Yanming Di

Department of Statistics
Oregon State University
235 Weniger Hall
Corvallis, OR 97331

Phone: 541-737-1992
Fax: 541-737-3489
Email: diy@stat.oregonstate.edu
Homepage: http://www.science.oregonstate.edu/~diy/

A. EDUCATION AND EMPLOYMENT INFORMATION

A.1. Education

2004–2009	PhD, Statistics, **University of Washington**, Seattle, WA
2001–2003	MS, Statistics, **Arizona State University**, Tempe, AZ
1997–2000	MS, Mathematics, **Tsinghua University**, Beijing, China
1993–1997	BS, Applied Mathematics, **Tsinghua University**, Beijing, China

A.2. Employment

Oregon State University

Fall 2015–present	Associate Professor of Statistics
Fall 2009–Summer 2015	Assistant Professor of Statistics
Fall 2010–present	Adjunct Faculty in the Department of Public Health (Biostatistics)
Fall 2011–present	Faculty member of the multidisciplinary Molecular and Cellular Biology (MCB) Graduate Program

University of Washington, Department of Statistics

2006–2009	Research Assistant (Advisor: Prof. Elizabeth Thompson)
Summer 2008	Instructor
Fall 2005	Research Assistant (Advisor: Prof. Michael Perlman)
2004–2005	Teaching Assistant

Brown University, Department of Mathematics

2003–2004	Teaching Assistant

Arizona State University, Department of Computer Science

Summer 2003	Research Assistant (Advisor: Prof. Arunaba Sen, Prof. Huan Liu)

Arizona State University, Department of Mathematics and Statistics

2001–2003	Teaching Assistant

B. TEACHING, ADVISING AND OTHER ASSIGNMENT

B.1. Instructional Summary

Credit Courses

Term	Course	Title	Enrollment
Winter 2015	ST 422/522	Introduction to Mathematical Statistics	41
Winter 2015	ST 201	Principles of Statistics	125
Winter 2014	ST 422/522	Introduction to Mathematical Statistics	37
Fall 2013	ST 421/521	Introduction to Mathematical Statistics	54
Winter 2013	ST 661	Advanced Theory of Statistics I	6
Winter 2013	ST 562	Theory of Statistics II	33
Fall 2012	ST 421/521	Introduction to Mathematical Statistics	43
Spring 2012	ST 563	Theory of Statistics III	23
Winter 2012	ST 562	Theory of Statistics II	28
Fall 2011	ST 421/521	Introduction to Mathematical Statistics	40
Spring 2011	ST 563	Theory of Statistics III	19
Fall 2010	ST 421/521	Introduction to Mathematical Statistics	14
Spring 2010	ST 563	Theory of Statistics III	24
Fall 2009	ST 421/521	Introduction to Mathematical Statistics	24

Curriculum Development

Curriculum committee service:

Winter 2014–present	Committee for Developing Online MS Program in Data Analytics
Winter 2014–present	Committee for Developing Certificate Program in Statistics for E-campus
Winter 2014–present	Task force for coordinating curriculum in biological informatics, computational and systems biology
Fall 2011–Summer 2013	Department of Statistics Curriculum Committee

Meetings and forums attended:

November 2013	Meeting to discuss undergraduate education in Bioinformatics and Genomics and the BRR option
March 2012	Discussion forum on curriculum with MCB faculty

Graduate Students

I am currently the major professor for 2 PhD students in Statistics. I have 1 PhD student who already completed our program. I have served as major professor for 5 MS students in Statistics for whom I supervised their master's projects. As major professor, I meet weekly with my students to discuss their course work and research. I particularly emphasize interdisciplinary training. All my PhD students have received scholarships to attend the Summer Institute in Statistical Genetics.

I am serving or have served as minor professor for 5 PhD students and 8 master's students seeking a minor degree in Statistics. I provided guidance and advice on statistical methods and data analyses for their projects.

I have served as committee member on 24 additional student committees (4 PhD and 20 MS, 21 in Statistics). During their courses of study, I attend committee meetings and regularly discuss with the students on their progress.

PhD Mentoring, Major Professor (3)

Name	Major	Year	Thesis Title
Wanli Zhang	Statistics	2014–present	
Bin Zhuo	Statistics	2013–present	
Gu Mi	Statistics	2011–2014	Statistical Analysis of RNA Sequencing Count Data

MS Mentoring, Major Professor (6)

Name	Major	Year	Master's Project Topic
Songqing Ye	Statistics	2016	WGCNA, Weighted Gene Co-expression Network Analysis
Wanli Zhang	Statistics	2014	Model-based clustering with measurement errors
Qian Wu	Statistics	2013	Disease Mapping and Spatial Regression
Chaozhe Liu	Statistics	2012	Statistical Modeling of RNA-Seq Data
Nathan Palmer	Statistics	2012	A Glance at Reliability: Cronbach's Alpha
Gu Mi	Statistics	2011	Length bias correction in gene-ontology analysis using logistic regression

Minor Professor (14, for students seeking a minor degree in Statistics)

Name	Degree	Major	College	Year
Jayda Spong	PhD	Chemistry	Science	
Sarah Guermond	MS	Integrative Biology	Science	
Zhaoen Dai	MEng	Mechanical, Industrial and Manufacturing Engineering	Engineering	2015
Weifeng Huang	MS	Mechanical Engineering	Engineering	2015
Sri Dattha R Palwai	MS	Industrial Engineering	Engineering	2015
Feifei Lei	MEng	Industrial Engineering	Engineering	2015
Nithin Saravana Marthandam	MS	Industrial Engineering	Engineering	2014
Xiong Xiao	MBA	Master of Business Administration	Business	2014
Le Sheng	MEng	Mechanical Engineering	Engineering	2014
Lian Ma	PhD	Pharmaceutical Sciences	Pharmacy	2013
Nima Dolatnia	MEng	Industrial Engineering	Engineering	2013
Behrouz Behmardi	PhD	Electrical Engineering	Engineering	2012

Qian Wu	PhD	Applied Economics		Agricultural Sciences	2012
Guomei Wang	PhD	Crop and Soil Science		Agricultural Sciences	2009 −2010

Other PhD Committees, Committe Member (4)

Name	Major	College	Year
Mitra Ansari	Molecular and Cellular Biology		2013–present
Shuping Jiang	Statistics	Science	2010–2013
David Dail	Statistics	Science	2010–2012
Yan Wang	EECS	Engineering	2009

Other MS Committees, Committee Member (22)

Name	Major	College	Year
Alyssa Pedersen	Statistics	Science	2016
Peter Rise	Statistics	Science	2016
Jie Hu	Statistics	Science	2015
John Orr	Statistics	Science	2015
Le Wang	Statistics	Science	2014
Lin Xiang	Statistics	Science	2014
Nima Dolatnia	Statistics	Science	2013
Hai Yang	Statistics	Science	2013
Danlin Wang	Statistics	Science	2013
Lujing Li	Statistics	Science	2013
Elise Korejwa	Statistics	Science	2013
Jianfei Zheng	Transportation Engineering	Engineering	2012
Ziyi Jin	Statistics	Science	2012
Lili Peng	Statistics	Science	2011
Rongrong Dong	Statistics	Science	2011
Guomei Wang	Statistics	Science	2011
Douglas Savin	Statistics	Science	2011
Hong Zhu	Statistics	Science	2011
Andy Olstad	Statistics	Science	2011
Kylene Hart	Statistics	Science	2011
Evercita Eugenio	Statistics	Science	2011
Shuping Jiang	Statistics	Science	2010

B.5. Other Assignments

35% of my appointment is statistical consulting and collaboration with researchers in the College of Agriculture Sciences (CAS). I have provided formal consulting to more than 20 researchers at OSU (listed below, most are associated with the CAS), as well as many informal short-term discussions and consultations with students and faculty members.

Dates	Project and collaborators/clients
Winter 2014 –present	"Differential Expression of Elaphocordyceps ophioglossoides grown on media simulating host and non-host tissues" Joey Spatafora, PhD, Associate Professor, Botany and Plant Pathology, Agricultural Sciences C. Alisha Quandt, PhD Candidate, Botany and Plant Pathology, Agricultural Sciences
Fall 2009–present	"Statistical methods for plant pathology" Jeff Chang, PhD, Assistant Professor, Botany and Pathology
Fall 2010–present	"Grape berry synchronization: Dissection of the genetic structure of a new ripening trait in grapes" Laurent Deluc, PhD, Assistant Professor, Horticulture
Summer 2015	"Impact of umbrella bloom on strawberry production" Cassie Bouska, PhD, Agriculture Extension Faculty, Coos/Curry OSU Extension Service, Agricultural Sciences
Summer 2014	"Role of Peroxisome Proliferator-Activated Receptor gamma on prevention/cure of Mastitis" Massimo Bionaz, PhD, Assistant Professor, Animal and Rangeland Sciences, Agricultural Sciences
Fall 2013	"Transcriptome analysis in resistant, susceptible, and hypersensitive potato varieties upon PVY inoculation by RNA-Seq" Aymeric Goyer, PhD, Assistant Professor, Botany and Plant Pathology, Agricultural Sciences
Fall 2012–Summer 2013	"Circadian clocks in drosophila" Jaga Giebultowicz, PhD, Associate Professor, Zoology, Science
Fall 2012	"Differential expression in a time course experiment with Douglas-fir trees" Richard Cronn, PhD, Courtesy Assistant Professor (Botany and Plant Pathology), PNW Research Station, USDA Forest Service Brian Knaus, PhD, Faculty Research Assistant, PNW Research Station, USDA Forest Service
Fall 2012–Spring 2013	"Expression of the simA cluster in Tolypocladium" Joey Spatafora, PhD, Associate Professor, Botany and Plant Pathology, Agricultural Sciences Kathryn Bushley, PhD student, Botany and Plant Pathology, Agricultural Sciences

Fall 2011–Summer 2014	"Responses of and interactions between nitrifying bacteria to environmental changes: a systems level approach" Daniel J. Arp, PhD, Professor, Botany and Plant Pathology, Agricultural Sciences
Fall 2011–Summer 2013	"Poplar methylome variation during in vitro regeneration" Steven H. Strauss, PhD, Professor, Forest Ecosystems and Society, Forestry Kelly Vining, PhD, Faculty Research Associate, Forest Ecosystems and Society, Forestry
Winter 2011–Fall 2011	"A genome-wide analysis of gene expression in response to perennial ryegrass toxicosis" Lixin Li, PhD, Faculty Research Assistant, Environmental and Molecular Toxicology
Fall 2011	"Canine cytotrophoblasts" Michelle Kutzler, PhD, Associate Professor, Animal Sciences Timothy Hazzard, PhD, Assistant Professor Sr. Research, Animal Sciences Elizabeth Fellows, MS, Master Student, Animal Sciences
Fall 2011	"Characterization of the plastome of Asclepias syriaca" Shannon Straub, PhD, Postdoctoral Scholar, Botany and Plant Pathology
Winter 2011	"Pathogen adaptation to novel hosts through introgressive hybridization" Michelle Steinauer, PhD, Research Assistant Professor, Biomedical Sciences, Veterinary Medicine Michael Blouin, PhD, Professor, Zoology
Fall 2010–Winter 2011	"Fruit set and berry development in commercially-grown blueberry cultivars" Emily Vollmer, MS, Faculty Research Assistant, Horticulture (joint with Cliff Pereira)
Summer 2010 –Winter 2011	"Studies on infectious diseases of salmonid fish" Sascha Hallett, PhD, Research Associate, Microbiology
Spring 2010	"Correlation between herd behavior and sedative effects of medication" Dawn Marie Sherwood, PhD, Assistant Professor, Animal Science
Spring 2010	"Fish migration behaviors" Ben Clemens, PhD candidate, Student, Fisheries and Wildlife
Winter 2010	"Gene translocation and genetic maps" Guomei Wang, PhD candidate, Student, Crop and Soil Science
Fall 2009	"Estimating contamination rates in ryegrasses" Dan Curry, Director of Seed Services, Crop and Soil Science (joint with Alix Gitleman)

C. SCHOLARSHIP AND CREATIVE ACTIVITY

C.1. Publications

(Underscored authors are my PhD students.)

Peer-reviewed Publications (Statistical Methodology)

17. **Di Y** (2015): "Single-gene negative binomial regression models for RNA-Seq data with higher-order asymptotic inference". *Statistics and Its Interface,* 8, 405-418. link

I used higher-order asymptotic techniques to improve the likedllihood ratio test for NB regression models where the dispersion is unknown. This case is technically more challenging than in Di et al. 2013, where the dispersion can be treated as known.

16. <u>Mi G</u>, **Di Y** (2015): "The Level of Residual Dispersion Variation and the Power of Differential Expression Tests for RNA-Seq Data". *PLoS ONE,* 10(4). link

I supervised Mi on this study. I provided guidance on all phases of this study: from methodology development, to study design, and to manuscript preparation and revision.

15. <u>Mi G</u>, **Di Y**, Schafer DW (2015): "Goodness-of-Fit Tests and Model Diagnostics for Negative Binomial Regression of RNA Sequencing Data". *PLoS ONE,* 10(3). link

Prof. Schafer and I co-supervised Mi on this study. I provided guidance on all phases of this study: from methodology development, to study design, and to manuscript preparation and revision.

14. Jiang Y, Emerson S, Wang L, Li L, **Di Y** (2014): "Family-based association test using normal approximation to gene dropping null distribution". *BMC Proceedings,* 8(Supp 1), S18. link

I was the corresponding author and led this study. I derived the analytical mean and variance of the score test statistic in gene-dropping simulations and provided insights into the gene-dropping test by decomposing the test statistic into a linkage component and an association component.

13. **Di Y**, Schafer DW, Emerson SC, Chang JH (2013): "Higher order asymptotics for negative binomial regression inferences from RNA-sequencing data". *Stat Appl Genet Mol Biol,* 12(1), 49-70. link

I was the lead author. I used higher-order asymptotic techniques (HOA) to improve the likedllihood ratio test for NB regression models. I demonstrated that the HOA-adjusted likelihood ratio test gives accurate Type-I errors even in small-sample situations.

12. <u>Mi G</u>, **Di Y**, Emerson SC, Cumbie JS, Chang JH (2012): "Length bias correction in gene-ontology analysis using logistic regression". *PLoS ONE,* 27(10), e46128. link

I supervised Mi on this study. I provided guidance on all phases of this study: methodology development, study design, and manuscript preparation and revision. We clarfied the issue of lenght-bias: gene transcript length becomes a confounding factor in gene-ontology (GO) analysis if it correlates with both GO category membership and the power of the DE test. We proposed to use logistic regression model to adjust for lenght-bias.

11. Kimbrel JA, **Di Y**, Cumbie JS, Chang JH (2011): "RNA-Seq for Plant Pathogenic Bacteria". *Genes,* 2(4), 689-705. link

This is a review paper. I contributed to the discussion of the technical and statistical challenges in the practical application of RNA-Seq for studying bacterial transcriptomes.

10. Cumbie JS, Kimbrel JA, **Di Y**, Schafer DW, Wilhelm LJ, Fox SE, Sullivan CM, Curzon AD, Carrington JC, Mockler TC, Chang JH (2011): "GENE-counter: a computational and statistical

pipeline for assessing RNA-seq data for genome-wide expression differences". *PLoS ONE,* 6(10), e25279. link

The paper introduces GENE-counter, a complete computational pipeline for analyzing RNA-Seq data for differential expression. My R package NBPSeq is one component in the pipeline. I also contributed to the discussion of and comparison with other statistical analysis tools available at the time.

9. Thomas A, Abel HJ, **Di Y**, Faye LL, Jin J, Liu J, Wu Z, Paterson AD (2011): "The impact of linkage disequilibrium on the identification of functional variants". *Genetic Epidemiology,* 35(Supp1), S115-S119. link

This paper summarized the contributions of Group 9 of Genetic Analysis Workshop 17. I contributed to the discussion of the effect of long range linkage disequilibrium on association tests.

8. **Di Y**, Mi G, Sun L, Dong R, Zhu H, Peng L (2011): "Power of Association Tests in the Presence of Multiple Causal Variants". *BMC Proceedings,* 5(Suppl 9), S63. link

I led this study. I organized and supervised the graduate students participants, filled them in on the background on statistical genetics. I designed the study, derived the theoretical results, and led the simulations. This study highlighted the often overlooked importance of long-range and weak linkage disequilibrium in genetic association studies.

7. **Di Y**, Schafer DW, Cumbie JS, Chang JH (2011): "The NBP Negative Binomial Model for Assessing Differential Gene Expression from RNA-Seq". *Statistical Applications in Genetics and Molecular Biology,* 10(1). link

I was the lead author. I developed the statistical software, performed the simulation studies, and drafted the paper. This paper is one of the early papers that raised the awareness of the issue of overdispersion in RNA-Seq data.

6. Marchani EE, **Di Y**, Choi Y, Cheung C, Su M, Boehm F, Thompson EA, Wijsman EM (2009): "Contrasting IBD estimators, association studies, and linkage analysis using the Framingham Heart Study data". *BMC Proceedings,* 3(Suppl 7), S102. link

I conducted family-based linkage analysis in this study.

5. **Di Y**, Thompson EA (2009): "Conditional tests for localizing trait genes". *Human Heredity,* 68(2), 139-150. link

I was the lead author. Prof. Thompson is my PhD thesis advisor. I developed a more robust and reliable approach for localizing trait genes in a family-based genetic linkage analysis.

4. Basu S, **Di Y**, Thompson EA (2008): "Exact trait-model-free tests for linkage detection in pedigrees". *Annals of Human Genetics,* 72(5), 676-682. link

I conducted power-simulations under marker model misspecification for the simulation-based test.

3. **Di Y**, Perlman MD (2008): "Detecting linear sequences and subsequences". *Journal of Statistical Planning and Inference,* 138(9), 2634-2648. link

I developed the computing algorithms and conducted simulations. Professor Perlman derived the theoretical results. We developed tests for detecting regular subsequences in a sequence of event occurence times.

2. Sung YJ, **Di Y**, Fu AQ, Rothstein JH, Sieh W, Tong L, Thompson EA, Wijsman EM (2007): "Comparison of multipoint linkage analyses for quantitative traits in the CEPH data: parametric LOD scores, variance components LOD scores, and Bayes factors". *BMC Proceedings,* 1(Suppl 1), S93.

I performed Bayesian oligogenic joint segregation and linkage analyses using Loki.

1. **Di Y**, Li H, Ramineni A, Sen A (2003): "Detecting hidden information in images: a comparative study". *Proceedings of Second Workshop on Privacy Preserving Data Mining,* .

All authors contributed equally. I used a logistic regression method to detect messages hidden in images by steganography.

Peer-reviewed Publications (Biological Subject Areas)

5. Quandt CA, **Di Y**, Elser J, Jaiswal P, Spatafora JW (2016): "Differential Expression of Genes Involved in Host Recognition, Attachment, and Degradation in the Mycoparasite Tolypocladium ophioglossoides". *G3: Genes, Genomes, Genetics,* 6(3), 731-741. link
I performed the differential expression anlaysis of the RNA-Seq data and provided inputs on other statistical analyses in the paper.

4. Li L, **Di Y**, Estill C, Duringer J, Blythe L, Galen A, Livesay S, Craig AM (2015): "Bovine hepatic differential gene expression in response to perennial ryegrass staggers". *World Mycotoxin Journal,* 8 (3), 351-360. link
I assisted in the statistical analysis of microarray data and provided advices on other statistical analyses in the paper.

3. Gouthu S, O'Neil ST, **Di Y**, Ansarolia M, Megraw M, Deluc LG (2014): "A comparative study of ripening among berries of the grape cluster reveals an altered transcriptional programme and enhanced ripening rate in delayed berries". *Journal of Experimental Botany,* 65(20), 5889-5902. link
I assisted in the statistical analysis of gene expression data and provided advices on other statistical analyses in the paper.

2. Vining K, Pomraning KR, Wilhelm LJ, Ma C, Pellegrini M, **Di Y**, Mockler TC, Freitag M, Strauss SH (2013): "Methylome reorganization during *in vitro* dedifferentiation and regeneration of *Populus* trichocarpa". *BMC Plant Biology,* 13(92). link
I contributed to the statistical analysis of DNA methylation data.

1. Bushley KE, Raja R, Jaiswal P, Cumbie JS, Nonogaki M, Boyd AE, Owensby CA, Knaus BJ, Elser J, Miller D, **Di Y**, McPhail KL, Spatafora JW (2013): "The Genome of Tolypocladium inflatum: Evolution, Organization, and Expression of the Cyclosporin Biosynthetic Gene Cluster". *PLoS Genetics,* 9(6), e1003496. link
I provided guidance and software for the statistical analysis and interpretation of RNA-Seq data.

Peer-reviewed Publications (Others)

1. **Di Y** (2001): "The class number of the field Q(sqrt(q1) , ..., sqrt(qt)) (t>=4) is Even". *Acta Mathematica Sinica,* 44(1), 125-130.
In algebraic number theory, class number is an important character of number fields.

Technical Reports

1. **Di Y** (2009): "t-separation and d-separation for directed acyclic graphs". *Tech Report 552, Department of Statistics, University of Washington,* .

C.2. Presentations

August 2015	"Identification of stably expressed genes from Arabidopsis RNA-Seq data" Topic Contributed Talk, Joint Statistical Meetings 2015, Seattle, WA
June 2015	"Identification of stably expressed genes from Arabidopsis RNA-Seq data" Invited talk, Joint 24th ICSA Applied Statistics Symposium and 13th Graybill Conference, Colorado State University, Fort Collins, CO

November 2014	"Statistical Issues in Differential Expression Analysis of RNA-Sequencing Data" Invited talk, ACMS Colloquium, University of Notre Dame, Notre Dame, IN
August 2014	"Visualizing a Large Number of Regression Models Fitted to RNA-Sequencing Data" Contributed poster, Joint Statistical Meetings 2014, Boston, MA
June 2014	"Power-Robustness Analysis of Statistical Models for RNA Sequencing Data" Contributed talk, 2014 Annual Meeting of the WNAR/IMS, University of Hawaii, Honolulu, Hawaii
August 2013	"Assessing Models of RNA-Sequencing Data" Contributed poster, Joint Statistical Meetings 2013, Montreal, Quebec, Canada
June 2013	"Goodness-of-Fit Tests and Model Diagnostics for Negative Binomial Regression of RNA-Seq Data" Contributed talk, 2013 Annual Meeting of the WNAR/IMS, UCLA, Los Angeles, CA
April 2013	"Higher Order Asymptotics for Negative Binomial Regression Inferences from RNA-Sequencing Data" Mathematical Biology Seminar, Oregon State University, Corvallis, OR
October 2012	"Randomized family-based association tests in the presence of linkage" Contributed poster, 2012 IGES (International Genetic Epidemiology Society) Conference, Stevenson, WA
August 2012	"Higher Order Asymptotics for Negative Binomial Regression Inferences from RNA-Sequencing Data" Contributed talk, Joint Statistical Meetings 2012, San Diego, CA
June 2012	"Higher Order Asymptotics for Negative Binomial Regression Inferences from RNA-Sequencing Data" Contributed talk, 2012 Annual Meeting of the WNAR/IMS, Colorado State University, Fort Collins, Colorado
April 2012	"Higher Order Asymptotics for Negative Binomial Regression Inferences from RNA-Sequencing Data" Department of Statistics Research Seminar, Oregon State University, Corvallis, OR
September 2011	"Regression Analysis of RNA-Seq Data" Invited talk, 2011 CGRB Fall Conference, Oregon State University, Corvallis, OR
June 2011	"Assessing Differential Gene Expression from RNA-Seq Data" RNA-Seq Workshop, Oregon State University, Corvallis, OR
October 2010	"Impact of LD on identification of functional variants (joint presentation of Group 9)" Genetic Analysis Workshop 17, Boston, MA
June 2010	"Randomized Family-Based Association Tests Conditional on Latent Inheritance Vector" Contributed talk, 2010 Annual Meeting of the WNAR/IMS, University of Washington, Seattle, WA

June 2009	"Family-based and population-based studies are complementary" Contributed talk, 2009 Annual Meeting of the WNAR/IMS, Portland State University, Portland, OR
April 2009	"Conditional tests for localizing trait genes" Job Candidate Talk, Department of Statistics, Oregon State University, Corvallis, OR
March 2009	"Conditional tests for localizing trait genes" Job Candidate Talk, School of Public Health (Biostatistics), Yale University, New Haven, CT
June 2008	"Conditional tests for localizing trait genes" Student paper competition, 2008 Annual Meeting of the WNAR/IMS, University of California, Davis, CA

C.3. Grants

Active

Title	**"Negative Binomial Regression and Higher-Order Asymptotic Inference for RNA-Seq Gene Expression Analysis"**
PI	Yanming Di
Source of Support	NIH/NIGMS
Total Award Amount	$567,599
Dates	07/01/2012-04/30/2016
Role	PI
Summary	The major goals of this proposal are to develop and implement nearly exact inferences for negative binomial regression analysis of RNA-Seq data. As the PI, Di's roles include coordinate the entire research project, develop, assess and implement the higher-order asymptotic (HOA) inference for negative binomial regression analysis, collaborate with co-investigators on developing interactive and dynamic visualization tools, integrating the statistical tools into an existing computational pipeline, and applying the developed tools and methods to analyze data sets.

Completed

Title	**"Responses of and interactions between nitrifying bacteria to environmental changes: a systems level approach"**
PI	Daniel J. Arp
Source of Support	DOE
Total Award Amount	$653,304
Amount to Di	$5,247
Dates	09/01/2011-08/31/2014
Role	Senior Personnel

Summary	The major goal of this project is to characterize the interaction between ammonia-oxidizing and nitrite-oxidizing bacteria. Di's role is to help analyze RNA-Seq data.
Title	**"CAREER: Investigating Dynamic Changes in the Plasma Membrane Proteome During Plant Immune Signaling"**
PI	Gitta Coaker
Source of Support	NSF/MCB
Total Award Amount	$726,338
Amount to Di	$0 (I participated during the no-cost extension period.)
Dates	04/01/2012-03/31/2013
Role	Senior Personnel
Summary	The major goals of this subaward are to use RNA-Seq to investigate host transcriptome changes during microbial infection and develop a network model that integrates proteome and transcriptome changes Di's role is to find clues from RNA-Seq data.

Pending

Title	**"Characterization of the seed influence on the ripening of fleshy fruits"**
PI	Laurent Deluc
Source of Support	NSF
Total Award Amount	$2,981,910
Dates	01/01/2015-12/31/2019
Role	Co-PI
Summary	We propose to study role of seed in determining the ripening of the pericarp tissue through an auxin signal.
Title	**"Systems biology investigation of conserved mechanisms of herpesvirus latency"**
PI	Ling Jin
Source of Support	NIH
Total Award Amount	$401,500
Dates	06/01/2015-05/31/2017
Role	Co-Investigator
Summary	We propose to combine the tools of systems biology and molecular virology to elucidate the herpesvirus latency program and to investigate the function of a previously identified candidate regulator within this program.

C.5. Other

Statistical Software

2011–present   NBPSeq
An R package for assessing differential gene expression from RNA-Sequencing data. I am the main developer and maintainer of this package. My PhD students and collaborators contributed codes and data sets to this package. This package is used in many published research papers.

2009   Civil
A program implementing conditional inheritance vector tests for genetic linkage analysis. The software has been released as part of the software package MORGAN 3. I was the main developer of this program.

2007   regsubseq
An R program for detecting regularities in event occurrence times. I was the main developer of this program.

D. SERVICE

D.1. University Service

University Committees

Jan 2014–present	Task force: Coordinating curriculum in biological informatics, computational and systems biology
Sep 2013–present	The Scientific Advisory Board for the Center for Genome Research and Biocomputing (CGRB)
Feb 2014–May 2014	Search Committee for the Director of MCB Program
Fall 2012–Spring 2013	CGRB Spring Conference organizing committee
Summer 2012–Winter 2013	Faculty Search Committee, Systems Biology of Host-Microbe Interactions, CGRB
Summer 2012–Winter 2013	Faculty Search Committee, Nucleic Acid Sequence Bioinformatics, CGRB
Summer 2012–Winter 2013	Biological Information and Genomics Cluster Hire Committee, CGRB
Fall 2011–Spring 2012	Task Force on Faculty Hiring Initiative in Biological Information and Genomics
Spring 2011	Search Committee for Assistant Professor (Systems Biology, CGRB)

Departmental Committees

Winter 2014–present	Committee for online MS program in Data Analytics
Winter 2014–present	Certificate Program for Ecampus Committee
Fall 2012–present	Graduate Committee
Fall 2011–present	Exam Committee
Winter 2015	Search Committee for Professor
Winter 2014	Search Committee for Assistant Professors
Summer 2013	Chair of the Search Committee for Instructor
Fall 2011–Summer 2013	Curriculum Committee
Spring 2011	Search Committee for Assistant Professors

D.2. Service to Profession

Referee for journals:

Journal	Dates
Nature Protocols	03/2013
Bioinformatics	10/2012
European Journal of Human Genetics	01/2012
Statistical Applications in Genetics and Molecular Biology	2009, 03/2011, and 03/2011
Statistical Science	2009

Organized invited sessions for WNAR 2013, ICSA 2014, ICSA 2015 conferences.

E. AWARDS

E.1. National and International

2008 Travel award, Algebraic Statistics Workshop (MSRI, UC Berkeley)