

## Pei Fen Kuan

---

CONTACT            Department of Statistics            Phone: (608)347-6226  
INFORMATION      1220 Medical Sciences Center        Email: kuanp@stat.wisc.edu  
                         1300 University Avenue, 1250D,      Web: <http://www.stat.wisc.edu/~kuanp>  
                         Madison, WI 53706  
                         USA

EDUCATION        **University of Wisconsin**, Madison, WI USA

Ph.D in Statistics (in progress). Expected graduation 2009. (GPA: 4.00/4.00)  
Dissertation title: Statistical methods for the analysis of genomic data from tiling arrays and next generation sequencing technologies.  
Preliminary examination committee: Sündüz Keleş, Michael Newton, Kam-Wah Tsui, Grace Wahba.

M.S. in Statistics, December 2006. (GPA: 4.00/4.00)

**National University of Singapore**, Singapore

B.S. in Science, May 2004 (GPA/CAP: 4.95/5.00)  
1<sup>st</sup> Major: Statistics    2<sup>nd</sup> Major: Applied Mathematics  
Thesis title: Genetic linkage analysis.

PUBLICATIONS    **Kuan, P.**, Chun, H., and Keleş, S. (2008). CMARRT: A tool for the analysis of ChIP-Chip data from tiling arrays by incorporating the correlation structure. *Pac Symposium of Biocomputing*, 515-526.

Wei, H.\*, **Kuan, P.\***, Tian, S., Yang, C., Nie, J., Sengupta, S., Ruotti, V., Jonsdottir, G.A., Keleş, S., Thomson, J.A., and Stewart, R. (2008). A study of the relationships between oligonucleotide properties and hybridization signal intensities from NimbleGen microarray datasets. *Nucleic Acids Research*, **36** (9): 2926-2938. \* *Joint first authors*

Zhang, X.D., **Kuan, P.**, Ferrer, M., Shu, X., Liu, Y.C., Gates, A.T., Kunapuli, P., Stec, E.M., Xu, M., Marine S.D., Holder, D.J., Strulovici, B., Heyse, J.F., and Espeseth, A.S. (2008). Hit selection with false discovery rate control in genome-scale RNAi screens. *Nucleic Acids Research*, **36** (14): 4667-4679.

Chung, N., Zhang, X.D., Kreamer, A., Locco, L., **Kuan, P.**, Bartz, S., Linsley, P.S., Ferrer, M., and Strulovici, B. (2008). Median absolute deviation to improve hit selection for genome-scale RNAi screens. *Journal of Biomolecular Screening*, **13** (2): 149-158.

**Kuan, P.**, Huebert, D., Gasch, A., and Keleş, S. (2008). A non-homogeneous hidden Markov model on first order differences for automatic detection of nucleosome positions. *Technical Report No.1150, University of Wisconsin-Madison*.

(submitted).

Peters, J., Mooney, R. A., **Kuan, P.**, Rowland, J., Keleş, S. and Landick, R. (2008). Transcription Termination Factor Rho Targets Stable RNAs. (submitted).

Meyer, M., Zella, L., Nerenz, R., **Kuan, P.**, Keleş, S. and Pike, J. (2008). The vitamin D analog gs1790 fails to induce hypercalcemia in vivo as a result of both gene-selective and receptor-specific actions in calcium-regulating tissues. (submitted).

**Kuan, P.**, Pan, G., Thomson, J., Stewart, R., and Keleş, S. (2009). A hierarchical semi-Markov model for detecting enrichment with application to ChIP-Seq experiments. *Technical Report No.1151, University of Wisconsin-Madison.* (submitted).

MANUSCRIPTS IN PREPARATION A robust permutation test for the analysis of tiling array data.

RESEARCH EXPERIENCE Predoctoral traineeship by Morgridge Institute for Research Support for Senior Graduate Students in Computation and Informatics in Biology and Medicine (Spring 2009)

Supervisor: Professor Sündüz Keleş, Department of Statistics and Department of Biostatistics and Medical Informatics, University of Wisconsin, Madison.

Research Assistant (Fall 2007-current)

Thesis advisor: Professor Sündüz Keleş, Department of Statistics and Department of Biostatistics and Medical Informatics, University of Wisconsin, Madison. Thesis research: Develop statistical methodologies for high-throughput genomic data generated from micro/tiling arrays and next generation sequencing technologies (ChIP-Chip and ChIP-Seq). Part of dissertation work is motivated by collaborations with the Thomson Lab (human embryonic stem cell data), Gasch Lab (yeast data), Landick Lab (e-coli data) and Pike Lab (mouse data) at University of Wisconsin, Madison.

Research Assistant (Spring 2005)

Supervisor: Professor Michael Newton, Department of Statistics and Department of Biostatistics and Medical Informatics, University of Wisconsin-Madison. Investigate and evaluate the performance of hierarchical mixture modeling against gene specific t-testing on a public data generated from HIV microarray experiment.

CONSULTING EXPERIENCE

Thomson Laboratory Bioinformatics Group (Fall 2006-current)

Principal Investigator: Professor James A. Thomson, Director of Regenerative Biology, Morgridge Institute for Research, Madison, WI.

Bioinformatics supervisor: Ron Stewart, Ph.D.

Provide statistical support in the analysis of ChIP-Chip and ChIP-Seq data.

TEACHING EXPERIENCE Teaching assistant (Fall 2004-Summer 2005)  
Department of Statistics, University of Wisconsin, Madison  
Organize and lead lecture discussion for two Statistics course: “Statistical Design of Experiments for Engineers” and “Introduction to Statistical Methods.”

WORK EXPERIENCE Summer Intern, 2007.  
Merck and Co. Inc, West Point, North Wales, PA.  
Supervisor: Xiaohua Douglas Zhang, Ph.D.  
Project: “A Bayesian inference in genome-scale high-throughput RNAi hit selection.” Develop a hierarchical modeling via a Bayesian framework for RNAi hit selection by incorporating priors derived from positive and negative control wells. Develop an R-package `BayesHTS` that implements the proposed methodology.

MEETINGS AND WORKSHOP Midwest Symposium of Computational Biology and Bioinformatics, Evanston, IL, 2007  
Poster “`CMARRT`: A tool for the analysis of ChIP-Chip data from tiling arrays by incorporating the correlation structure.”

Summer internship workshop at Merck and Co. Inc, 2007  
Oral presentation on “`BayesHTS`: A Bayesian approach with FDR control in genome-wide RNAi screens hit selection.”

Pacific Symposium of Biocomputing, Hawaii, 2008  
Oral presentation on “`CMARRT`: A tool for the analysis of ChIP-Chip data from tiling arrays by incorporating the correlation structure.”

WNAR Meetings, Davis, CA, 2008  
Invited talk on “Statistical Issues in the Analysis of ChIP Experiments Measuring Nucleosome Occupancy.”

Midwest Symposium of Computational Biology and Bioinformatics, Urbana-Champaign, IL, 2008  
Poster “A non-homogeneous hidden Markov model on first order differences for automatic detection of nucleosome positions.”

Genomic Science Training Program Annual Retreat, Madison, WI, 2008  
Poster “A non-homogeneous hidden Markov model on first order differences for automatic detection of nucleosome positions.”

Second Midwest Statistics Research Colloquium, Chicago, IL, 2009  
Invited talk on “A Hierarchical Semi-Markov Model for Detecting Enrichment with Application to ChIP-Seq Experiments.”

HONORS AND AWARDS Conferences  

- 2<sup>nd</sup> poster prize in Midwest Symposium of Computational Biology and Bioinformatics

formatics, Evanston, IL, 2007.

- Graduate Student Invited Talks for Second Midwest Statistics Research Colloquium, Chicago, IL, 2009 (2 graduate students selected out of > 30 submissions).

University of Wisconsin, Madison

- Shapiro Fellowship, Department of Biostatistics, Fall 2005.
- Best teaching assistant in Department of Statistics for 2005/2006.
- Teaching fellow nominee for College of Letter and Science, 2006.
- Morgridge Institute for Research predoctoral award for senior graduate student in Computation and Informatics in Biology and Medicine, Spring 2009.

National University of Singapore

- Dean's List for 7 out of 8 semesters.
- Valedictorian for university commencement main ceremony July 6, 2004.
- Overall best student for Faculty of Science.
- Winner of Lee Kuan Yew Gold Medal.
- Best student for Department of Statistics and Applied Probability.
- Winner of Saw Swee Hock Gold Medal.

MEMBERSHIP      American Statistical Association (ASA)  
International Biometric Society (IBS)

SKILLS            Language: fluent in written and spoken English and Malay, fluent in spoken Chinese (Mandarin, Taiwanese Minnan and Cantonese).

Computer:

- Software: Author of R package CMARRT (Correlation, Moving Average, Robust and Rapid method on Tiling array).
- Programming Language and Statistical Packages: Proficient in SAS, Splus/R, Minitab, SPSS.
- Some experience with Java, C, C++ and Matlab.
- Others: Proficient in Latex, Emacs, Unix/Linux, Windows, Excel, Word and PowerPoint.

REFERENCES      Available upon request.

UPDATED ON  
01/29/2009