

## Jean Yee Hwa Yang

School of Mathematics and Statistics,  
Carslaw Building F07,  
University of Sydney  
NSW 2006, Australia

Phone: (61 2) 9351-3012

Mobile: (61 4) 31-260-701

Email : [jean.yang@sydney.edu.au](mailto:jean.yang@sydney.edu.au)

Homepage: <http://www.maths.usyd.edu.au/u/jeany>

### RESEARCH INTERESTS

My research concerns the application of statistics to problems in biological and medical research. In particular, my research work has centered on the development of statistical methodology and software for the design and analysis of high-throughput biotechnological data such as next generation sequencing, mass spectrometry and microarrays. In addition, she is also developing methods for integrating these omics data with other biological metadata such as gene annotation, sequence information and clinical data.

### PROFESSIONAL EXPERIENCE

**Associate Professor**, January 2013 to present.

School of Mathematics and Statistics, University of Sydney, Australia.

**Senior Lecturer**, January 2008 to December 2012.

School of Mathematics and Statistics, University of Sydney, Australia.

**Lecturer**, July 2005 to December 2009.

School of Mathematics and Statistics, University of Sydney, Australia.

**Assistant Professor**, Dec 2003 to Apr 2005.

Lung Biology Center, Department of Medicine and Center for Bioinformatics and Molecular Biostatistics, University of California, San Francisco.

**Postdoctoral Scientist**, Aug 2002 to Nov 2003.

Center for Bioinformatics and Molecular Biostatistics, Division of Biostatistics, University of California, San Francisco.

Research with Professor Mark Segal on various statistical issues relating to gene expression data. In addition provide statistical consulting for the Functional Genomics Core Group in the Sandler Center for Basic Research in Asthma.

**Assistant Statistician**, Jan 1997- Jul 1997.

Image analysis group, Mathematical and Information Sciences, Commonwealth Scientific and Industrial Research Organisation, Australia.

**EDUCATION**

- **Ph.D. in Statistics**, University of California, Berkeley, 2002.  
Dissertation: *Statistical methods in the design and analysis of gene expression data from cDNA microarray experiment*.  
Committee: Terence P. Speed, Chair; David A. Freedman; John Ngai.
- **M.A. in Statistics**, University of California, Berkeley, 1999.
- **B.Sc in Mathematics and Statistics**, with First Class Honours and Medal, University of Sydney, Australia, 1996. Honours project: *Bootstrap confidence intervals* with Prof. John Robinson.

**GRANTS**

1. ARC Discovery grant (DP130100488, 2013-2015)  
*Vertically integrated statistical modelling in multi-layered omics studies*.  
Investigators: **Yang, Y.H.**, Mueller, S., Mann, G.
2. ARC Discovery grant (DP120102849, 2012-2014)  
*Heads or tails which did echinoderms lose in the evolution of radial symmetry?*  
Investigators: Byrne, M. **Yang, Y.H.**, Wray, G.
3. ARC Future Fellow (FT0991918, 2010-2013)  
*New statistical methods for identifying microRNA regulatory network*.  
Investigator: **Yang, Y.H.**
4. ARC Linkage grant (LP0990067, 2009-2011)  
*Functional Genomics to Predict and Enhance Response to Interferon*.  
Investigators: Prof J George, **Yang, Y.H.**, McKay, F.C., Suppiah, V., Booth, D.R., Stewart, G. and Bahlo, M.
5. ARC discovery grant (DP0984267, 2009-2011)  
*Statistical and computational methods using a multiscale approach for protein identification and quantification*  
Investigators: **Yang, Y.H.**, Arthur, J. and Sewart, M.
6. ARC discovery grant (DP0770395, 2007-2009)  
*Statistical methods and tools for integrative microarray analysis*.  
Investigator: **Yang, Y.H.**
7. University of Sydney R&D Grants (2007)  
*Statistical methods for identifying regulatory microRNA from gene expression data*.  
Investigator: **Yang, Y.H.**
8. Univeristy of Sydney Sesqui Teaching Equipment Fund (2006)  
*Advanced Bioinformatics Teaching Facility*.  
Investigators: Arthur, J., Triccas, J., Jermiin, L., Charleston, M., **Yang, J.** and Britton, W.

---

**PUBLICATIONS**

1. Mann GJ, Pupo GM, Campain AE, Carter CD, Schramm SJ, Pianova S, Gerega SK, De Silva C, Lai K, Wilmott JS, Synnott M, Hersey P, Kefford RF, Thompson JF, **Yang YH**, Scolyer RA. BRAF mutation, NRAS mutation, and the absence of an immune-related expressed gene profile predict poor outcome in patients with stage III melanoma. *J Invest Dermatol*. 2013 Feb;133(2):509-17. PubMed PMID: 22931913.
2. Yang P, Ma J, Wang P, Zhu Y, Zhou BB, **Yang YH**. Improving X!Tandem on peptide identification from mass spectrometry by self-boosted Percolator. *IEEE/ACM Trans Comput Biol Bioinform*. 2012 Sep-Oct;9(5):1273-80. PubMed PMID: 22689082.
3. Schramm SJ, Campain AE, Scolyer RA, **Yang YH**, Mann GJ. Review and cross-validation of gene expression signatures and melanoma prognosis. *J Invest Dermatol*. 2012 Feb;132(2):274-83. PubMed PMID: 21956122.
4. Jayaswal V, Lutherborrow M, **Yang YH**. Measures of association for identifying microRNA-mRNA pairs of biological interest. *PLoS One*. 2012;7(1):e29612. PubMed PMID: 22253745.
5. Yang P, Ho JW, **Yang YH**, Zhou BB. Gene-gene interaction filtering with ensemble of filters. *BMC Bioinformatics*. 2011;12 Suppl 1:S10. PubMed PMID: 21342539.
6. Wu JQ, Dwyer DE, Dyer WB, **Yang YH**, Wang B, Saksena NK. Genome-wide analysis of primary CD4+ and CD8+ T cell transcriptomes shows evidence for a network of enriched pathways associated with HIV disease. *Retrovirology*. 2011;8:18. PubMed PMID: 21410942.
7. Schramm SJ, Campain AE, Scolyer RA, **Yang YH**, Mann GJ. Review and cross-validation of gene expression signatures and melanoma prognosis. *J Invest Dermatol*. 2011 Feb;132(2):274-83. PubMed PMID: 21956122.
8. Peters T, Bulger DW, Loi TH, **Yang JY**, Ma D. Two-step cross-entropy feature selection for microarrays-power through complementarity. *IEEE/ACM Trans Comput Biol Bioinform*. 2011 Jul-Aug;8(4):1148-51. PubMed PMID: 21321369.
9. Marques FZ, Campain AE, Tomaszewski M, Zukowska-Szzechowska E, **Yang YH**, Charchar FJ, Morris BJ. Gene expression profiling reveals renin mRNA overexpression in human hypertensive kidneys and a role for microRNAs. *Hypertension*. 2011 Dec;58(6):1093-8. PubMed PMID: 22042811.

10. Marques FZ, Campain AE, Davern PJ, **Yang YH**, Head GA, Morris BJ. Global identification of the genes and pathways differentially expressed in hypothalamus in early and established neurogenic hypertension. *Physiol Genomics*. 2011 Jun 28;43(12):766-71. PubMed PMID: 21487032.
11. Marques FZ, Campain AE, Davern PJ, **Yang YH**, Head GA, Morris BJ. Genes influencing circadian differences in blood pressure in hypertensive mice. *PLoS One*. 2011;6(4):e19203. PubMed PMID: 21541337.
12. Lutherborrow M, Bryant A, Jayaswal V, Agapiou D, Palma C, **Yang YH**, Ma DD. Expression profiling of cytogenetically normal acute myeloid leukemia identifies microRNAs that target genes involved in monocytic differentiation. *Am J Hematol*. 2011 2011;86(1):2-11. PubMed PMID: 20981674.
13. ayaswal V, Lutherborrow M, Ma DD, **Yang YH**. Identification of microRNA-mRNA Modules using Microarray Data. *BMC Genomics*. 2011 Mar 6;12(1):138. PubMed PMID: 21375780.
14. Wang P, Yang P, Arthur J, **Yang JY**. A dynamic wavelet-based algorithm for pre-processing tandem mass spectrometry data. *Bioinformatics*. 2010 Sep 15;26(18):2242-9. PubMed PMID: 20628072.
15. Marques FZ, Campain AE, **Yang YH**, Morris BJ. Meta-analysis of genome-wide gene expression differences in onset and maintenance phases of genetic hypertension. *Hypertension*. 2010 Aug;56(2):319-24. PubMed PMID: 20585107.
16. Lutherborrow M, Bryant A, Jayaswal V, Agapiou D, Palma C, **Yang YH**, Ma DD. Expression profiling of cytogenetically normal acute myeloid leukemia identifies microRNAs that target genes involved in monocytic differentiation. *Am J Hematol*. 2010 Jan;86(1):2-11. PubMed PMID: 20981674.
17. Karim RZ, Gerega SK, **Yang YH**, Spillane A, Carmalt H, Scolyer RA, Lee CS. p16 and pRb immunohistochemical expression increases with increasing tumour grade in mammary phyllodes tumours. *Histopathology*. 2010 Jun;56(7):868-75. PubMed PMID: 20497245.
18. Jensen SO, Apisiridej S, Kwong SM, **Yang YH**, Skurray RA, Firth N. Analysis of the prototypical *Staphylococcus aureus* multiresistance plasmid pSK1. *Plasmid*. 2010 Nov;64(3):135-42. PubMed PMID: 20547176.

19. Collinson EJ, Wimmer-Kleikamp S, Gerega SK, **Yang YH**, Parish CR, Dawes IW, Stocker R. The yeast homolog of heme oxygenase-1 affords cellular antioxidant protection via the transcriptional regulation of known antioxidant genes. *J Biol Chem*. 2010 Jan 21;286(3):2205-14. PubMed PMID: 21081499.
20. Campaign A, **Yang YH**. Comparison study of microarray meta-analysis methods. *BMC Bioinformatics*. 2010;11:408. PubMed PMID: 20678237.
21. Reijo Pera RA, DeJonge C, Bossert N, Yao M, **Hwa Yang JY**, Asadi NB, Wong W, Wong C, Firpo MT. Gene expression profiles of human inner cell mass cells and embryonic stem cells. *Differentiation*. 2009 Jul;78(1):18-23. PubMed PMID: 19398262.
22. Karim RZ, Li W, Sanki A, Colman MH, **Yang YH**, Thompson JF, Scolyer RA. Reduced p16 and increased cyclin D1 and pRb expression are correlated with progression in cutaneous melanocytic tumors. *Int J Surg Pathol*. 2009 Oct;17(5):361-7. PubMed PMID: 19666944.
23. Karim RZ, Gerega SK, **Yang YH**, Spillane A, Carmalt H, Scolyer RA, Lee CS. Phyllodes tumours of the breast: a clinicopathological analysis of 65 cases from a single institution. *Breast*. 2009 Jun;18(3):165-70. PubMed PMID: 19329316.
24. Karim RZ, Gerega SK, **Yang YH**, Horvath L, Spillane A, Carmalt H, Scolyer RA, Lee CS. Proteins from the Wnt pathway are involved in the pathogenesis and progression of mammary phyllodes tumours. *J Clin Pathol*. 2009 Nov;62(11):1016-20. PubMed PMID: 19861560.
25. Jayaswal V, Lutherborrow M, Ma DD, **Hwa Yang Y**. Identification of microRNAs with regulatory potential using a matched microRNA-mRNA time-course data. *Nucleic Acids Res*. 2009 May;37(8):e60. PubMed PMID: 19295134.
26. Wu JQ, Dwyer DE, Dyer WB, **Yang YH**, Wang B, Saksena NK. Transcriptional profiles in CD8+ T cells from HIV+ progressors on HAART are characterized by coordinated up-regulation of oxidative phosphorylation enzymes and interferon responses. *Virology*. 2008 Oct 10;380(1):124-35. PubMed PMID: 18692859.
27. Lewis CC, **Yang JY**, Huang X, Banerjee SK, Blackburn MR, Baluk P, McDonald DM, Blackwell TS, Nagabhushanam V, Peters W, Voehringer D, Erle DJ. Disease-specific gene expression profiling in multiple models of lung disease. *Am J Respir Crit Care Med*. 2008 Feb 15;177(4):376-87. PubMed PMID: 18029791.

28. Kechris K, **Yang YH**, Yeh RF. Prediction of alternatively skipped exons and splicing enhancers from exon junction arrays. *BMC Genomics*. 2008;9:551. PubMed PMID: 19021909.
29. Xiao Y, Segal MR, **Yang YH**, Yeh RF. A multi-array multi-SNP genotyping algorithm for Affymetrix SNP microarrays. *Bioinformatics*. 2007 Jun 15;23(12):1459-67. PubMed PMID: 17459966.
30. Woodruff PG, Boushey HA, Dolganov GM, Barker CS, **Yang YH**, Donnelly S, Ellwanger A, Sidhu SS, Dao-Pick TP, Pantoja C, Erle DJ, Yamamoto KR, Fahy JV. Genome-wide profiling identifies epithelial cell genes associated with asthma and with treatment response to corticosteroids. *Proc Natl Acad Sci U S A*. 2007 Oct 2;104(40):15858-63. PubMed PMID: 17898169.
31. Winn VD, Haimov-Kochman R, Paquet AC, **Yang YJ**, Madhusudhan MS, Gormley M, Feng KT, Bernlohr DA, McDonagh S, Pereira L, Sali A, Fisher SJ. Gene expression profiling of the human maternal-fetal interface reveals dramatic changes between midgestation and term. *Endocrinology*. 2007 Mar;148(3):1059-79. PubMed PMID: 17170095.
32. Williams EO, Xiao Y, Sickles HM, Shafer P, Yona G, **Yang JY**, Lin DM. Novel subdomains of the mouse olfactory bulb defined by molecular heterogeneity in the nascent external plexiform and glomerular layers. *BMC Dev Biol*. 2007;7:48. PubMed PMID: 17506891.
33. **Yang YH**, Xiao Y, Segal MR. Identifying differentially expressed genes from microarray experiments via statistic synthesis. *Bioinformatics*. 2005 Apr 1;21(7):1084-93. PubMed PMID: 15513985.
34. Xiao Y, **Yang YH**, Burckin TA, Shiue L, Hartzog GA, Segal MR. Analysis of a Splice Array Experiment Elucidates Roles of Chromatin Elongation Factor Spt4-5 in Splicing. *PLoS Comput Biol*. 2005 Sep 16;1(4):e39. PubMed PMID: 16172632.
35. Xiao Y, Segal MR, **Yang YH**. Stepwise normalization of two-channel spotted microarrays. *Statistical Applications in Genetics and Molecular Biology*. 2005 Feb, 7, 2005;4(1):Article 4.
36. Woodruff PG, Koth LL, **Yang YH**, Rodriguez MW, Favoreto S, Dolganov GM, Paquet AC, Erle DJ. A distinctive alveolar macrophage activation state induced by cigarette smoking. *Am J Respir Crit Care Med*. 2005 Dec 1;172(11):1383-92. PubMed PMID: 16166618.

37. Kuperman DA, Lewis CC, Woodruff PG, Rodriguez MW, **Yang YH**, Dolganov GM, Fahy JV, Erle DJ. Dissecting asthma using focused transgenic modeling and functional genomics. *J Allergy Clin Immunol*. 2005 Aug;116(2):305-11. PubMed PMID: 16083784.
38. Gonzalez R, **Yang YH**, Griffin C, Allen L, Tigue Z, Dobbs L. Freshly isolated rat alveolar type I cells, type II cells, and cultured type II cells have distinct molecular phenotypes. *Am J Physiol Lung Cell Mol Physiol*. 2005 Jan;288(1):L179-89. PubMed PMID: 15447939.
39. Barker CS, Griffin C, Dolganov GM, Hanspers K, **Yang JY**, Erle DJ. Increased DNA microarray hybridization specificity using sscDNA targets. *BMC Genomics*. 2005 Apr 22;6(1):57. PubMed PMID: 15847692.
40. Rodriguez MW, Paquet AC, **Yang YH**, Erle DJ. Differential gene expression by integrin beta 7+ and beta 7- memory T helper cells. *BMC Immunol*. 2004 Jul 5;5:13. PubMed PMID: 15236665.
41. Lin DM<sup>+</sup>, **Yang YH**<sup>+</sup>, Scolnick JA, Brunet LJ, Marsh H, Peng V, Okazaki Y, Hayashizaki Y, Speed TP, Ngai J. Spatial patterns of gene expression in the olfactory bulb. *Proc Natl Acad Sci U S A*. 2004 Aug 24;101(34):12718-23. PubMed PMID: 15304640.
42. Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S, Ellis B, Gautier L, Ge Y, Gentry J, Hornik K, Hothorn T, Huber W, Iacus S, Irizarry R, Leisch F, Li C, Maechler M, Rossini AJ, Sawitzki G, Smith C, Smyth G, Tierney L, **Yang JY**, Zhang J. Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol*. 2004;5(10):R80. PubMed PMID: 15461798.
43. **Yang YH**, Thorne N. Normalization for Two-color cDNA Microarray Data. *Science and Statistics*. In: Goldstein D, editor. *A Festschrift for Terry Speed, IMS Lecture Notes, Monograph Series*. 402003. p. 403–18.
44. Smyth GK, **Yang YH**, Speed T. Statistical issues in cDNA microarray data analysis. *Methods Mol Biol*. 2003;224:111-36. PubMed PMID: 12710670.
45. Erle DJ, **Yang YH**. Asthma investigators begin to reap the fruits of genomics. *Genome Biol*. 2003;4(11):232. PubMed PMID: 14611649.
46. Diaz E, **Yang YH**, Ferreira T, Loh KC, Okazaki Y, Hayashizaki Y, Tessier-Lavigne M, Speed TP, Ngai J. Analysis of gene expression in the developing mouse retina. *Proc Natl Acad Sci U S A*. 2003 Apr 29;100(9):5491-6. PubMed PMID: 12702772.

47. **Yang YH**, Speed T. Design issues for cDNA microarray experiments. *Nat Rev Genet.* 2002 Aug;3(8):579-88. PubMed PMID: 12154381.
48. **Yang YH**<sup>+</sup>, Dudoit S<sup>+</sup>, Luu P<sup>++</sup>, Lin DM, Peng V, Ngai J, Speed TP. Normalization for cDNA microarray data: a robust composite method addressing single and multiple slide systematic variation. *Nucleic Acids Res.* 2002 Feb 15;30(4):e15. PubMed PMID: 11842121.
49. **Yang YH**, Buckley MJ, Dudoit S, Speed T. Comparison of methods for image analysis on cDNA microarray data. *Journal of Computational and Graphical Statistics* 2002;11:108–36.
50. Speed T, **Yang YH**. Direct versus indirect designs for cDNA microarray experiments. *Sankhya Series A, Part 3.* 2002;64:706–20.
51. Dudoit S<sup>+</sup>, **Yang YH**<sup>++</sup>, Callow MJ, Speed T. Statistical methods for identifying differentially expressed genes in replicated cDNA microarray experiments. *Statistical Sinica.* 2002;12:111–39.
52. Diaz E, Ge Y, **Yang YH**, Loh KC, Serafini TA, Okazaki Y, Hayashizaki Y, Speed TP, Ngai J, Scheffele P. Molecular analysis of gene expression in the developing pontocerebellar projection system. *Neuron.* 2002 Oct 24;36(3):417-34. PubMed PMID: 12408845.
53. **Yang YH**, Buckley MJ, Speed TP. Analysis of cDNA microarray images. *Brief Bioinform.* 2001 Dec;2(4):341-9. PubMed PMID: 11808746.
54. Buckley MJ, **Yang JY**. Regularised shortest-path extraction. *Pattern Recognition Letters.* 1997;8:621–9.

+ indicate joint-first authorship of the paper.

## PRESENTATIONS, CONFERENCES AND WORKSHOPS

### Selected Invited Presentations

- *Integrating mRNA and miRNA expression data*, EMBL PhD course, Melbourne, July 2013.
- *Integration of high throughput biological data*, Meeting the challenges of high dimension, Institute for Mathematical Sciences, Singapore, Oct 2012.
- *An integrative analysis of miRNA-Seq and mRNA-Seq data*, WNAR-Graybill 2012 conference, Fort Collins, June 2012.



- 
- *Statistical analysis of quantitative proteomics data*, Annual Symposium, Science at the Shine Dome, Canberra, 7 May 2010.
  - *Statistical analysis of quantitative proteomics data*, Annual Symposium, Science at the Shine Dome, Canberra, 7 May 2010.
  - *Identification of candidate microRNA using matched mRNA-miRNA time course data*, Emerging Statistical Challenges in Genome and Translational Research, Banff International Research Station, June, 2008.
  - *Identification of candidate microRNA using matched mRNA-miRNA time course data*, Australian Statistical Conference, Melbourne, Sep, 2008.
  - *Analysis of human alternative splices predicted from exon junction arrays*, CMIS, CSIRO, Australia, 3 July, 2006.
  - *Case Studies in the Design and Analysis of Microarray Experiment.*, Computational Biology Group, Cambridge, UK, 12 Sep, 2004.
  - *Analyzing Data from Splice Array Experiment.*, Institute for Mathematical Sciences, National University of Singapore, 9 Jan, 2004.
  - *Statistical Issues in the Design and Analysis of Microarray Experiment.*, Workshop on Microarray Data Analysis, Academia Sinica, Taiwan, 26 Nov, 2003.
  - *Analyzing DNA Microarray Data using Bioconductor.*, Third Annual Ontario Microarray Network Technology Symposium, Toronto, Canada, 2–13 Nov, 2003. *Statistical Issues in the design of microarray experiments.*, Interface 2003: Security and Infrastructure Protection, Salt Lake City, Utah, 12–15, March, 2003.
  - *Identifying differentially expressed genes and clusters of genes in cDNA Microarray experiment*, International Congress on Differentiation, Molecular and Cell Biology, Gold Coast, Australia, 24–28 September 2000.

**Short courses and workshops** 12 co-designed and co-delivered, including: Joint Statistical Meetings, Minneapolis, Aug 2005 (2 days; partially funded); Royal Institute of Technology, Albanova University Centre, Stockholm, Oct 2004 (four days; fully funded); Institute for Mathematical Sciences, National University of Singapore, Jan 2004 (four days; fully funded); Health Canada, Ottawa, April 2003 (two days, fully funded); Fields Institute, Toronto, May 2002 (one day, partially funded). After our initial successes in 2001/02, these events became increasingly popular, attracting audiences from 20 to several hundred university and industry researchers in biomedical fields.

---

## TEACHING

**Formal Teaching** Range of undergraduate and postgraduate teaching including “Applied statistics for Bioinformatics”, “Bioinformatics”, “Applied Statistics” and statistics component for undergraduate pharmacy courses.

### Current Honors / Postgraduate Supervision

Dario Strbenac (PhD); Supervisor.

Kaushala Jayawardana (PhD); Supervisor.

Shila Ghazanfar (PhD); Supervisor.

Ellis Patrick (PhD); Supervisor.

Tram Doan (PhD); Associate supervisor.

Rebecca Barter (Honors); Supervisor.

Narelle Quinon (Honors); Supervisor.

### Past Postgraduate Supervision

Anna Campain (PhD, completed); Supervisor.

Pengyi Yang (PhD); Associate supervisor.

Emi Tanaka (PhD); Associate supervisor.

Selma Klanten (Master Research project); Supervisor.

Yu-ping Wang (Master Research project); Supervisor.

Xiaodan (Alice) Dong (Master Research project); Supervisor.

## AWARDS

- Evelyn Fix Memorial Medal, UC Berkeley, 2002.
- Student Presentation award - International Conference on Statistics, Combinatorics and Related Areas, 2001.
- Student Presentation award - Joint Conference of Australasian Biometrics and New Zealand Statistical Association, 2001.
- Statistical Society of Australia Honours Scholarship, 1996.
- Dean’s honour list, 1994-95.
- Sydney University Science Scholarship, 1993.

## REFEREE SERVICE

*Journals Associate Editor*, BMC Bioinformatics (since 2010); Biostatistics (2005-2009).

*Regular reviewer* of manuscripts, Journal of the American Statistical Association (since 02), Bioinformatics (since 02), Nucleic Acids Research (since 02), Biometric (since 05), JCGS (since 07)

## PROFESSIONAL MEMBERSHIPS

Statistical Society of Australia

Australasian Microarray Associated Technologies Association

## COMPUTING EXPERIENCE

– Member of the core team in the Bioconductor project, an open source and open development software project for the analysis of genomic and other biological data.

– Languages/Software: C, R/S-Plus, Matlab, Mathematica, L<sup>A</sup>T<sub>E</sub>X, HTML.