Chaolong Wang

CONTACT INFORMATION	Harvard School of Public Health Bldg. 2-451 Biostatistics 655 Huntington Avenue Boston, MA 02115, USA	*Office Phone*: (617) 432-4918 *Email*: chaolong@umich.edu *Email*: cwang@hsph.harvard.edu `http://www.umich.edu/~chaolong`

EDUCATION

August 2012: **Ph.D. in Bioinformatics**
University of Michigan, Ann Arbor, Michigan, USA
- Committee co-chairs: *Drs. Noah Rosenberg* and *Michael Boehnke*
- Research areas: population genetics, statistical genetics

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

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

September - December 2005: Exchange Student, Physics
Hong Kong University of Science & Technology, Hong Kong, China

HONORS AND AWARDS

From 2003 to Present
- Stellar Abstract Award in the 6th Annual PQG Conference, Program in Quantitative Genomics, Harvard University, 2012
- International Student Research Fellowship, Howard Hughes Medical Institute, 2011-2012 ($43,000/year for Ph.D. training)
- Rackham Predoctoral Fellowship, University of Michigan, 2011-2012 (declined)
- DeLill Nasser Award for Professional Development in Genetics, Genetics Society of America, 2011
- Rackham Conference Travel Grant, University of Michigan, 2009, 2010, 2011
- Departmental Fellowship (Biomedical Sciences), University of Michigan, 2008
- Second Prize in the Contemporary Undergraduate Mathematical Contest in Modeling, China Society for Industrial and Applied Mathematics, 2007
- Second Prize in the Undergraduate Contest of Mathematical Modeling and Computer Application, Peking University, 2007
- Award for Social Work, Peking University, 2007
- May Fourth Scholarship, Peking University, 2006
- Dean's List, Hong Kong University of Science & Technology, 2005
- Distinguished Freshman Scholarship, Peking University, 2004
- Third Prize in the 20th Chinese Physics Olympiad National Final Contest, 2003

TEACHING EXPERIENCE

Winter 2011, Graduate Student Instructor, University of Michigan, Ann Arbor
 BIOSTAT 646 - High throughput molecular genomic and epigenomic data analysis

RESEARCH EXPERIENCE

September 2012 - Present: Biostatistics, Harvard University, Boston
 Postdoctoral research fellow, under *Drs. Xihong Lin* and *Liming Liang*
 Statistical methods for genetic mapping of complex traits and disease

February 2012 - August 2012: Bioinformatics, University of Michigan, Ann Arbor
 Research, under *Drs. Gonçalo Abecasis* and *Sebastian Zöllner*
 Estimating individual ancestry using next generation sequencing

January 2009 - August 2012: Bioinformatics, University of Michigan, Ann Arbor
Ph.D. dissertation research, under *Dr. Noah Rosenberg*
Statistical methods for analyzing human genetic variation in diverse populations

September 2008 - December 2008: Bioinformatics, University of Michigan, Ann Arbor
Research rotation, under *Dr. Xiaoxia (Nina) Lin*
Mathematical modeling of the genetic regulatory network of *G1-S* phase transition in yeast cell cycle

July - August 2006, 2007, 2008: Physics, Hong Kong University of Science & Technology
Junior research assistant, under *Dr. Kwok-Yip Szeto*
Topology and dynamics of complex networks, bioinformatics

September 2007 - June 2008: Peking University, Beijing
Bachelor thesis research, under *Fangting Li*
Modeling positive feedbacks in transcriptional regulation of genetic networks

PUBLICATIONS   **C Wang***, X Zhan*, J Bragg-Gresham, D Stambolian, E Chew, K Branham, J Heckenlively, RS Fulton, RK Wilson, ER Mardis, A Swaroop, S Zöllner, GR Abecasis (2013). Estimating individual ancestry using next generation sequencing. Under review by *Nature Genetics*. (* Joint first author)

X Zhan*, DE Larson*, **C Wang***, DC Koboldt, Y Sergeev, RS Fulton, LL Fulton, CC Fronick, K Branham, J Bragg-Gresham, G Jun, Y Hu, HM Kang, D Liu, M Othman, M Brooks, R Ratnapryia, A Boleda, RK Wilson, J Heckenlively, E Chew, D Stambolian, ER Mardis, A Swaroop, GR Abecasis (2013). Targeted sequencing, augmented with public resources, identifies a rare complement 3 allele associated with AMD. Under review by *Nature Genetics*. (* Joint 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 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.

SOFTWARE	*TRACE* (fasT and Robust Ancestry Coordinate Estimation)

TRACE (fasT and Robust Ancestry Coordinate Estimation)
This program uses SNP genotypes to *trace* individual ancestry in a PCA space constructed by a reference sample set. (Wang *et al.* 2013, in preparation).

LASER (Locating Ancestry from SEquence Reads)
This program estimates individual ancestry by directly analyzing shortgun sequence reads without calling genotypes (Wang *et al.* 2013, submitted to *Nature Genetics*).

MicroDrop
This program can estimate and correct for allelic dropout in microsatellite data without requiring replicate genotypes (Wang *et al.* 2012, *Genetics* **192**: 651-669).

INVITED TALKS   May 2012, Department of Biology, Stanford University, Stanford, California.

May 2012, Center for Theoretical Evolutionary Genomics, University of California, Berkeley, California.

March 2012, Department of Biostatistics, Harvard University, Boston, Massachusetts.

March 2012, Center for Computational Molecular Biology, Brown University, Providence, Rhode Island.

February 2012, Center for Statistical Genomics and Proteomics, Yale University, New Haven, Connecticut.

CONFERENCE
PRESENTATIONS

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)**

TJ Pemberton, **C Wang**, JZ Li, NA Rosenberg. Inference of unexpected genetic relatedness among HapMap Phase III individuals. *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)**

L Huang, **C Wang**, NA Rosenberg. Evaluating the relationship between imputation error and statistical power in genetic association studies. *The American Society of Human Genetics 59th Annual Meeting*, Honolulu Hawaii, USA, October 2009. **(Poster)**

MEMBERSHIP

Member, American Society of Human Genetics
Member, American Statistical Association
Member, International Biometric Society, ENAR
Associate, Program in Quantitative Genomics, Harvard School of Public Health

PROFESSIONAL ACTIVITIES

Organizer, PQG Short Courses, Harvard School of Public Health, 2012-2013
Committee member, HHMI Alumni Network in Boston Region, 2013
Reviewer for peer-reviewed journals, including *Annals of Human Geneitcs, BMC Bioinformatics, Genetics, Human Biology, Molecular Biology and Evolution, PLoS ONE*

COMPUTING SKILLS

Programming Languages: C/C++, Perl, R, MATLAB, Shell Scripting
Operating Systems: Linux/Unix, Mac OS X, Microsoft Windows
Applications: LaTeX, Microsoft Office, Adobe Illustrator, HTML
Genetics Software: PLINK, GATK, SAMtools, IGV, ADMIXTURE, CLUMPP, ms