

Curriculum Vitae

Lixia Yao

1400 E 55th place, Apt 415S
Chicago, IL 60637
(646) 300-5036
ly2142@columbia.edu
www.dbmi.columbia.edu/~liy7001

Education

- Ph.D** **Columbia University**, New York, NY, USA
Biomedical Informatics, May 2010 (expected)
Dissertation: **Towards Computational Drug Discovery** Advisor: **Andrey Rzhetsky**
- M.S.** **National University of Singapore**, Singapore
Computational Science, Jul 2004
Thesis: **Inhibitor Prediction by Machine Learning Approaches** Advisor: **Yuzong Chen**
- B.Eng** **Dalian University of Technology**, Dalian, P.R. China
Chemical Engineering, Jul 2002

Publications

- (1) A probabilistic semantic similarity measure for medical concepts. **Yao L**, and Rzhetsky A. manuscript in preparation.
- (2) Benchmarking ontologies: bigger or better? **Yao L**, Divoli A, Mayzus I, Evans JA, and Rzhetsky A. manuscript under review.
- (3) Novel opportunities for computational biology and sociology in drug discovery. **Yao L**, Evans JA, and Rzhetsky A. *Trends in Biotechnology*, 27(9): 531-540. (2009)
- (4) Quantitative systems-level determinants of human genes targeted by successful drugs. **Yao L**, and Rzhetsky A. *Genome Research*, 18(2): 206-213. (2008)
- (5) Does drug-target have a likeness? Chen X, Fang Y, **Yao L**, Chen Y, Xu H. *Methods of Information in Medicine*, 46(3):360-6. (2007)
- (6) Internet resources for proteins associated with drug therapeutic effects, adverse reactions, and ADME. Ji Z, Sun L, Chen X, Zheng C, Yao L, Han L, Cao Z, Wang J, Yeo WK, Cai C, and Chen Y. *Drug Discovery Today*, 8(12): 526-529. (2003).
- (7) KDBI: kinetic data of bio-molecular interactions database. Ji Z, Chen X, Zheng C, **Yao L**, Han L, Yeo WK, Chung P, Puy H, Tay YT, Muhammad A, and Chen Y. *Nucleic. Acids. Research*, 31(1): 255-257. (2003).

Presentations

- (1) "Computational opportunities in drug discovery and development" Invited talk at the IBM T. J. Watson Research Center, Yorktown Heights, NY, Apr 2010.
- (2) "Towards mining drug-phenotype data: context-based calculation of probabilistic semantic similarity" Oral presentation at Pacific Symposium on Biocomputing (PSB), Big Island, Hawaii, Jan 2010.
- (3) "Computational drug discovery: analysis of drug targets and others" Poster presentation at NIH,

Bethesda, Nov 2009

- (4) “Quantitative systems-level determinants of drug targets” Poster presentation at International Conference on Intelligent Systems for Molecular Biology (ISMB), Toronto, Canada, Jul 2008.

Honors and Awards

- Travel award, Pacific Symposium on Biocomputing (PSB) Jan 2010
- National Graduate Student Research Festival, NIH Bethesda Nov 2009
One of three awarded nation-wide from 659 applications
- Annual student research award, Nucleic Acid Research and Oxford Journals Dec 2008
- Best poster award on the 4th International Society for Computational Biology (ISCB) Student Council Symposium Jul 2008
- Travel fellowship award, the 16th International Conference on Intelligent Systems for Molecular Biology (ISMB) Jul 2008
- Travel award for workshop on Search and Knowledge Building for Biological Databases, Institute for Pure and Applied Mathematics (IPAM) at UCLA Nov 2007
- First-year Fellowship, Graduate School of Arts and Sciences, Columbia University 2005-2006

Other Professional Experience

05/07-08/07 GlaxoSmithKline, Collegeville, PA

As a summer intern, I worked on a translational medicine project which is about identifying and validating biomarkers from gene expression data for different cancer types. I integrated the microarray data from in-house cancer cell line samples with and without certain drug treatments and then used nested ANOVA, t-test and certain ranking method to find a list of genes with differentiated expression pattern. In the end I evaluated the results on an external dataset from patients in clinics.

01/07-05/07 Columbia University, New York, NY

Teaching assistant of Computational Biology and Bioinformatics II

Teaching assistant of Biological Sequence Analysis

05/06-04/06 Columbia University, New York, NY

Peer advisor (volunteer), International Students and Scholars Office

08/04-05/05 Rensselaer Polytechnic Institute, Troy, NY

Teaching assistant of Organic Chemistry Laboratory I & II

01/03-12/03 National University of Singapore, Singapore

Mentor of Undergraduate Research Opportunities Programme for 2nd year college students

Mentor of Science Research Program for senior high school students

Teaching assistant of CZ1102 Problem Solving and Computation in C

Technical Skills

- Proficient at Python, Matlab, Perl, MySQL, UNIX/Linux OS and HTML & CSS
- Experienced in extra large dataset processing, cluster computing, C, and Fortran
- Familiar with molecular modeling/bioinformatics techniques (e.g., QSAR, docking, sequence alignment) homology modeling, major biomedical ontologies and databases (e.g., UMLS, SNOMED, dbSNP, Gene Ontology, DrugBank, GEO), natural language processing tools (e.g., MetaMap, Stanford pos tagger), and network analyzing and visualizing tools (e.g., Cytoscape)