Nicholas Furlotte

Curriculum Vitae

Education

2008–2013 **PhD. Computer Science**, *The University of California*, Los Angeles.

2007–2008 **Masters Bioinformatics**, *The University of Memphis*, Memphis, TN.

2001–2005 **Bachelors Computer Science**, *The University of Memphis*, Memphis, TN.

PhD. Thesis

Title *Computational Genetics Approaches for the Dissection of Complex Traits*

Supervisors Professor Eleazar Eskin

Description Developed multiple computational approaches to increase statistical power to discover genetic variants associated with complex traits with a focus on linear-mixed models.

Masters Thesis

Title *Literature-based evaluation of microarray normalization procedures*

Supervisors Professor Ramin Homayouni & Robert Williams (University of Tennessee)

Description Developed a computational method to evaluate the efficacy of microarray normalization techniques based on literature-derived gene-gene correlations.

Experience

2016–Present **Senior Scientist Statistical Genetics, Health R&D**, 23andMe, Mountain View, CA.

Developing the next generation of data-driven consumer genetics health products using statistics and machine learning and one of the largest collections of genetic and phenotypic data.

2013–2016 **Statistical Geneticist**, 23andMe, Mountain View, CA.

Developed algorithms for genetics-based disease risk prediction. Drove a new product development paradigm based on the use of internal scientific data. Helped to create a new group (Health R&D) focused on creating the next iteration of 23andMe products.

2008–2013 **Graduate Research Assistant**, UCLA, Los Angeles.

2012 **Research Intern**, Microsoft, Los Angeles, Supervised by David Heckerman..

Developed a mixed-model-based method for quantifying the uncertainty in heritability estimation.

899 W. Evelyn – Mountain View, CA 94043

✉ *nick.furlotte [at] 23andMe [dot] com* • *whatmind.com*

2010–2012 **Computational Biology and Math Tutor**, UCLA, Los Angeles.

2008 **Software Engineer and Research Assistant**, University of Tennessee Health Science Center, Memphis, TN.
Developed a web-based platform for collecting clinical research data and performing analyses.

2007–2008 **Graduate Research Assistant**, The University of Memphis, Memphis, TN.

2006 **Partner and Lead Software Developer**, Shadespot.com, Memphis, TN.
Built a backend and frontend web store.

2005 **Undergraduate Research Assistant**, The University of Memphis, Memphis, TN.
Bioinformatics research with Dr. Max Garzon.

2003–2006 **Software Engineer**, Apley Inc., Memphis, TN.
Built front end and back end web tools based on MySQL, perl, python and other open source technologies.

2001 **Computer Technician**, Network and Computer Services (NCS), Memphis, TN.

Research Interests

- Computational and statistical genetics
- Complex trait analysis and risk prediction
- Machine Learning
- Statistics and Data Science
- Healthcare and Public Health

Awards

2012 Northrup-Grumman Outstanding Graduate Student Research

2005 Graduated Magna Cum Laude

Grants

2012–2013 Neurobehavioral Genetics Training Grant

2008–2010 Genomic Analysis Training Grant

Publications

A. V. Jones, J. R. Hockley, C. Hyde, D. Gorman, A. Sredic-Rhodes, J. Bilsland, G. McMurray, **Furlotte, N. A.**, Y. Hu, D. A. Hinds, P. J. Cox, and S. Scollen. Genome-wide association analysis of pain severity in dysmenorrhea identifies association at chromosome 1p13.2, near the nerve growth factor locus. *Pain*, Jul 2016.

P. Gormley et al. Meta-analysis of 375,000 individuals identifies 38 susceptibility loci for migraine. *Nature Genetics*, 48(8):856–866, Aug 2016.

A. Okbay et al. Genome-wide association study identifies 74 loci associated with educational attainment. *Nature*, 533(7604):539–542, May 2016.

A. Okbay et al. Genetic variants associated with subjective well-being, depressive symptoms, and neuroticism identified through genome-wide analyses. *Nature Genetics*, 48(6):624–633, Jun 2016.

J. H. Sul, M. Bilow, W. Y. Yang, E. Kostem, **Furlotte, N.**, D. He, and E. Eskin. Accounting for Population Structure in Gene-by-Environment Interactions in Genome-Wide Association Studies Using Mixed Models. *PLoS Genetics*, 12(3):e1005849, Mar 2016.

L. D. Orozco, M. Morselli, L. Rubbi, W. Guo, J. Go, H. Shi, D. Lopez, **Furlotte, N. A.**, B. J. Bennett, C. R. Farber, A. Ghazalpour, M. Q. Zhang, R. Bahous, R. Rozen, A. J. Lusis, and M. Pellegrini. Epigenome-wide association of liver methylation patterns and complex metabolic traits in mice. *Cell Metabolism*, 21(6):905–917, Jun 2015.

Furlotte, N. A. and E. Eskin. Efficient multiple-trait association and estimation of genetic correlation using the matrix-variate linear mixed model. *Genetics*, 200(1):59–68, May 2015.

C. L. Campbell, **Furlotte, N. A.**, N. Eriksson, D. Hinds, and A. Auton. Escape from crossover interference increases with maternal age. *Nature Communications*, 6:6260, 2015.

Furlotte, N. A., D. Heckerman, and C. Lippert. Quantifying the uncertainty in heritability. *J. Human Genetics*, 59(5):269–275, May 2014.

D. He, **Furlotte, N. A.**, F. Hormozdiari, J. W. Joo, A. Wadia, R. Ostrovsky, A. Sahai, and E. Eskin. Identifying genetic relatives without compromising privacy. *Genome Research*, 24(4):664–672, Apr 2014.

E. Y. Kang, B. Han, **Furlotte, N.**, J. W. Joo, D. Shih, R. C. Davis, A. J. Lusis, and E. Eskin. Meta-analysis identifies gene-by-environment interactions as demonstrated in a study of 4,965 mice. *PLoS Genetics*, 10(1):e1004022, Jan 2014.

A. Ghazalpour et al. Hybrid mouse diversity panel: a panel of inbred mouse strains suitable for analysis of complex genetic traits. *Mammalian Genome*, 23(9-10):680–692, Oct 2012.

Furlotte, N. A., E. Eskin, and S. Eyheramendy. Genome-wide association mapping with longitudinal data. *Genetic Epidemiology*, 36(5):463–471, Jul 2012.

Furlotte, N. A., E. Y. Kang, A. Van Nas, C. R. Farber, A. J. Lusis, and E. Eskin. Increasing association mapping power and resolution in mouse genetic studies through the use of meta-analysis for structured populations. *Genetics*, 191(3):959–967, Jul 2012.

T. M. Keane et al. Mouse genomic variation and its effect on phenotypes and gene regulation. *Nature*, 477(7364):289–294, Sep 2011.

A. Ghazalpour et al. Comparative analysis of proteome and transcriptome variation in mouse. *PLoS Genetics*, 7(6):e1001393, Jun 2011.

Furlotte, N. A., H. M. Kang, C. Ye, and E. Eskin. Mixed-model coexpression: calculating gene coexpression while accounting for expression heterogeneity. *Bioinformatics*, 27(13):i288–294, Jul 2011.

L. Xu, **Furlotte, N.**, Y. Lin, K. Heinrich, M. W. Berry, E. O. George, and R. Homayouni. Functional cohesion of gene sets determined by latent semantic indexing of PubMed abstracts. *PLoS ONE*, 6(4):e18851, 2011.

D. He, F. Hormozdiari, **Furlotte, N.**, and E. Eskin. Efficient algorithms for tandem copy number variation reconstruction in repeat-rich regions. *Bioinformatics*, 27(11):1513–1520, Jun 2011.

C. R. Farber, B. J. Bennett, L. Orozco, W. Zou, A. Lira, E. Kostem, H. M. Kang, **Furlotte, N.**, A. Berberyan, A. Ghazalpour, J. Suwanwela, T. A. Drake, E. Eskin, Q. T. Wang, S. L. Teitelbaum, and A. J. Lusis. Mouse genome-wide association and systems genetics identify Asxl2 as a regulator of bone mineral density and osteoclastogenesis. *PLoS Genetics*, 7(4):e1002038, Apr 2011.

D. He, **Furlotte, N.**, and E. Eskin. Detection and reconstruction of tandemly organized de novo copy number variations. *BMC Bioinformatics*, 11 Suppl 11:S12, 2010.

T. Viangteeravat, I. M. Brooks, W. J. Ketcherside, R. Houmayouni, **Furlotte, N.**, S. Vuthipadadon, and C. S. McDonald. Biomedical Informatics Unit (BMIU): Slim-prim system bridges the gap between laboratory discovery and practice. *Clin Transl Sci*, 2(3):238–241, Jun 2009.

A. Kirby, H. M. Kang, C. M. Wade, C. Cotsapas, E. Kostem, B. Han, **Furlotte, N.**, E. Y. Kang, M. Rivas, M. A. Bogue, K. A. Frazer, F. M. Johnson, E. J. Beilharz, D. R. Cox, E. Eskin, and M. J. Daly. Fine mapping in 94 inbred mouse strains using a high-density haplotype resource. *Genetics*, 185(3):1081–1095, Jul 2010.

B. J. Bennett, C. R. Farber, L. Orozco, H. M. Kang, A. Ghazalpour, N. Siemers, M. Neubauer, I. Neuhaus, R. Yordanova, B. Guan, A. Truong, W. P. Yang, A. He, P. Kayne, P. Gargalovic, T. Kirchgessner, C. Pan, L. W. Castellani, E. Kostem, **Furlotte, N.**, T. A. Drake, E. Eskin, and A. J. Lusis. A high-resolution association mapping panel for the dissection of complex traits in mice. *Genome Research*, 20(2):281–290, Feb 2010.

T. Viangteeravat, I. M. Brooks, E. J. Smith, **Furlotte, N.**, S. Vuthipadadon, R. Reynolds, and C. S. McDonald. Slim-prim: a biomedical informatics database to promote translational research. *Perspect Health Inf Manag*, 6:6, 2009.

V. Phan and **Furlotte, N. A.** Motif Tool Manager: a web-based framework for motif discovery. *Bioinformatics*, 24(24):2930–2931, Dec 2008.

Computing skills

Stats/Machine Learning

- numpy, scipy, pandas, patsy, skLearn, statsmodels, matplotlib, seaborn, R

- Familiarity with Theano and tensorflow as well as libraries like Lasagne and Keras

Programming, Infrastructure and Web Tech

- python, perl, C/C++,R
- Linux, AWS
- MySQL, HTML, javascript, css

Software

pyLMM A lightweight linear mixed model solver for GWAS. http://github.com/nickFurlotte/pylmm

mvLMM Matrix-variate linear mixed-model - an efficient method for performing multiple trait genome wide association analysis. http://genetics.cs.ucla.edu/mvLMM/

longGWAS a method for performing genome-wide association analysis with longitudinal phenotypes. http://genetics.cs.ucla.edu/longGWAS/

MMC Mixed-model coexpression - a method for performing analyzing gene coexpression that is robust to statistical confounding. http://genetics.cs.ucla.edu/mmc/