[P17] Automated Generation of 2D-Pharmacophores from Large Datasets

Jean-Philippe Métivier[1,2], Antoine Abel[1], Alban Lepailleur[1], Bertrand Cuissart[2], Ronan Bureau[1]

[1]*CERMN (CNRS 3038 INC3M), Université de Caen Normandie, Caen, France*
[2]*GREYC (CNRS UMR 6072), Université de Caen Normandie, Caen, France*

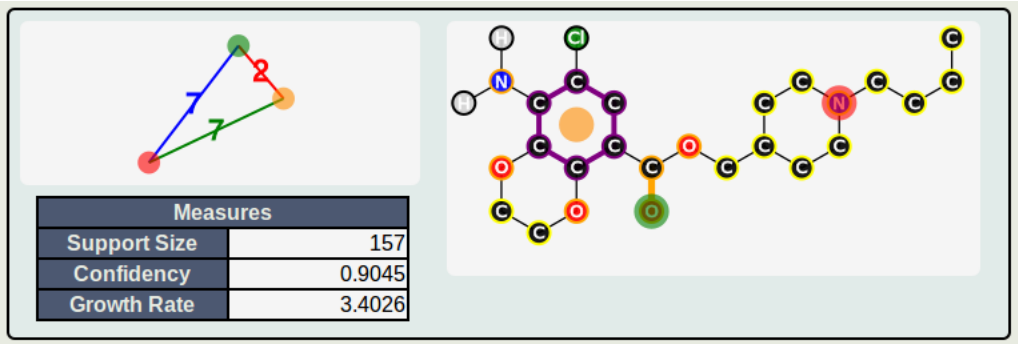
Introduced by M. Kier in the late sixties, pharmacophores have been widely used in chemoinformatics, from supporting query in molecular databases to guiding drug design. Usually, a pharmacophore is computed through the alignment of a small set of carefully selected active molecules [1]. Even if it is able to exhibit highly relevant pharmacophores, such a computation requires an important investissement from experts. This work presents a method to compute 2D-pharmacophores from a large set of molecules.

Our procedure follows three steps: 1) molecules are labelled with pharmacophoric features (HBA, HBD, ArR, Hyd, Pos, Neg) and topological distances between these feautres, 2) recurring situations are computed together with their distance constraints, 3) discovered 2D-Pharmacophores are ranked according to their quality.

Features are annotated using OpenBabel [2] and PHASE definitions [3]. Once these annotations are done, topological distances (i.e. the number of bonds betwen annotated pharmacophoric features) are simply computed using Dijkstra's algorithm.

To avoid the search of unfrequent pharmacophores, the list of all frequent enough feature patterns is extracted using an APRIORI algorithm [4]. Then, each feature pattern is turned into 2D-Pharmacophores by adding topological distance constraints. A feature pattern with a complete description of its inter-feature distances is a 2D-pharmacophores (see Figure).

To evaluate the quality of 2D-pharmacophores, several measures are available to experts: the support size (i.e. the number of molecules exhibiting a pharmacophore), the confidency (i.e. the ratio of active molecules fitting a pharmacophore), the growth rate (i.e. an adjusted confidency). Futhermore, all 2D-pharmacophores, their supporting molecules, and measures are reported in webpages to allow a quick analysis of the discovered knowledge.



Bibliography:

[1] A. R. Leach, V. J. Gillet, R. A. Lewis, R. Taylor, J, Med. Chem., 53 (2010) 539-558
[2] N. M. O'Boyle, M. Banck, C. A. James, C. Morley, T. Vandermeersch, G. R. Hutchison, J. Cheminfo., 3 (2011) 33
[3] S.L. O Dixon, A.M. Smondyrev, E.H. Knoll, S.N. Rao, D.E. Shaw, R;A. Friesner, J. Comput. Aided Mol. Des., 20 (2006) 647-671
[4] R, Agrawal, R. Srikant, Proceedings of 20th International Conference on Very Large Data Bases, (1994) 487-499