KA YEE YEUNG, PH.D.

206-732-6156 kayee@u.washington.edu

SCIENTIFIC RESEARCH EXPERIENCE

- Extensive independent research experience in computational biology.
- Experience as the Principal Investigator on NIH-funded research grants (R01, administrative supplements and K25).
- Proficient in the preparation, application and management of research grants.
- Diverse research expertise spanning multiple disciplines, including computational biology, computer science, statistics, and biology.
- Collaboration with biomedical and computational scientists to identify research problems, formulate the scientific framework and engineer practical solutions.
- Experience leading research projects, mentoring graduate students, postdoctoral fellows and working with software developers.
- Open-minded approach to problem solving.

RESEARCH INTERESTS

- Development of algorithms and tools for the inference of predictive network models and gene signatures by integrating heterogeneous high-throughput data sources.
- Design of systems biology approaches and software tools to identify biologically meaningful biomarkers from high-throughput data.
- Customization and validation of data mining and pattern recognition methods applied to high-throughput biological data.
- Collaborative research work spanning both computational and biomedical disciplines.

EDUCATION

1998-2001	Ph.D. Computer Science, University of Washington, Seattle, WA Advisor: Professor Walter L. Ruzzo.
1996-1998	M.S. Computer Science, University of Washington, Seattle, WA Advisor: Professor Richard Karp.
1995-1996	M. Mathematics in Computer Science, University of Waterloo, Ontario, Canada Advisor: Professor Ian Munro.
1992-1995	B. Mathematics in Computer Science and Actuarial Science from University of Waterloo, Ontario, Canada

RESEARCH POSITIONS

6/04-present	**Research Assistant Professor**, Department of Microbiology, University of Washington, Seattle, WA
1/02-5/04	**Research Scientist/Senior Fellow**, Department of Microbiology, University of Washington, Seattle, WA Mentor: Associate Professor Roger Bumgarner

PUBLICATIONS

Original Journal Publications

1. Raftery A.E., Niu X., Hoff P. and **Yeung K.Y.** Fast inference for the latent space network model using a case-control approximate likelihood. *Journal of Computational and Graphical Statistics*, in press. On-line version http://amstat.tandfonline.com/doi/abs/10.1080/10618600.2012.679240

2. Lo K., Raftery A.E., Dombek K.M., Zhu J., Schadt E.E., Bumgarner R.E. and **Yeung K.Y**. Integrating external biological knowledge in the construction of regulatory networks from time-series expression data. *BMC Systems Biology* 2012, 6:101.
3. **Yeung K.Y.**, Gooley T.A., Zhang A., Raftery A.E., Radich J.P., and Oehler V.G. Predicting relapse prior to transplantation in chronic myeloid leukemia by integrating expert knowledge and expression data. *Bioinformatics* 2012, 28(6): 823-830.
4. **Yeung K.Y.**, Dombek K.M., Lo K., Mittler J.E., Zhu J., Schadt E.E., Bumgarner R.E. and Raftery A.E. Construction of regulatory networks using time series microarray data in genotyped yeast segregants. PNAS 2011, 108(48): 19436 - 41.
5. Zarbl H., Gallo M.A., Glick J., **Yeung K.Y.** and Vouros P. The vanishing zero revisited: Thresholds in the age of genomics. *Chemico-Biological Interactions 2010*, 184(1-2): 273-8.
6. Oehler VG*, **Yeung KY***, Choi E, Bumgarner RE, Raftery AE and Radich. "The derivation of diagnostic markers of chronic myeloid leukemia progression from microarray data". *Blood 2009*, 114: 3292-3298. *Co-first authors.
7. Annest A., Bumgarner R.E., Raftery A.E., **Yeung K.Y**. Iterative Bayesian Model Averaging: a method for the application of survival analysis to high-dimensional microarray data. *BMC Bioinformatics* 2009, 10: 72.
8. Chu VT, Gottardo R, Raftery AE, Bumgarner RE and **Yeung KY**. MeV+R: Using MEV as a GUI for Bioconductor applications. *Genome Biology* 2008, 9:R118.
9. Liu X, Sivaganesan S, **Yeung KY**, Guo J, Bumgarner RE and Medvedovic M. Bayesian Context-specific infinite mixture model for clustering of gene expression profiles accross diverse microarray datasets. *Bioinformatics* 2006, 22: 1737-1744
10. Gottardo R, Raftery AE, **Yeung KY** and Bumgarner RE. Bayesian robust inference for differential gene expression in cDNA microarrays with multiple samples. *Biometrics* 2006*,* 62: 10-18.
11. Gottardo R, Raftery AE, **Yeung KY** and Bumgarner RE. Robust estimation of cDNA microarray intensities with replicates. *Journal of the American Statistical Association* 2006 101: 30-40.
12. Li Q, Fraley C, Bumgarner RE, **Yeung KY** and Raftery AE. Donuts, scratches and blanks: Robust model-based segmentation of microarray images. *Bioinformatics* 2005 21: 2875 - 2882.
13. **Yeung KY**, Bumgarner RE and Raftery AE. Bayesian model averaging: development of an improved multi-class, gene selection and classification tool for microarray data. *Bioinformatics* 2005 21: 2394-2402.
14. Vanasse GJ, Winn RK, Rodov S, Zieske AW, Li JT, Tupper JC, Peters MA, **Yeung KY**, and Harlan JM. Bcl-2 overexpression leads to increases in suppressor of cytokine signaling-3 expression in B cells and de novo follicular lymphoma. *Molecular Cancer Research* 2004, 2: 620-631.
15. **Yeung KY**, Medvedovic M, Bumgarner RE. From co-expression to co-regulation: how many microarray experiments do we need? *Genome Biology* 2004, 5:R48.
16. Medvedovic M, **Yeung KY**, Bumgarner RE. Bayesian mixture model based clustering of replicated microarray data. *Bioinformatics* 2004, 20:1222-1232.
17. **Yeung KY**, Bumgarner RE. Multi-class classification of microarray data with repeated measurements: application to cancer. *Genome Biology* 2003, 4:R83.
18. **Yeung KY**, Medvedovic M, Bumgarner RE. Clustering Gene-Expression Data with Repeated Measurements. *Genome Biology* 2003, 4:R34.
19. Barrett MT, **Yeung KY**, Ruzzo WL, Hsu L, Blount PL, Sullivan R, Zarbl H, Delrow J, Rabinovitch PS, Reid BJ. Transcriptional Analyses of Barrett's Metaplasia and Normal Upper GI Mucosae. *Neoplasia* 2002, 4(2):121-128.

20. **Yeung KY**, Fraley C, Murua A, Raftery AE, Ruzzo WL. Model-based Clustering and Data Transformations for Gene Expression Data. *Bioinformatics* 2001, 17:977-987.
21. **Yeung KY**, Ruzzo WL. Principal Component Analysis for Clustering Gene Expression Data. *Bioinformatics* 2001, 17:763-774.
22. **Yeung KY**, Haynor DR, Ruzzo WL. Validating Clustering for Gene Expression Data. *Bioinformatics* 2001, 17:309-318.

Book Chapters and Review Articles

23. **Yeung K.Y**. Discovery of expression signatures in chronic myeloid leukemia by Bayesian Model Averaging. Book chapter in *Statistical Diagnostics of Cancer: Genetics and Genomics*, in press.
24. **Yeung K.Y**. Bayesian model averaging for biomarker discovery from genome-wide microarray data. Chapter 2 in *A Practical Guide to Bioinformatics Analysis*, 2010.
25. Bumgarner RE and **Yeung KY**. "Methods for the inference of biological pathways and neworks" in *Computational Systems Biology*. Methods in Molecular Biology. 2009; 541:225-45.
26. **Yeung KY**, Bumgarner RE. Pattern recognition in expression data. *Recent Developments in Nucleic Acids Research* 2004, 1: 333-354.
27. **Yeung KY**. "Clustering or automatic class discovery: non-hierarchical, non-SOM" in *A practical approach to microarray data analysis*. Kluwer Academic Publisher 2003, Chapter 16.

Conference Proceeding

28. Karp RM, Stoughton R, **Yeung KY.** Algorithms for Choosing Informative Differential Gene Expression Experiments. *Proceedings of the 3rd annual international conference on computational biology (RECOMB)* 1999, pp. 208-217.

Meeting Abstracts

29. **Yeung K.Y.**, Dombek K.M., Lo K., Mittler J.E., Zhu J., Schadt E.E., Bumgarner R.E. and Raftery A.E. Construction of regulatory networks using time series microarray data in genotyped yeast segregants. *20th Annual International Conference on Intelligent Systems and Molecular Biology (ISMB Highlight Track*), July 2012.
30. Oehler V.G., **Yeung K.Y.**, Zhang A., Gooley T.A., and Radich J. P. Differential Gene Expression Associated with Chronic Myeloid Leukemia (CML) Progression Predicts Relapse and Survival Prior to Allogeneic Transplantation In Chronic Phase CML Patients. *Blood (ASH Annual Meeting Abstracts)*, Nov 2010; 116: 3507.
31. **Yeung K.Y.**, Oehler V.G., Choi E., Bumgarner R.E., Raftery A.E. and Radich J.P. Derivation of Diagnostic Gene Predictors for the Progression of Chronic Myeloid Leukemia from Microarray Data and Independent PCR Validation. *Blood (ASH Annual Meeting Abstracts*), Nov 2008; 112: 3211.

PROFESSIONAL SKILLS & EXPERIENCE

Scientific Writing and Presentation Experience

- Proven track record of scientific manuscript publications.
- Experience with research grant preparation and application.
- Extensive experience with scientific presentations. Presented at international scientific meetings and conferences.
- Excellent reviews from guest lectures, invited lectures and workshops.
- Excellent communication and presentation skills.

Technical Skills

- Programming experience in C, C++, Java, Perl, R, Splus, Matlab and Maple.

- Scientific writing using Latex, Bibtex, Microsoft Word and EndNote.
- Experience using both commercial and academic data analysis tools, including Bioconductor packages, R, Splus, Matlab, Maple, Rosetta Resolver, SpotFire, and GeneTraffic.
- Experience contributing to publicly available open-source software (including Bioconductor packages).
- Conversant with various operating systems, including Linux, Unix, Mac OS X, Windows; and standard software applications

Other Professional Experience

- Reviewer for various scientific journals, including Bioinformatics, BMC Bioinformatics, Genome Biology, Nucleic Acids Research, Applied Genomics and Proteomics, Biometrics, Biostatistics, the Computer Journal, Pattern Recognition Letters, Statistical Applications in Genetics and Molecular Biology, Journal of Bioinformatics and Computational Biology.
- Adhoc reviewer for NIH and the University of Washington Royalty Research Fund.
- Invited lectures at a bioinformatics summer school in 2004 (the Advanced School in Biomedicine and Bioinformatics, Lipari, Italy). http://lipari.cs.unict.it/bio-info/index.htm
- Experience with developing and managing research budget.
- Talk and poster presentations at scientific meetings.

SPECIAL AWARDS & HONORS

2004 **Fast moving front in Computer Science**, Institute of Scientific Information (ISI) Essential Science Indicators. My first-authored publication titled "*Model-based clustering and data transformations for gene expression data"* (Bioinformatics 2001) was selected to be one of the most highly cited recent papers in Computer Science (Jan 2004). http://www.esi-topics.com/fmf/2004/january04-KaYeeYeung.html

2002 **Fast breaking paper in Computer Science**, Institute of Scientific Information (ISI) Essential Science Indicators. My first-authored publication titled "*Validating clustering for gene expression data"* (Bioinformatics 2001) was selected to be one of the most highly cited recent papers in Computer Science (Dec 2002). http://www.esi-topics.com/fbp/comments/december02-WalterLRuzzo.html

1996 **Alan George Student Leadership Award**, University of Waterloo, Ontario, Canada

1995 **NSERC Scholarship**, National Sciences and Engineering Research Council, Canada. The Natural Sciences and Engineering Research Council of Canada (NSERC) provides financial support to high-calibre scholars who are engaged in postgraduate programs in the natural sciences or engineering.

1995 **Provost Scholarship**, University of Waterloo, Ontario, Canada

1995 **ITRC Scholarship**, University of Waterloo, Ontario, Canada

1995 Honorable Mention, **Outstanding Female Undergraduate Student Competition**, Computing Research Association (CRA). CRA's Outstanding Undergraduate Award program recognizes undergraduate students who show outstanding research potential in an area of computing research. http://www.cra.org/Activities/awards/undergrad/