazimine presented a cytotoxic effect on the parasites (not shown).

## DISCUSSION

### Comparison With Previous Studies

Whereas studies to exploit polyamine transporters as molecular drug targets for the development of new trypanocidal agents

are at an early stage, some interesting considerations may be outlined from the considering the present study and the few previous reports describing the screening of new inhibitors of *T. cruzi* putrescine uptake. First, some of the hits identified in earlier studies have shown to be more effective against *T. cruzi* trypomastigotes than against other stages of the parasite (Reigada et al., 2017, 2018). For instance, the reported IC<sub>50</sub> of isotretinoin against trypomastigotes (130 nM) is about 230-fold lower than the one against epimastigotes (Reigada et al., 2017). The opposite was observed in this study for cinnarizine, since epimastigotes



were more sensitive to the drug than trypomastigotes. There are some possible explanations to these varying degrees of sensitivity to putrescine uptake inhibition across *T. cruzi* stages. It is possible that different forms of the parasite resort to different primary transport mechanisms of polyamines (each of them with different drug specificities) (Seguel et al., 2016). A similar possibility has been suggested in *Leishmania*, where it has been observed

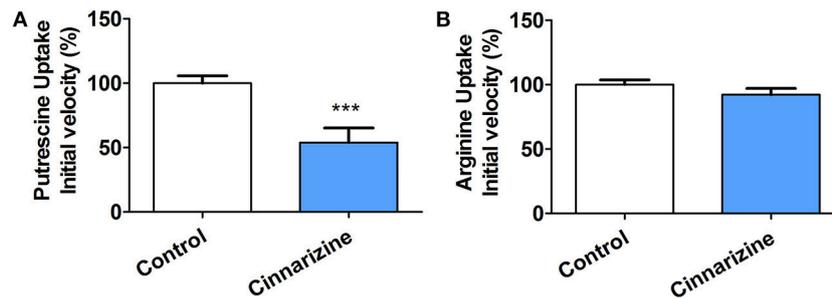
that amastigotes and promastigotes use different transporters for polyamine uptake (Müller et al., 2001; Colotti and Ilari, 2011).

The comparison of the present study with previous *in silico*-guided drug repurposing campaigns targeting *T. cruzi* polyamine uptake is hindered by the fact that very few earlier studies exist and that different virtual screening approaches have been applied in them. The first report of a virtual screening application to identify putrescine uptake inhibitors came from Alberca et al. who back in 2016 used an ensemble of six linear classifiers and identified three novel confirmed hits (out of five hits submitted to experimental screening): sertaconazole, triclabendazole, and paroxetine (Alberca et al., 2016). The same 6-model ensemble was later applied in parallel with a similarity-based screen and in sequence with target-based screening, with one out of four hits validating their predicted activity *in vitro* (Dietrich et al., 2018). For their part, Reigada et al. used a combination of similarity-based virtual screening and molecular docking, obtaining two experimentally confirmed hits (isotretinoin and acitretin) out of a total of three tested compounds (Reigada et al., 2017). The statistical comparison of the true PPV (confirmed hits over total number of hits tested) in the aforementioned virtual screening campaigns has no point due to the small number of hits tested in each occasion. Nevertheless, such PPV was in all cases, including the current study, well above the median value of hit rate in virtual screening studies (13%), as reported in a critical literature analysis of virtual screening results published between 2007 and 2011 (Zhu et al., 2013).

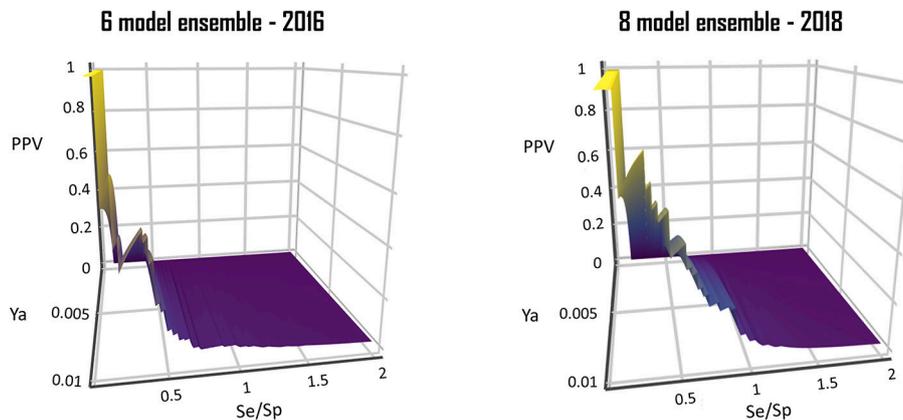
From a visual comparison of the (theoretic) PPV surfaces of the 6-model ensemble reported by Alberca et al. (2016) and the currently reported 8-model ensemble, it is clear that the 8-model ensemble reported here shows a more robust and predictable behavior in the *Se/Sp* range that goes from 0.3 to 1.0. (Figure 10).

## The Power of Ensemble Learning

The DUD-E database used as a validation tool in our retrospective virtual screening campaign confirmed that retrospective screening is the most challenging task (besides prospective screening) for a model conceived for *in silico* screening applications, as in the case of individual classifiers, the



**FIGURE 9** | Effect of 50  $\mu$ M cinnarizine and clofazimine on putrescine (A) and arginine (B) uptake in *T. cruzi* epimastigotes. Values are expressed as % mean  $\pm$  SD in comparison with control. Statistical analysis was performed by one-way ANOVA test followed by a *post-hoc* Dunnett's multiple comparison test (\*\*\*)  $p < 0.005$ .



**FIGURE 10** | Comparison of the PPV surface of the 8-model ensemble of classifiers reported here with that of the 6-model ensemble reported back in 2016.

AUROC tends to fall sharply when progressing from the training set to the DUD-E database. In fact, whereas an astonishing 82% of the individual models achieved remarkable AUROCs (above 0.8) when confronted to the training set, the percentage of individual models that obtains such achievement against the DUD-E falls to only 25%.

Echoing previous studies, though, the ensemble learning approximation led to more robust results, improving accuracy in the predictions and generalization (which, in our case, is reflected by an improved behavior on the test set and the DUD-E database).

## Do We Have Good Repurposed Candidates?

Despite two out of four hits did not confirm inhibitory effects on putrescine uptake, all of them did confirm their trypanocidal effects using concentrations in the low  $\mu$ M range. However, does this mean that they are good candidates for drug repurposing?

Not necessarily. Cinnarizine, meclizine, and butoconazole all belong to therapeutic classes that have shown potential as antichagasic therapies in the past: calcium channel blockers (Engel et al., 2010; Benaim and Paniz Mondolfi, 2012; Planer et al., 2014), antihistaminic drugs (Engel et al., 2010; Planer et al., 2014; Lara-Ramirez et al., 2017) and antifungals (Lepesheva et al., 2011).

Butoconazole, however, is not a straightforward candidate for repositioning, since it is only used topically as antifungal. Accordingly, most of the advantages of drug repurposing will be lost if the second indication (in this case, Chagas' disease chemotherapy) requires systemic administration (Oprea and Overington, 2015). Furthermore, our experimental evidence of cytotoxic activity discourages further investigation. The advantages of drug repositioning will also be mostly lost if the dose required for the second indication is above the dose range clinically used for the original one (Oprea and Overington, 2015), as would probably be the case for cinnarizine and meclizine, which are currently administered in low daily doses. This is especially true in the case of cinnarizine, which showed only weak activity against the clinically relevant forms of *T. cruzi*. The scenario is further complicated by the fact that free drug exposure will be diminished by plasma protein binding: in the case of cinnarizine, for instance, around 90% of the drug in plasma is bound to plasma proteins.

The previous analysis does not imply that is impossible to efficaciously repurpose meclizine (which showed a very interesting effect against amastigotes, inhibiting 75% growth at only 5  $\mu$ M concentration) or that both cinnarizine and meclizine are not useful as starting points for hit-to-lead programs but, in any case, many of the advantages of the strategy (i.e., bypassing some preclinical and clinical studies) will probably be lost.

## CONCLUSIONS

A cascade virtual screening approach comprising an ensemble of ligand-based classifiers and structure-based screening has been applied, identifying two new inhibitors of putrescine uptake in *T. cruzi* and reporting, for the first time, the trypanocidal effects of butoconazole (and antifungal) and cinnarizine and meclizine, two antihistaminic agents of the diphenylmethylpiperazine group commonly used to treat motion sickness and balance disorders. Interestingly, neither cinnarizine nor clofazimine modified arginine uptake by another member of the putative amino acid transporter (TcPAT) family.

This is, to our knowledge, the first report implementing PPV surface analysis to select the score value to be applied in a virtual screening campaign.

## ETHICS STATEMENT

All animal procedures were approved by institutional regulations of the Committee for the Care and Use of Laboratory Animals of the Universidad de Buenos Aires (Buenos Aires, Argentina) in accordance with government regulations [SENASA, resolution No. RS617/2002, Argentina]. All procedures were performed in accordance with institutional safety procedures.

## AUTHOR CONTRIBUTIONS

LA and JM have built and validated (*in silico*) ligand-based models and model ensembles to identify putrescine uptake inhibitors, under the supervision of AT. They later performed the first *in silico* screening step and data analysis related to ligand-based VS. RD and PP have applied structure-based filters

(molecular docking) in the virtual screening campaign. MS, MR, LF, AP, and CM have performed the *in vitro* assays against the different *Trypanosoma cruzi* stages. LA and MR have performed the transport assays. They work in CA and CC labs, under their supervision. LA, MS, CC, CA, PP, and AT have also been actively engaged in manuscript preparation and discussion. All authors took part in the result discussion for the revised version.

## FUNDING

Authors thank ANPCyT (PICT 2013-0520) and UNLP (Incentivos X785) for funding their research. All authors hold either permanent positions or fellowships at the Argentinean National Council of Science and Technological Research (CONICET).

## ACKNOWLEDGMENTS

All authors but MS hold fellowships or permanent positions at CONICET. Support was received from the National University of La Plata (UNLP) [grant X785], the Agencia Nacional de Promoción Científica y Tecnológica (ANPCyT) [grant PICT2013-0520].

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fcimb.2018.00173/full#supplementary-material>

**Data Sheet S1** | Inactive dataset compounds.

**Data Sheet S2** | Active dataset compounds.

## REFERENCES

- Alberca, L. N., Sbaraglini, M. L., Balcazar, D., Fraccaroli, L., Carrillo, C., Medeiros, A., et al. (2016). Discovery of novel polyamine analogs with anti-protozoal activity by computer guided drug repositioning. *J. Comput. Aided Mol. Des.* 30, 305–321. doi: 10.1007/s10822-016-9903-6
- Andrews, K. T., Fisher, G., and Skinner-Adams, T. S. (2014). Drug repurposing and human parasitic protozoan diseases. *Int. J. Parasitol. Drugs Drug. Resist.* 4, 95–111. doi: 10.1016/j.ijpddr.2014.02.002
- Bellera, C. L., Balcazar, D. E., Alberca, L., Labriola, C. A., Talevi, A., and Carrillo, C. (2013). Application of computer-aided drug repurposing in the search of new cruzipain inhibitors: discovery of amiodarone and bromocriptine inhibitory effects. *J. Chem. Inf. Model.* 53, 2402–2408. doi: 10.1021/ci400284v
- Bellera, C. L., Sbaraglini, M. L., Balcazar, D. E., Fraccaroli, L., Vanrell, M. C., Casassa, A. F., et al. (2015). High throughput drug repositioning for the discovery of new treatments for Chagas disease. *Mini. Rev. Med. Chem.* 15, 182–193. doi: 10.2174/138955751503150312120208
- Benaim, G., and Paniz Mondolfi, A. E. (2012). The emerging role of amiodarone and dronedarone in Chagas disease. *Nat. Rev. Cardiol.* 9, 605–609. doi: 10.1038/nrcardio.2012.108
- Bermudez, J., Davies, C., Simonazzi, A., Real, J. P., and Palma, S. (2016). Current drug therapy and pharmaceutical challenges for Chagas disease. *Acta Trop.* 156, 1–16. doi: 10.1016/j.actatropica.2015.12.017
- Browne, A. J., Guerra, C. A., Alves, R. V., da Costa, V. M., Wilson, A. L., Pigott, D. M., et al. (2017). The contemporary distribution of *Trypanosoma cruzi* infection in humans, alternative hosts and vectors. *Sci. Data.* 4:170050. doi: 10.1038/sdata.2017.50
- Carbonneau, M. A., Granger, E., Raymond, A. J., and Gagnon, G. (2016). Robust multiple-instance learning ensembles using random subspace instance selection. *Patt. Recogn.* 58, 83–99. doi: 10.1016/j.patcog.2016.03.035
- Carrillo, C., Canepa, G. E., Algranati, I. D., and Pereira, C. A. (2006). Molecular and functional characterization of a spermidine transporter (TcPAT12) from *Trypanosoma cruzi*. *Biochem. Biophys. Res. Commun.* 344, 936–940. doi: 10.1016/j.bbrc.2006.03.215
- Carrillo, C., Cejas, S., González, N. S., and Algranati, I. D. (1999). *Trypanosoma cruzi* epimastigotes lack ornithine decarboxylase but can express a foreign gene encoding this enzyme. *FEBS Lett.* 454, 192–196. doi: 10.1016/S0014-5793(99)00804-2
- Carrillo, C., Cejas, S., Huber, A., González, N. S., and Algranati, I. D. (2003). Lack of arginine decarboxylase in *Trypanosoma cruzi* epimastigotes. *J. Eukaryot. Microbiol.* 50, 312–316. doi: 10.1111/j.1550-7408.2003.tb00141.x
- Colotti, G., and Ilari, A. (2011). Polyamine metabolism in *Leishmania*: from arginine to trypanothione. *Amino acids* 40, 269–285. doi: 10.1007/s00726-010-0630-3
- Díaz, M. V., Miranda, M. R., Campos-Estrada, C., Reigada, C., Maya, J. D., Pereira, C. A., et al. (2014). Pentamidine exerts *in vitro* and *in vivo* anti *Trypanosoma cruzi* activity and inhibits the polyamine transport in *Trypanosoma cruzi*. *Acta Trop.* 134, 1–9. doi: 10.1016/j.actatropica.2014.02.012
- Dietrich, R. C., Alberca, L. N., Ruiz, D. M., Palestro, P. H., Carrillo, C., and Talevi, A., et al (2018). Identification of cisapride as new inhibitor of putrescine

- uptake in *Trypanosoma cruzi* by combined ligand- and structure-based virtual screening. *Eur. J. Med. Chem.* 149, 22–29. doi: 10.1016/j.ejmech.2018.02.006
- Doak, A. K., Wille, H., Prusiner, S. B., and Shoichet, B. K. (2010). Colloid formation by drugs in simulated intestinal fluid. *J. Med. Chem.* 53, 4259–4265. doi: 10.1021/jm100254w
- Ekins, S., Williams, A. J., Krasowski, M. D., and Freundlich, J. S. (2011). In silico repositioning of approved drugs for rare and neglected diseases. *Drug Discov. Today* 16, 298–310. doi: 10.1016/j.drudis.2011.02.016
- El Habib Daho, M., and Amine Chikh, M. (2015). Combining bootstrapping samples, random subspaces and random forests to build classifiers. *J. Med. Imaging Health Inf.* 5, 539–544. doi: 10.1166/jmihi.2015.1423
- Engel, J., Ang, K. K., Chen, S., Arkin, M. R., McKerrow, J. H., and Doyle, P. S. (2010). Image-based high-throughput drug screening targeting the intracellular stage of *Trypanosoma cruzi*, the agent of Chagas' disease. *Antimicrob. Agents Chemother.* 54, 3326–3334. doi: 10.1128/AAC.01777-09
- Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., Bourne, P. E., et al. (2000). The protein data bank. *Nucleic Acid Res.* 28, 235–242. doi: 10.1093/nar/28.1.235
- Fernández, M., Becco, L., Correia, I., Benítez, J., Piro, O. E., Echeverría, G. A., et al. (2013). Oxidovanadium(IV) and dioxovanadium(V) complexes of tridentate salicylaldehyde semicarbazones: searching for prospective antitrypanosomal agents. *J. Inorg. Biochem.* 127, 150–160. doi: 10.1016/j.jinorgbio.2013.02.010
- Ferreira, L. G., and Andricopulo, A. D. (2016). Drug repositioning approaches to parasitic diseases: a medicinal chemistry perspective. *Drug Discov. Today* 21, 1699–1710. doi: 10.1016/j.drudis.2016.06.021
- Hasne, M. P., Soysa, R., and Ullman, B. (2016). The *Trypanosoma cruzi* diamine transporter is essential for robust infection of mammalian cells. *PLoS ONE* 11:e0152715. doi: 10.1371/journal.pone.0152715
- World Health Organization (2015). Chagas disease in Latin America: an epidemiological update based on 2010 estimates. *Wkly. Epidemiol.* 90, 33–44.
- Jin, G., and Wong, S. T. (2014). Toward better drug repositioning: prioritizing and integrating existing methods into efficient pipelines. *Drug Discov. Today* 19, 637–644. doi: 10.1016/j.drudis.2013.11.005
- Kaiser, M., Mäser, P., Tadoori, L. P., Ioset, J.-R., and Brun, R. (2015). Antiprotozoal activity profiling of approved drugs: a starting point toward drug repositioning. *PLoS ONE* 10:e0135556. doi: 10.1371/journal.pone.0135556
- Klug, D. M., Gelb, M. H., and Pollastra, M. P. (2016). Repurposing strategies for tropical disease drug discovery. *Bioorg. Med. Chem. Lett.* 1, 2569–2576. doi: 10.1016/j.bmcl.2016.03.103
- Lara-Ramírez, E. E., López-Cedillo, J. C., Noguera-Torres, B., Kashif, M., Garcia-Perez, C., Bocanegra-García, V., et al. (2017). An *in vitro* and *in vivo* evaluation of new potential trans-sialidase inhibitors of *Trypanosoma cruzi* predicted by a computational drug repositioning method. *Eur. J. Med. Chem.* 132, 249–261. doi: 10.1016/j.ejmech.2017.03.063
- Law, V., Knox, C., Djoumbou, Y., Jewison, T., Guo, A. C., Liu, Y., et al. (2014). DrugBank 4.0: shedding new light on drug metabolism. *Nucleic Acids Res.* 42, D1091–1097. doi: 10.1093/nar/gkt1068
- Lepesheva, G. I., Villalta, F., and Waterman, M. R. (2011). Targeting *Trypanosoma cruzi* Sterol 14 $\alpha$ -Demethylase (CYP51). *Adv. Parasitol.* 75, 65–87. doi: 10.1016/B978-0-12-385863-4.00004-6
- Lipinski, C. A., Lombardo, F., Dominy, B. W., and Feeney, P. J. (2001). Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug. Deliv. Rev.* 46, 3–26. doi: 10.1016/S0169-409X(00)00129-0
- Melo, F., and Sali, A. (2007). Fold assessment for comparative protein structure modeling. *Protein Sci.* 16, 2412–2426. doi: 10.1110/ps.072895107
- Min, S. H. (2016). A genetic algorithm-based heterogeneous random subspace ensemble model for bankruptcy prediction. *Int. J. Appl. Eng. Res.* 11, 2937–2931.
- Miranda, C. G., Solana, M. E., Curto Mde, L., Lammel, E. M., Schijman, A. G., and Alba Soto, C.D. (2015). A flow cytometer-based method to simultaneously assess activity and selectivity of compounds against the intracellular forms of *Trypanosoma cruzi*. *Acta Trop.* 152, 8–16. doi: 10.1016/j.actatropica.2015.08.004
- Monge, A., Arrault, A., Marot, C., and Morin-Allory, L. (2006). Managing, profiling and analyzing a library of 2.6 million compounds gathered from 32 chemical providers. *Mol. Divers.* 10, 389–403. doi: 10.1007/s11030-006-9033-5
- Morillo, C. A., Marin-Neto, J. A., Avezum, A., Sosa-Estani, S., Rassi, A. Jr., Rosas, F., et al. (2015). BENEFIT Investigators. Randomized trial of benznidazole for chronic Chagas' cardiomyopathy. *N. Engl. J. Med.* 373, 1295–1306. doi: 10.1056/NEJMoa1507574
- Mosmann, T. (1983). Rapid colorimetric assay for cellular growth and survival: application to proliferation and cytotoxicity assays. *J. Immunol. Methods* 65, 55–63. doi: 10.1016/0022-1759(83)90303-4
- Müller, S., Coombs, G. H., and Walter, R. D. (2001). Targeting polyamines of parasitic protozoa in chemotherapy. *Trends Parasitol.* 17, 242–249. doi: 10.1016/S1471-4922(01)01908-0
- Mysinger, M. M., Carchia, M., Irwin, J. J., and Shoichet, B. K. (2012). Directory of useful decoys, enhanced (DUD-E): better ligands and decoys for better benchmarking. *J. Med. Chem.* 55, 6582–6594. doi: 10.1021/jm300687e
- Novick, P. A., Ortiz, O. F., Poelman, J., Abdulhay, A. Y., and Pande, V. S. (2013). SWEETLEAD: an *in silico* database of approved drugs, regulated chemicals, and herbal isolates for computer-aided drug discovery. *PLoS ONE* 8:e79568. doi: 10.1371/journal.pone.0079568
- Nunes, M. C., Dones, W., Morillo, C. A., Encina, J. J., and Ribeiro, A. L. (2013). Council on chagas disease of the interamerican society of cardiology. Chagas disease. An overview of clinical and epidemiological aspects. *J. Am. Coll. Cardiol.* 62, 767–776. doi: 10.1016/j.jacc.2013.05.046
- Oprea, T. I., and Overington, J. P. (2015). Computational and practical aspects of drug repositioning. *Drug Repurp. Rescue Reposit.* 1, 28–35. doi: 10.1089/drrr.2014.0009
- Perez-Lamas, C., and Lopez-Bigas, N. (2011). Gitools: analysis and visualization of genomic data using interactive heat-maps. *PLoS ONE* 6:e19541. doi: 10.1371/journal.pone.0019541
- Planer, J. D., Hulverson, M. A., Arif, J. A., Ranade, R. M., Don, R., and Buckner, F. S. (2014). Synergy testing of FDA-approved drugs identifies potent drug combinations against *Trypanosoma cruzi*. *PLoS Negl. Trop. Dis.* 8:e2977. doi: 10.1371/journal.pntd.0002977
- Rassi, A. Jr., Rassi, A., and Marin-Neto, J. A. (2010). Chagas disease. *Lancet* 375, 1388–1402. doi: 10.1016/S0140-6736(10)60061-X
- Reigada, C., Phanstiel, O., Miranda, M. R., and Pereira, C. A. (2018). Targeting polyamine transport in *Trypanosoma cruzi*. *Eur. J. Med. Chem.* 147, 1–6. doi: 10.1016/j.ejmech.2018.01.083
- Reigada, C., Valera-Vera, E. A., Sayé, M., Errasti, A. E., Avila, C. C., Miranda, M. R., et al. (2017). Trypanocidal effect of isotretinoin through the inhibition of polyamine and amino acid transporters in *Trypanosoma cruzi*. *PLoS Negl. Trop. Dis.* 11:e0005472. doi: 10.1371/journal.pntd.0005472
- Robin, X., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sanchez, J. C., et al. (2011). pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics* 12:77. doi: 10.1186/1471-2105-12-77
- Sbaragli, M. L., Vanrell, M. C., Bellera, C. L., Benaim, G., Carrillo, C., Talevi, A., et al. (2016). Neglected tropical protozoan diseases: drug repositioning as a rational option. *Curr. Top. Med. Chem.* 16, 2201–2222. doi: 10.2174/1568026616666160216154309
- Seguel, V., Castro, L., Reigada, C., Cortes, L., Díaz, M. V., Miranda, M. R., et al. (2016). Pentamidine antagonizes the benznidazole's effect *in vitro*, and lacks of synergy *in vivo*: Implications about the polyamine transport as an anti-*Trypanosoma cruzi* target. *Exp. Parasitol.* 171, 23–32. doi: 10.1016/j.exppara.2016.10.007
- Shen, M. Y., and Sali, A. (2006). Statistical potential for assessment and prediction of protein structures. *Protein Sci.* 15, 2507–2524. doi: 10.1110/ps.062416606
- Soysa, R., Venselaar, H., Poston, J., Ullman, B., and Hasne, M. P. (2013). Structural model of a putrescine-cadaverine permease from *Trypanosoma cruzi* predicts residues vital for transport and ligand binding. *Biochem. J.* 452, 423–432. doi: 10.1042/BJ20130350
- Stanaway, J. D., and Roth, G. (2015). The burden of chagas disease. *Estim. Chall. Glob. Heart.* 10, 139–144. doi: 10.1016/j.jgheart.2015.06.001
- Sykes, M. L., and Avery, V. M. (2013). Approaches to protozoan drug discovery: phenotypic screening. *J. Med. Chem.* 56, 7727–7740. doi: 10.1021/jm4004279
- Toropova, A. P., and Toropov, A. A. (2017). CORAL: binary classifications (active/inactive) for drug-induced liver injury. *Toxicol. Lett.* 268, 51–57. doi: 10.1016/j.toxlet.2017.01.011
- Truchon, J. F., and Bayly, C. L. (2007). Evaluating virtual screening methods: good and bad metrics for the “early recognition” problem. *J. Chem. Inf. Model.* 47, 488–508. doi: 10.1021/ci600426e

- Veber, D. F., Johnson, S. R., Cheng, H. Y., Smith, B. R., Ward, K. W., and Kopple, K. D. (2002). Molecular properties that influence the oral bioavailability of drug candidates. *J. Med. Chem.* 45, 2615–2623. doi: 10.1021/jm020017n
- Ventura-Garcia, L., Roura, M., Pell, C., Posada, E., Gascón, J., Aldasoro, E., et al. (2013). Socio-cultural aspects of chagas disease: a systematic review of qualitative research. *PLoS Negl. Trop. Dis.* 7:e2410. doi: 10.1371/journal.pntd.0002410
- Vyskovsky, R., Schwarz, D., Janousova, E., and Kasperek, T. (2016). “Random subspace ensemble artificial neural networks for first-episode Schizophrenia classification,” in *Proceedings of the 2016 Federated Conference on Computer Science and Information Systems* (Gdansk: FedCSIS), 317–321.
- Webb, B., and Sali, A. (2016). Comparative protein structure modeling using MODELLER. *Curr. Protoc. Bioinform.* 54, 5.6.1–5.6.0. doi: 10.1002/0471250953.bi0506s15
- Zhang, Q., and Muegge, I. (2006). Scaffold hopping through virtual screening using 2D and 3D similarity descriptors: ranking, voting, and consensus scoring. *J. Med. Chem.* 49, 1536–1548. doi: 10.1021/jm050468i
- Zhu, T., Cao, S., Su, P. C., Patel, R., Shah, D., Chokshi, H. B., et al. (2013). Hit identification and optimization in virtual screening: practical recommendations based upon a critical literature analysis. *J. Med. Chem.* 56, 6560–6572. doi: 10.1021/jm301916b

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Alberca, Sbaraglini, Morales, Dietrich, Ruiz, Pino Martínez, Miranda, Fraccaroli, Alba Soto, Carrillo, Palestro and Talevi. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.