



SciForum  
MOL2NET

## An approach toward the identification of new antileishmaniasic compounds.

Naiví Flores Balmaseda<sup>1</sup>, Orlando Álvarez González<sup>1</sup> and Juan A. Castillo-Garit<sup>1,2,\*</sup>

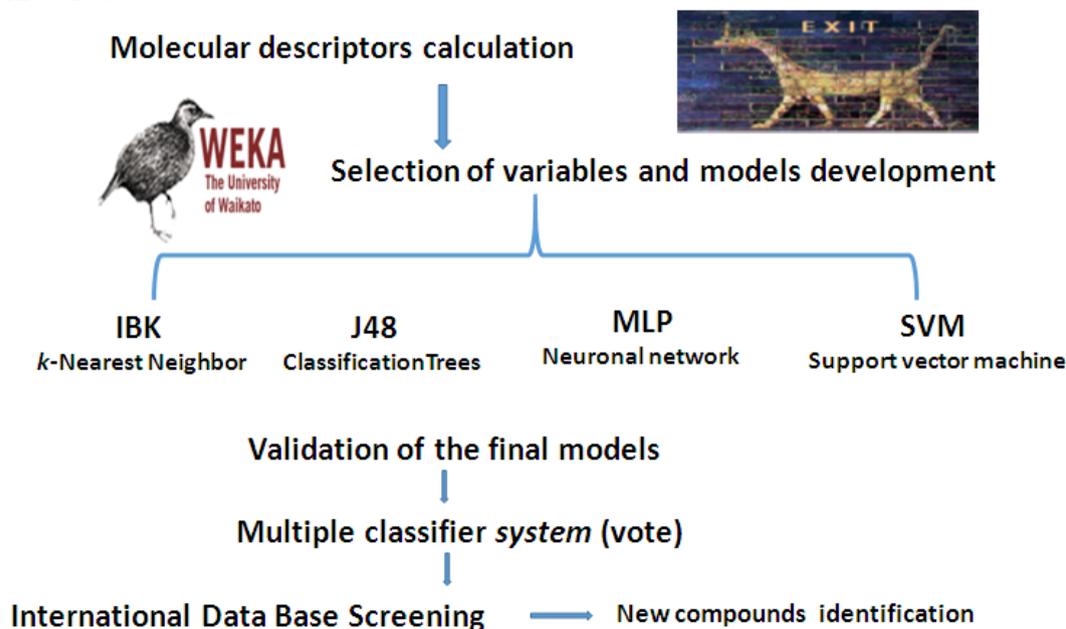
<sup>1</sup> Unit of Computer-Aided Molecular “Biosilico” Discovery and Bioinformatic Research (CAMD-BIR Unit), Facultad de Química-Farmacía, Universidad Central “Marta Abreu” de Las Villas, Santa Clara 54830, Villa Clara, Cuba.

<sup>2</sup> Unidad de Toxicología Experimental, Universidad de Ciencias Médicas “Dr. Serafín Ruiz de Zárate Ruiz”, Santa Clara 50200, Villa Clara, Cuba.

\* Corresponding author at: Unidad de Toxicología Experimental, Universidad de Ciencias Médicas “Dr. Serafín Ruiz de Zárate Ruiz”, Santa Clara 50200, Villa Clara, Cuba. E-mail addresses: [jacgarit@yahoo.es](mailto:jacgarit@yahoo.es), [juancgarit@infomed.sld.cu](mailto:juancgarit@infomed.sld.cu) Tel.: +53 42 273236

**Abstract:** Herein we present results of a quantitative structure–activity relationship (QSAR) study to identify new antileishmaniasic compounds (*Leishmania amazonensis*) by using a set of more than 2000 0D-2D Dragon’s molecular descriptors and machine learning techniques. A data set of organic chemicals, with antileishmaniasic activity against promastigote forms of the parasite, is used to develop four QSAR models based on *k*-nearest neighbors, Support Vector Machine, Multi-Layer Perceptron and classification tree techniques. External validation procedures were developed to demonstrate the predictive power of the models. Promastigote’s models correctly classify more than 89% chemicals in both training and external prediction groups, respectively. In addition to the individual techniques an assembled system of majority vote was personalized with the aim of improving the results of the obtained models. To identify new compounds with potential activity against this parasite, a virtual screening was performed using DrugBank international database. There were identified more than five hundred new potential antileishmaniasic compounds. The current results constitute a step forward in the search for efficient ways to discover new antileishmaniasic lead compounds.

**Keywords:** antileishmaniasic, *Leishmania amazonensis*, machine learning, virtual screening.

**Graphical Abstract:****Introduction:**

Leishmaniasis, is a disease caused by obligate intracellular protozoa of the genus *Leishmania*, is an old but largely unknown disease that afflicts the World's poorest populations [1]. It presents a broad spectrum of clinical forms and is transmitted to humans and animals through the bite of insects of the *Psychodidae* family [2]. There have been reported by WHO (World Health Organization) more than 20 species of *Leishmania* and between them *Leishmania amazonensis* is of vital importance for the American continent because it is the cause of a wide variety of clinical manifestations, some of them potentially fatal [3-4].

"*In silico*" methods are useful tools for screening chemicals, especially in early stages of the drug discovery process [5-8]. In the last two decades these studies have played a fundamental role in the development of a number of drugs that are currently on the market [9].

**Materials and Methods:**

All the compounds included in the research were gathered from published in PubChem bioassays. We select specifically studies carried out against promastigotes of *Leishmania amazonensis*. Different researchers have reported

them and publish it in the last years, in several journals with high impact on the Web of Sciences. To verify the structural diversity of the compounds of the database a Cluster Analysis (CA) implemented in the software STATISTICA 8.0 was performed [10].

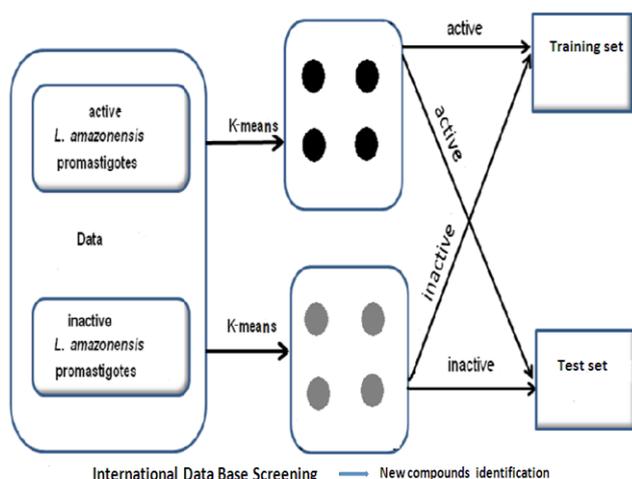
Models by using k-nearest neighbors (IBK), classification trees (J48), artificial neural network (MLP for its acronym MultiLayer Perceptron) and support vector machine (SMO for Sequential Minimal Optimization) techniques were obtained for promastigote form of the parasite, with the employee of WEKA software [11]. External validation procedures were developed to demonstrate the predictive power of the four resultants models. Virtual screening was performed using DrugBank international database.

**Results and Discussion:**

A new database of antileishmaniasis compounds with a high degree of structural variability was performed. The parameterization of the structures was carried out using 2489 molecular descriptors 0D-2D implemented in the DRAGON software. WEKA's selection procedures were used to obtain a subset of variables for models development.

Active and inactive compounds were divided into different subsets using k-MCA so we could obtain Training and Test sets following the procedure shown in the Figure 1.

Promastigote's models correctly classify more than 82% and 80% of chemicals in both training and external prediction groups, respectively. Figure 2 shows the Accuracy percentages obtained for the models SVM results the higher accuracy model followed by IBK and MLP respectively.



**Figure 1** Training and Test Sets obtention procedure.

The external validation of the four models for promastigotes using a new set of 22 compounds previously evaluated in PubChem bioassays showing positive results. In addition to the individual techniques an assembled system of majority voting was personalized with the aim of improving the results of the obtained models [12].

## Conclusions

With the use of artificial intelligence techniques we develop four validated models with good statistical parameters. The obtained models were able to identify new compounds with potential activity against promastigote forms of *L. amazonensis* through virtual screening of databases. This work constitutes a useful tool in the search of new leading compounds against this parasite.

**Conflicts of Interest:** "The authors declare no conflict of interest"

## References:

- [1] Castillo-Garit, J.A.; Abad, C.; Rodriguez-Borges, J.E.; Marrero-Ponce, Y.; Torrens, F. A Review of QSAR studies to Discover New Drug-like Compounds Actives Against Leishmaniasis and Trypanosomiasis. *Curr. Top. Med. Chem.*, **2012**, *12* (8), 852-865.



**Figure 2.** Accuracy percentages obtained for training and test series on final models

To identify new compounds with potential activity against promastigotes forms of *L. amazonensis* a virtual screening using DrugBank international database was performed. There were identified more than five hundred new potential antileishmaniasis compounds. The use of the assembly by the majority vote enabled us to reduce the screening compounds identified as potentially active. These compounds can be experimentally evaluated to corroborate their activity against *L. amazonensis* with favorable repercussion in a time saving and use of chemical reagents, so the current results constitute a step forward in the search for efficient ways to discover new antileishmaniasis lead compounds.

- [2] Organización Panamericana de la Salud Leishmaniasis: Informe Epidemiológico en las Américas. Available in <http://www.paho.org/> (accessed on 12 sept 2016).
- [3] Montalvo Álvarez, A.M. Leishmaniasis: Interesting features on exotic parasitism for Cuba. *Rev. Cub. Higiene Epidemiol.*, **2010**, *48*, 78-92.
- [4] Ramos, J.M.; Segovia, M. Estado actual del tratamiento farmacológico de la leishmaniasis. Available from: [http://www.seq.es/seq/html/revista\\_seq/0197/rev1.html](http://www.seq.es/seq/html/revista_seq/0197/rev1.html).
- [5] Castillo-Garit, J.A.; Del Toro-Cortés, O.; Kouznetsov, V.V.; Puentes, C.O.; Romero Bohórquez, A.R.; Vega, M.C.; Rolón, M.; Escario, J.A.; Gómez-Barrio, A.; Marrero-Ponce, Y.; Torrens, F.; Abad, C. Identification In Silico and In Vitro of Novel Trypanosomicidal Drug-Like Compounds. *Chem. Biol. Drug Des.*, **2012**, *80* (1), 38-45.
- [6] Castillo-Garit, J.A.; Marrero-Ponce, Y.; Barigye, S.J.; Medina-Marrero, R.; Bernal, M.G.; Vega, J.M.G.d.l.; Torrens, F.; Arán, V.J.; Pérez-Giménez, F.; García-Domenech, R.; Acevedo-Barrios, R. In silico Antibacterial Activity Modeling Based on the TOMOCOMD-CARDD Approach. *J. Braz. Chem. Soc.*, **2015**, *26*, 1218-1226.
- [7] Marrero-Ponce, Y.; Martínez-Albelo, E.; Casañola-Martín, G.; Castillo-Garit, J.; Echeverría-Díaz, Y.; Zaldivar, V.; Tygat, J.; Borges, J.; García-Domenech, R.; Torrens, F.; Pérez-Giménez, F. Bond-based linear indices of the non-stochastic and stochastic edge-adjacency matrix. 1. Theory and modeling of *ChemPhys* properties of organic molecules. *Mol. Div.*, **2010**, *14* (4), 731-753.
- [8] Le-Thi-Thu, H.; Canizares-Carmenate, Y.; Marrero-Ponce, Y.; Torrens, F.; Castillo-Garit, J.A. Prediction of Caco-2 Cell Permeability Using Bilinear Indices and Multiple Linear Regression. *Lett. Drug Des. Discov.*, **2016**, *13* (2), 161-169.
- [9] Geary, T.G.; Sakanari, J.A.; Caffrey, C.R. Anthelmintic Drug Discovery: Into the Future. *J. Parasitol.*, **2015**, *101* (2), 125-133.
- [10] *STATISTICA (data analysis software system)*, StatSoft, Inc. version 6. [www.statsoft.com](http://www.statsoft.com): Tulsa, 2001.
- [11] Witten, I.H.; Frank, E., *Data Mining: Practical Machine Learning Tools and Techniques*. Second ed.; Morgan Kaufmann Publishers: **2005**.
- [12] Dietterich, T.G., Ensemble methods in machine learning. In *International workshop on multiple classifier system*, Springer Berlin Heidelberg, **2000**; pp 1-15.

© **2015** by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions defined by MDPI AG, the publisher of the Sciforum.net platform. Sciforum papers authors the copyright to their scholarly works. Hence, by submitting a paper to this conference, you retain the copyright, but you grant MDPI AG the non-exclusive and unrevocable license right to publish this paper online on the Sciforum.net platform. This means you can easily submit your paper to any scientific journal at a later stage and transfer the copyright to its publisher (if required by that publisher). (<http://sciforum.net/about> ).