

CURRICULUM VITAE

Name: Di WU
Position: Assistant Professor
Current Address: Department of Periodontology, School of Dentistry
Department of Biostatistics, Gillings School of Global Public Health
UNC-Chapel Hill, Carolina Health Informatics Program
Koury Oral Health Sciences Building, Room 4504
385 S. Columbia Street, CB#7455
Chapel Hill, NC 27599-7455, USA
Telephone: + 9195373277 (office)
E-mail: dwu@unc.edu
URL: <http://statistics.fas.harvard.edu/people/di-wu>

RESEARCH INTEREST:

Statistical bioinformatics and biostatistics for preprocess and integration of high dimensional bio-medical data. Including microarray data and sequencing data analysis, pathway analysis, GWAS, eQTL and GTEx data analysis, epigenomics, system biology, gene set tests, gene expression pattern discovery, big data integration, classification, linear model, empirical Bayes approach, random effects models, Bayesian approach, MCMC, statistical bioinformatics software, breast cancer, lung cancer, prostate cancer, stem cell research, autoimmune diseases, oral biology, drug repurposing, health informatics, microbiome data analysis

EDUCATION:

- Ph.D.: Bioinformatics Division, Walter and Eliza Hall Institute of Medical Research,
University of Melbourne, VIC, Australia, Feb 2007- Jan 2011
Supervisors: Professor Gordon Smyth and Professor Terry Speed
Topic: Finding hidden relationships between gene expression profiles with application to breast cancer biology
- M.S.: Biostatistics Division, Department of Epidemiology and Biostatistics, School of Medicine,
Case Western Reserve University, Cleveland, OH, USA, July 2003~May 2006
Topic: Interaction of drug (Methotrexate/Mercaptopurine) responses on gene expression profiles of acute lymphoblastic leukemia cells
Supervisor: A/Professor Tomas Radivoyevitch
- B.S.: Department of Biotechnology,
School of Life Science & Biotechnology, Shanghai Jiao Tong University,
Shanghai, China, Aug1994 ~ June 1998

WORKING EXPERIENCE:

April 2015~ Now
Title: Assistant Professor

Department of Periodontology, UNC-Chapel Hill School of Dentistry
Joint with Department of Biostatistics, UNC Gillings School of Global Public Health, Carolina
Health Informatics Program

April 2011~ April 2015

Title: Postdoctoral Fellow

Statistics Department, Harvard University, Jun S. Liu group

Biostatistics, Dana Farber Cancer Institute, Rafael Irizarry group

Division of genetics, Harvard Medical School, Brigham and Women's Hospital,
Robert Plenge group

Affiliated to Broad Institute of Harvard and MIT

Projects: Gene-gene interaction in GWAS and eQTL data, Bayesian Gene set test, small RNA array data normalization, high-throughput drug screen, drug repurposing, disease risk prediction

March 2007 ~ Jan 2011 (casual)

Title: Biostatistician

Center for Cancer Research,

Monash Institute of Medical Research, Australia

Projects: microarray data analysis of Intrauterine Growth Restriction (IUGR) project, aCGH data analysis, CpG Methylation array data analysis, miRNA array data analysis, limiting dilution data analysis, a random effect model of cell migration data, sample size calculation, longitudinal data analysis

June 2006 ~ February 2007

Title: Biostatistician

Center for Cancer Research,

Monash Institute of Medical Research, Australia

Projects: cost effectiveness analysis using WINBUGS and STATA, survival data analysis, microarray data analysis in interferon treated renal cell carcinoma (RCC) and interferon treated macrophage, combining array CGH data and gene expression data, associating breast cancer expression data with clinical outcomes

January 2004 ~ June 2006

Title: Research Assistant 3

Department of Pathology, Case Western Reserve University, Cleveland, OH, US

Supervised by Professor Alan Tartakoff

Project: Single promoting/inhibiting target identification, experimental design and microarray data analysis

June 2003 ~ January 2004

Title: Research Assistant 2

Case Comprehensive Cancer Center, Case Western Reserve University, Cleveland, OH, US

Supervised by Dr. Mark Jackson

Project: Signal transduction in breast cancer

June 2001 ~ June 2003

Title: Research Assistant 2, Research Assistant 1,
Department of Pathology, Case Western Reserve University, Cleveland, OH, US
Supervised by Dr. Qingzhong Kong

Project: Sporadic prion disease study

June 1998 ~ March 2001

Title: Research Assistant,
Shanghai Institute of Biochemistry, Chinese Academy of Science

Project: Gene therapy

TEACHING AND MENTORSHIP:

August 2016 ~ now

Academic mentor for Public Health Informatics (PHI) master student through Carolina Health Informatics Program (CHIP), UNC

2015~Now

MS Dissertation Committee, Oral microbiome changes associated with fixed prosthodontic restoration, Sarah Kay Youny Lee, Dental School, UNC

2015

Direction of Master's Research, The analgesic effect of a new ibuprofen formulation on odontogenic pain, Tanjit S. Taggar, co-advise with Asma Khan, Dental School, UNC

2016

PhD Dissertation Committee, Doug Wilson, advisor Wei Sun and Joe Ibrahim, Biostatistics Department UNC

September 2012 ~ April 2015

Mentor graduate students in the program "Harvard Graduate Women in Science and Engineering"

July 2011 ~ September 2011

Project: Lung cancer subtype classification using gene set tests
Supervise undergraduate summer student from University of Washington
Statistics Department, Harvard University, Cambridge, MA, USA

February 2010 ~ May 2010

Demonstrator of high school students in Victoria
Gene Technology Access Centre
University High School, Parkville, VIC, Australia

February 2008 ~ June 2008

Tutor for MPH students
Subject: 505-101 Introduction to Biostatistics
Melbourne School of Population Health, University of Melbourne, VIC, Australia

July 2005 ~ June 2006

Supervise undergraduate students
Department of Pathology
Case Western Reserve University, Cleveland, OH, USA

AWARDS AND SCHOLARSHIPS:

2012: Travel Award, Calls for Talents, Shandong Province, China
2011: 2010 Chinese Government Award for Outstanding Self-Financed students Abroad (\$5,000)
2010: Edith Moffat Overseas Travel Scholarship, Walter and Eliza Hall Institute
2007~2010: Australian Postgraduate Award, University of Melbourne
2009: Student Travel Bursary, AMATA Conference, Sydney, Australia
2007: Student Travel Bursary, MGED and AMATA Conference, Brisbane, Australia
2007: Travel Scholarships, ICE-EM Australian Graduate School in Mathematics
1995: Scholarship for Excellence, Shanghai Jiao Tong University
1994: The Shu-Ermer Scholarship for Excellence, Shenyang

PROGRAMMING AND OPERATING SYSTEMS:

R, S-plus, SAS, SPSS, STATA, WINBUGS
C++, Java, Fortran, Python, Perl

SOFTWARE DEVELOPMENT:

roast function for the rotation gene set tests in limma package
camera function for the correlation adjusted mean rank gene set analysis in limma package

JOURNAL REFEREE:

Biometrics
Nucleic Acids Research
PLoS Computational Biology
Bioinformatics
Statistics in Medicine
Briefings in Bioinformatics
BMC Bioinformatics

Statistics in Biosciences
PLoS One
'Genomics, Proteomics and bioinformatics-Elsevier'
Biomarkers in Cancer
Evolutionary Bioinformatics
IEEE/ACM Transactions on Computational Biology and Bioinformatics
Combinatorial Chemistry & High Throughput Screening
International Journal of Dentistry
Lead Guest Editor of Cancer Informatics journal supplement (Libertas Academica)
JASA

GRANT REVIEW:

Australian National Health and Medical Research Council (2014)

COMMITTEE SERVICE

The UNC Biostatistics Computing Committee

SOCIETY MEMBERSHIP

2015-present American Statistical Association

GRANTS AND FELLOWSHIPS:

2011: PI, Australian National Health and Medical Research Council (NHMRC) Early Career Overseas

Fellowship, for 4 years since 2012 March, total \$340,000

2015: co-Investigator with Asma Khan. Project Sponsor: American Association of Endodontists (AAE)

RAMSeS Project Number: A16-0983. Project Title: Molecular Markers in the Gingival Crevicular Fluids of Teeth with Normal and Inflamed Pulp

PUBLICATIONS:

1. Li G, Cunin P, **Wu D**, Diogo D, Yang Y, Okada Y, Plenge R, Nigrovic P. 2016. The Rheumatoid Arthritis Risk Variant CCR6DNP Regulates CCR6 via PARP-1. *PLoS Genet* 12(9): e1006292.
2. Morelli T, Moss K, Beck J, Preisser JS, **Wu D**, Divaris K, Offenbacher S. Derivation and Validation of the Periodontal and Tooth Profile Classification System for Patient Stratification. *Journal of Periodontology*. Accepted on 09-Aug-2016.
3. Leong YA, Chen Y, Ong HS, **Wu D**, et al, Belz GT, Lewin SR, Kallies A, Yu D. 2016. CXCR5+ follicular cytotoxic T cells control viral infection in B cell follicles. Accepted. *Nature Immunology*, July 29, 2016, doi:10.1038/ni.3543

4. Sun X, Dalpiaz D, **Wu D**, Liu JS, Zhong W, Ma P. 2016. Statistical Inference for Time Course Statistical inference for time course rna-seq data using a negative binomial mixed-effect model. *BMC Bioinformatics*. DOI 10.1186/s12859-016-1180-9
5. Zhang S, Divaris K, Moss K, Yu N, Barros S, Marchesan J, Morelli T, Agler C, Kim SJ, **Wu D**, North KE, Beck J, and Offenbacher S. 2016. The Novel ASIC2 Locus Is Associated with Severe Gingival Inflammation, *JDR Clinical and Translational Research*. Published online before print April 20, 2016, doi: 10.1177/2380084416645290
6. **Wu D**, Gantier MP. Normalization of Affymetrix miRNA Microarrays for the Analysis of Cancer Samples. *Methods in molecular biology (Clifton, N.J.)* 05/2015; DOI:10.1007/7651_2015_239
7. Xu H, Xiao T, Chen CH, Li W, Meyer CA, Wu Q, **Wu D**, Cong L, Zhang F, Liu JS, Brown M, Liu XS. 2015. Sequence determinants of improved CRISPR sgRNA design. *Genome Research* 06/2015; DOI:10.1101/gr.191452.115
8. Worley MJ, Liu S, Hua Y, et al., **Wu D** et al., Ng SW. 2015. Molecular Changes in Endometriosis-Associated Ovarian Clear Cell Carcinoma. *European journal of cancer* 06/2015; 360. DOI:10.1016/j.ejca.2015.05.011
9. Ng AP, Hu Y, Metcalf D, Hyland CD, Ierino H, Phipson B, **Wu D**, et al. 2015. Early Lineage Priming by Trisomy of Erg Leads to Myeloproliferation in a Down Syndrome Model. *PLoS Genet* 11(5): e1005211
10. Ritchie ME, Phipson B, **Wu D**, Hu Y, Law CW, Shi W, Smyth G. 2015. Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acid Res.* 01/2015; 43(7). DOI:10.1093/nar/gkv007
11. Tartakoff AM and **Wu D**. 2014. The Axis of Progression of Disease. *Cancer Informatics*. 2014;13(S6). (<http://www.la-press.com/the-axis-of-progression-of-disease-article-a4464>)
12. Okada Y, **Wu D**, Trynka G, et al., Plenge R. 2013. Genetics of rheumatoid arthritis contributes to biology and drug discovery. *Nature*. doi:10.1038/nature12873. Epub 2013 Dec 25
13. **Wu D**, Pang Y, Wilkerson MD, Wang D, Hammerman PS and Liu JS. 2013. Gene expression data integration for squamous cell lung cancer subtypes reveals drug sensitivity. *British Journal of Cancer*. doi: 10.1038/bjc.2013.452. Epub 2013 Sep 3.
14. **Wu D**, Hu Y, Tong S, Williams BRG, Smyth GK and Gantier MP. 2013. The use of miRNA microarrays for the analysis of cancer samples with global miRNA decrease. *RNA*. 19(7):876-88.
15. Li G, Diogo D, **Wu D**, et al, Plenge PM. Human genetics guides a high-throughput drug screen of the CD40 signalling pathway in rheumatoid arthritis. *PLoS Genetics*. May;9(5):e1003487. doi: 10.1371/journal.pgen.1003487. Epub 2013 May 16.
16. Woollard D, Opekin K, Coso S, **Wu D**, Baldwin M, Williams E. 2012. Differential expression of VEGF ligands and receptors in prostate cancer. *The Prostate*. doi: 10.1002/pros.22596. Article first published online: Oct 4 2012
17. **Wu D**, Smyth GK. 2012. CAMERA: a competitive gene set test accounting for inter-gene correlation. Online method. *Nucleic Acid Res*. doi: 10.1093/nar/gks461, First published online: May 25, 2012
18. **Wu D**, Wang D. 2012. Application of Advanced Gene Set Tests in Breast Cancer Research. *Chinese Journal of Bioinformatics*. 10(2):92-95
19. Asselin-Labat ML, Sutherland KD, Vaillant F, Gyorki DE, **Wu D**, Holroyd S, Breslin K, Ward T, Shi W, Bath ML, Deb S, Fox SB, Smyth GK, Lindeman GJ, Visvader JE. 2011. Gata-3 negatively regulates the tumor-initiating capacity of mammary luminal progenitor cells and targets the putative tumor suppressor caspase-14. *Mol Cell Biol*. 31(22):4609-4622

13. **Wu D**, Lim E, Vaillant F, Asselin-Labat ML, Visvader JE, Smyth GK. 2010. ROAST: rotation gene set tests for complex microarray experiments. *Bioinformatics*. 26(17):2176-82
14. Asselin-Labat ML, Vaillant F, Sheridan J, Pal B, **Wu D**, Simpson ER, Yasuda H, Smyth G, Martin TJ, Lindeman GJ, Visvader JE. 2010. Control of mammary stem cell function by steroid hormone signaling. *Nature*. 465(7299):798-802
 - Accompanying News and Views in Nature
 - Cited in Nature Reviews Cancer and A-IMBN (Nature Asia-Pacific) as a ‘Research Highlight’
13. Lim E*, **Wu D***, Pal B*, Bouras T, Asselin-Labat ML, Vaillant F, Yagita H, Lindeman GJ, Smyth GK, Visvader JE. 2010. Transcriptome analyses of mouse and human mammary cell sub-populations reveals multiple conserved genes and pathways. *Breast Cancer Res*. 12(2):R21
 *these authors contributed equally
 - Accompanying Editorial
 - Faculty 1000 Medicine Recommended
14. Lim E, Vaillant F, **Wu D**, Forrest NC, Pal B, Hart AH, Asselin-Labat ML, Gyorki DE, Ward T, Partanen A, Feleppa F, Huschtscha LI, Thorne HJ, Fox SB, Yan M, French JD, Brown MA, Smyth GK, Visvader JE, Lindeman GJ. 2009. Aberrant luminal progenitors as the candidate target population for basal tumor development in brca1 mutation carriers. *Nat Med*, 15(8):907–13
 - Accompanying News and Views in Nature Medicine, Cell Stem Cell
 - Cited in Nature and Nature Reports Stem Cells as a ‘Research Highlight’
 - Cited ‘Must Read’ by Faculty 1000 Medicine
 - Listed in Nature Medicine’s landmark papers ‘Classic Collection’
15. Gargett CE, Schwab KE, Zillwood RM, Nguyen HPT, **Wu D**. 2009. Isolation and culture of epithelial progenitors and mesenchymal stem cells from human endometrium. *Biol Reprod* 80(6):1136-1145.
 - Cited in Biology of Reproduction as a ‘Highlight’
 - Faculty 1000 Medicine Recommended
16. **Wu D**, Townsley E, Tartakoff AM. 2007. Covert genetic selections to optimize phenotypes. *PLoS ONE*. 2 (11):e1200.
17. Huang S, Liang J, Zheng M, Li X, Wang M, Wang P, Vanegas D, **Wu D**, Chakraborty B, Hays AP, Chen K, Chen SG, Booth S, Cohen M, Gambetti P, Kong Q. 2007. Inducible overexpression of wild-type prion protein in the muscles leads to a primary myopathy in transgenic mice. *Proc Natl Acad Sci* 104(16):6800-5
18. Kong Q, Huang S, Zou W, Vanegas D, Wang M, **Wu D.**, Yuan J, Bai H, Zheng M, Deng H, Chen K, Jenny AL, O’Rourke K, Belay ED, Schonberger LB, Petersen RB, Sy M-S, Chen G, and Gambetti P. 2005. Chronic Wasting Disease of Elk: Transmissibility to humans examined by transgenic mouse models. *J. Neuroscience*. 25(35) :7944-9
19. Kong Q, Huang S, **Wu D.**, Wang M, Vanegas D, Bai H, Deng H, Chen K, Zou W, Jenny AL, O’Rourke K, Sy M, Chen SG, and Gambetti P. 2004. Transmissibility of chronic wasting disease of elk and deer to humans. *J. Neuropathology and Experimental Neurology*. 63: 515.

DRAFTS AVAILABLE:

1. **Wu D**, Hu Y, Ritchie M, Pang Y, Liu JS, Smyth GK. 2016 Rotation Gene Set Enrichment Analysis (ROMER). This method has been used in >5 publications.

PATENTS:

- ‘Gene expression profiles and uses therefor’ WEHI, Geoffrey LINDEMAN, Jane Visvader, Gordon Smyth, Di Wu. PCT number pending. Priority Date: 6th May 2009. Provisional patent Application No. AU2009901989. Human breast stem and luminal progenitor cells

CONFERENCE ORGANIZATION and UNIVERSITY SERVICE:

- Session Chair, International Chinese Statistical Association (ICSA), June 2014. Portland, USA
- Judge for ASHG’s 2014 DNA Day Essay Contest
- Harvard Postdoc Advisory Board, since 2013
- Session Chair, International Chinese Statistical Association (ICSA), June 2012. Boston, USA
- Coordinator, 14th Sino-American Pharmaceutical Professionals Association New England (SAPANE) Annual Conference, June 2012. Boston, USA
- Coordinator, the 3rd Australia-China Biomedical Research Conference (ACABS), Oct 2011. Melbourne, Australia.

INVITED INSTITUTE LEVEL SEMINARS:

- Dental Research Day
- A mixture model for contamination detection in target DNaseq, Nov 2015, UNC Lineberger Comprehensive Cancer Center
- Cancer personalized medicine, UNC health informatics program, Oct 2015
- RNAseq data analysis, UNC genetics, Aug 2015
- Genomic data integration to discover/repurpose drugs for complex diseases, School of Dentistry, University of North Carolina at Chapel Hill September 2014
- Genomic data integration to discover/repurpose drugs for complex diseases, Cleveland Clinic Foundation, June 2014
- Genomic data integration to discover/repurpose drugs for complex diseases, Massachusetts General Hospital, Harvard Medical School, Apr 2014
- Genomic data integration to discover/repurpose drugs for complex diseases, Biostatistics Department, University of North Carolina at Chapel Hill, Feb 2014
- Genomic data based drug discovery/repurposing, Biostatistics Department, MD Anderson Cancer Center, Houston, Feb 2014
- Genomic data based drug discovery/repurposing. Ontario Institute for Cancer Research. University of Toronto. Nov 2013
- Integrating GWAS data with drug information for drug repurposing. Program in Quantitative Genomics, Harvard School of Public Health. Boston. Apr 2013
- Gene set tests in breast cancer, stem cell and drug repurposing. Institute for Stem Cell Biology and Regenerative Medicine, Stanford University, Palo Alto, Jan 2013
- CD40 pathway analysis, Shandong Academy of Sciences, Jinan, China, Nov 2012
- shRNA data analysis using limma. Broad Institute, Project Achilles Group. Feb 2012
- shRNA data analysis using limma. Dana Farber Institute, Matthew Meyerson group. Jan 2012. Boston, USA.
- Gene set testing, stem cells and breast cancer. The high dimension data seminar in Biostatistics Department, Harvard University. Nov 2011, Boston, USA.

- Investigation of gene-gene interaction in GWAS data and eQTL data. WEHI. Oct 2011, Melbourne, Australia.
- Analysis of transcriptional signatures reveals the cell of origin for breast cancer subtypes. National Information and Communications Technology Australia (NICTA). August 13, 2010. Melbourne, Australia
- Analysis of transcriptional signatures reveals the cell of origin for breast cancer subtypes. The Eskitis Institute for Cell and Molecular Therapies. July 1, 2010. Brisbane, Australia
- Analysis of transcriptional signatures reveals the cell of origin for breast cancer subtypes. July - August, 2010. USA (Harvard University, Broad Institute, Dana Farber Cancer Institute, John Hopkins University, MD Anderson Cancer Center, Stanford University)

CONFERENCE PRESENTATIONS:

- **Invited Talk:** A mixture model for contamination detection in target DNaseq, the 3rdTaihu International Statistics Forum. July 2016. Shanghai, China
- **Talk:** ROMER, ranked based rotation gene set test, JSM 2014 - Joint Statistical Meetings. July 2014. Boston, USA
- **Workshop talk:** Introduction to novel gene set tests. Bioconductor Annual Meeting. July 2014. Boston, USA
- **Invited Speaker:** Data integration for drug repurposing. EITC. July 2014. Cambridge, MA
- **Invited Speaker:** Data integration for drug repurposing. Symposium-International Chinese Statistical Association (ICSA). June 2014. Portland, OR
- **Poster:** GTEX retreat, June 2013. Cambridge, MA
- **Invited Speaker:** Novel gene set tests in breast cancer and stem cell research. Symposium-International Chinese Statistical Association (ICSA). June 2012. Cambridge, MA
- **Poster:** Bayesian Gene Set Test. Eastern North American Region/International Biometric Society Meeting (ENAR). April 1 2012. Washington DC USA
- **Invited Speaker:** Gene set testing, stem cells and Breast cancer. EITC-YIC. August 18, 2011, Boston, USA
- **Poster:** ROAST: rotation gene set tests for complex microarray experiments. 18th Annual International Conference on Intelligent Systems for Molecular Biology (ISMB). July 11-13, 2010. Boston, USA
- **Poster:** Patterns across data sets: finding the cell origin of basal-like breast tumours. Bioinformatics Australia. October 29-30, 2009. Melbourne, Australia
- **Poster:** Patterns across data sets: finding the cell origin of basal-like breast tumours. The 9th Annual Australian Microarray and Associated Technologies Association (AMATA-9). October 18-21, 2009. Sydney, Australia
- **Talk:** roast, a gene set testing method for laboratory generated microarray data. The 8th Annual Australian Microarray and Associated Technologies Association (AMATA-8). November 11-15, 2008. Dunedin, New Zealand
- **Talk:** Residual space permutation for gene set testing in designed microarray experiments. Australian Statistical Conference. June 30 to July 3, 2008. Melbourne, Australia
- **Talk:** Gene set enrichment tests in microarray experiments with small sample sizes.

The 10th annual Microarray and Gene Expression Data Society meeting (MGED-10) and the 7th Annual Australian Microarray and Associated Technologies Association (AMATA-7). September 3-5, 2007. Brisbane, Australia

- **Talk:** An empirical Bayes approach used in microarray data analysis.
Graduate Statistics Course of Australian Mathematical Sciences Institute (AMSI).
July, 2007. Brisbane, Australia.

MEETING ATTENDANCE:

- JSM 2015 - Joint Statistical Meetings. July 2015. Seattle, USA
- American Society of Human Genetics Annual Meeting, Nov 2013. Boston, USA.
- The 25th New England Statistics Symposium, April 16, 2011. Connecticut, USA.
- Epigenetics 2009 Australian Scientific Conference, December 1-4, 2009. Melbourne, Australia.