

REVIEW OUTLINE (papers outlined in the talk are highlighted below)

- **Introduction**
 - References 1, 2, 10, 12, 18, 19, 22, 24, 29, 30
- 1. **Generic Scoring Functions: ranking diverse protein-ligand structures**
 - References 1, 2, 4, 10, 11, 15, 16, 17, 20, 25, 27, 32
- 2. **Family-specific Scoring Functions: ranking protein-ligand structures of a target**
 - References 3, 4, 11, 13, 14, 17, 20, 21, 23, 28
- 3. **Family-specific Scoring Functions: ranking docking poses of a target**
 - References 6, 8, 26, 32
- 4. **Interpreting Scoring Functions: variable importance and feature selection**
 - References 1, 10, 11, 15, 21, 33
- 5. **Virtual Screening with ML Scoring Functions: regression**
 - References 4, 9, 13, 31 (with generic SF), 32
- 6. **Virtual Screening with ML Scoring Functions: binary classification**
 - References 3, 4, 5, 6, 8 (check refs 27-32 in that paper), 9, 13, 14, 32
- **Future prospects**
 - Each student to write 2-3 paragraphs from their respective sections here.

NB> Student 1 (sections 1&4), student 2 (sections 2&3) and student 3 (sections 5&6).

RELEVANT LITERATURE

1. Ballester PJ, Mitchell JBO. A machine learning approach to predicting protein-ligand binding affinity with applications to molecular docking. *Bioinformatics*. 2010 May 1;26(9):1169–75.
2. Ballester PJ, Mitchell JBO. Comments on ‘Leave-Cluster-Out Cross-Validation Is Appropriate for Scoring Functions Derived from Diverse Protein Data Sets’: Significance for the Validation of Scoring Functions. *J Chem Inf Model*. 2011 May 31
<http://www.ncbi.nlm.nih.gov/pubmed/21591735>
3. Kinnings SL, Liu N, Tonge PJ, Jackson RM, Xie L, Bourne PE. A Machine Learning-Based Method To Improve Docking Scoring Functions and Its Application to Drug Repurposing. *Journal of Chemical Information and Modeling*. 2011 Feb 28;51(2):408–19. [there is also a short erratum]
4. Durrant JD, McCammon JA. NNScore 2.0: A Neural-Network Receptor–Ligand Scoring Function. *J. Chem. Inf. Model*. 2011;51(11):2897–903.
5. Durrant JD, McCammon JA. NNScore: A Neural-Network-Based Scoring Function for the Characterization of Protein–Ligand Complexes. *Journal of Chemical Information and Modeling*. 2010 Oct 25;50(10):1865–71.
6. Sato T, Honma T, Yokoyama S. Combining Machine Learning and Pharmacophore-Based Interaction Fingerprint for in Silico Screening. *J. Chem. Inf. Model*. 2010 Jan 25;50(1):170–85.
7. Das S, Krein MP, Breneman CM. Binding Affinity Prediction with Property-Encoded Shape Distribution Signatures. *Journal of Chemical Information and Modeling*. 2010 Feb 22;50(2):298–308.
8. Ding B, Wang J, Li N, Wang W. Characterization of Small Molecule Binding. I. Accurate Identification of Strong Inhibitors in Virtual Screening. *J. Chem. Inf. Model*. [Internet]. 2012 Dec 21 [cited 2013 Jan 10]; Available from: <http://dx.doi.org/10.1021/ci300508m>

9. Li L, Li J, Khanna M, Jo I, Baird JP, Meroueh SO. Docking to Erlotinib Off-Targets Leads to Inhibitors of Lung Cancer Cell Proliferation with Suitable in Vitro Pharmacokinetics. *ACS Med. Chem. Lett.* 2010;1(5):229–33.
10. Deng W, Breneman C, Embrechts MJ. Predicting Protein–Ligand Binding Affinities Using Novel Geometrical Descriptors and Machine-Learning Methods. *Journal of Chemical Information and Computer Sciences.* 2004 Mar 1;44(2):699–703.
11. Sottriffer CA, Sanschagrin P, Matter H, Klebe G. SFCscore: scoring functions for affinity prediction of protein-ligand complexes. *Proteins.* 2008 Nov 1;73(2):395–419.
12. Cheng T, Li Q, Zhou Z, Wang Y, Bryant SH. Structure-Based Virtual Screening for Drug Discovery: a Problem-Centric Review. *The AAPS Journal.* 2012 Jan 27 [Review] <http://www.springerlink.com/content/577448371305305q/>
13. Li L, Wang B, Meroueh SO. Support Vector Regression Scoring of Receptor–Ligand Complexes for Rank-Ordering and Virtual Screening of Chemical Libraries. *Journal of Chemical Information and Modeling* 2011 Jul 14; <http://dx.doi.org/10.1021/ci200078f>
14. Li L, Khanna M, Jo I, Wang F, Ashpole NM, Hudmon A, et al. Target-Specific Support Vector Machine Scoring in Structure-Based Virtual Screening: Computational Validation, In Vitro Testing in Kinases, and Effects on Lung Cancer Cell Proliferation. *J. Chem. Inf. Model.* 2011;51(4):755–9.
15. Ballester, P.J., Schreyer, A., Blundell, T.L.: 'Fooled by uncertainty: the exploitation of structural and interaction data to predict intermolecular binding affinity'. (In Review)
16. Artemenko N. Distance Dependent Scoring Function for Describing Protein–Ligand Intermolecular Interactions. *Journal of Chemical Information and Modeling.* 2008 Mar 1;48(3):569–74.
17. Zhang S, Golbraikh A, Tropsha A. Development of Quantitative Structure–Binding Affinity Relationship Models Based on Novel Geometrical Chemical Descriptors of the Protein–Ligand Interfaces. *Journal of Medicinal Chemistry.* 2006 May 1;49(9):2713–24.
18. Sottriffer C. Scoring Functions for Protein–Ligand Interactions. In: Gohlke H, editor. *Protein–Ligand Interactions* [Internet]. Wiley-VCH Verlag GmbH & Co. KGaA; 2012 p. 237–63. [Review] <http://onlinelibrary.wiley.com/doi/10.1002/9783527645947.ch12/summary>
19. Ma D-L, Chan DS-H, Leung C-H. Drug repositioning by structure-based virtual screening. *Chem. Soc. Rev.* 2013 Jan 4 [Review] <http://pubs.rsc.org/en/content/articlelanding/2013/cs/c2cs35357a>
20. Cheng T, Li X, Li Y, Liu Z, Wang R. Comparative Assessment of Scoring Functions on a Diverse Test Set. *Journal of Chemical Information and Modeling.* 2009 Apr 27;49(4):1079–93.
21. Amini A, Shrimpton PJ, Muggleton SH, Sternberg MJE. A general approach for developing system-specific functions to score protein-ligand docked complexes using support vector inductive logic programming. *Proteins.* 2007 Oct;69(4):823–31.
22. Lahti JL, Tang GW, Capriotti E, Liu T, Altman RB. Bioinformatics and Variability in Drug Response: A Protein Structural Perspective. *J. R. Soc. Interface* 2012 May 2 ; [Review] <http://rsif.royalsocietypublishing.org/content/early/2012/05/01/rsif.2011.0843>
23. Kramer C, Gedeck P. Leave-Cluster-Out Cross-Validation Is Appropriate for Scoring Functions Derived from Diverse Protein Data Sets. *Journal of Chemical Information and Modeling.* 2010 Nov 22;50(11):1961–9.

24. Baum B, Muley L, Smolinski M, Heine A, Hangauer D, Klebe G. Non-additivity of functional group contributions in protein-ligand binding: a comprehensive study by crystallography and isothermal titration calorimetry. *J. Mol. Biol.* 2010 Apr 9;397(4):1042–54.
25. Ballester, P.J. (2012): 'Machine Learning Scoring Functions based on Random Forest and Support Vector Regression'. *Lecture Notes in Bioinformatics* 7632, Springer, 14-25.
26. Hongjian Li H, Ballester PJ, Leung K-S, Wong Man-Hon: 'istar: A Web Platform for Large-Scale Online Protein-Ligand Docking' (In preparation) <http://istar.cse.cuhk.edu.hk/idock/>
27. Ouyang X, Handoko SD, Kwok CK. CSCORE: A SIMPLE YET EFFECTIVE SCORING FUNCTION FOR PROTEIN–LIGAND BINDING AFFINITY PREDICTION USING MODIFIED CMAC LEARNING ARCHITECTURE. *Journal of Bioinformatics and Computational Biology.* 2011;09(supp01):1.
28. Ashtawy HM, Mahapatra NR. (2012): A Comparative Assessment of Ranking Accuracies of Conventional and Machine-Learning-Based Scoring Functions for Protein-Ligand Binding Affinity Prediction. *IEEE/ACM Transactions on Computational Biology and Bioinformatics.* 5555;99(1).
29. Hecht D, Fogel GB. Computational Intelligence Methods for Docking Scores. *Current Computer - Aided Drug Design.* 2009;5(1):56–68. [Review]
30. Huang S-Y, Grinter SZ, Zou X. Scoring functions and their evaluation methods for protein-ligand docking: recent advances and future directions. *Phys Chem Chem Phys.* 2010 Oct 28;12(40):12899–908. [Review]
31. Ballester, P.J., Mangold, M., Howard, N.I., Marchese-Robinson, R.L., Abell, C., Blumberger, J. and Mitchell, J.B.O. (2012): 'Hierarchical virtual screening for the discovery of new molecular scaffolds in antibacterial hit identification'. *Journal of the Royal Society Interface* 9: 3196-3207
32. Li G-B, Yang L-L, Wang W-J, Li L-L, Yang S-Y. ID-Score: A New Empirical Scoring Function Based on A Comprehensive Set of Descriptors Related to Protein-Ligand Interactions. *J. Chem. Inf. Model.* 2013 Feb 9 <http://dx.doi.org/10.1021/ci300493w>
33. Sugiyama M, Imajo K, Otaki K, Yamamoto A. Semi-Supervised Ligand Finding Using Formal Concept Analysis. *Information and Media Technologies.* 2012;7(3):928–37.