Chaolong Wang, Ph.D.

Curriculum Vitae updated on September 24, 2017

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`

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

05/2016-Now: **Adjunct Assistant Professor**, Centre for Computational Biology
Duke-NUS Medical School, Singapore

09/2012-12/2014: **Research Fellow**, Department of Biostatistics
Harvard T.H. Chan School of Public Health, Harvard University, Boston, Massachusetts, USA
- Postdoctoral research on statistical genetics with *Xihong Lin* and *Liming Liang*

Education

08/2012: **Ph.D. in Bioinformatics**
University of Michigan, Ann Arbor, Michigan, USA
- Ph.D. advisor: *Noah Rosenberg* (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 *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
- Fellowship (Program in Biomedical Sciences), University of Michigan, 2008-2009
- May Fourth Scholarship, Peking University, 2006

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

Publications

$^{\#}$ indicates corresponding author; * indicates co-first author.

J Dou*, B Sun*, X Sim, JD Hughes, DF Reilly, ES Tai, J Liu, **C Wang**$^{\#}$. Estimation of kinship coefficient in structured and admixed populations using sparse sequencing data. *PLOS Genetics*. (in press)

J Liu, X Wan, **C Wang**, C Yang, X Zhou, C Yang (2017). LLR: A latent low-rank approach to colocalizing genetic risk variants in multiple GWAS. *Bioinformatics*. (in press)

H Liu, Z Wang, Y Li, G Yu, X Fu, ..., **C Wang**, ..., S Chen, J Liu, F Zhang (2017). Genome-wide analysis of protein-coding variants in leprosy. *Journal of Investigative Dermatology*. (in press)

Z Wang, BC Henn, **C Wang**, Y Wei, L Su, R Sun, H Chen, PJ Wagner, Q Lu, X Lin, R Wright, D Bellinger, M Kile, M Mazumdar, MM Tellez-Rojo, L Schnaas, DC Christiani (2017). Genome-wide gene by lead exposure interaction analysis identifies UNC5D as a candidate gene for neurodevelopment. *Environmental Health*, **16**: 81.

D Taliun[#], S Chothani, S Schönherr, L Forer, M Boehnke, GR Abecasis, **C Wang**[#] (2017). LASER server: ancestry tracing with genotypes or sequence reads. *Bioinformatics*, **33**: 2056-2058.

W Dai, M Yang, **C Wang**, T Cai (2017). Sequence robust association test (SRAT) for familial data. *Biometrics*, **doi**: 10.1111/biom.12643. (in press)

X Wang, Z Zhang, M Nathan, T Cai, S Lee, **C Wang**, TW Yu, CA Walsh, X Lin (2017). Rare variant association test in family based sequencing studies. *Briefings in Bioinformatics*, **doi**: 10.1093/bib/bbw083. (in press)

BE Cade, H Chen, AM Stilp, KJ Gleason, T Sofer, ..., **C Wang**, PC Zee, CL Hanis, SR Sunyaev, SR Patel, CC Laurie, X Zhu, R Saxena, X Lin, S Redline (2016). Genetic Associations with Obstructive Sleep Apnea Traits in Hispanic/Latino Americans. *American Journal of Respiratory and Critical Care Medicine*, **194**: 886-897.

H Chen[*], **C Wang**[*], MP Conomos, AM Stilp, Z Li, T Sofer, AA Szpiro, W Chen, JM Brehm, JC Celedon, S Redline, GJ Papanicolaou, TA Thornton, CC Laurie, K Rice, X Lin (2016). Control for population structure and relatedness for binary traits in genetic association studies via logistic mixed models. *American Journal of Human Genetics*, **98**: 653-666.

X Lin, S Lee, M Wu, **C Wang**, H Chen, Z Li, X Lin (2016). Test for rare variants by environment interactions in sequencing association studies. *Biometrics*, **72**: 156-164.

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*, **96**: 926-937.

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.

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.

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 & *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

SEEKIN: a C++ program for estimating pairwise genetic relatedness in structured and admixed populations using sparse sequencing data. (Dou *et al.* 2017, *PLOS Genet*)
LASER: a package written in C++ for estimating individual ancestry using either sequencing reads or genotyping data. (Wang *et al.* 2014, *Nat Genet*; Wang *et al.* 2015, *Am J Hum Genet*)
LASER Server: a web server based on the LASER method to estimate individual ancestry. (`https://laser.sph.umich.edu`; Taliun *et al.* 2017, *Bioinformatics*)
GMMAT: an R package to perform efficient genome-wide association tests based on generalized linear mixed models. (Chen*, Wang* *et al.* 2016, *Am J Hum Genet*)
MicroDrop: a C++ program for estimating and correcting for allelic dropout in microsatellite data without replicated genotyping. (Wang *et al.* 2012, *Genetics*)

Membership

American Society of Human Genetics (ASHG)
International Genetic Epidemiology Society (IGES)

Professional Activities

2015-Now: **Member**, Committee for the GIS Research Pipeline Development Team
2013-2014: **Member**, Organizing Committee of the HHMI Alumni Network in Boston Region
2012-2013: **Organizer**, PQG Short Courses, Harvard School of Public Health
2017: **Reviewer** for Student Paper Award Competition, American Statistical Association (ASA) Section on Statistics in Genomics and Genetics
2009-Now: **Reviewer** for *Annals of Human Genetics*, *Bioinformatics*, *Biometrics*, *BMC Bioinformatics*, *BMC Genetics*, *Clinical Chemistry*, *Gene*, *Genetics*, *Genome Research*, *Human Biology*, *Human Heredity*, *Journal of Human Genetics*, *Molecular Biology and Evolution*, *Molecular Ecology*, *PLoS ONE*, *Proceedings of the National Academy of Sciences (PNAS)*, *Scientific Reports*