Journal of Cheminformatics

ORAL PRESENTATION Open Access

Winnow based identification of potent hERG inhibitors *in silico*: comparative assessment on different datasets

Richard L Marchese Robinson[1*], Robert C Glen[1], John BO Mitchell[2]

From 7th German Conference on Chemoinformatics: 25 CIC-Workshop
Goslar, Germany. 6-8 November 2011

Due to the potentially lethal effects of potent hERG inhibition, *in silico* approaches which can identify potent (IC_{50} < 1 μM) inhibitors are of considerable interest to the pharmaceutical industry [1]. We present recent work [2] in which *in silico* binary classifiers were trained to discriminate potent inhibitors from compounds exhibiting weaker (IC_{50} ≥ 1 μM) inhibition. Initial models were based on a version of the memory efficient Winnow algorithm [3]. These initial models were generated using various descriptor sets. The descriptor set yielding the best cross-validated initial Winnow model was used to build models using each of Winnow, Random Forest and Support Vector Machine. Analysis of the contributions of different substructural and physiochemical features in the final Winnow models indicates they may be interpreted, albeit with caution. All final models were externally validated, with no algorithm consistently outperforming the others. These approaches were directly compared, on various datasets, to those proposed by Thai and Ecker [4] and by Dubus et al. [5]. The results indicate that the Winnow models are competitive with earlier approaches proposed in the literature. The findings also emphasise a potential difficulty when seeking to estimate the predictive power of *in silico* models on small quantities of data: model performance may vary considerably, particularly when training and validating on different datasets.

Author details
[1]Unilever Centre for Molecular Science Informatics, Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge, CB2 1EW, UK.

[2]EaStCHEM School of Chemistry and Biomedical Sciences Research Centre, University of St Andrews, North Haugh, St Andrews, KY16 9ST,UK.

Published: 1 May 2012

References

1. Gavaghan CL, Hasselgren Arnby C, Blomberg N, Strandlund G, Boyer S: **Development, interpretation and temporal evaluation of a global QSAR of hERG electrophysiology screening data.** *J Comput-Aided Mol Des* 2007, **21**:189-206.
2. Marchese Robinson RL, Glen RC, Mitchell JBO: **Development and Comparison of hERG Blocker Classifiers: Assessment on Different Datasets Yields Markedly Different Results.** *Mol Inf* 2011, **30**:443-458.
3. Nigsch F, Mitchell JBO: **How To Winnow Actives from Inactives: Introducing Molecular Orthogonal Sparse Bigrams (MOSBs) and Multiclass Winnow.** *J Chem Inf Model* 2008, **48**:306-318.
4. Thai KM, Ecker GF: **A binary QSAR model for classification of hERG potassium channel blockers.** *Bioorg Med Chem* 2008, **16**:4107-4119.
5. Dubus E, Ijjaali I, Petitet F, Michel A: **In Silico Classification of hERG Channel Blockers: a Knowledge-Based Strategy.** *ChemMedChem* 2006, **1(6)**:622-630.

doi:10.1186/1758-2946-4-S1-O6
Cite this article as: Marchese Robinson *et al.*: Winnow based identification of potent hERG inhibitors *in silico*: comparative assessment on different datasets. *Journal of Cheminformatics* 2012 **4** (Suppl 1):O6.

* Correspondence: rlm46@cam.ac.uk
[1]Unilever Centre for Molecular Science Informatics, Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge, CB2 1EW, UK
Full list of author information is available at the end of the article