

Curriculum Vitae

Nengjun Yi, Ph.D.

Professor

The Sir David Cox Endowed Professorship
The Section on Statistical Genetics
Department of Biostatistics
The University of Alabama at Birmingham
Birmingham, AL 35294-0022

Phone: (205) 934-4924

Fax: (205) 975-2540

Email: nyi@uab.edu

EDUCATION AND TRAINING

- Postdoctoral Research Associate, Statistical genetics/genomics, University of California at Riverside, July 1998 - December 2001. (With Prof. Shizhong Xu).
- Ph.D. July 1996. Quantitative and Population Genetics, Nanjing Forestry University, P. R. China.
- M.S. July 1988. Biostatistics, Zhejiang University, P. R. China.
- B.S. July 1983. Mathematics, Shaoyang Normal University, P. R. China.

CONTINUING EDUCATION

- Short course on Bayesian Modeling, Inference and Prediction. Presenter: David Draper, University of California at Santa Cruz. March 18th, 2005.
- Bayesian statistics and Markov chain Monte Carlo algorithms in genetics, Iowa State University, June 20-30, 2000.
- Advanced SAS course, Department of Statistics, Beijing University, P. R. China, July 20-August 2, 1995.
- Mathematics, Department of Mathematics, Hunan Teacher University, P. R. China, July 1985-July 1986.

CURRENT APPOINTMENTS

- Professor (with tenure), the Section on Statistical Genetics, Department of Biostatistics, University of Alabama at Birmingham.
- The Sir David Cox Endowed Professorship in Biostatistics in the Department of Biostatistics in the School of Public Health, University of Alabama at Birmingham. April 2011 – present.
- Senior Scientist of Nutrition Obesity Research Center (NORC), University of Alabama at Birmingham, June 2003 - present.
- Senior Scientist of the UAB Comprehensive Cancer Center, the Cancer Control and Population Science Program, University of Alabama at Birmingham, 2009 - present

PREVIOUS APPOINTMENTS

- Associate Professor (with tenure), the Section on Statistical Genetics, Department of Biostatistics, University of Alabama at Birmingham, August 2005 – September 2012.
- Assistant Professor of Biostatistics and Statistical Genetics, the Section on Statistical Genetics, Department of Biostatistics, University of Alabama at Birmingham, July 2004-August 2005.
- Research Assistant Professor of Biostatistics and Statistical Genetics, the Section on Statistical Genetics, Department of Biostatistics, University of Alabama at Birmingham, May 2002-July 2004.
- Assistant Professor of Biostatistics and Statistical Genetics, Creighton University, January 2002-May 2002.
- Postdoctoral Research Associate, Statistical genetics/genomics, University of California, Riverside, July 1998-December 2001. (With Prof. Shizhong Xu).
- Associate Professor of Statistics and Genetics, Nanjing Forestry University, P. R. China, April 1996-June 1998.
- Assistant Professor of Mathematics, Nanjing Forestry University, P. R. China, August 1988-March 1996.
- Teacher of Mathematics, Chengbu High School, Chengbu, Hunan Province, P. R. China, August 1983-July 1985.

HONORS AND AWARDS

- 2014 Best Paper Award by Science Unbound Foundation, USA.
- The Sir David Cox Endowed Professorship in Biostatistics in the UAB Department of Biostatistics in the School of Public Health, appointed by The University of Alabama System Board of Trustees in April 8, 2011.
- The 2011 Graduate Dean's Award for Excellence in Mentorship, University of Alabama at Birmingham
- On the basis of my substantial service to peer review at the National Institutes of Health, I have been given special dispensation from NIH to submit grants at any time without respect to deadline from OCTOBER 01, 2010 to SEPTEMBER 30, 2011.
- CME Credit for JAMA Review. My peer review has met the criteria of quality and timeliness required to claim credit. 3/22/2010.
- Birmingham Top Ten Outstanding Chinese, Awarded by Birmingham Chinese Association. July 2009.
- The 2007 recipient of the Distinguished Faculty Investigator Award, School of Public Health, University of Alabama at Birmingham
- 2004 Named New Investigator, Clinical Nutrition Research Center, University of Alabama at Birmingham
- 2004 Best Paper Award by Science Unbound Foundation, USA.
- Best Paper Award by Zhejiang science and technology commission, 1992.
- National Natural Science Foundation of China Grant "Statistical methods for mapping quantitative trait loci in forest trees", 1998-2000.
- National Natural Science Foundation of China Grant "Construction of genetic linkage map and mapping quantitative trait loci in Chinese Fir and Poplar", 1998-2000.
- National Natural Science Foundation of China Grant "Genetic analysis of forest germplasm resources for important trees species of *Castanopsis* special in China", 1998-2000.

PROFESSIONAL MEMBERSHIPS

- The American Statistical Association, 2005-present.
- The American Society of Human Genetics, 2010-present.

- NAASO, North American Association for the Study of Obesity, 2004-present.
- International Biometric Society, 2000-present.
- International Society for Bayesian Analysis, 2000-present.

JOURNAL EDITORIAL BOARD

- Associate Editor for *Genetics*, 2012 - present
- Associate Editor for *BMC Genetics*, 2009 - present
- Associate Editor for *International Journal of Plant Genomics*, 2009 - present
- Associate Editor for *Journal of Biometrics & Biostatistics*, 2010 – present
- Associate Editor for *Review of Bioinformatics & Biometrics*, 2012 – present
- Associate Editor for *Journal of Applied Bioinformatics & Computational Biology*, 2012 – present
- Review Editor for *Frontiers in Statistical Genetics and Methodology*, 2010- present
- Associate Editor for *Genetics Research*, 2009 – 2012

RESEARCH AREAS/INTERESTS AND ACHIEVEMENTS

My research areas and interests include:

- **Statistical Genetics:** genetic association studies of common and rare variants in human populations, QTL mapping for complex traits in animal and plant experimental crosses, statistical analysis of genetic interactions (gene-gene and gene-environment interactions), predictive and prognostic modeling of complex diseases and traits
- **Statistical Precision or Personalized Medicine:** developing hierarchical GLMs and survival models for predicting disease risk or survival using large-scale clinical and molecular profiling data and longitudinal measurements
- **Microbiome/Metagenomics Data Analysis:** developing hierarchical zero-inflated models and algorithms for microbiome association studies and for microbiome-based predictive modeling
- **Bayesian Methods:** Bayesian hierarchical modeling, deterministic and MCMC algorithms, and Bayesian software (RStan and WinBUGS)
- **Computer Software Development:** developing R packages for implementing our own and existing statistical methods
- **Genetics and Genomics of Human Diseases:** Obesity, Cancers, etc.

I have developed Bayesian models and Markov chain Monte Carlo (MCMC) algorithms for detecting multiple quantitative trait loci (QTL) and genetic interactions (G x G and G x E) for complex continuous, binary and ordinal phenotypes in experimental crosses. I have also developed various Bayesian hierarchical models (including hierarchical generalized linear models (GLMs), Cox survival models, and zero-inflated (negative binomial, Poisson, Gaussian) models) and fast deterministic algorithms for jointly analyzing numerous clinical, environmental and genetic variables for identifying important predictors and predicting disease outcomes. I have implemented our Bayesian multiple QTL mapping methods in R package qtlbim (www.qtlbim.org), and our Bayesian hierarchical methods in R package BhGLM (<http://www.ssg.uab.edu/bhglm>).

I have published a number of methodological papers in prestigious journals such as *PLoS Genetics*, *PNAS*, *Genetics*, *Bioinformatics*, *Genetic Epidemiology*, and *Statistical Applications in Genetics and Molecular Biology*, review papers in *Current Opinion in Genetics & Development*, *Nature Genetics*, and *Genetics Research*, and applied papers in journals such as *International Journal of Obesity*, *Journal of Lipid Research*, *BMC Genomics*, *Critical Care*, *British Journal of Cancer*, and *Molecular Cancer Research*. I have completed two NIH R01 grants as PI, entitled “*Bayesian methods for mapping complex epistatic genes*” and “*Bayesian Methods for Genome-Wide Interacting QTL Mapping*”, respectively, and multiple NIH grants as Co-Investigator. I have served as Associate Editor for several journals.

PEER-REVIEWED PUBLICATIONS

1. Yang, Y., H. Pei and N. Yi, 1991 Shortcut sampling inspection plans and their OC functions. *J. Biomathematics*, 6(4): 49-52.
2. Yi, N., 1992 A study of precision of several sampling plans. *J. Biomathematics*, 7(1): 38-42.
3. Yi, N., Y. Yang *et al*, 1992 Mean precision of multi-stage sampling. *J. Biomathematics*, 7(4): 169-171.
4. Yang, Y., H. Pei, N. Yi, and Y. Gao, 1994 Optimum allocations of sample size in two-stage sampling plan. *J.App.Math.*, 9(4): 46-50.
5. Yi, N., J. Shi and M. Wang, 1995 Multivariate analysis of genetic differentiation in forest populations. *J. Nanjing For. Uni.*, 4: 7-14.
6. Yi, N., 1995 Comparison of 3P sampling with PSWOR sampling. *J.Nanjing For.Uni.*, 19(1): 89-102.

7. **Yi, N.** and Y. Yang, 1996 Mean precision of the ratio-type estimator in the two-stage sampling. *J. Biomathematics*, 11(1): 28-31.
8. **Yi, N.**, J. Shi and M. Wang, 1996 Genetic diversity and multilocus structure in forest trees. *Chinese Biodiversity*, 4(3): 153-159.
9. **Yi, N.** "A study on population genetic structure in forest trees", Ph.D. Dissertation, Nanjing Forestry University, 1996.
10. Yin, T, Y. Sun, **N. Yi**, X. Li, M. Huang and M. Wang, 1998 Genome fingerprinting analysis in *Populus Deltoides*. *Acta Botanica Sinica*, 40(8):778-780.
11. **Yi, N.**, J. Shi and M. Wang, 1998 Estimation, variance and optimal sampling of multilocus gene diversity. *J. Biomathematics*, 14: 23-30.
12. **Yi, N.**, J. Shi and M. Wang, 1998 Linkage disequilibrium in populations of *Cunninghamia lanceolata* (Lamb.) Hook from The People's Republic of China. *J. Biomathematics*, 15: 11-19.
13. **Yi, N.**, T. Yin, M. Huang, and L. Zhu, 1998 Mapping quantitative trait loci in forest trees. *Progress in Biotechnology (Chinese)*, 18(3): 19-25.
14. **Yi, N.**, T. Yin, Z. Han, M. Huang, and H. Cao, 2000 Genetic variation of RAPD markers in a disease resistant seed orchard of *Pinus Elliottii* Engelm. *Scientia Silvae Sinicae* 36: 51-53.
15. **Yi, N.**, 2000 A random model approach to mapping quantitative trait loci using Megagametophyte in Conifers. *Scientia Silvae Sinicae*. 37: 36-41.
16. Li Mei, Shi Jisen, Gan Siming, He Zhenxiang, Li Li, **N. Yi**, 2001 Correlation between RAPD based parental genetic distance and filial performance of Chinese fir, *Forest Research*. 14 (1):35-40 (Chinese)
17. Li Mei, Shi Jisen, He Zhenxiang, **N. Yi**, 2001 Study on molecular genetic variation of superior trees in Chinese fir, *SCIENTIA SILVAE SINICAE*. 37(4):137-141.
18. **Yi, N.** and S. Xu, 1999 Mapping quantitative trait loci for complex binary traits in outbred populations. *Heredity* 82: 668-676.
19. **Yi, N.** and S. Xu, 1999 A random approach to mapping quantitative trait loci for complex binary traits in outbred populations. *Genetics* 153: 1029-1040.
20. Xu, S. and **N. Yi**, 2000 Mixed model analysis of quantitative trait loci. *Proceedings of the National Academy of Sciences of the USA*. 97: 14542-14547.
21. **Yi, N.** and S. Xu, 2000 Bayesian mapping of quantitative trait loci for complex binary traits. *Genetics* 155: 1391-1403.

22. **Yi, N.** and S. Xu, 2000 Bayesian mapping of quantitative trait loci under the identity-by-descent-based variance component model. *Genetics* 156: 411-422.
23. **Yi, N.** and S. Xu, 2001 Bayesian mapping of quantitative trait loci under complicated mating designs. *Genetics* 157: 1759-1771.
24. Xu, S. and **Yi, N.**, 2001 Genetic mapping for complex traits using Bayesian statistics. In “International symposium on mapping and identification of genes for complex polygenic traits and diseases and symposium on application of biological high technology and its products”. Changsha, China.
25. **Yi, N.** and Xu, S., 2002 Mapping quantitative trait loci with epistatic effects. *Genetical Research* 79: 185-198.
26. **Yi, N.** and S. Xu, 2002 Linkage analysis of quantitative trait loci in multiple line crosses. *Genetica* 114: 217-320. [PMID: 12206360]
27. **Yi, N.**, V. George and D. B. Allison, 2003 Stochastic search variable selection for mapping multiple quantitative trait loci. *Genetics* 164: 1129-1138. [PMID: 12871920 PMCID: PMC1462611]
28. **Yi, N.**, S. Xu and D. B. Allison, 2003 Bayesian model choice and search strategies for mapping interacting quantitative trait loci. *Genetics* 165: 867-883. [PMID: 14573494 PMCID: PMC1462771]
29. Xu, S., **N. Yi**, D. Burke, A. Galeki and R. A. Miller, 2003 An EM algorithm for mapping binary disease loci: application to fibrosarcoma in a four-way cross mouse family. *Genetical Research* 82: 127-138. [PMID: 14768897]
30. Hsu, H., Zhang, H., Li, L., **N. Yi**, Yang, P., Wu, Q., Wu, Y., Sun, S., Renda, J., Xu, X., Yang, X., Lu, L., Van Zant, G., Williams, R. W., Allison, D. B., & Mountz, J. D., 2003 Age-related Thymic Involution in C57BL/6J X DBA/2J Recombinant Inbred Mice Fits A Negative Exponential Regression Model and Maps to Mouse Chromosome 9. *Genes and Immunity* 4: 402-10.
31. **Yi, N.** and Xu, S., 2003 Designs and methods to detect QTL for production traits based on random genetic models. In “*Poultry breeding and biotechnology*”, edited by W. Muir and S. Aggrey. CABI Publishing, Wallingford, UK.
32. G. Churchill (**N. Yi** among many other authors), 2004 The collaborative cross, a community resource for the genetic analysis of complex traits. *Nature Genetics* 11: 1133-1137. [PMID: 15514660]
33. **Yi, N.**, 2004 A unified Markov chain Monte Carlo framework for mapping multiple quantitative trait loci. *Genetics* 167: 967-975. [PMID:15238545 PMCID: PMC1470906]

34. **Yi, N.**, A. Diament, S. Chiu, J. Fisler and C. Warden, 2004 Epistatic interaction between two nonstructural loci on chromosomes 7 and 3 influences hepatic lipase activity in BSB mice. *J Lipid Res* 2004; 45:2063-70. [PMID: 15314098]
35. **Yi, N.**, Adam Diament, Sally Chiu, Janis Fisler and Craig Warden, 2004 Characterization of Epistasis influencing complex spontaneous obesity in the BSB model. *Genetics* 167: 399-409. [PMID: 15166164 PMCID: PMC1470871]
36. **Yi, N.**, S. Xu, V. George and D. B. Allison, 2004 Mapping multiple quantitative trait loci for complex ordinal traits. *Behavior Genetics* 34: 3-15. [PMID: 14739693]
37. Warden, C., **N. Yi** and J. Fisler, 2004 Epistasis among genes is a universal phenomenon in obesity: evidence from rodent models. *Nutrition* 20: 74-77. [PMID: 14698018]
38. Huang-Ge Zhang, Hui-Chen Hsu, **N. Yi**, PingAr Yang, Qi Wu, David B. Allison and John D. Mountz, 2004 Identification of Multiple Genetic Loci that Regulate Adenovirus Gene Therapy. *Gene Therapy* 11: 4-14.
39. Beasley, T. M., D. Yang, **N. Yi**, D. C. Bullard, C. I. Amos, S. Xu, and D. B. Allison, 2004 A class of combined covariance-based and marginal-based tests for quantitative trait loci in experimental crosses. *Genetics Selection Evolution* 36: 601-619. [PMID: 15496283 PMCID: PMC2697196]
40. Farahani, P., J. S. Fisler, H. Wong, A. L. Diament, **N. Yi** and C. H. Warden, 2004 Reciprocal hemizygoty analysis of mouse hepatic lipase (*Lipc*) reveals influence on obesity. *Obesity Research* 2: 292-306. [PMID: 14981222]
41. C. Hansen, **N. Yi**, Y-M Zhang, S. Xu, Gavora J., and H. Chen, 2005 Identification of QTL for production traits in chicken. *Animal Biotechnology* 16: 67-79. [PMID: 15926264]
42. John D. Mountz, PingAr Yang, Qi Wu, Juling Zhou, Angela Fitzgerald, Jennifer Allen, **N. Yi**, Lu Lu, Robert W. Williams, and Hui-Chen Hsu, 2005 Genetic segregation of spontaneous erosive arthritis and generalized autoimmune disease in BXD2 recombinant inbred strain of mice. *Scand J Immunology* 61: 128-138.
43. Huang-Ge Zhang, Katherine A. High, Qi Wu, PingAr Yang, Alex Schlachterman, Shaohua Yu, **N. Yi**, Hui-Chen Hsu, and John D. Mountz, 2005 Genetic Analysis of the Antibody Response to AAV2 and Factor IX. *Molecular Therapy* 11(6):866-74. [PMID: 15922957]
44. Maria De Luca, **N. Yi**, David B. Allison, Jeff Leips, and Douglas M. Ruden, 2005 Mapping Quantitative Trait Loci affecting variation in *Drosophila* triacylglycerol storage. *Obesity Research* 13: 1596-1606. [PMID: 16222063]
45. Joseph R. Vasselli, Richard Weindruch, Stephen B. Heymsfield, F. Xavier Pi-Sunyer, Carol N. Boozer, **N. Yi**, Chenxi Wang, and David B. Allison, 2005 'Intentional' Weight Loss Reduces

- Mortality Rate in a Rodent Model of Dietary Obesity. *Obesity Research* 13:693-702. [PMID: 15897478]
46. **Yi, N.**, B. Yandell, G. Churchill, D. Allison, E. J. Eisen and D. Pomp, 2005 Bayesian model selection for genome-wide epistatic analysis. *Genetics* 170: 1333-1344. [PMID: 15911579 PMCID: PMC1451197]
 47. Ruth, J.H., Amin, M.A., Woods, J.M., He, X., Samuel, S.L., **Yi, N.**, Haas, C.S., Koch, E.A., and Bullard, D.C., 2005 Accelerated Development of Arthritis in Mice Lacking Endothelial Selectins. *Arthritis Research & Therapy* 7: R959-R970. [PMID: 16207337 PMCID: PMC1257424]
 48. Yang, R., **N. Yi** and S. Xu, 2006 Box-Cox transformation for QTL mapping. *Genetica* 128:133–143. [PMID: 17028946]
 49. **Yi, N.**, Denise K. Zinniel, Kyoungmi Kim, Eugene J. Eisen, Alfred Bartolucci, David B. Allison and Daniel Pomp, 2006 Bayesian analysis of multiple epistatic QTL models for body weight and body composition in Mice. *Genetical Research* 87: 45-60. [PMID: 16545150 PMCID: PMC5002393]
 50. Solomon K Musani, Huang-Ge Zhang, Hui-Chen Hsu, **N. Yi**, David B Allison, and John D. Mountz, 2006 Principal Component Analysis of Quantitative Trait Loci for Adenovirus. *Hereditas* 143: 189-197. [PMID: 17362354]
 51. Daniel Shriner, Solomon K Musani, and **N. Yi**, 2007 Statistical methods for multiple QTL mapping in experimental cross. In “Current Topics in Human Genetics: studies in complex diseases” ed. Hong-Wen Deng. World Scientific Publishing Co. Pte. Ltd, Singapore.
 52. Hui-Chen Hsu, Lu Lu, **N. Yi**, G. V. Zant, R. W. Williams and J. D. Mountz, 2007 Quantitative Trait Loci (QTL) Mapping in Aging Systems. In “Methods in Molecular Biology: Biological Aging: Methods and Protocols”, edited by Tollefsbol, Trygve O.. Humana Press.
 53. S. K. Musani, D. Shriner, N. Liu, R. Feng, C. S. Coffey, **N. Yi**, H. K. Tiwari, & D. B. Allison, 2007 Detection of Gene \times Gene Interactions in Association Studies of Human Data. *Human Heredity* 63(2):67-84. [PMID: 17283436]
 54. Brian S. Yandell, Tapan Mehta, Samprit Banerjee, Daniel Shriner, Ramprasad Venkataraman, Jee Young Moon, W. Whipple Neely, Hao Wu, Randy von Smith and **Nengjun Yi**, 2007 R/qtlbim: QTL with Bayesian Interval Mapping in Experimental Crosses. *Bioinformatics* 23: 641-634. [PMID: 17237038 PMCID: PMC4995770]
 55. **Yi, N.**, D. Shriner, S. Banerjee, Tapan Mehta, D. Pomp, and Brian S. Yandell, 2007 An efficient Bayesian model selection approach for interacting QTL models with many effects. *Genetics* 176: 1865–1877. [PMID: 17483424 PMCID: PMC1931520]

56. **Yi, N.**, S. Banerjee, D. Pomp, and B. S. Yandell, 2007 Bayesian analysis of genome-wide interacting QTL for ordinal traits. *Genetics* 176: 1855–1864. [PMID: 17507680 PMCID: PMC1931535]
57. Hui-Chen Hsu, Lu Lu, **N. Yi**, G. V. Zant, R. W. Williams and J. D. Mountz, 2007 Quantitative Trait Loci (QTL) Mapping in Aging Systems. *Methods Mol Biol.* 371:321-48. [PMID: 17634591]
58. **Yi, N.** and D. Shriner, 2008 Advances in Bayesian multiple QTL mapping in experimental designs. *Heredity* 100: 240-252. [PMID: 17987056 PMCID: PMC5003624]
59. Shriner, D. and **N. Yi**, 2008 Deviance information criterion (DIC) in Bayesian multiple QTL mapping. *Computational Statistics and Data Analysis.* doi:10.1016/j.csda.2008.01.016. PMID: 20161286 PMCID: PMC2682718]
60. **Yi, N.** and S. Xu, 2008 Bayesian LASSO for quantitative trait loci mapping. *Genetics* 179: 1045–1055. [PMID: 18505874 PMCID: PMC2429858]
61. Banerjee, S., B.S. Yandell, and **N Yi**, 2008 Bayesian QTL mapping for multiple traits. *Genetics* 179: 2275-2289. [PMID: 18689903 PMCID: PMC2516097]
62. **Yi, N.**, S. Ding, S. W. Keith, C. S. Coffey, and D. B. Allison, 2008 Bayesian analysis of the effect of intentional weight loss on mortality rate. *International Journal of Body Composition Research* 6: 185-192. [PMID: 25285041 PMCID: PMC4181669]
63. Badu, G., D. Pomp, D. Shriner, D. B. Allison, and **N. Yi**, 2009 Genetic influences on growth and body composition in mice: multilocus interactions. *International Journal of Obesity* 33, 89–95. [PMID: 18982013 PMCID: PMC3206648]
64. **Yi, N.** and Banerjee, S. 2009 Hierarchical generalized linear models for genome-wide interacting QTL mapping. *Genetics* 181: 1101-1113. [PMID: 19139143 PMCID: PMC2651046]
65. Richard J. Reynolds, James M. Kelley, Laura B. Hughes, **N. Yi**, and S. Louis Bridges, Jr. 2009 Genetic association of htSNPs across the major histocompatibility complex with rheumatoid arthritis in an African American population. *Genes and Immunity* 11(1):94-7 doi:10.1038/gene.2009.69 [PMID: 19741715 PMCID: PMC2809137]
66. Boris Pasche, Kari B. Wisinski, Maureen Sadim, Virginia Kaklamani, Michael Pennison, Qinghua Zeng, Naresh Bellam, Jacquelyn Zimmerman, **Nengjun Yi**, Kui Zhang, John Baron, Daniel O. Stram, M. Geoffrey Hayes. 2010. Constitutively decreased *TGFBR1* allelic expression is a common finding in colorectal cancer and is associated with three *TGFBR1* SNPs. *Journal of Experimental & Clinical Cancer Research* **29**:57 doi:10.1186/1756-9966-29-57 [PMID: 20500843 PMCID: PMC2890549]
67. G. A. Ankra-Badu, D Shriner, E. LE Bihan-Duval, S. Mignon-Grasteau, F. Pitel, C. Beaumont, M. J. Duclos, J. Simon, T. E. Porter, A. Vignal, L. A. Cogburn, D. B. Allison, **N. Yi**, S. E. Aggrey. 2010. Mapping main, epistatic and sex-specific QTL for body composition in a chicken population divergently selected for low or high growth rate. *BMC Genomics* 2010, **11**:107.

[PMID: 20149241 PMCID: PMC2830984]

68. Jun Li, Richard Renolds, Daniel Pomp, David B. Allison, **Nengjun Yi**. 2010. Mapping Interacting QTL for Count Phenotypes Using Hierarchical Poisson and Binomial Models: an Application to Reproductive Traits in Mice. *Genetics Research* 92: 13–23. [PMID: 20199696 PMCID: PMC2938180]
69. Boris Pasche, **N Yi** 2010 Candidate gene association studies: Successes and failures. *Current Opinion in Genetics & Development* 20: 257–261. [PMID: 20417090 PMCID: PMC2885524]
70. **Yi, N.**, Virginia Kaklamani, and Boris Pasche 2011 Bayesian Analysis of Genetic Interactions in Case-Control Studies, With Application to Adiponectin Genes and Colorectal Cancer Risk. *Annals of Human Genetics* 75: 90–104. (Article first published online: doi: 10.1111/j.1469-1809.2010.00605.x) [PMID: 20846215 PMCID: PMC3005151]
71. **Yi, N.** 2010 Statistical Analysis of Genetic Interactions. *Genetics Research* 92: 443-459. [PMID: 21429274 PMCID: PMC3203544]
72. **Yi, N.** and Degui Zhi 2011 Bayesian Analysis of Rare Variants in Genetic Association Studies. *Genetic Epidemiology*. 35: 57-69. [PMID: 21181897 PMCID: PMC3200544]
73. Virginia Kaklamani, **N. Yi**, K. Zhang, Maureen Sadim, Kenneth Offit, Carole Oddoux, Harry Ostrer, Christos Mantzoros, Boris Pasche. 2011 Polymorphisms of *ADIPOQ* and *ADIPOR1* and prostate cancer risk. *Metabolism* 60(9):1234-43. [PMID: 21397927 PMCID: PMC3134585]
74. Jun Li, Kui Zhang, and **Nengjun Yi** 2011 A Bayesian hierarchical model for detecting haplotype-haplotype and haplotype-environment interactions in genetic association studies. *Human Heredity* 71(3): 148-60 (DOI: 10.1159/000324841). [PMID: 21778734 PMCID: PMC3153342]
75. Lin WY, Zhang B, **Yi N**, Gao G, Liu N 2011 Evaluation of Pooled Association Tests for Rare Variants Identification with the Genetic Analysis Workshop 17 Data. *BMC Proceedings* 5 (Suppl 9): S118.
76. Virginia Kaklamani, **Nengjun Yi**, Maureen Sadim, Kalliopi Siziopikou, Kui Zhang, Yianfei Xu, Sarah Tofilon, Surbhi Agarwal, Boris Pasche and Christos Mantzoros 2011 The role of the fat mass and obesity associated gene (FTO) in breast cancer risk. *BMC Medical Genetics*, 12:52, doi:10.1186/1471-2350-12-52. [PMID: 21489227 PMCID: PMC3089782]
77. John J Arcaroli, Nianjun Liu, **Nengjun Yi**, and Edward Abraham 2011 Association between IL-32 and outcome in infection-associated acute lung injury. *Critical Care* 15:R138. [PMID: 21649914 PMCID: PMC3219007]
78. Wu G, **Yi N**, Absher D, Zhi D (2011) Statistical Quantification of Methylation Levels by Next-Generation Sequencing. *PLoS ONE* 6(6): e21034. doi:10.1371/journal.pone.0021034. [PMID: 21698242 PMCID: PMC3115964]
79. Chen, G. B., K. H. Ingram, G. de los Campos, **N. Yi**, X. Y. Lou, D. Pomp, D. B. Allison. (2011) A two-step modeling strategy for testing and estimating genetic susceptibility to the ill-effects of adiposity: Illustration in an outbred F2 mice population. *International Journal of Obesity Supplements*, S2:21.

80. **Yi, N.**, N. Liu, D. Zhi, J. Li 2011 Hierarchical Generalized Linear Models for Multiple Groups of Rare and Common Variants: Jointly Estimating Group and Individual-Variants Effects. *PLoS Genetics* 7(12): e1002382. doi:10.1371/journal.pgen.1002382. [PMID: 22144906 PMCID: PMC3228815]
81. Lin WY, Zhang B, **Yi N**, Gao G, Liu N. 2011 Evaluation of pooled association tests for rare variant identification. *BMC Proc.* 2011 Nov 29; 5 Suppl 9: S118. [PMID: 22373333 PMCID: PMC3287842]
82. Reng-Yun Liu, Ping Chen, Zhe Lei, Jingcheng Miao, **Nengjun Yi**, Kui Zhang, Boris Pasche, Hong-Tao Zhang 2012 Association between *IL6* -174G/C and cancer: a meta-analysis of 103,827 individuals. *Experimental and Therapeutic Medicine* 3(4): 655-664. [PMID: 22969947 PMCID: PMC3438703]
83. Jacquelyn W. Zimmerman, Michael J. Pennison, Ivan Brezovich, **Nengjun Yi**, Celeste T. Yang, Ryne Ramaker, Devin Absher, Richard M. Myers, Niels Kuster, Frederico P. Costa, Alexandre Barbault, Boris Pasche 2012 Cancer cell proliferation is inhibited by specific modulation frequencies. *British Journal of Cancer* 106: 307–313. doi:10.1038/bjc.2011.523. [PMID: 22134506 PMCID: PMC3261663]
84. Ma, Shuangge, Jian Huang, Yang Xie, **Nengjun Yi** 2012 Identification of Breast Cancer Prognosis Markers using Integrative Sparse Boosting. *Methods of Information in Medicine* 51(2):152-61. doi: 10.3414/ME11-02-0019. [PMID: 22344268 PMCID: PMC3598607]
85. Wan-Yu Lin, **Nengjun Yi**, Degui Zhi, Kui Zhang, Guimin Gao, Hemant K. Tiwari, and Nianjun Liu 2012 Haplotype-based methods for detecting uncommon causal variants with common SNPs. *Genetic Epidemiology*. DOI: 10.1002/gepi.21650. [PMID: 22706849 PMCID: PMC3513398]
86. Jun Chen, Reng-Yun Liu, Lixin Yang, Jun Zhao, Xueying Zhao, Daru Lu, **Nengjun Yi**, Baohui Han, Xiao-Feng Chen, Kui Zhang, Jun He, Zhe Lei, Yifeng Zhou, Boris Pasche, Xiangdong Li, Hong-Tao Zhang 2012 A two-SNP IL-6 promoter haplotype is associated with increased lung cancer risk. *Journal of Cancer Research and Clinical Oncology* 139: 231-242. [PMID: 23052692 PMCID: PMC4535449]
87. Ana I. Vazquez, M.S. Gustavo de los Campos, Yann C. Klimentidis, Guilherme J.M. Rosa, Daniel Gianola, **Nengjun Yi**, David B. Allison 2012 A Comprehensive Genetic Approach for Improving Prediction of Skin Cancer Risk in Humans. *Genetics* 192, 1493–1502. [PMID: 23051645 PMCID: PMC3512154]
88. Banerjee, S. and **Yi, N.**, 2012 Identifying QTL for Multiple Complex Traits in Experimental Crosses. In “the Methods in Molecular Biology series” ed. Scott Rifkin. Springer. *Methods Mol Biol.* 2012;871:205-25. doi: 10.1007/978-1-61779-785-9_11. [PMID: 22565839 PMCID: PMC5003623]
89. **Yi, N.** and S. Ma 2012 Hierarchical Shrinkage Priors and Model Fitting Algorithms for High-dimensional Generalized Linear Models. *Statistical Applications in Genetics and Molecular Biology* 11 (6): 1544-6115. DOI: 10.1515/1544-6115.1803. [PMID:23192052 PMCID: PMC3658361]

90. Lin WY, Zhang B, **Yi N**, Gao G, Liu N. Evaluation of pooled association tests for rare variant identification. *BMC Proc.* 2011 Nov 29;5 Suppl 9:S118. doi: 10.1186/1753-6561-5-S9-S118. [PMID: 22373333 PMCID: PMC3287842]
91. Wan-Yu Lin, **Nengjun Yi**, Xiang-Yang Lou, Degui Zhi, Kui Zhang, Guimin Gao, Hemant K. Tiwari, and Nianjun Liu 2013 Haplotype Kernel Association Test as a Powerful Method to Identify Chromosomal Regions Harboring Uncommon Causal Variants. *Genetic Epidemiology* 37(6):560-70. [PMID:23740760 PMCID: PMC4116485]
92. Mallick, H. and **Nengjun Yi** 2013 Bayesian Methods for High Dimensional Linear Models. *Journal of Biometrics & Biostatistics* S1: 005. doi:10.4172/2155-6180.S1-005. [PMID: 24511433 PMCID: PMC3914549]
93. Mallick, H. and **Nengjun Yi** 2013 Hierarchical Models for Genetic Association Studies. *Journal of Biometrics & Biostatistics* 4: e124. doi:10.4172/2155-6180.1000e124
94. Peng Li, Hemant K. Tiwari, Wan-Yu Lin, David B. Allison, Wendy K. Chung, Rudolph L. Leibel, **Nengjun Yi**, Nianjun Liu 2013 Genetic Association Analysis of 30 Genes Related to Obesity in European-American Population. *International Journal of Obesity* doi: 10.1038/ijo.2013.140. [PMID: 23900445 PMCID: PMC3909018]
95. Sangeetha Reddy, Maureen Sadim, Jun Li, **Nengjun Yi**, Surbhi Agarwal, Christos S Mantzoros, and Virginia Kaklamani 2013 Clinical and Genetic Predictors of Weight Gain in Patients Diagnosed with Breast Cancer. *British Journal of Cancer* 109: 872–881. [PMID: 23922112 PMCID: PMC3749587]
96. **Nengjun Yi**, Shizhong Xu, Xiang-Yang Lou, and Himel Mallick 2014 Multiple Comparisons in Genetic Association Studies: A Hierarchical Modeling Approach. *Statistical Applications in Genetics and Molecular Biology* 2014 Feb 1;13(1):35-48. doi: 10.1515/sagmb-2012-0040 [PMID: 24259248 PMCID: PMC5003626]
97. Sarah A. Jackson, Andrew A. Davis BS, Jun Li, **Nengjun Yi**, Shelley R. McCormick, Carly Grant, Taya Fallen, Beth Crawford, Kate Loranger, Jennifer Litton, Banu Arun, Kimberly Vande Wydeven, Amer Sidani, Katie Farmer, Merideth Sanders, Kent Hoskins, Robert Nussbaum, Laura Esserman, Judy E. Garber, Virginia G. Kaklamani and the Northwestern Cancer Genetics Group 2014 Characteristics of individuals with breast cancer rearrangements in BRCA1 and BRCA2. *Cancer* 120(10): 1557-64. [PMID: 24522996]
98. Qi Yan, Hemant K. Tiwari, **Nengjun Yi**, Wan-Yu Lin, Guimin Gao, and Nianjun Liu 2014 Kernel-Machine Testing Coupled with a Rank-Truncation Method for Genetic-Pathway Analysis. *Genetic Epidemiology* 38:447–456. [PMID:24849109 PMCID: PMC4073214]
99. Charles C. Weige, Marc R. Birtwistle, Himel Mallick, **Nengjun Yi**, Zuzana Berrong, Emily Cloessner, Keely Duff, Josephine Tidwell, Megan Clendenning, Brent Wilkerson4, Christopher Farrell, Fred Bunz, Hao Ji, Michael Shtutman, Kim E. Creek, Carolyn E. Banister, Phillip J. Buckhaults 2014 Transcriptomes and shRNA Suppressors in a TP53 Allele-specific Model of

Early-onset Colon Cancer in African Americans. *Molecular Cancer Research* DOI: 10.1158/1541-7786. [PMID: 24743655 PMCID: PMC4101030]

100. Himel Mallick and **Nengjun Yi** 2014 A new Bayesian LASSO. *Statistics and Its Interface* 7(4): 571-582. [PMCID: PMC4996624]
101. Yan, Q., Hemant K. Tiwari, **Nengjun Yi**, Guimin Gao, Wan-Yu Lin, Xiang-Yang Lou, and Nianjun Liu 2015 A Sequence Kernel Association Test for Dichotomous Traits in Family Samples under a Generalized Linear Mixed Model. *Human Heredity* 79(2):60-68. [PMID: 25791389 PMCID: PMC4825859]
102. Yan, Q., Daniel E. Weeks, Hemant K. Tiwari, **Nengjun Yi**, Kui Zhang, Guimin Gao, Wan-Yu Lin, Xiang-Yang Lou, Wei Chen, and Nianjun Liu 2016 Rare-Variant Kernel Machine Test for Longitudinal Data for Population and Family Samples. *Human Heredity* 80(3):126-138. [PMID: 27161037 PMCID: PMC4940283]
103. Lei Zhang, Yue-Ping Shen, Wen-Zhu Hu, Shu Ran, Yong Lin, Shu-Feng Lei, Christopher J. Papasian, **Nengjun Yi**, Yu-Fang Pei 2016 A new method for estimating effect size distribution and heritability from genome-wide association summary results. *Human Genetics* 135(2):171-84. [PMID: 26661625]
104. Xinyan Zhang, Yan Li, Zaixiang Tang, and **Nengjun Yi** 2016 Statistical Methods in Precision Medicine: Employing Systems Biology for Cancer Survival Prediction (invited review). *Journal of Biometrics & Biostatistics* 7: 281. doi:10.4172/2155-6180.1000281
105. Yan Li, Xinyan Zhang, Tomi Akinyemiju, Akinyemi I. Ojesina, Jeff M. Szychowski, Bo Xu, and **Nengjun Yi** 2016 A two-stage approach for combining gene expression and mutation with clinical data improves survival prediction in myelodysplastic syndromes and ovarian cancer. *Journal of Bioinformatics and Genomics* 1(1): 1-9.
106. Xinyan Zhang, Himel Mallick, and **Nengjun Yi** 2016 Zero-inflated Negative Binomial Regression for Differential Abundance Testing in Microbiome Studies. *Journal of Bioinformatics and Genomics* 1(2): 1-9.
107. Hong Zhu, Wei Xia, Xing-Bo Mo, Xiang Lin, Ying-Hua Qiu, **Nengjun Yi**, Yong-Hong Zhang, Fei-Yan Deng, Shu-Feng Lei 2016 Gene-based Genome-wide Association Analysis in Europeans and Asians identified novel genes for rheumatoid arthritis. *PLOS ONE* 11(11): e0167212. doi:10.1371/journal.pone.0167212
108. Xinyan Zhang, Yan Li, Omotomilayo F. Akinyemiju, Akinyemi I. Ojesina, Phillip Buckhaults, Nianjun Liu, Bo Xu, and **Nengjun Yi** 2017 Pathway-Structured Predictive Model for Cancer Survival Prediction: A Two-Stage Approach. *Genetics* 205: 89-100. [PMID: 28049703 PMCID: PMC5223526]

109. Zaixiang Tang, Yueping Shen, Xinyan Zhang, **Nengjun Yi** 2017 The Spike-and-Slab Lasso Generalized Linear Models for Prediction and Associated Genes Detection. *Genetics* 205: 77-88. [PMID: 27799277 PMID: PMC5223525]
110. Xinyan Zhang, Himel Mallick, Xiangqin Cui, Andrew K. Benson, and **Nengjun Yi** 2017 Negative Binomial Mixed Models for Analyzing Microbiome Count Data. *BMC Bioinformatics* 18(1):4 [PMID: 28049409 PMID: PMC5209949]
111. Himel Mallick and **Nengjun Yi** 2017 The Bayesian Group Bridge for Bi-level Variable Selection. *Computational Statistics and Data Analysis* (in press)
112. Zaixiang Tang, Qinghua Zeng, Yan Li, Xinyan Zhang, Jinlu Ma, Mark J Suto, Bo Xu, **Nengjun Yi** 2017 Development of a Radiosensitivity Gene Signature for Patients with Soft Tissue Sarcoma. *Oncotarget* (in press)
113. Zaixiang Tang, Yueping Shen, Xinyan Zhang, **Nengjun Yi** 2017 The Spike-and-Slab lasso Cox Model for Survival Prediction and Associated Genes Detection. *Bioinformatics*
114. Veturi Y., **Yi N.**, Huang W., Vazquez A.I., and de los Campos G. 2016 Estimating correlations of marker effects between Caucasians and African Americans by modeling genotype-by-race interactions. *Genetics*
115. Ranjit Kumar, **Nengjun Yi**, Degui Zhi, Peter Eipers, Kelly T. Goldsmith, Paula Dixon, David K. Crossman, Michael R. Crowley, Elliot J. Lefkowitz, J. Martin Rodriguez and Casey D. Morrow 2016 Identification of Donor Microbe Species that Colonize and Persist Long Term in the Recipient After Fecal Transplant for Recurrent *Clostridium difficile*. *Biofilms and Microbiome*
116. Himel Mallick and **Nengjun Yi** 2016 Bayesian bridge regression. *Journal of Applied Statistics*
117. Wei Xia; Xiao-Wei Zhu; Xin-Bo Mo; Long-Fei Wu; Jian Wu; Yu-Fan Guo; Ke-Qin Zeng; Ming-Jun Wang; Xiang Lin; Ying-Hua Qiu; Lan Wang; Pei He; Fang-Fei Xie; Peng-Fei Bing; Xin Lu; Yao-Zhong Liu; **Nengjun Yi**; Fei-Yan Deng; Shufeng Lei, 2016 Integrative multi-omics analysis revealed SNP-lncRNA-mRNA (SLM) networks in human peripheral blood mononuclear cells. *PLOS Genetics*
118. Zaixiang Tang, Qinghua Zeng, Yan Li, Xinyan Zhang, Mark J Suto, Bo Xu, **Nengjun Yi** 2016 A Radiosensitivity Gene Signature for Predicting Treatment Response for Patients with Soft Tissue Sarcoma.

BOOK CHAPTER

1. **Yi, N.** and Xu, S., 2003 Designs and methods to detect QTL for production traits based on random genetic models. In “*Poultry breeding and biotechnology*”, edited by W. Muir and S. Aggrey. CABI Publishing, Wallingford, UK.

2. Hui-Chen Hsu, Lu Lu, **N. Yi**, G. V. Zant, R. W. Williams and J. D. Mountz, 2007 Quantitative Trait Loci (QTL) Mapping in Aging Systems. In “Methods in Molecular Biology: Biological Aging: Methods and Protocols”, edited by Tollefsbol, Trygve O.. Humana Press.
3. Daniel Shriner, Solomon K Musani, and **N. Yi**, 2007 Statistical methods for multiple QTL mapping in experimental cross. In “Current Topics in Human Genetics: studies in complex diseases” ed. Hong-Wen Deng. World Scientific Publishing Co. Pte. Ltd, Singapore.
4. Banerjee, S. and **Yi, N.**, 2012 Identifying QTL for Multiple Complex Traits in Experimental Crosses. In “the Methods in Molecular Biology series” ed. Scott Rifkin. Springer.

SOFTWARE RELEASES

1. **R/qtlbim**: QTL Bayesian interval mapping. (B.S. Yandell and **N. Yi**, free release on September 2006 through CRAN) (www.qtlbim.org)

R/qtlbim is an extensible, interactive environment for the Bayesian Interval Mapping of QTL, built on top of R/qtl (Broman *et al.* 2003), providing Bayesian analysis of multiple interacting quantitative trait loci (QTL) models for continuous, binary and ordinal traits in experimental crosses. It includes several efficient Markov chain Monte Carlo (MCMC) algorithms for evaluating the posterior of genetic architectures, *i.e.* the number and locations of QTL, their main and epistatic effects, and gene-environment interactions. R/qtlbim provides extensive informative graphical and numerical summaries, and model selection and convergence diagnostics of the MCMC output, illustrated through the vignette, example and demo capabilities of R (R Development Core Team 2006). The package is freely available from cran.r-project.org.

2. **R/BhGLM**: Bayesian hierarchical GLMs and Survival Models, with application to Genetics and Epidemiology (<http://www.ssg.uab.edu/bhglm/>)

R/BhGLM provides functions for setting up and fitting various Bayesian hierarchical models (generalized linear models (GLMs), Cox survival models, negative binomial models, zero-inflated models, ordered logistic or probit regressions, truncated regressions, and conditional logistic models), for numerically and graphically summarizing the fitted models, for evaluating the predictive performance, and for hypothesis testing using Wald-type tests and score tests. Several types of priors on the coefficients can be used: double-exponential, Student-t, mixture double-exponential, mixture normal and generalized double Pareto. The methods can be used to analyze

not only general data but also large-scale genomic data (i.e., detecting disease-associated genes or variants and predicting phenotypes).

MANUSCRIPTS UNDER REVIEW

- Costa, Frederico P., de Oliveira, André Cosme, Meirelles, Roberto, Machado, Marcel C.C., Zanesco, Tatiana, Surjan, Rodrigo, Chammas, Maria Cristina, Souza Rocha, Manoel, Morgan, Desiree, Cantor, Alan, **Yi, Nengjun**, Yang, Celeste T., Bottger, Brad, Bomholt, Fin, Kuster, Niels, Barbault, Alexandre, Pasche Boris. Treatment of advanced hepatocellular carcinoma with intrabuccally-administered amplitude-modulated electromagnetic fields. (in review)
- Monique Hinchcliff, Michael J Pennison, Jacquelyn W. Zimmerman, Naresh Bellam, Qinghua Zeng, Chiang-Ching Huang, Richard Pope, Maureen Sadim, Wendy Wolf, Jeffrey Edberg, Robert Kimberly, Kui Zhang, Jun Li, **Nengjun Yi**, Maureen D. Mayes, John Varga, Boris Pasche. Constitutively decreased *TGFBR1* signaling and risk for systemic sclerosis. JAMA (in review)
- Tomas Kirchhoff, Justin Rendleman, Khedoudja Nafa, Virginia Kaklamani, Maureen Sadim, Nengjun Yi, Kui Zhang, Naresh Bellam, Kenneth Offit, Boris Pasche. *TGFBR1*6A* and susceptibility to lymphoma.
- Jun Li and Nengjun Yi. A Bayesian Hierarchical Generalized Linear Model for Identifying Multiple Interacting Genes in Family-Based Case-Control Studies.
- Michael J. Pennison, Qinghua Zeng, Naresh Bellam, Jacquelyn W. Zimmerman, Lakisha Moore-Smith, Minghui Wang, John C. Henegan, Phillip J. Buckhaults, Maureen Sadim, Virginia Kaklamani, Noralane Lindor, John L. Hopper, Loïc Le Marchand, Steven Gallinger, Polly A. Newcomb, Robert Haile, John A. Baron, Daniel O. Stram, Hong-Tao Zhang, Kui Zhang, Nengjun Yi, Boris Pasche 2016 *TGFBR1* and colorectal cancer risk. Nature (in review)
- Himel Mallick and Nengjun Yi 2016 Bayesian bridge regression
- Hong Zhu, Yu-Fan Guo, Wei Xia, Xing-Bo Mo, Nengjun Yi, Shu-Feng Lei 2015 Gene-based Genome-wide Association Analysis in Europeans and Asians identified novel genes for rheumatoid arthritis. Arthritis Care & Research
- Xinyan Zhang, Yan Li, Omotomilayo F. Akinyemiju, Akinyemi I. Ojesina, Phillip Buckhaults, Bo Xu, and Nengjun Yi 2016 Pathway-Structured Predictive Model for Cancer Survival Prediction: A Two-Stage Approach.
- Yan Li, Xinyan Zhang, Tomi Akinyemiju, Akinyemi I. Ojesina, Jeff M. Szychowski, Bo Xu, and Nengjun Yi 2016 A two-stage approach for combining gene expression and mutation with clinical data improves survival prediction in myelodysplastic syndromes and ovarian cancer.

- Zaixiang Tang, Yueping Shen, Xinyan Zhang, Nengjun Yi 2016 The Spike-and-Slab lasso Cox Model for Survival Prediction and Associated Genes Detection.
- Zaixiang Tang, Yueping Shen, Xinyan Zhang, Nengjun Yi 2016 The Spike-and-Slab Lasso Generalized Linear Models for Prediction and Associated Genes Detection
- Xinyan Zhang, Himel Mallick, and Nengjun Yi 2016 Zero-inflated Negative Binomial Regression for Differential Abundance Testing in Microbiome Studies.
- Xinyan Zhang, Himel Mallick, Xiangqin Cui, Andrew K. Benson, and Nengjun Yi 2016 Negative Binomial Mixed Models for Analyzing Microbiome Count Data
- Veturi Y., Yi N., Huang W., Vazquez A.I., and de los Campos G. 2016 Estimating correlations of marker effects between Caucasians and African Americans by modeling genotype-by-race interactions

RESEARCH GRANTS

Pending

1 R21 CA190936-01 (Yi)

NIH

Hierarchical Models for Analyzing Gene-Environment (G x E) Interactions

This project plans to develop novel statistical methods and computer software for jointly analyzing various environmental factors, main effects of genetic variants, and gene by environment (G x E) interactions for various diseases or traits.

Role: Principal investigator

1 R01 HG008739-01 (Yi)

NIH

Bayesian Methods for Incorporating Biological Information into Prognostic and Predictive Modeling

This project plans to develop a series of novel Bayesian hierarchical models for prognosis and prediction, to release user-friendly software for public use, and to fully analyze public datasets from The Cancer Genome Atlas (TCGA).

Role: Principal investigator

NSCOR: Genome instability and breast cancer risks after HZE radia (Xu)

NASA

Role: Project Principal Investigator

NIH R01 (Parant)

Deciphering how sister chromatid cohesion dysfunction accelerates tumorigenesis

Role: Co-Investigator

NIH R03 (Liu)

High-dimensional statistical genetic methods for dental caries and orofacial clefts.

Role: Co-Investigator

Active

NIH 2 RO1 GM069430 (Yi) 06/01/10 – 05/30/2016 5.76 cal
NIH/NIGMS \$ 231,057 (1st yr direct) \$ 1233237 (total)
Bayesian Methods for Genome-Wide Interacting QTL Mapping (score 2.2, percentile 10%)
To develop and evaluate Bayesian hierarchical models and model selection methods for mapping interacting QTL.
Role: Principal investigator

5P30-AI027767 (Saag) 06/15/09 – 05/31/15 0.6 cal
NIH/NIAIDS \$4,114,842
UAB Center for AIDS Research
The primary purpose of this project is to generate interdisciplinary AIDS research efforts.
Role: Co-Investigator

1U01CA158428-01A1 (Buckhaults) 07/01/13 - 06/30/17 0.72 cal
NIH/National Cancer Institute \$243,172
A Genetic Model for Early-Onset Breast and Colon Cancer in African Americans African. This work will formally prove whether or not specific DNA sequences, which are more common in African Americans than in Caucasian Americans, can predispose individuals to early onset breast or colon cancer, regardless of their race. If the hypothesis put forth in this proposal is validated, these results will provide the basis for identifying individuals that are at high risk for developing breast or colon cancer, and justify deploying more aggressive screening guidelines for those harboring high-risk DNA sequences.
Role: Co-Investigator

1R03DE024198-01A1 (Liu) 09/01/2015-08/31/2017 0.6 cal
NIH/NIDCR
Secondary analysis and risk prediction of dental caries and oral clefts
The goal of this project is to identify new genetic variants associated with dental caries and orofacial clefts by analyzing existing GWAS data using novel statistical methods, and to develop and validate high-dimensional, family-based Bayesian models to predict risks of dental caries and oral clefts.
Role: Co-Investigator

Completed

1R01 GM069430 (Yi) 06/01/05 – 05/30/2010 6.0 cal
NIH/NIGMS
Bayesian methods for mapping complex epistatic genes (score 130, percentile 0.7%)
This project is to develop Bayesian model selection methods and computer software for mapping epistatic genes in experimental crosses.
Total cost: \$1,150,175

Role: Principal investigator

Pilot-feasibility study (Yi) 06/01/03 – 06/01/05 1.2 cal

Clinical Nutrition Research Center, UAB

Mapping complex epistatic genes for obesity in experimental designs

This study is to develop statistical methods and algorithms for mapping multiple and epistatic genes for obesity in experimental crosses.

Role: Principal investigator

2 R01 CA112520-06A2 (Pasche) 12/01/09 – 11/30/14 0.6 cal

NIH

\$1,967,476 (total direct)

TGF-beta polymorphisms and breast cancer in families

Role: Co-investigator

1 R01 CA137000-01A1 (Pasche) 12/01/09 – 11/30/14 0.36 cal

NIH

\$1,250,000 (total direct)

TGFBR1 Signaling in Colorectal Cancer

Role: Co-investigator

NIH R01GM077490 (Allison) 9/01/07 - 08/31/11 3.48 cal

NIH/NIGMS

Genome-wide Structured Association Testing & Regional Admixture Mapping

Developing, evaluating, and applying enhanced SAT and RAM methods in the context of massive scale genome-wide association studies.

Role: Co-investigator

1R01 DK067487-01A1 (Allison) 01/01/2005 – 12/31/2008 1.8 cal

NIH

Effects of intentional weight loss on mortality rate

This project is to develop statistical methods for studying effects of intentional weight loss on mortality rate.

Role: Co-investigator

R01 HL80812-01 (Maria De Luca) 10/01/2004-09/30/2008 1.8 cal

NIH

QTL mapping age-related changes in lipid storage

This project is to identify QTL for age-related change in lipid storage in *Drosophila*.

Role: Co-investigator

R01 ES012933-01 (Douglas Ruden) 09/24/2004-08/31/2009 0.6 cal

NIH

QTL AND MICROARRAY MAPPING LEAD SENSITIVITY GENES

The major goals of this project are to use QTL mapping microarrays to identify lead sensitivity genes in *Drosophila*

Role: Co-investigator

VA Merit Review (John Mountz) 10/01/04-09/30/09 1.2 cal

Autoantibody Induced Arthritis and Autoimmune Disease in BXD2 Mice

The major goals of this project are to autoantibody induced arthritis and autoimmune disease using BXD2 mice.

Role: Co-investigator

R01 ES09912 (Amos) 7/1/02 – 6/30/07 1.2 cal
NIH

Positional Gene Identification of Complex Traits

This project is to develop statistical methods for mapping genes for complex traits in humans.

Role: Co-Investigator

R01 RR17009-01 (Bullard) 4/1/02 - 3/31/07
NIH

Genetic Mouse Models for Chronic Inflammatory Disease

This project will use the MRL/MpJ-Faslpr model system to further define the ICAM-1-dependent pathways responsible for mediating vasculitis.

Role: Co-Investigator

RO1 AG11653-08 (Mountz) 02/29/02 - 02/29/07
NIH

Defective Fas Apoptosis During Aging

This study focuses on identification of populations of age-responder (A-R) and age non-responder (A-NR) T cells and apoptosis pathways.

Role: Co-Investigator

R21 ES 11751 (Douglas Ruden) 03/15/02 - 03/14/04
NIEHS

Developmental Toxicology of Lead, Mercury, and Cadmium

The major goals of this project are to use microarrays to identify protective genes induced by *Drosophila* larvae exposed to heavy metals.

Role: Co-Investigator

NIH R01 CA112520 (Pasche) 09/07/05-08/31/09

TGF- β polymorphisms and breast cancer in families

To assess the association between the *TGFBR1**6A and the *TGFBI* T29C variants and familial breast cancer in 5,357 discordant sibling case-control sets for whom DNA and high-quality breast cancer risk factor data are available.

Role: Co-Investigator

NIH R01 CA108741 (Pasche) 02/24/06-01/31/09

TGF- β polymorphisms and colon cancer risk

To assess the association between the *TGFBR1**6A and the *TGFBI* T29C variants and familial colorectal cancer in 4,208 sibling case-control pairs from the NCI-sponsored familial colon cancer registry for whom DNA and high-quality colorectal cancer risk factor data are available.

Role: Co-Investigator

CURRENT TEACHING

- BST 740, Bayesian analysis. University of Alabama at Birmingham, Fall 2004, 2006, 2008, 2010, 2012, 2014, 2016.
- BST 741, Advanced Bayesian Analysis II. University of Alabama at Birmingham, Fall 2011, 2013, 2015.
- BST 760, Generalized Linear/Mixed Models. University of Alabama at Birmingham, Spring 2016.

PREVIOUS TEACHING

- BST 795, Special topics: Advanced Bayesian Analysis. University of Alabama at Birmingham, Fall, 2007, 2009.
- Faculty at the 1st Annual NSF-funded short course on statistical genetics & genomics. July 23, 2008.
- Guest Lecturer for HMG 703 “Quantitative Methods”. University of Alabama at Birmingham, April 2008
- Guest Lecturer for HMG 703 “Quantitative Methods”. University of Alabama at Birmingham, April 2007
- Guest Lecturer for MGE 702 “Advanced Human Genetics”, Statistical Approaches to Genetic Crosses, May 3, 2005.
- Faculty at the 3th Annual NIH-funded Short Course on Statistical Genetics for Obesity and Nutrition Researchers, Birmingham, AL, February 6-9, 2003.
- Quantitative Genetics: Bayesian analysis of QTL mapping in line crosses. University of California, Riverside, January 2001. Organized by Prof. Xu.
- Quantitative Genetics: Bayesian analysis of QTL mapping in complex pedigrees. University of California, Riverside, January 2001. Organized by Prof. Xu.
- Applied Statistics, Population Genetics, Quantitative Genetics, Forest Genetics and Tree Breeding. Nanjing Forestry University, P. R. China, August 1996-June 1998.
- Mathematical statistics, Multivariate statistics. Nanjing Forestry University, P. R. China, August 1988-July 1993.

STUDENTS AND POST-DOCTORAL FELLOW:

Post-doctoral Fellow:

2003 – 2005, Kyoungmi Kim (Co-mentor with Grier Page) – Assistant Professor, UC David

2005 - 2008, Daniel Shriner (Co-mentor with David Allison) – Research Fellow, Center for Research on Genomics and Global Health, NIH

2007 - 2009, Georgina Badu (Co-mentor with David Allison) – Research Associate, Department of Twin Research and Genetic Epidemiology, King's College London

2008 – 2010, Richard Reynolds (Co-mentor with Lou Bridges) – Assistant Professor, Division of Clinical Immunology and Rheumatology, University of Alabama at Birmingham

2010 – 2012, Ana Vazquez (Co-mentor with David Allison) - Assistant Professor, Department of Biostatistics, University of Alabama at Birmingham

2012 - 2013, Jun Li

Graduate Students:

2003 – 2008, Samprit Banerjee – Assistant, Associate Professor at Division of Biostatistics and Epidemiology, Department of Public Health, Weill Medical College of Cornell University

2008 – 2011, Jun Li – Post-doc fellow at the section on statistical genetics, University of Alabama at Birmingham

2008 – 2013, Tapan Metha – Assistant Professor, University of Alabama at Birmingham

2011 – 2015, Himel Mallick – Postdoc researcher, Harvard University

2013 – present, Xinyan Zhang (Doctoral candidate)

2013 – 2016, Peng Qu (Mater candidate)

2013 – 2016, Qingrui Meng (Mater candidate)

MASTER/PHD COMMITTEE CHAIR OR MEMBERSHIP

Samprit Benerjee, PhD dissertation, Bayesian genome-wide QTL mapping for multiple traits, 2008 (Committee Chair)

Jun Li, PhD dissertation, Bayesian Hierarchical Generalized Linear Models for Detecting (Rare)

Haplotype-haplotype and Haplotype-environment Interactions in Genetic Association Analysis, 2011 (Committee Chair)

Scott Keith, PhD dissertation, Free-knot splines and bootstrapping for nonlinear modeling in complex samples, 2008

Fenglong Xie, Master dissertation, Are SNPs in FCAR associated with SLE? 2008

Guodong Wu, Master dissertation, Statistical Quantification of Methylation Levels by Next Generation Sequencing, 2010

Guodong Wu, PhD dissertation, Statistical Analysis in Genomic Study, 12-03-2013

Qi Yan, PhD, PhD dissertation, Statistical Methods for Set-Based Association Tests in Genetic Studies, 01-2014

Committee member of Honors Thesis Proposal Committee, Ali Fry, Department of Mathematics, University of Alabama, 12-12-2014

Himel Mallick, PhD dissertation, Some contributions to Bayesian regularization methods with applications to genetics and clinical trials, 05-05-2015 (Committee Chair)

Qingrui Meng, Master dissertation, Survival Prediction with Stratified Cox Proportional Hazards Models of Primary Breast Cancer in African American versus White Women; SEER data 2010-2012, 04-15-2016 (Committee Chair)

Peng Qu, Master dissertation, Analysis of Risk Factors on Obesity/Overweight among Adolescents and Young Adults based on 1999-2012 NHANES Data. 04-18-2016 (Committee Chair)

Amanda L. Hall, Master dissertation, A Comparison of Immune Repertoire Diversity Profiling Techniques in the Descriptive Analysis of Patients Diagnosed with Inflammatory Bowel Disease. 06-08-2016.

Holly Hartman, Master dissertation, A Novel Approach to Analyzing High Throughput ELISA Data in Absence of a Standard. 06-07-2016.

Xinyan Zhang, PhD Proposal, statistical methods in cancer survival prediction and microbiome data analysis. 04-26-2016.

Seth T Lirette, PhD Proposal, Approaching brain perfusion data from a statistical perspective. 08-22-2016.

Fenglong Xie, PhD Proposal, Comparative effectiveness of tocilizumab on risk of cardiovascular disease among patients with rheumatoid arthritis. 08-19-2016.

Yan Li, PhD Proposal, Sample size re-estimation and sample size reallocation procedures in randomized clinical trials with multiple arms. 11-16-2016.

Yogasudha Veturi, PhD Dissertation, Methods for the analysis of genetic differences in ethnicity and sex for complex human traits. 11-18-2016 (Committee chair)

INVITED PRESENTATIONS

- Bayesian analysis of quantitative trait loci for complex traits in complicated mating designs. Pioneer Hi-Bred International, Inc. April 26, 2001.
- Bayesian analysis of quantitative trait loci for complex traits. Department of Biostatistics, the University of Alabama at Birmingham. July 12, 2001.
- Bayesian analysis of quantitative trait loci with epistatic effects. Plant and Animal Genome VIII, San Diego, California, January 2002.
- Mapping epistatic genes for complex traits. The conference "Statistical Integration of Genetic Information Across Data Domains: Biomedical, Agricultural, and Comparative Genomics", University of Alabama, Birmingham, AL. December 2nd, 2002.
- Bayesian Model Selection and Epistasis. Third Annual Short Course on Statistical Genetics for Obesity & Nutrition Researchers. University of Alabama, Birmingham, AL. February 6-9, 2003.
- Epistasis influences complex spontaneous obesity in the BSB model. NAASO 2003 Annual Scientific Meeting. Ft. Lauderdale, FL. October 11-15, 2003.
- The practical implementation of Bayesian model selection for identifying epistatic genes in experimental crosses. Plant and Animal Genome XII, San Diego, California, January 2004.
- Analysis of genome-wide epistasis via Bayesian model selection. Spring meeting of International Biometric Society, Eastern North American Region, Pittsburgh, PA, March 28-31, 2004.
- Bayesian model selection for mapping complex epistatic genes in experimental designs. Complex Trait Consortium, The Jackson Lab, July 6-9, 2004.
- Mapping complex epistatic genes for complex traits in experimental designs. Department of Biostatistics and Epidemiology, Florida International University, Feb 11, 2005.
- Mapping complex epistatic genes for obesity in experimental designs. Department of Nutrition, University of Alabama at Birmingham, March 15, 2005.

- Genetic Architecture of the Developmental Trajectory for Body Weight in a Cross of High Growth and Wild Strains of Mice. NAASO 2005 Annual Scientific Meeting. Vancouver, British Columbia, Canada. October 15-19, 2005.
- Bayesian Model selection for mapping interacting QTL. SCMA 2005 / FIM XII. Auburn University, Auburn, AL. December 3-4, 2005.
- Genome-wide interacting QTL analysis for complex traits in experimental crosses. Department of Biochemistry and Microbiology, Marshall University, December 2005.
- Efficient MCMC algorithms for mapping genome-wide interacting QTL. Spring meeting of International Biometric Society, Eastern North American Region, Tampa, FL, March 26-29, 2006.
- Bayesian multiple QTL analysis. Yangzhou University, China. June 26, 2006.
- Bayesian multiple QTL analysis. Nanjing Agricultural University, China. June 29, 2006.
- Bayesian multiple QTL analysis. Nanjing Forest University, China. July 3, 2006.
- Bayesian multiple QTL analysis. Shanghai Jiaotong University, China. July 29, 2006.
- Genome-wide interacting QTL analysis for obesity in experimental crosses. Monell Chemical Senses Center, Philadelphia. October 3, 2006.
- Bayesian multiple interacting QTL analysis. Department of Statistics and Bioinformatics Research Center, NCSU, Raleigh NC. March 29-30, 2007.
- Recent advances in Bayesian multiple QTL analysis. Lunch & Med Stat Seminar CCC Biostat & Bioinfo Unit & SOPH Biostat Dep, UAB. Feb 11, 2008.
- Large-scale hierarchical generalized linear models for interacting QTL mapping. Spring meeting of International Biometric Society, Eastern North American Region, Arlington, Virginia, March 16-19, 2008.
- Bayesian hierarchical generalized linear models for interacting QTL mapping. 1st Annual NSF-funded short course on statistical genetics & genomics. July 23, 2008.
- Hierarchical generalized linear models for multiple-SNP analysis. JSM Washington DC. Aug 2009.
- Detecting and characterizing multiple interacting genes in cancer case-control studies. 13th annual research retreat. UAB comprehensive cancer center. October 28, 2009.

- Statistical methods for finding the missing heritability of complex diseases. Genetics and Genomics Seminar Series, UAB school of medicine. Oct 8th, 2010.
- Hierarchical generalized linear models for finding the missing heritability of complex diseases/traits. Nanjing Agricultural University, China. July, 2011.
- Hierarchical generalized linear models for finding the missing heritability of complex diseases/traits. Yangzhou University, China. July, 2011.
- Hierarchical generalized linear models for finding the missing heritability of complex diseases/traits. Nanjing Forest University, China. July, 2011.
- Hierarchical GLMs for multiple groups of rare and common genetic variants. International Conference of Biometrics & Biostatistics, Omaha, NE. March 2012.
- Bayesian hierarchical GLMs, with application to genetic data analysis. Guangzhou University, China. June 2013.
- Bayesian hierarchical GLMs for genetic association studies. University of North Carolina at Chapel Hill. October 2, 2013.
- Hierarchical models for post-GWAS. Nanjing Agricultural University, China. July, 2014.
- Bayesian hierarchical generalized linear models (BhGLM) for genetic data analysis. Nanjing Agricultural University, China. July, 2014.
- Hierarchical models for post-GWAS. Suzhou University, China. August, 2014.
- Bayesian hierarchical GLMs for genetic association studies. Department of Mathematics, University of Alabama at Birmingham. November 21, 2014.
- Bayesian Methods for Incorporating Biological Information into Prognostic and Predictive Modeling. Soochow University, China. Jan 15, 2015.
- Bayesian Data Analysis. Soochow University, China. Jan 2015.
- LeClerc Rotary Lecture: Bayesian hierarchical GLMs for genetic association studies. University of Maryland, April 8, 2015.
- LeClerc Rotary Lecture: GLMs, hierarchical GLMs, R package BhGLM. University of Maryland, April 9, 2015.
- Statistical Methods for Precision medicine. Soochow University, China. June 24 – July 5, 2015.
- Bayesian hierarchical models for prognostic prediction using large-scale molecular data. ICASA, Shanghai, China. July 7, 2015.

- Bayesian hierarchical models for high-dimensional molecular data. Yangzhou University, China. December 21, 2015.
- Bayesian statistical methods for prognostic prediction of diseases using clinical and molecular data. Suzhou University, China. December 25-29, 2015.
- Penalized and Hierarchical Models for Prognostic Prediction Using Molecular Data. Fudan University, China. December 31, 2015.
- Statistical methods for microbiome data analysis. The Samuel Roberts Noble Foundation, Ardmore, OK. April 29, 2016.
- Bayesian hierarchical models with application to genomics and epidemiology. 3rd Short Course R25DK099080 The Mathematical Sciences in Obesity. June 13-17, 2016. University of Alabama at Birmingham.
- Statistical methods for microbiome data analysis. Suzhou University, China. July 02, 2016.
- Statistical methods for high-dimensional data and precision medicine. Soochow University, China. Aug 04 – 06, 2016.
- Microbiome/Metagenomics, NHGRI Short Course, University of Alabama at Birmingham. December 14, 2016

PEER REVIEW ACTIVITIES

Referee for Journals:

Genetics, American Journal of Human Genetics, Genetics Research, Genetica, Heredity, Trends in genetics, Human Heredity, Bioinformatics, Obesity Reviews, Genetics Selection Evolution, Genes and Immunity, BMC genomics, BMC Medical Genetics, Mammalian Genome, PLoS Genetics, PLoS One, Theoretical and Applied Genetics, Nature Genetics, Journal of the Royal Statistical Society, Annals of Human Genetics, BMC Genetics, Journal of Heredity, Computational Statistics and Data Analysis, Annals of Applied Statistics, Biometrics, JAMA, Genetic Epidemiology, Journal of Biopharmaceutical Statistics

Book review:

Springer (2006, 2007, 2008)

SAGE (2009)

NIH STUDY SECTION AND REVIEW OF GRANT PROPOSALS

- Genetic Variation and Evolution Study Section, NIH (July 2005)
- GCAT Study Section, NIH (March 2-3 2006)
- ZRG1 GGG-A, Center for Scientific Review Special Emphasis Panel, NIH (March 06-07, 2007)
- ZRG1 HOP-Q, Center for Scientific Review Special Emphasis Panel, NIH (May 16, 2008)
- ANR GENOPAT 2008, The French National Research Agency (ANR) (May 2008)
- ANR GENOMICS 2008, The French National Research Agency (ANR) (June 2008)
- ANR GENOPAT 2009, The French National Research Agency (ANR) (April 2009)
- ZRG1 PSE-03M, Center for Scientific Review Special Emphasis Panel, NIH (April 2009)
- ZRG1 PSE-02M, Center for Scientific Review Special Emphasis Panel, NIH (May 2009)
- ZRG1 GGG-F, Center for Scientific Review Special Emphasis Panel, NIH (June 2009)
- ZRG1 PSE-C, Challenge Grants Panel, NIH (June 2009)
- ZRG1 PSE-J, Challenge Grants Panel, NIH (June 2009)
- ZAT1 SM, Omics and Variable Responses to CAM (R21, R01), NIH (December 2009)
- External Peer Review, Cancer Research UK (Feb. 2010)
- ETMRC - CSO Research Grant, Scotland (Jan. 2010)
- RC4, Molecular Genetics A Study Section, NIH (May 2010)
- Morehouse School of Medicine/Tuskegee University/UAB Comprehensive Cancer Center Partnership, (June 2010)
- ZRG1 PSE-B 04 M, Genetics and Epidemiology of Chronic Diseases, NIH (Jan. 2011)
- ANR GEDA 2011, The French National Research Agency (ANR) (April 2011)
- ANR GWIS-AM 2011, The French National Research Agency (ANR) (April 2011)
- GCAT Study Section, NIH (October 6-7 2011)
- ANR JCJC SVSE 1 2012, The French National Research Agency (ANR) (March 2012)
- 2012 Collaborative Funding Grant (CFG) Program, North Carolina Biotechnology Center (March 2012)
- KNOD Study Section, NIH (May 31 – June 1 2012)

LOCAL SERVICE

- Associate director of postdoctoral fellows in the Section on Statistical Genetics, department of Biostatistics, UAB, 2011- 2013.
- The graduate program committee's administrative subcommittee, 2006-2010.
- Organizing the monthly journal club meeting at the Section on Statistical Genetics, Department of Biostatistics at UAB, from 2003-2005.
- Judge, UAB Post-doctoral Research Day, 2005.
- Grader for qualify exams at Department of Biostatistics, 2005–2011.