Chaolong Wang, Ph.D.

Curriculum Vitae updated on May 11, 2015

Contact Information

Genome Institute of Singapore
60 Biopolis Street, Genome 02-01
Singapore 138672, Singapore

Phone (office): (+65) 6808 8341
Email: wangcl@gis.a-star.edu.sg
Website: `http://chaolongwang.github.io`

Personal Information

Citizenship: People's Republic of China
Languages: English, Chinese Mandarin, Cantonese, Hokkien (Min Nan)

Research Interest

Population genetics, statistical genetics/genomics, disease gene mapping, integrative genomics, next-generation sequencing data analysis, high dimensional data analysis

Professional Appointments

01/2015-Now: **Principal Investigator**, Computational & Systems Biology
Genome Institute of Singapore, A*STAR, Singapore

09/2012-12/2014: **Research Fellow**, Department of Biostatistics
Harvard University, Boston, Massachusetts, USA
- Postdoctoral research on statistical genetics with *Dr. Xihong Lin*

Education

08/2012: **Ph.D. in Bioinformatics**
University of Michigan, Ann Arbor, Michigan, USA
- Ph.D. advisor: *Dr. Noah Rosenberg* (at Stanford University since 07/2011)
- Thesis: Statistical methods for analyzing human genetic variation in diverse populations
- 01/2012-08/2012: "Post"-doctoral research on statistical genetics with *Dr. Gonçalo Abecasis*

04/2011: **M.A. in Statistics** and **M.S. in Bioinformatics**
University of Michigan, Ann Arbor, Michigan, USA

07/2008: **B.S. in Physics**
Peking University, Beijing, China

Selected Awards

- Charles J. Epstein Trainee Award semifinalist, American Society of Human Genetics, 2013
- Stellar Abstract Award in the 6th Annual PQG Conference, Harvard University, 2012
- HHMI International Student Research Fellowship, Howard Hughes Medical Institute, 2011-2012
- Rackham Predoctoral Fellowship, University of Michigan, 2011-2012 (declined)
- DeLill Nasser Award, Genetics Society of America, 2011
- Departmental Fellowship (Biomedical Sciences), University of Michigan, 2008
- May Fourth Scholarship, Peking University, 2006
- Dean's List, Hong Kong University of Science & Technology, 2005

Teaching Experience

06/2014 & 12/2014: Sequence Analysis Workshop Instructor, University of Michigan
Lecture and hands-on practical on Estimates of Genetic Ancestry

10/2013: PQG Short Course Lecturer, Harvard School of Public Health
Statistical methods for ancestry inference with applications to disease gene mapping

01/2011-04/2011: Graduate Student Instructor, University of Michigan
BIOSTAT 646 - High throughput molecular genomic and epigenomic data analysis

Manuscripts submitted

H Chen*, **C Wang***, MP Conomos, AM Stilp, Z Li, T Sofer, AA Szpiro, W Chen, JM Brehm, JC Celedon, SS Redline, GJ Papanicolaou, TA Thornton, CC Laurie, K Rice, X Lin. Control for population structure and relatedness for binary traits in genetic association studies using logistic mixed models. Submitted. (* Co-first author)

X Lin, S Lee, M Wu, **C Wang**, H Chen, Z Li, X Lin. Test for rare variants by environment interactions in sequencing association studies. In revision.

Publications

C Wang, X Zhan, L Liang, GR Abecasis, X Lin (2015). Improved ancestry estimation for both genotyping and sequencing data using projection Procrustes analysis and genotype imputation. *American Journal of Human Genetics*, in press.

C Wang*, X Zhan*, J Bragg-Gresham, HM Kang, D Stambolian, E Chew, K Branham, J Heckenlively, The FUSION Study, RS Fulton, RK Wilson, ER Mardis, X Lin, A Swaroop, S Zöllner, GR Abecasis (2014). Ancestry estimation and control of population stratification for sequence-based association studies. *Nature Genetics*, 46:409-415. (* Co-first author)

X Zhan*, DE Larson*, **C Wang***, DC Koboldt, Y Sergeev, 52 other coauthors, ER Mardis, A Swaroop, GR Abecasis (2013). Identification of a rare coding variant in Complement 3 associated with age-related macular degeneration. *Nature Genetics* **45**: 1375-1379. (* Co-first author)

C Wang, KB Schroeder, NA Rosenberg (2012). A maximum-likelihood method to correct for allelic dropout in microsatellite data with no replicate genotypes. *Genetics* **192**: 651-669.

C Wang, S Zöllner, NA Rosenberg (2012). A quantitative comparison of the similarity between genes and geography in worldwide human populations. *PLoS Genetics* **8**: e1002886. [Featured in *Science* 337: 1151, 2012]

TJ Pemberton, **C Wang**, JZ Li, NA Rosenberg (2010). Inference of unexpected genetic relatedness among individuals in HapMap Phase III. *American Journal of Human Genetics* **87**: 457-464. [Featured in *Am. J. Hum. Genet.* 87: 447-448, 2010 and *Genetics* 186(2): NP, 2010]

C Wang, ZA Szpiech, J Degnan, M Jakobsson, TJ Pemberton, JA Hardy, AB Singleton, NA Rosenberg (2010). Comparing spatial maps of human population-genetic variation using Procrustes analysis. *Statistical Applications in Genetics and Molecular Biology* **9**: 13.

JT Mosher, TJ Pemberton, K Harter, **C Wang**, EO Buzbas, P Dvorak, C Simon, SJ Morrison, NA Rosenberg (2010). Lack of population diversity in commonly used human embryonic stem-cell lines. *New England Journal of Medicine* **362**: 183-185. [Featured in *Nature* 462: 945, 2009]

NM Kopelman, L Stone, **C Wang**, D Gefel, MW Feldman, J Hillel, NA Rosenberg (2009). Genomic microsatellites identify shared Jewish ancestry intermediate between Middle Eastern and European populations. *BMC Genetics* **10**: 80.

L Huang, **C Wang**, NA Rosenberg (2009). The relationship between imputation error and statistical power in genetic association studies in diverse populations. *American Journal of Human Genetics* **85**: 692-698. [Featured in *Am. J. Hum. Genet.* 85: 539-540, 2009 and *Nat. Rev. Genet.* 10: 817, 2009]

CL Wang, KW Au, CK Chan, HW Lau, KY Szeto (2008). Detecting hierarchical organization in complex networks by nearest neighbor correlation. *Studies in Computational Intelligence* **129**: 487-494 (Conference Proceedings of NICSO 2007).

Software Developed

LASER: a package written in C++ for estimating individual ancestry using either sequencing reads or gentyping data. (Wang *et al.* 2014, *Nat. Genet.*; Wang *et al.* 2015, *AJHG*)

MicroDrop: a C++ program for estimating and correcting for allelic dropout in microsatellite data without replicated genotyping. (Wang *et al.* 2012, *Genetics*).

Conference Presentations

C Wang, Han Chen, X Lin. Control of population stratification in family data using pedigree information and ancestry principal components. *The American Society of Human Genetics 64th Annual Meeting*, San Diego, USA, October 2014. **(Poster)**

C Wang, L Liang, GR Abecasis, X Lin. Tracing individual ancestry in a principal components space. *The American Society of Human Genetics 63rd Annual Meeting*, Boston, USA, October

2013. **(Platform talk, Trainee Award Semifinalist)**

C Wang, X Zhan, S Zöllner, GR Abecasis. Estimating individual ancestry using next generation sequencing. *CSHL Meeting "The Biology of Genomes"*, Cold Spring Harbor, New York, USA, May 2013. **(Poster)**

C Wang, X Zhan, S Zöllner, GR Abecasis. Estimating individual ancestry using next generation sequencing. *HHMI Janelia Conference "Biological Sequence Analysis and Probabilistic Models"*, Ashburn, Virginia, USA, March 2013. **(Poster and blitz talk)**

C Wang, X Zhan, S Zöllner, GR Abecasis. Estimating individual ancestry from extremely low-coverage sequencing data. *The 6th Annual PQG Conference "Sequencing and Complex Traits: beyond 1000 Genomes"*, Boston, Massachusetts, USA, November 2012. **(Platform talk, Stellar Abstract Award)**

C Wang, NA Rosenberg. A quantitative comparison of the similarity between genes and geography in worldwide human populations. *The 12th International Congress of Human Genetics & The American Society of Human Genetics 61st Annual Meeting*, Montreal, Canada, October 2011. **(Poster)**

C Wang, KB Schroeder, NA Rosenberg. A maximum likelihood genotype imputation method to correct for allelic dropout in microsatellite data. *ENAR 2011 Spring Meeting*, Miami, Florida, USA, March 2011. **(Platform talk)**

C Wang, KB Schroeder, NA Rosenberg. A genotype imputation method for allelic dropout in microsatellite data. *The American Society of Human Genetics 60th Annual Meeting*, Washington DC, USA, November 2010. **(Poster)**

C Wang, NA Rosenberg. Comparing spatial maps of genetic variation and geographic locations using Procrustes analysis. *EMBO/EMBL Symposium "Human Variation: Cause and Consequence"*, Heidelberg, Germany, June 2010. **(Poster)**

C Wang, ZA Szpiech, J Degnan, M Jakobsson, TJ Pemberton, JA Hardy, AB Singleton, NA Rosenberg. Comparing spatial maps of population-genetic variation using Procrustes analysis. *The American Society of Human Genetics 59th Annual Meeting*, Honolulu, Hawaii, USA, October 2009. **(Poster)**

Membership

2009-Now: American Society of Human Genetics
2010-2011: American Statistical Association
2010-2011: International Biometric Society, ENAR

Professional Activities

2013-2014: **Member**, Organizing Committee of the HHMI Alumni Network in Boston Region
2012-2013: **Organizer**, PQG Short Courses, Harvard School of Public Health
2009-Now: **Referee** for peer-reviewed journals, including *Annals of Human Geneitcs*, *BMC Bioinformatics*, *BMC Genetics*, *Clinical Chemistry*, *Genetics*, *Human Biology*, *Journal of Human Genetics*, *Molecular Biology and Evolution*, *PLoS ONE*, *PNAS*, *Scientific Reports*

Computing Skills

Programming: C/C++, Perl, R, MATLAB, Shell Scripting
Operating Systems: Linux/Unix, Mac OS X, Microsoft Windows
Applications: LaTeX, Microsoft Office, Adobe Illustrator, HTML

References

Available upon request.