Lingling An

1177 E. 4th ST, Shantz bld, Room 403
University of Arizona
Tucson, AZ, 85721

Phone (520) 621-1248
anling@email.arizona.edu

Homepage: http://cals.arizona.edu/~anling/
Research group: http://cals.arizona.edu/~anling/sbg/research.html

EMPLOYMENT

Associate Professor	Dept. of Ag & Biosystems Engr., Univ. of Arizona	2015-present
Associate Professor	College of Public Health, Univ. of Arizona	2016-present
Assistant Professor	Dept. of Ag & Biosystems Engr., Univ. of Arizona	2008-2015
Faculty member	Graduate Interdisciplinary Program in Statistics	2008-present
Faculty member	Bio5 Institute	2009-present
Research Assistant	Department of Statistics, Purdue University	2002-2008

EDUCATION

Ph.D. in Statistics	Purdue University	8/2008
M.S. in Mathematical Statistics	Purdue University	5/2003

RESEARCH INTERESTS

- Statistical bioinformatics
- Metagenomics
- Next generation sequencing data analysis
- Pattern recognition

REFEREED JOURNAL PUBLICATIONS

2015

1. Ban Y, **An L**, Jiang H. Investigating microbial co-occurrence patterns based on metagenomic compositional data. ***Bioinformatics***. doi: 10.1093/bioinformatics/btv364

2. Sohn M, Du R, **An L***, A robust approach for identifying differentially abundant features in metagenomic samples. ***Bioinformatics*** 31: 2269-75 (*: corresponding author)

3. Drewry J, Choi C, **An L**, Gharagozloo P. A computational fluid dynamics model of algal growth: development and validation. ***Transactions of the American Society of Agricultural and Biological Engineers.*** 58:2, 203-213

4. Yigiter A, Chen J, **An L**, Danacioglu N. An on-line CNV detection method for short sequencing reads. ***Journal of Applied Statistics***. 42:7, 1556–1571

2014

5. Jiang H, **An L**, Baladandayuthapani V, Auer P. (2014) Classification, predictive modelling, and statistical analysis of cancer data. Cancer informatics 01/2014; 13(Suppl 2):1-3. DOI: 10.4137/CIN.S19328
6. Du R, Mercante D, **An L***, Fang Z*. A statistical approach to correcting cross-annotations in a metagenomic functional profile. ***Journal of Biometrics & Biostatistics***. 5:208 (*: co-corresponding author)
7. Pookhao N, Sohn M, Li Q, Jenkins I, Du R, Jiang H, **An L***. A two-stage statistical procedure for feature selection and comparison in functional analysis of metagenomes. ***Bioinformatics*** 31:158-165 (*: corresponding author**)**
8. Sohn M, **An L*,** Pookhao N, Li Q. Accurate Estimation of Genome Relative Abundance for Closely Related Species in a Metagenomic Sample. ***BMC Bioinformatics*** 15:242 (*: corresponding author)
9. **An L***, Pookhao N, Jiang H, Xu J. Statistical approach of functional profiling for a microbial community. ***PLoS ONE*** 9(9): e106588 (*: corresponding author)

2013

10. Piegorsch W$^{\#}$, **An L**$^{\#}$, Wickens A, West W, Peña E, Wu W. Information-theoretic model-averaged benchmark dose analysis in environmental risk assessment. ***Environmetrics*** 24:143-157 ($^{\#}$: co-first authors)
11. Kadiyala V, Patrick N, Mathieu W, Jaime-Frias R, *Pookhao N*., **An L**, Smith C. Class I Lysine Deacetylases Facilitate Glucocorticoid-Induced Transcription. ***Journal of Biological Chemistry*** 288: 28900-12
12. Tamimi E, Murat K, Choi C, **An L**. Analysis of Microclimate Uniformity in a Naturally Vented Greenhouse with High Pressure Fogging System. ***Transactions of the American Society of Agricultural and Biological Engineers*****.** 56(3): 1241-1254

2012

13. Jiang H$^{\#}$, **An L**$^{\#}$, Lin SM, Feng G, Qiu Y. A Statistical Framework for Accurate Taxonomic Assignment of Metagenomic Sequencing Reads. ***PLoS ONE*** 7(10): e46450. doi:10.1371/journal.pone.0046450 ($^{\#}$: co-first authors)
14. **An L** and Doerge RW. Dynamic Clustering of Gene Expression. ***ISRN Bioinformatics***. Vol. 2012, Article ID 537217. Doi:10.5402/2012/537217
15. West W, Piegorsch W, Peña E, **An L**, Wu W, Wickens A, Xiong H, and Chen W. The Impact of Model Uncertainty on Benchmark Dose Estimation. ***Environmetrics***. 23(8): 706–716

2011

16. McDowell E, Kapteyn J, Schmidt A, Li C, Kang J, Descour A, Shi F, Larson M, Schilmiller A, **An L**, Howe G, Jones A, Pichersky E, Soderlund C, Gang D. Comparative analysis of *Solanum* glandular trichome types. ***Plant Physiology,*** 155: 524-539
17. Niu Y, Hao N, and **An L***. Detection of functional rare variants using Group ISIS. ***BMC Proceedings*** **5**(Suppl 9):S108 (*: corresponding author)
18. Zeng L. **An L**, Wu X. Modeling Drug-Carrier Interaction in the Drug Release fromNanocarriers. ***Journal of Drug Delivery***. Volume 2011, Article ID 370308

2010

19. Long A, Mahapatra C, Leung H, Woodruff E, Rohrbough J, **An L**, Doerge RW, Metzstein M, Shino S, Pak W, Broadie K. The nonsense-mediated decay (NMD) pathway is critical for maintaining synapse architecture and synaptic vesicle cycle efficacy. ***Journal of Cell Science**,* 123: 3303-3315. PMID: 20826458

20. Riddle N, Jiang H, **An L**, Doerge RW, Birchler J. Gene expression analysis at the intersection of ploidy and hybridity in maize. ***Theoretical and Applied Genetics**.* 120(2):341-353.

21. Story D, Kacira M, Kubota C, Akoglu A, **An L**. Plant Nutrient Deficiency Detection using Automated Morphology Based Sensing in Controlled Environments. ***Computers and Electronics in Agriculture**,* 74: 238-243

2009

22. Liu S, Kim H, Chen J, and **An L**. Visualizing Desirable Patient Healthcare Experience. ***Health Marketing Quarterly***. 27: 116-130. PMID: 20155554

2008

23. Long A, Kim E, Leung H, Woodruff E, **An L**, Doerge RW, Pak W, and Broadie K. Presynaptic calcium channel localization and calcium-dependent synaptic vesicle exocytosis regulated by the fuseless protein. ***Journal of Neuroscience***. 28:3668-3682, PMID: 18385325

24. Leung H, Tseng-Crank J, Kim E, Mahapatra C, Zhou Y, **An L**, Doerge RW, and Pak W. DAG Lipase Activity Is Necessary for TRP Channel Regulation in Drosophila Photoreceptors. ***Neuron***. 58:884-896, PMID: 18579079

25. Spach K, **An L**, Blake M, Blankenhorn E, Bunn J, Doerge RW, McElvany B, Noubade R, Tung K, and Teuscher C. Age, season and stress influence EAE sexual dimorphism in SJL/J mice. ***Journal of Neuroimmunology***. 302(2): 212

2007

26. Zhao J, Wang J, **An L**, Doerge R W, Chen ZJ, Grau CR, Meng J, and Osborn TC. Analysis of Gene Expression Profiles in Response to *Sclerotinia Sclerotiorum* in *Brassica Napus*. ***Planta**.* 227(1):13-24

BOOK CHAPTER UNDER DEVELOPMENT

An L and Jiang H. "Metagenomics" in "*No Boundary Thinking in Bioinformatics*" edited by Huang X and J Moore.

GRANTS

2015-2017 Misdiagnosis Of Speech Sound Disorders In Latino Children**, NIH, R21,** PI: Fabiano-Smith, direct cost $225,885, My role: **co-investigator**

2012-2017 Southwest Environmental Health Sciences Center**, NIH,** PI: Lau, total cost $8.3M, My role: **statistician**

2012-2016 ATD: Statistical methods for functional metagenomics in biothreat detection, National Science Foundation (**NSF)**, *Division of Mathematical Sciences*, total cost $722,979. My role: **PI.**

2010-2014 ATD: A Mixture Modeling Framework for Statistical Identification of Multiple Genomes in Metagenomics Samples. National Science Foundation (**NSF)**, *Division of Mathematical Sciences*, total cost $725,009. My role: **co-PI.**

SOFTWARE/R PACKAGES DEVELOPED IN MY LAB

RAIDA	Ratio Approach for Identifying Differentially Abundant features across different microbial conditions
ENNB:	A two-stage statistical procedure for feature selection and comparison in functional analysis of metagenomes
metaFunction:	R package for statistical profiling functions in a microbial community
FunctionSIM:	(Java software) A sequencing simulator for functional metagenomics
TAEC:	R package for Taxonomic Analysis by Elimination and Correction on closely related species
TAMER:	R package for accurate taxonomic assignment of metagenomic sequencing reads

CONFERENCE/ SYMPOSIUM PRESENTATIONS:

2015 **An L**, Sohn M.B. Du, R. A Robust Approach For Identifying Differentially Abundant Features In Metagenomic Samples. *NSF ATD workshop*, Arlington, VA, July, 2015.

An, L., Sohn, M.B. and Du, R. A Robust Approach For Identifying Differentially Abundant Features In Metagenomic Samples. *International Human Microbiome Congress*. Luxembourg, Mar. 2015

Ban, Y., **An, L**., and Jiang, H. Investigating microbial co-occurrence patterns based on metagenomics compositional data. *International Human Microbiome Congress*. Luxembourg, Mar. 2015

2014 **An L**, Sohn M, Pookhao N, Li Q. Accurate Estimation of Genome Relative Abundance for Closely Related Species in a Metagenomic Sample. *Joint Statistical meeting*, Boston, MA. Aug, 2014.

Du R, **An L**. A New Normalization Method on Metagenomic Sequencing Data. *Joint Statistical meeting, Boston*, MA. Aug, 2014.

2013 **An L**, Pookhao N, Jiang H, Xu J. Statistical Methods on Functional Analysis of Metagenomes. *ENAR meeting. The Eastern North American Region of The International Biometric Society.* Orlando, FL, Mar 2013.

An L, Pookhao N, Jiang H, Xu J. Statistical Methods on Functional Analysis of Metagenomes. *WNAR meeting. The Western North American Region of The*

International Biometric Society. Los Angeles, CA, Jun 2013

2012 **An L,** H. Jiang, N. Pookao, I. Jakins. A comparison of statistical methods for comparative metagenomic analysis. *Joint Statistical Meeting*, San Diego, CA, Aug 2012.

An L, H. Jiang. Statistical Identification of Multiple Genomes in a Metagenomic Sample. *Annual Conference of Institute Biological Engineering*. Indianapolis, IN, Mar 2012

Jiang H, **An L**. Taxonomic Assignment of Metagenomic Sequencing Reads. *Joint Statistical Meeting*, San Diego, CA, Aug 2012.

An L, Jiang H. Statistical Identification of Multiple Genomes in a Metagenomic Sample. *Conference on New Statistical Methods for Next Generation Sequencing Data Analysis*. Demoise, IW, May, 2012.

2011 Jiang H, **An L**, Lin S, Feng G. Qiu Y. Estimating relative abundance of multiple genomes in a metagenomic sample. *Statistical Methods for Very Large Datasets Conference*, Baltimore, Maryland, June 2011.

2010 Jiang H. and **An L**. Statistical issues for analysis of metagenomics sequencing data. *Applied Statistics Symposium of International Chinese Statistical Association*, Indianapolis, IN, June 2010

Schafer A, Cárdenas-Mora J, Plachot C, Stanley J, Stegelman D, Lipka A, Baumann D, Hodges K, McDole K , **An L**, Doerge RW, and Lelièvre S. CDS1: A new marker for nutritional effects on breast cancer development. *International Symposium on Breast Cancer Prevention*, West Lafayette, IN, October 2010.

2009 **An L**. Dynamic clustering of time series gene expression across multiple conditions. *Joint Statistical Meeting*, Washington DC, Aug. 2009.

An L, Peng F, and Lu Z. Effect of sampling rate and time duration of noisy time series with application in microarray experiment. *13th Annual International Conference on Research in Computational Molecular Biology*, Tucson, AZ, May 2009.

2008 Plachot J. Stanley A, **An L**, Adissu H, Doerge RW, and Lelievre S. Disruption of apical polarity in mammary epithelium is associated with the transcriptional down-regulation of genes linked to cellular homeostasis and breast cancer. *Conference of American Society for Cell Biology*, San Diego, CA. Dec. 2008.

Zhang S, **An L**, Rau A, Livermore Auer P, Doerge RW and Foster DR. Alterations in the Expression of Intestinal Transporters and Metabolic Enzymes in Cultured Intestinal Cells treated with Green Tea Extract. *American College of Clinical Pharmacy Meeting*. Louisville, KT, Oct. 2008.

An L, Doerge RW. Dynamic Clustering of Time Series Gene Expression. *Conference on Applied Statistics in Agriculture*, Manhattan, KS, May 2008.

2006 **An L**, Riddle N, Birchler J and Doerge RW. Clustering Gene Expression: Maize Ploidy Series with Repeated Measurements. *Conference on Applied Statistics in Agriculture*, Manhattan, KS, April 2006.

2005 **An L,** Wang J, Zhao J, Osborn T, Chen CJ and Doerge RW. Identification of Differentially Expressed Genes of Brassica napus and the Effect of Unbalanced Gene Replication. *Conference on Applied Statistics in Agriculture*, Manhattan, KS, May 2005.

SEMINAR/COLLOQUIUM TALKS

2014 Computational and statistical challenges in analysis of metagenomes based on next-generation sequencing data, College of Pharmacy, University of Arizona, Tucson, AZ, Nov, 2014

Computational and statistical challenges in analysis of metagenomes based on next-generation sequencing data. Colloquium of Statistics GIDP, University of Arizona, Tucson, AZ, Dec, 2014

A Two-Stage Statistical Procedure For Feature Selection And Feature Comparison In Functional Analysis of Metagenomes. Department of Statistics, Northwestern University, April 2014

A Two-Stage Statistical Procedure For Feature Selection And Feature Comparison In Functional Analysis of Metagenomes. Department of Statistics, Purdue University, April 2014

2013 Introduction to metagenomics. Seminar of Statistics on Next Generation Sequencing (SONGS), University of Arizona, Mar 2013

Statistical methods on functional metagenomics. Department of Math & Statistics, University of Missouri at Kansas City, Apr 2013

2012 Statistical and Computational Challenges for Metagenomics Analysis Based on Next-generation Sequencing Data. Quantitative Biology Colloquium, Math Department, University of Arizona, Feb 2012

2010 Statistical issues in analysis of metagenomics sequencing data. Biostatistics Seminar, College of Public Health, University of Arizona. Nov 2010

Who is there and what are they doing? Graduate seminar at Ag & Biosystems Engineering, University of Arizona. April, 2010

2008 Dynamic Clustering of Time Series Gene Expression.

- Biostatistics Seminar & Stat GIDP colloquium, University of Arizona, Tucson, AZ. Oct 2008
- Department of Mathematics and Statistics, University of Colorado Denver. Denver, CO. Mar 2008
- Department of Agricultural and Biosystems Engineering, University

of Arizona. Tucson, AZ. Feb 2008

- Department of Mathematics and Computer Science, Clark University. Worcester, MS. Feb 2008
- Department of Statistics, Purdue University, West Lafayette, IN, April 2008

2005 Introduction to microarry analysis with Bioconductor: Affymetrix data. STATCOM consulting service, Purdue University, West Lafayette, IN. Nov 2005

Differential expression analysis of Affymetrix data using Bioconductor. STATCOM consulting service, Purdue University, West Lafayette, IN. Nov 2005

JOURNAL REFEREEING

- Bioinformatics
- BMC Bioinformatics
- BMC Plant Biology
- Cancer Informatics
- Journal of Applied Biostatistics
- Journal of the American Statistical Association
- Journal of Natural Resources and Life Sciences Education
- Molecular Biology and Evolution
- Risk Analysis
- PLoS ONE
- Statistics in Medicine
- Theoretical and Applied Genetics

OTHER PROFESSIONAL ACTIVITIES

- Session Chair, “Phylogenetics and Genetics” at *Joint Statistical Meeting (JSM)*, Boston, MA, Aug 2014
- Organizer for the invited session “Methods for Sequence-based Metagenomic Analysis” *The International Biometric Society (IBS) - WNAR annual meeting*, Los Angles, CA, Jun 2013
- Organizer for the invited session “Statistical and Computational Challenges for Metagenomic Sequencing Data Analysis” at *Applied Statistics Symposium of International Chinese Statistical Association*, Baltimore, MD, Jun 2013
- Session Chair "Applications of High-Dimensional Regression" which was supported by the Biometrics Section of American Statistics Association. *Joint Statistical Meetings*, San Diego, CA, Aug, 2012
- Session Chair, "Genomics, Including Copy-Number Variants and ChIP" at *Joint Statistical Meeting (JSM)*, Vancouver, Canada, Aug 2010
- Organizer and Chair of the invited session, "Statistical challenges in next-generation

sequencing data analysis" at *Applied Statistics Symposium of International Chinese Statistical Association*, Indianapolis, IN, June 2010

HONORS/AWARDS

- ASABE Superior Paper Award, American Society of Agricultural and Biological Engineers, ASABE/CSBE, Montreal, Canada (2014).
- Alpha Epsilon, Agricultural Engineering Honor Society (2011-present).
- I.W. Burr Award, Purdue University, Department of Statistics for excellence in dissertation research, teaching, and consulting as a graduate student at Purdue (2009)
- Faculty of 1000 Biology. Article published on *Neuron* selected for Faculty of 1000 Biology which is an expert guide to the most important advances in biology (2008) http://f1000biology.com/home
- Travel grant from Conference on Applied Statistics in Agriculture, Kansas State University (2005, 2006, 2008)
- Travel grant, Purdue University, Graduate Student Government (2006)

TEACHING EXPERIENCE

Instructor for the following courses:

- STAT/MATH 517A **Advanced Statistical Regression Analysis** (Fall 2008)
- MCB/GENE/ ECOL 416A/516A **Statistical Bioinformatics and Genomic Analysis** (Spring 2010, Fall 2011, Spring 2014)
- RNR/ABE 613 **Applied Biostatistics** (Spring 2011, Fall 2012, Fall 2013, Fall 2014)
- STAT/MATH 571B **Design of Experiments** (Spring 2011, Spring 2012, Spring 2013, Spring 2014, Spring 2015)
- STAT 574S **Survey Sampling** (Fall 2010, Fall 2012, Spring 2015)

STUDENTS/POSTDOC MENTOREE

PhD students:

- Michael Sohn, PhD in Statistics, May 2015
    - Dissertation title "*Novel Computational And Statistical Approaches In Metagenomic Studies*"
    - Current position: postdoc in University of Pennsylvania
- Naruekamol Pookhao, PhD in Ag & Biosystems Engineering, May 2014
    - Dissertation title "*Statistical Methods For Functional Metagenomic Analysis Based On Next-Generation Sequencing Data*"
    - Current position: Director of The Institute of Research Development and Innovation for Industry, The Federation of Thai Industries

MS students:

- Sara Ziebell (MS in Statistics, Dec, 2015)
    - Thesis title: *“A Powerful Correlation Method for Microbial Co-occurrence Networks”*
- Ahmad Hakeem Abdul Wahab, MS in Statistics, Aug 2015
    - Thesis title "*Statistical Discovery of Biomarkers in Metagenomics*"
    - Current position: Ph. D student in Department of Statistics at Purdue University
- Isaac Jenkins, MS in Statistics, Aug 2012
    - Thesis title "*Beta Regression And Comparative Metagenomics*"
    - Current position: Biostatistician in Fred Hutchinson Cancer Research Center

Postdocs:

- Ruofei Du (2014-2015)
    - Current position: Research Scientist at University of New Mexico Cancer Center

Other students:

- Dan Luo (PhD in Biostatistics, in progress)
- Meng Lu (PhD in Statistics, in progress)
- Nick Lytal (PhD in Statistics, in progress)
- Kyle Carter (PhD in Statistics, in progress)