MOL2NET, **2020**, 6, **ISSN**: 2624-5078,
http://sciforum.net/conference/mol2net-06/chembiomol-01

CHEMBIOMOL-01: Chem. Biol. & Med. Chem. Workshop, Rostock, Germany-Galveston, Texas, USA, 2020

MOL2NET, International Conference Series on Multidisciplinary Sciences

PTMLIF model of Metabolic Reaction Networks and ChEMBL Antibacterial Compounds

Karel Diéguez-Santana [a, b], Gerardo M. Casañola-Martin [c,d], James R Green [d], and Bakhtiyor Rasulev [d]

[a] *Departamento Ciencias de la Vida, Universidad Estatal Amazónica, km 2 1/2 Vía Tena, Puyo, Pastaza, Ecuador.*

[b] *Department of Organic Chemistry II, University of the Basque Country UPV/EHU, 48940, Leioa, Spain.*

[c] *Department of Systems and Computer Engineering, Carleton University, K1S 5B6, Ottawa, ON, Canada.*

[d] *Department of Coatings and Polymeric Materials, North Dakota State University, Fargo, ND, 58102, USA*

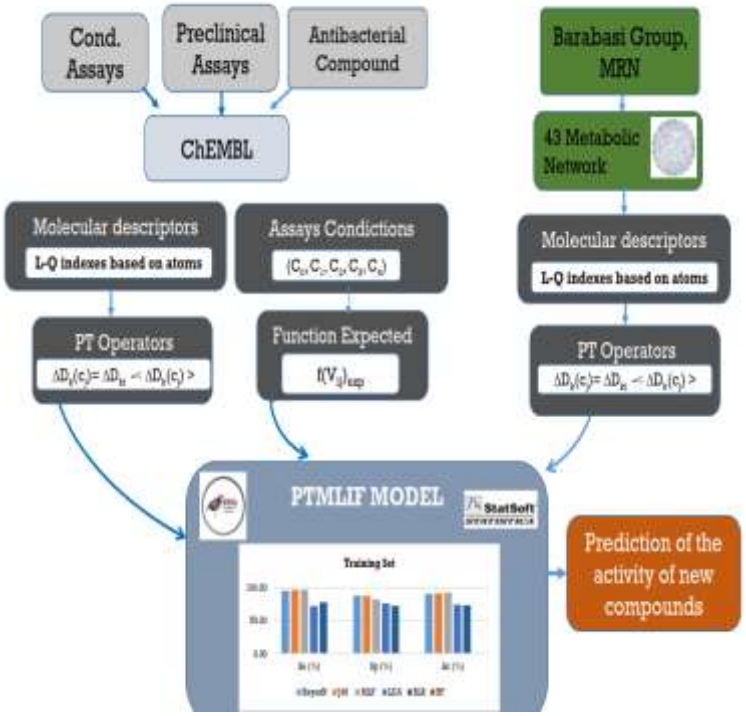
*** Corresponding authors:** K. Dieguez-Santana (karel.dieguez.santana@gmail.com) and G. M. Casañola-Martin gmaikelc@gmail.com)

.

.

MOL2NET, **2020**, 6, **ISSN**: 2624-5078,
http://sciforum.net/conference/mol2net-06/chembiomol-01

Graphical Abstract



Abstract.

Antimicrobial resistance has prompted research and development of new antibiotic treatments. Efforts to discover new drugs with antibacterial activity have generated large data sets from multiple preclinical trials with different experimental conditions. Predicting the activity of new chemical compounds on pathogenic microorganisms with different Metabolic Reaction Networks (MRNs) has become an important objective in the field. PTMLIF (Perturbation Theory, Machine Learning and Information Fusion) models are the combination of perturbation theory with machine learning and information fusion. In this document, we merge >100000 preclinical antibacterial assays from the ChEMBL database with the structural information for >40 MRNs of different microorganisms reported by the Barabási group. Non-linear PTMLIF models were applied applying Random Forest (RF), J48-decision tree, and Bayesian Network (BN) algorithms. BN, and RF models presented better results. specificity (>88%), sensitivity (>95%), AUROC (>95%). and accuracy (~90%), In this work we also demonstrated the power of information fusion of experimental characteristics of drugs/compounds and MRN for the prediction of antibacterial activity of chemical compounds.

Keywords: Machine Learning; Perturbation Theory; Antibacterial activity; Information Fusion; Bayesian Network; Random Forest.

References .

1. Jeong, H.; Tombor, B.; Albert, R.; Oltval, Z.N.; Barabásl, A.L. The large-scale organization of metabolic networks. *Nature* **2000**, *407*, 651-654, doi:10.1038/35036627.
2. Zaengle-Barone, J.M.; Jackson, A.C.; Besse, D.M.; Becken, B.; Arshad, M.; Seed, P.C.; Franz, K.J. Copper Influences the Antibacterial Outcomes of a β-Lactamase-Activated Prochelator against Drug-Resistant Bacteria. *ACS Infectious Diseases* **2018**, *4*, 1019-1029, doi:10.1021/acsinfecdis.8b00037.
3. Bediaga, H.; Arrasate, S.; González-Díaz, H. PTML Combinatorial Model of ChEMBL Compounds Assays for Multiple Types of Cancer. *ACS Combinatorial Science* **2018**, *20*, 621-632, doi:10.1021/acscombsci.8b00090.
4. Casañola-Martin, G.M.; Le-Thi-Thu, H.; Pérez-Giménez, F.; Marrero-Ponce, Y.; Merino-Sanjuán, M.; Abad, C.; González-Díaz, H. Multi-output model with Bob-Jenkins operators of linear indices to predict multi-target networks of ubiquitin-proteasome system inhibitors. *Mol. Divers.* **2015**, *19*, 347-356.
5. Dieguez-Santana, K.; Pham-The, H.; Rivera-Borroto, O.M.; Puris, A.; Le-Thi-Thu, H.; Casanola-Martin, G.M. A Two QSAR Way for Antidiabetic Agents Targeting Using α-Amylase and α-Glucosidase Inhibitors: Model Parameters Settings in Artificial Intelligence Techniques. *Letters in Drug Design & Discovery* **2017**, *14*, 862-868, doi:10.2174/1570180814666161128121142.
6. Diéguez-Santana, K.; Rivera-Borroto, O.M.; Puris, A.; Pham-The, H.; Le-Thi-Thu, H.; Rasulev, B.; Casañola-Martin, G.M. Beyond Model Interpretability using LDA and Decision Trees for α-Amylase and α-Glucosidase Inhibitor Classification Studies. *Chemical Biology & Drug Design* **2019**, 10.1111/cbdd.13518, doi:10.1111/cbdd.13518.
7. Nocedo-Mena, D.; Cornelio, C.; Camacho-Corona, M.d.R.; Garza-González, E.; Waksman de Torres, N.; Arrasate, S.; Sotomayor, N.; Lete, E.; González-Díaz, H. Modeling Antibacterial Activity with Machine Learning and Fusion of Chemical Structure Information with Microorganism Metabolic Networks. *Journal of Chemical Information and Modeling* **2019**, *59*, 1109-1120, doi:10.1021/acs.jcim.9b00034.
8. Speck-Planche, A.; Kleandrova, V.V.; Cordeiro, M.N.D.S. Chemoinformatics for rational discovery of safe antibacterial drugs: Simultaneous predictions of biological activity against streptococci and toxicological profiles in laboratory animals. *Bioorganic and Medicinal Chemistry* **2013**, *21*, 2727-2732, doi:10.1016/j.bmc.2013.03.015.